JMIR Human Factors

Contents

Original Papers

Original Paper

Perspectives of Patients and Professionals on Information and Education After Myocardial Infarction With Insight for Mixed Reality Implementation: Cross-Sectional Interview Study

Alexander D Hilt[1*], MD; Kevin Mamaqi Kapllani[2*], MEng; Beerend P Hierck[3], MD, PhD; Anne C Kemp[2], MEng; Armagan Albayrak[2], MEng, PhD; Marijke Melles[2], MEng, PhD; Martin J Schalij[1], MD, PhD; Roderick W C Scherptong[1], MD, PhD

[1]Department of Cardiology, Leiden University Medical Center, Leiden, Netherlands
[2]Faculty of Industrial Design Engineering, Delft University of Technology, Delft, Netherlands
[3]Department of Anatomy and Embryology, Leiden University Medical Center, Leiden, Netherlands
[*]these authors contributed equally

Corresponding Author:
Roderick W C Scherptong, MD, PhD
Department of Cardiology
Leiden University Medical Center
Albinusdreef 2
Leiden, 2333 ZA
Netherlands
Phone: 31 715262020
Email: r.w.c.scherptong@lumc.nl

Abstract

Background: Patient education is crucial in the secondary prevention of cardiovascular disease. Novel technologies such as augmented reality or mixed reality expand the possibilities for providing visual support in this process. Mixed reality creates interactive digital three-dimensional (3D) projections overlaying virtual objects on the real-world environment. While augmented reality only overlays objects, mixed reality not just overlays but anchors virtual objects to the real world. However, research on this technology in the patient domain is scarce.

Objective: The aim of this study was to understand how patients perceive information provided after myocardial infarction and examine if mixed reality can be supportive in this process.

Methods: In total, 12 patients that experienced myocardial infarction and 6 health care professionals were enrolled in the study. Clinical, demographic, and qualitative data were obtained through semistructured interviews, with a main focus on patient experiences within the hospital and the knowledge they gained about their disease. These data were then used to map a susceptible timeframe to identify how mixed reality can contribute to patient information and education.

Results: Knowledge transfer after myocardial infarction was perceived by patients as too extensive, not personal, and inconsistent. Notably, knowledge on anatomy and medication was minimal and was not recognized as crucial by patients, whereas professionals stated the opposite. Patient journey analysis indicated the following four critical phases of knowledge transfer: at hospital discharge, at the first outpatient visit, during rehabilitation, and during all follow-up outpatient visits. Important patient goals were understanding the event in relation to daily life and its implications on resuming daily life. During follow-up, understanding physical limitations and coping with the condition and medication side effects in daily life emerged as the most important patient goals. The professionals' goals were to improve recovery, enhance medication adherence, and offer coping support.

Conclusions: There is a remarkable difference between patients' and professionals' goals regarding information and education after myocardial infarction. Mixed reality may be a practical tool to unite perspectives of patients and professionals on the disease in a more even manner, and thus optimize knowledge transfer after myocardial infarction. Improving medication knowledge seems to be a feasible target for mixed reality. However, further research is needed to create durable methods for education on medication through mixed reality interventions.

XSL•FO
RenderX

KEYWORDS

human factors; myocardial infarction; mixed reality; patient education; patient experience; PROM

Introduction

Coronary artery disease is a major cause of mortality in developed countries, leading to roughly 1.5 million deaths annually worldwide [1,2]. Improvements in early recognition of the disease and treatment have significantly decreased the mortality rate after myocardial infarction over the last few decades [3]. However, increased complexity in treatment and long-term care makes educating patients about their disease a challenge for health care professionals. Guiding patients through complex terminology, pathophysiological concepts, and extensive treatment options in a limited time frame is a stressful and demanding process for both health care professionals and patients [4].

Improvements have been made regarding patient information and education through extensive written information, informational videos, or digitalized "how does it look" visual models [5-8]. Attempts at improving education in patients following myocardial infarction are scarce and have mainly focused on care processes and anatomical knowledge [9-11]. With rapid development of new technologies such as virtual reality [12] or more recent mixed reality modalities [3], patient information and education approaches have also been changing [13-15]. Mixed reality creates interactive digital three-dimensional (3D) projections that are viewed through a head-mounted display such as Microsoft HoloLens.

With the introduction of this new technology, the possibilities to support daily care increase, in particular regarding improvements in anatomical knowledge. However, this adds another layer of complexity to the care process. The question therefore remains as to how to best establish the added value of implementing a new technology such as mixed reality in the educational process on a patient level.

To optimize the process of patient information and education after myocardial infarction, information should add to the sustainability of health and disease prevention [16]. The latter aspect is a particular cornerstone of myocardial infarction care [1]. Toward this end, the aim of this study was to assess how patients perceive patient information and education resources offered after myocardial infarction without the use of a mixed reality app. A secondary aim was to identify targets for mixed reality within the domain of patient information and education after myocardial infarction.

Methods

Design

This was a cross-sectional interview study. Ethical approval for the project was obtained through the local medical ethics committee of Leiden University Medical Center (protocol number P18.132).

Study Population

Twelve consecutive patients who visited the dedicated outpatient clinic for patients after myocardial infarction were asked to participate in the study. The patients were at various stages in their recovery, ranging between 1 and 12 months after the initial myocardial infarction. In addition, two cardiologists, two nurse specialists, one psychologist, and one sexologist were included in the study to obtain the professional stakeholders' point of view. Demographic data such as age, gender, occupation, and time of interviewing (1, 3, 6, or 12 months after myocardial infarction) were collected. Additionally, clinical demographics such as comorbidities (smoking, hypertension, diabetes mellitus), initial diagnosis (ST-elevation myocardial infarction [STEMI] or nonST-elevation myocardial infarction [NSTEMI]), culprit lesion of the myocardial infarction, maximum troponin levels at admission, and left ventricular ejection fraction (LVEF%) at hospital discharge were collected from the electronic medical record.

Semistructured Interviews and Questionnaires

In line with existing value-based health care literature, generic Patient Reported Outcome Measure tools were used in the current study [17]. First, we evaluated whether patients felt that the information provided during clinical care was sufficient, if they understood what medications they were taking, and the purpose of the medication. Second, we assessed the extent of knowledge the patients had about their disease and the effect on cardiac function.

We conducted semistructured interviews to assess patients' knowledge about personal myocardial infarction characteristics. A list of questions (Multimedia Appendix 1) was used to conduct the interviews. The first part of the interview included questions related to social and demographic factors. The second part of the interview consisted of questions related to myocardial infarction-specific knowledge. The last part of the interview included the Generic Short Patient Experiences Questionnaire (GS-PEQ). This questionnaire was originally developed to be used in multiple health care settings to evaluate the patient experience through standardized questions in addition to other qualitative measures such as semistructured interviews [18]. According to the aim of this study, the GS-PEQ was used to gain insight into patients' opinions about their experience during clinical care.

Since one of the core features of mixed reality is visualizing complex 3D models to interact with, it is relevant to understand if patients have a basic understanding of cardiac anatomy. Therefore, the level of knowledge about coronary artery disease was tested. Two forms were used: one that showed a representation of the coronary arteries, in which the patients could label the vessels that were occluded/obstructed in their case (Multimedia Appendix 2), and the other included two diagrams representing the simplified cardiac anatomy of the heart on which patients could label the area affected and how it is related with pump function, if applicable (Multimedia

Appendix 3). All interviews were audio-recorded and subsequently transcribed.

Semistructured Interview With Professionals

To gain insights into the process and map the professionals' perspective on information provision during the patient journey, semistructured interviews were conducted with professionals engaged in the treatment of patients with myocardial infarction. A list of questions was used to guide the interview (Multimedia Appendix 4), which were adapted according to the specific professional activities. The main focus of the interviews was to identify the materials professionals use to interact with patients, the dynamics of the consultations they conduct, and how and when they consider the need to educate patients.

Analysis

Content Analysis

Content analysis was used to structure all of the qualitative data from the interviews, which were summarized through descriptive statistics and examples of general comments. Numerical data are presented as means (SD) and categorical data are presented as proportions. GS-PEQ outcomes were used to structure the patient journey (see further description below); these outcomes were then used for the establishment of themes relevant to both professionals and patients.

Patient Journey Analysis

A patient experience journey was created via a standardized approach to analyze the patient experience within the dedicated care track of myocardial infarction treatment, with specific attention paid to knowledge transfer between professionals and patients [19]. For this purpose, the patients underwent observations during outpatient visits at our department, and were then interviewed subsequently with the researchers and were asked to fill out questionnaires consecutively.

Patient journey mapping is a frequently used method among design engineers, but is relatively new in the medical domain. This approach combines several methods to best understand the patient's experience by dividing the management of a specific condition, or process such as education, into a series of consecutive steps or events [19]. The mapping is performed using data collected from semistructured interviews, questionnaires, and observations. Combining these data, the result of the final patient journey offers a description of the dedicated care track as seen by professionals and experienced by the patient. In this study, the patient journey analysis included

descriptions of the main event (myocardial infarction), acute treatment and total duration of treatment, the environment in which treatment takes place, and interactions with professionals. Importantly, this analysis can highlight the key points of knowledge transfer, materials of interaction, patient concerns, patient goals, professional goals, and guide eventually possible mixed reality interventions throughout the patient experience when treated for myocardial infarction.

Results

Demographics of the Study Population

A total of 12 patients and 6 professionals were interviewed in this study. There were 9/12 (75%) and 3/6 (50%) men in the patient and professional group, respectively. The average age of the patients and health care professionals was 62.7 (SD 10.4) years and 43.2 (SD 9.6) years, respectively. Among the patients, there were 2/12 (17%) current smokers, and the remaining 10 (83%) had stopped smoking after myocardial infarction. Six (50%) patients suffered from hypertension and 2/12 (16%) had diabetes. The majority of patients (10/12, 83%) suffered from a STEMI, with a common culprit vessel being the left anterior descending artery (6 patients, 50%, Table 1). The average LVEF at discharge after myocardial infarction was 49.8% (SD 6.8%) and the average maximum troponin release was 8140.3 ng/L (SD 13.623).

General Experience

Six (50%) patients (all men) indicated that the information shared (written or spoken, presented in analog or digital format) was too extensive and repetitive, whereas one male patient stated that more information was needed. Overall, the patients indicated that clinicians were able to provide them with sufficient care, specifically regarding information on their diagnosis. However, 9/12 (75%; 2 women, 7 men) patients noted that they were not involved in specific decisions regarding their treatment process. Only one male patient reported that the given treatment was incorrect according to his own judgment (Table 2).

From the professionals' perspective, optimal timing for information exchange is perceived at the first visit at 1 month after myocardial infarction (6/6, 100%). All professionals (6/6, 100%) also stated that they wish to educate patients in a understandable and complete manner, although the timeframe is perceived to be too short in the outpatient setting.

Table 1. Demographic overview of the patients.

Sex	Age (years)	Profession	Interview time after MI[a]	Smoking	HT[b]	DB[c]	LVEF (%)[d]	Tmax[e] (ng/L)	Type of MI	Culprit vessel
Female	61	Administrative assistant	1 month	Stopped	Yes	Yes	39	10,553	STEMI[f]	LAD[g]
Male	62	Lawyer	1 month	Stopped	No	No	58	50,000	STEMI	LAD
Male	74	Vice principal	3 months	Stopped	No	No	58	1160	STEMI	RCA[h]
Male	52	Manager	3 months	Stopped	Yes	No	58	1504	NSTEMI[i]	RCA
Male	57	Foreman	6 months	Yes	Yes	No	48	8389	STEMI	LAD
Female	63	Nurse	6 months	Stopped	Yes	No	58	20	STEMI	RCA
Male	54	Engineer	6 months	Stopped	No	No	44	5659	STEMI	LAD
Male	56	Information technology consultant	6 months	Stopped	Yes	No	50	5308	STEMI	RCA
Male	64	Dentist	12 months	Yes	No	Yes	49	3990	STEMI	LAD
Male	78	Architect	12 months	Stopped	No	No	45	2078	NSTEMI	D1[j]
Male	82	Truck driver	12 months	Stopped	No	No	48	8406	STEMI	RCx[k]
Female	49	Housewife	12 months	Stopped	Yes	No	42	622	STEMI	LAD

[a]MI: myocardial infarction.

[b]HT: hypertension.

[c]DB: diabetes mellitus.

[d]LVEF: left ventricular function at infarction.

[e]Tmax: maximal troponin release.

[f]STEMI: ST-elevation myocardial infarction.

[g]LAD: left anterior descending artery.

[h]RCA: right coronary artery.

[i]NSTEMI: nonST-elevation myocardial infarction

[j]D1: diagonal branch.

[k]RCx: circumflex artery.

Table 2. Generic Short Patient Experiences Questionnaire (GS-PEQ) (N=12).

Question	Agree, n (%)
Did the clinician talk to you in a way that was easy to understand?	12 (100)
Do you have confidence in the clinicians' professional skill?	9 (75)
Did you get sufficient information about your diagnosis?	11 (92)
Did you perceive the treatment as adapted to your situation?	11 (92)
Were you involved in decisions regarding your treatment?	3 (25)
Did you perceive the institution's work to be well organized?	12 (100)
Did you have to wait before you were admitted for services at the institution?	12 (100)
Overall, was the help and treatment you received at the institution satisfactory?	11 (92)
Did you benefit from the care given at the institution?	11 (92)
Do you believe that you were in any way given incorrect treatment?	1 (8)

Medication Usage

Six of the 12 (50%; 2 women, 4 men) patients were unaware of the type of medication they were taking and its purpose. In addition, 10/12 (83%; 3 women, 7 men) patients considered the medication to influence their recovery in a negative manner.

From the professionals' perspective, written and hand-drawn educational information were stated as the most frequently used materials for both providing medication information and anatomical knowledge transfer (6/6, 100%), followed by video (3/6, 50%) and Microsoft PowerPoint presentations (1/6, 17%). Accurate insight on medication ("what medication do you use and why?") among patients was perceived to be poor by

professionals; 4/6 (85%) of the professionals stated that they frequently encounter this problem in the outpatient setting. The professionals equally stated a desire to educate patients on the cardioprotective function as completely as possible (6/6, 100%).

Anatomical Knowledge

Regarding anatomical knowledge, 4/12 (33%; 1 woman, 3 men) patients were aware of the culprit vessel (Figure 1, Table 3) and 4/12 (33%; 1 woman, 3 men) knew the affected site (Figure 2, Table 4). Only 2/12 (17%, both men) patients knew the area of

the heart that was affected by the culprit lesion: 10/12 (83%; 3 women, 7 men) patients had no knowledge of the relationship between the diseased (culprit) vessel and the effect on their heart. Six (50%; 1 woman, 5 men) of the patients noted that this type of information was not relevant to them. Examples of comments given by patients are shown in Textbox 1.

All professionals (6/6, 100%) stated that there should be more time available to educate patients on an anatomical understanding of myocardial infarction.

Figure 1. Representation of the coronary arteries. Patients were asked the following: "Could you please tick on the boxes which of your arteries have been affected, if any? Also, on the left illustration, draw the parts affected after the myocardial infarction."
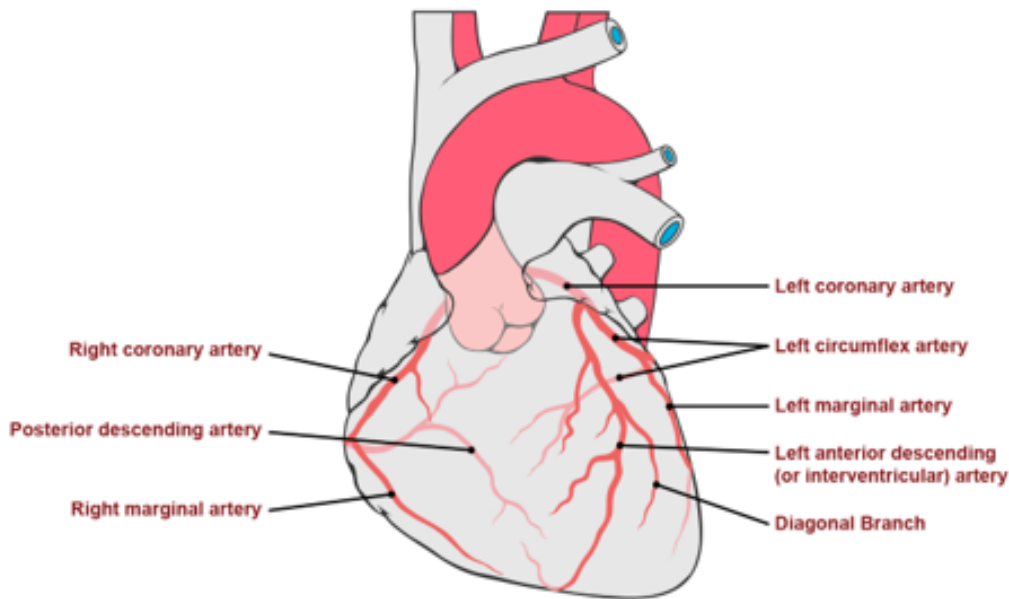
Table 3. Culprit lesion knowledge (also see Figure 1).

Patient	Culprit lesion	Correctly shown in figure?
1	LAD[a]	No
2	LAD	Yes
3	RCA[b]	No
4	RCA	No
5	LAD	No
6	RCA	No
7	LAD	Yes
8	RCA	Yes
9	LAD	No
10	D1[c]	Yes
11	RCx[d]	No
12	LAD	No

[a]LAD: left anterior descending artery.

[b]RCA: right coronary artery.

[c]D1: left anterior descending artery diagonal branch.

[d]RCx: circumflex artery.

XSL•FO
RenderX

Figure 2. Representation of heart blood circulation (left) and the main parts of the heart (right). Patients were asked: "Could you please tick on the boxes corresponding to the parts of your heart that have been affected, if any? Also, draw the affected parts on the left illustration."
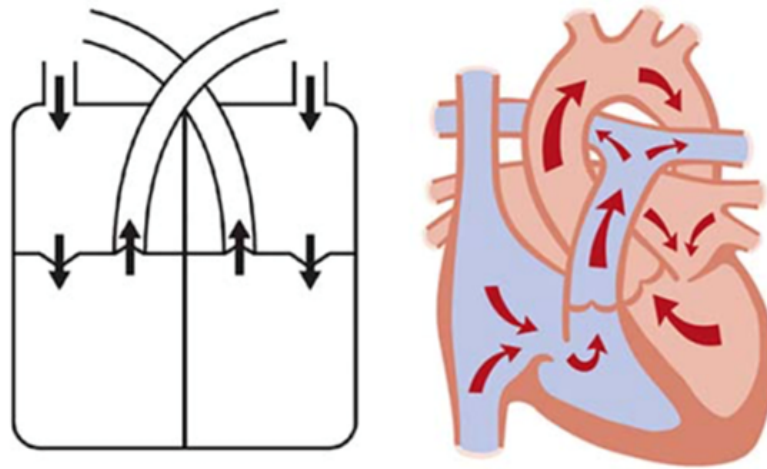

Table 4. Affected site knowledge (see Figure 2).

Patient	Correct site shown
1	No
2	No
3	No
4	No
5	Yes
6	No
7	Yes
8	Yes
9	No
10	No
11	No
12	Yes

Textbox 1. Example patient comments related to information exchange with professionals.

- Overall information exchange

"Too much information to comprehend at once"

"I really don't need to know all what they tell me"

"I really wanted to know way more than they tell me"

- Medication-related information

"I have no idea what I am taking"

"I am in charge over my body and I want to live a great life without medication"

"So many pills! That is a big problem for me, but what can I do?"

"I have different kind of colors and sizes, don't know what they do"

Patient Experience Journey: Care Track and Opportunities for Mixed Reality

Mixed Reality Information Exchange Goals

Figure 3 shows the key elements regarding knowledge transfer after myocardial infarction, and Multimedia Appendix 5 provides a full overview of the patient journey. The patient journey includes the goals of both patients and professionals at each step of the care track. Key points regarding information transfer were assessed at hospital discharge, during the first outpatient visit, and during the rehabilitation initiation. Information exchange during these phases is currently performed using drawings, the postmyocardial infarction care track information booklet, and videos (Figure 3).

Figure 3. Overview of patient experience regarding knowledge transfer and mixed reality (MR) possibilities.
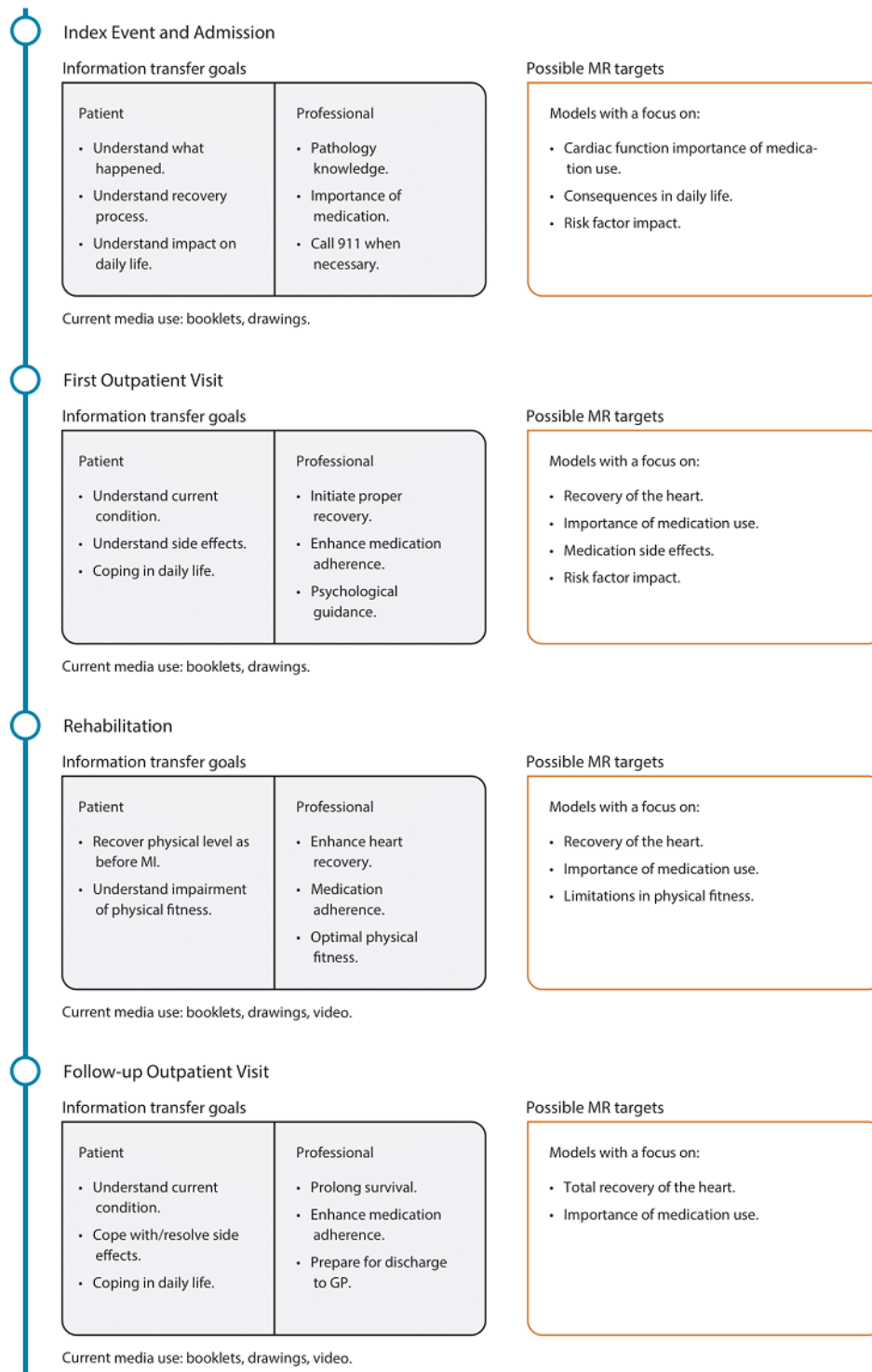
Discharge and First Outpatient Visit

Patient goals at discharge were understanding what happened, what the current condition is, and how it affects their daily life. Professionals focus on describing the event, relating it to risk factors, and stressing the importance of seeking attention when similar symptoms that may indicate a myocardial infarction are experienced.

Goals at the first outpatient visit were the same as those at discharge with the addition of understanding the side effects of medications as well as coping with the disease in daily life. Professionals focus on optimal recovery through optimal medication adherence, stressing the importance of rehabilitation and providing psychological guidance when needed. Mixed reality can help to visually support the patient's clinical state when they leave the hospital, as well as stressing the importance of medication, risk factor impacts such as smoking, and possible side effects of medication that are to be expected (Figure 3).

Rehabilitation and Outpatient Follow-Up

During rehabilitation, patient goals focus on physical fitness in terms of understanding the impairment of the disease and reaching the premyocardial infarction level of fitness. Professionals focus on increasing physical fitness through exercise and support recovery by stressing medication adherence.

During outpatient follow-up, patient goals focus on adjusting to the current health condition in daily life and understanding the potential side effects that may occur. Professionals focus mainly on prolonging survival by optimizing medication adherence and lifestyle as well to prepare patients for eventual discharge to the family physician. Mixed reality can visually support the physical condition of patients by showing the current state of heart function and its effect on physical fitness, along with the state of recovery of the heart and highlighting the long-term importance of medication on survival.

Discussion

Principal Findings

Overall, the results of the current study demonstrate that patients and clinical staff have very different opinions about the overall information shared during outpatient clinical visits, anatomical knowledge, and medication. First, patients reported that the information shared was too extensive and superfluous, whereas staff members stated a desire to share more information. Second, patients perceived medication as a hurdle toward their recovery, whereas professionals viewed the medication as an important part of their recovery. Third, the anatomical knowledge of patients was minimal regarding the culprit lesion and its effect on cardiac function. The patient journey in this regard showed that patients transition from a state of uncertainty to a state of confidence; however, the lack of knowledge remains and reassurance by health care providers is regarded as important.

Patient Information Education After Myocardial Infarction

Throughout the year following myocardial infarction, patients see roughly 4 clinical specialists and often also see a psychologist or sexologist, all of whom elaborate on the same concept of myocardial infarction. However, our outcomes suggest that patient knowledge of simple anatomical and physiological concepts of heart disease remains minimal. Furthermore, patients regard medication as a hurdle toward recovery although it is the hallmark of secondary prevention in cardiovascular care.

Scott et al [20] found that patients ranked explanation of anatomical and pathophysiological concepts as well as medication information at high importance after myocardial infarction; however, the effect of teaching these aspects to patients regarding their long-term survival is not known. It is also questionable if teaching of these concepts is essential to reach the goal of preventing new myocardial infarction, and evidence in this regard is lacking.

Our patients received identical information after myocardial infarction; however, they seem to have gained little understanding from this education, and mainly perceived the information provided as too extensive, which was not considered to be in line with their own goals. Therefore, our study highlights room for improvement in patient information education after myocardial infarction.

Professional goals (prevention of new myocardial infarction) and patient goals (living a normal life) differ to a striking degree (Figure 3). Although the necessity of teaching anatomical and pharmacological concepts might be debatable, patient care regarding information exchange should be in line with the goals of patients to support patient-centered care [17]. To unite these goals, the interaction between a patient and professional needs to be assessed and reevaluated based on the results of our study. When this information exchange is goal-oriented, learning and adoption of new information will be more effective, as stated by the cognitive load theory proposed by Sweller [21]. This theory states that the methods of information exchange should promote a low extraneous cognitive load (ie, presentation of information). Conventional methods (ie, booklets) create high levels of extraneous load, whereas visual methods create a low extraneous load [21]. Therefore, use of a mixed reality app might effectively aid in generating a low extraneous load and offer a new method of learning. This warrants further research, particularly if implementation of mixed reality for patient information education can lead to improvement of medication adherence.

Identifying Targets for Mixed Reality

As seen in the patient journey analysis, there are certain points at which mixed reality may provide solutions in patient information exchange. Certain targets might provide less information, but will nonetheless be aligned with actual patient data, including guidance on the effect of medication on their current health condition.

Mixed reality has been recently popularized by the development of Google Glass and subsequently Microsoft HoloLens, released in March 2016 [14]. HoloLens can project interactive 3D images in the field of vision of the user and recognize the environment owing to the presence of four environment-sensing cameras, a depth camera, and a light sensor. Apart from recognizing the

environment, HoloLens also memorizes it, thereby reducing the time required for the next interaction. HoloLens can also recognize human gestures to enable interaction and teamwork around the same projected objects owing to integration of human understanding software such as spatial sound, gaze tracking, gesture input, and voice support [22].

Table 5 provides an overview of the different types of media available for mixed reality and their usability, along with a summary of usability and capabilities. The main capabilities of HoloLens to be considered in the outpatient setting are: (i) recognize and interact with the environment, to choose the best environment for the interventions and base the design accordingly; (ii) project 3D images that can rotate, scale, or move; and (iii) encourage teamwork by enabling doctors and patients to collaborate through synchronization of doctor and patient images in space, giving them an opportunity to collaboratively study the model. Through these capabilities, mixed reality creates new ways of collaboration between the patient and professional. Recent studies have tested mixed reality for medical training [15] and as a surgical assistive technology

[23]. For medical students, especially those with lower visual-spatial abilities, mixed reality was shown to significantly improve 3D knowledge acquisition [24]. However, no apps currently exist that use mixed reality specifically to educate myocardial infarction patients or to improve their experience during the treatment after myocardial infarction.

Our results indicate that mixed reality may be of aid in compiling patient-specific data in one model such as a simplified model of the heart and coronary anatomy using radiographic and ultrasound data. This may be used at the end of the hospital stay when patients are fit to go home, and when uncertainties are present. A mixed reality intervention at discharge can provide a crude overview of myocardial infarction and the importance of medication and education on minimizing risk factors such as smoking. This technology can be used consecutively throughout all outpatient visits, compiling cardiac function in the model and thereby offering the possibility to use one model consecutively. Furthermore, mixed reality can be used to explain the effects of medication on long-term survival.

Table 5. Media types and usefulness in patient education.

Usability and capability	Mixed Reality (HoloLens)	Augmented Reality	Virtual Reality	Video	Text and Images
Interaction between two or more users	Full	Partial	Partial	No	No
Movement	Yes	Yes	Yes	Partial	No
Environment aware	Yes	Partial	No	No	No
Device needed	Yes	Yes (phone)	Yes (phone)	Yes (TV, computer, or phone)	No

Medication as a Specific Target for Mixed Reality

The patients included in our study perceived medication as a hurdle toward recovery. They indicated that this is mainly coupled to side effects but also that the beneficial effects are unclear (despite all information provided). Optimal medical therapy after myocardial infarction is the cornerstone of cardioprotective care and is essential in preventing new events [12]. This has been stressed by both the European and American cardiology societies [25]. However, nonadherence to medication is a common problem [26]. Through the years, attempts have been made to improve medication nonadherence; however, it remains a challenge to create sustainable interventions [27].

The patient journey analysis suggested that reassurance is important for patients to understand their condition such as whether or not they are physically fit. Clear explanation of medication benefits on their health and daily life may resolve the lack of understanding of medication effects and potentially lower the need for reassurance.

Tailoring education to patient-specific features and needs such as medication adherence seems to be effective, which has been proposed in other studies. Nieuwkerk et al [28] demonstrated that by clarifying the effect and importance of statins visually, low-density lipoprotein cholesterol levels can be reduced along with an increase in the intake of statins. A randomized study conducted by Jones et al [29] in 2015 showed that providing

visual education after myocardial infarction improved illness and medication perceptions in the intervention group. A similar approach may be feasible in patients after myocardial infarction that are offered a new form of education through mixed reality. A model could be developed, not focusing on anatomy per se but rather on statin use and the effect on the patient's cardiovascular health, such as by demonstrating atherosclerosis in coronary vessels, which is targeted by statin therapy [1]. The effect of such a mixed reality intervention could be measured according to assessing medication beliefs and illness perceptions.

Further research is needed to test our assumptions. Importantly, the implementation and evaluation of a mixed reality app in the elderly should be undertaken. Along with an aging population, potential users will be between 60 and 80 years old, which is accompanied by different forms of disabilities (ie, impaired vision, hearing, or cognitive function) that can complicate use. However, mixed reality seems to be an accessible and feasible tool in the elderly, as highlighted by Rohrbach et al [30] in patients with Alzheimer disease. Since patients with Alzheimer disease comprise a complex patient group, it is feasible to assume that patients with no cognitive impairments might also benefit from mixed reality apps.

In this era of rapidly evolving technology that brings new opportunities regarding patient information education, it is important to thoroughly evaluate how these technologies can

be used in a changing medical setting and with what goal in mind, especially given the sparsity of research on the topic.

Limitations

There are certain limitations to our study. First, all interviews were conducted in a group of patients and professionals belonging to a single hospital. Using a different group of professionals and patients from different hospitals and social backgrounds, different outcomes may be generated concerning patient information education. Second, and following this point, the small study size could have led to overestimating the assumptions such as the problems patients have with medication. Further investigation on this subject is therefore warranted. Third, observational interview studies have inherent biases (such as responder bias or social desirability bias). This can also be corrected using a larger-scale study.

Conclusion

We identified a remarkable difference between the goals of patients and health care professionals regarding information and education after myocardial infarction. Mixed reality may be a practical tool to unite the perspectives of patients and professionals on the disease in a more even manner, and thus optimize knowledge transfer after myocardial infarction. Medication understanding seems to be a feasible target for mixed reality. However, further research is needed to develop durable methods for education on medication through mixed reality.

Authors' Contributions

AH, KK, RS, AA, and MM conceived of the study. AH, KK, RS, AA, and MM developed the theory, and AH, KK, and CK performed the data collection and analysis. BH, RS, AA, MM, and MS supervised the research and critically reviewed the findings. AH and KK drafted the manuscript. All authors discussed the results and contributed to the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Patient interviews at the outpatient clinic visit (translated in English). The following questions were used to obtain information about the patient, the crucial points through the journey, and their anatomy knowledge.
[DOCX File , 14 KB - humanfactors_v7i2e17147_app1.docx]

Multimedia Appendix 2
Images used during interviews: coronary anatomy.
[DOCX File , 452 KB - humanfactors_v7i2e17147_app2.docx]

Multimedia Appendix 3
Images used during interviews: understanding heart anatomy and function.
[DOCX File , 444 KB - humanfactors_v7i2e17147_app3.docx]

Multimedia Appendix 4
Questions used in clinical staff interviews.
[DOCX File , 13 KB - humanfactors_v7i2e17147_app4.docx]

Multimedia Appendix 5
Full patient journey (added as a vector file).
[DOCX File , 620 KB - humanfactors_v7i2e17147_app5.docx]

References

1.   Roffi M, Patrono C, Collet J, Mueller C, Valgimigli M, Andreotti F, ESC Scientific Document Group. 2015 ESC Guidelines for the management of acute coronary syndromes in patients presenting without persistent ST-segment elevation: Task Force for the Management of Acute Coronary Syndromes in Patients Presenting without Persistent ST-Segment Elevation of the European Society of Cardiology (ESC). Eur Heart J 2016 Jan 14;37(3):267-315. [doi: 10.1093/eurheartj/ehv320] [Medline: 26320110]

2. Amsterdam E, Wenger N, Brindis R, Casey D, Ganiats T, Holmes D, et al. 2014 AHA/ACC Guideline for the Management of Patients with Non-ST-Elevation Acute Coronary Syndromes: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol 2014 Dec 23;64(24):e139-e228 [FREE Full text] [doi: 10.1016/j.jacc.2014.09.017] [Medline: 25260718]

3. Nabel EG, Braunwald E. A tale of coronary artery disease and myocardial infarction. N Engl J Med 2012 Jan 05;366(1):54-63. [doi: 10.1056/NEJMra1112570] [Medline: 22216842]

4. Westermann GM, Verheij F, Winkens B, Verhulst FC, Van Oort FV. Structured shared decision-making using dialogue and visualization: a randomized controlled trial. Patient Educ Couns 2013 Jan;90(1):74-81. [doi: 10.1016/j.pec.2012.09.014] [Medline: 23107362]

5. Kaphingst KA, Persky S, McCall C, Lachance C, Loewenstein J, Beall AC, et al. Testing the effects of educational strategies on comprehension of a genomic concept using virtual reality technology. Patient Educ Couns 2009 Nov;77(2):224-230 [FREE Full text] [doi: 10.1016/j.pec.2009.03.029] [Medline: 19409749]

6. Becker A, Herzberg D, Marsden N, Thomanek S, Jung H, Leonhardt C. A new computer-based counselling system for the promotion of physical activity in patients with chronic diseases--results from a pilot study. Patient Educ Couns 2011 May;83(2):195-202. [doi: 10.1016/j.pec.2010.05.024] [Medline: 20573467]

7. Harless WG, Zier MA, Harless MG, Duncan RC, Braun MA, Willey S, et al. Evaluation of a virtual dialogue method for breast cancer patient education. Patient Educ Couns 2009 Aug;76(2):189-195. [doi: 10.1016/j.pec.2009.02.006] [Medline: 19321289]

8. Zissiadis Y, Harper E, Kearney E. Impact of more intensive written information in patients having radical radiation therapy: results of a prospective randomized phase III trial. Radiother Oncol 2010 Aug;96(2):254-258. [doi: 10.1016/j.radonc.2010.06.002] [Medline: 20598391]

9. Kuo MC, Chang P. Developing a visualized patient-centered, flow-based and objective-oriented care path of cardiac catheterization examination. Stud Health Technol Inform 2009;146:879-880. [Medline: 19593031]

10. Otto CM. Communicating with our patients for shared decision making. Heart 2018 Mar 27;104(6):451-453. [doi: 10.1136/heartjnl-2018-313114] [Medline: 29487173]

11. Kaufman BG, Kim S, Pieper K, Allen LA, Gersh BJ, Naccarelli GV, et al. Disease understanding in patients newly diagnosed with atrial fibrillation. Heart 2018 Mar 08;104(6):494-501 [FREE Full text] [doi: 10.1136/heartjnl-2017-311800] [Medline: 28790169]

12. Task Force Members, Montalescot G, Sechtem U, Achenbach S, Andreotti F, Arden C, ESC Committee for Practice Guidelines, Document Reviewers, et al. 2013 ESC guidelines on the management of stable coronary artery disease: the Task Force on the management of stable coronary artery disease of the European Society of Cardiology. Eur Heart J 2013 Oct;34(38):2949-3003. [doi: 10.1093/eurheartj/eht296] [Medline: 23996286]

13. Bork F. [Interactive augmented reality systems : Aid for personalized patient education and rehabilitation]. Unfallchirurg 2018 Apr;121(4):286-292. [doi: 10.1007/s00113-018-0458-y] [Medline: 29383388]

14. Biglino G, Capelli C, Wray J, Schievano S, Leaver L, Khambadkone S, et al. 3D-manufactured patient-specific models of congenital heart defects for communication in clinical practice: feasibility and acceptability. BMJ Open 2015 Apr 30;5(4):e007165-e007165 [FREE Full text] [doi: 10.1136/bmjopen-2014-007165] [Medline: 25933810]

15. Wang S, Parsons M, Stone-McLean J, Rogers P, Boyd S, Hoover K, et al. Augmented Reality as a Telemedicine Platform for Remote Procedural Training. Sensors (Basel) 2017 Oct 10;17(10):2294 [FREE Full text] [doi: 10.3390/s17102294] [Medline: 28994720]

16. Porter ME. What is value in health care? N Engl J Med 2010 Dec 23;363(26):2477-2481. [doi: 10.1056/NEJMp1011024] [Medline: 21142528]

17. Fitzpatrick R, Davey C, Buxton M, Jones D. Evaluating patient-based outcome measures for use in clinical trials. Health Technol Assess 1998;2(14):1-74 [FREE Full text] [Medline: 9812244]

18. Sjetne IS, Bjertnaes OA, Olsen RV, Iversen HH, Bukholm G. The Generic Short Patient Experiences Questionnaire (GS-PEQ): identification of core items from a survey in Norway. BMC Health Serv Res 2011 Apr 21;11(1):88 [FREE Full text] [doi: 10.1186/1472-6963-11-88] [Medline: 21510871]

19. Trebble TM, Hansi N, Hydes T, Smith MA, Baker M. Process mapping the patient journey: an introduction. BMJ 2010 Aug 13;341(aug13 1):c4078-c4078. [doi: 10.1136/bmj.c4078] [Medline: 20709715]

20. Scott J, Thompson DR. Assessing the information needs of post-myocardial infarction patients: a systematic review. Patient Educ Couns 2003 Jun;50(2):167-177. [doi: 10.1016/s0738-3991(02)00126-x] [Medline: 12781932]

21. Sweller J. Cognitive Load During Problem Solving: Effects on Learning. Cognitive Sci 1988;12(2):257-285. [doi: 10.1207/s15516709cog1202_4]

22. Chen H, Lee A, Swift M, Tang J. 3D Collaboration Method over HoloLens™ and Skype™ End Points. In: Proceedings of the 3rd International Workshop on Immersive Media Experiences.: Association for Computing Machinery; 2015 Oct Presented at: ImmersiveME '15; October 2015; Brisbane, Australia p. 27-30. [doi: 10.1145/2814347.2814350]

23. Sauer IM, Queisner M, Tang P, Moosburner S, Hoepfner O, Horner R, et al. Mixed Reality in Visceral Surgery: Development of a Suitable Workflow and Evaluation of Intraoperative Use-cases. Ann Surg 2017 Nov;266(5):706-712. [doi: 10.1097/SLA.0000000000002448] [Medline: 28767561]

24. Bogomolova K, van der Ham IJM, Dankbaar ME, van den Broek WW, Hovius SE, van der Hage JA, et al. The Effect of Stereoscopic Augmented Reality Visualization on Learning Anatomy and the Modifying Effect of Visual-Spatial Abilities: A Double-Center Randomized Controlled Trial. Anat Sci Educ 2019 Dec 30. [doi: 10.1002/ase.1941] [Medline: 31887792]

25. Fihn S, Gardin J, Abrams J, Berra K, Blankenship J, Dallas A, American College of Cardiology Foundation. 2012 ACCF/AHA/ACP/AATS/PCNA/SCAI/STS guideline for the diagnosis and management of patients with stable ischemic heart disease: executive summary: a report of the American College of Cardiology Foundation/American Heart Association task force on practice guidelines, and the American College of Physicians, American Association for Thoracic Surgery, Preventive Cardiovascular Nurses Association, Society for Cardiovascular Angiography and Interventions, and Society of Thoracic Surgeons. Circulation 2012 Dec 18;126(25):3097-3137. [doi: 10.1161/CIR.0b013e3182776f83] [Medline: 23166210]

26. Naderi SH, Bestwick JP, Wald DS. Adherence to drugs that prevent cardiovascular disease: meta-analysis on 376,162 patients. Am J Med 2012 Sep;125(9):882-887. [doi: 10.1016/j.amjmed.2011.12.013] [Medline: 22748400]

27. Briesacher BA, Andrade SE, Fouayzi H, Chan KA. Comparison of drug adherence rates among patients with seven different medical conditions. Pharmacotherapy 2008 Apr;28(4):437-443 [FREE Full text] [doi: 10.1592/phco.28.4.437] [Medline: 18363527]

28. Nieuwkerk PT, Nierman MC, Vissers MN, Locadia M, Greggers-Peusch P, Knape LP, et al. Intervention to improve adherence to lipid-lowering medication and lipid-levels in patients with an increased cardiovascular risk. Am J Cardiol 2012 Sep 01;110(5):666-672. [doi: 10.1016/j.amjcard.2012.04.045] [Medline: 22621795]

29. Jones ASK, Ellis CJ, Nash M, Stanfield B, Broadbent E. Using Animation to Improve Recovery from Acute Coronary Syndrome: A Randomized Trial. Ann Behav Med 2016 Feb 23;50(1):108-118. [doi: 10.1007/s12160-015-9736-x] [Medline: 26497696]

30. Rohrbach N, Gulde P, Armstrong AR, Hartig L, Abdelrazeq A, Schröder S, et al. An augmented reality approach for ADL support in Alzheimer's disease: a crossover trial. J Neuroeng Rehabil 2019 Jun 03;16(1):66 [FREE Full text] [doi: 10.1186/s12984-019-0530-z] [Medline: 31159816]

Abbreviations

3D: three-dimensional
GS-PEQ: Generic Short Patient Experiences Questionnaire
LVEF%: left ventricular ejection fraction
NSTEMI: nonST-elevation myocardial infarction
STEMI: ST-elevation myocardial infarction

Original Paper

Development and Usability Testing of a Web-Based and Therapist-Assisted Coping Skills Program for Managing Psychosocial Problems in Individuals With Hand and Upper Limb Injuries: Mixed Methods Study

Folarin Omoniyi Babatunde[1], MSc; Joy MacDermid[1,2,3], PhD; Ruby Grewal[2,4], MD, MSc; Luciana Macedo[1], PhD; Mike Szekeres[2], PhD

[1]School of Rehabilitation Science, Institute of Applied Health Science, McMaster University, Hamilton, ON, Canada

[2]Roth McFarlane Hand and Upper Limb Centre, St. Joseph's Health Care London, Western University, London, ON, Canada

[3]Department of Physical Therapy, Western University, London, ON, Canada

[4]Schulich School of Medicine and Dentistry, Western University, London, ON, Canada

Corresponding Author:
Folarin Omoniyi Babatunde, MSc
School of Rehabilitation Science
Institute of Applied Health Science
McMaster University
1400 Main Street West, Room 403
Hamilton, ON, L8S 1C7
Canada
Phone: 1 9055259140 ext 22867
Email: babatufo@mcmaster.ca

Abstract

Background: Ineffective coping has been linked to prolonged pain, distress, anxiety, and depression after a hand and upper limb injury. Evidence shows that interventions based on cognitive behavioral therapy (CBT) may be effective in improving treatment outcomes, but traditional psychological interventions are resource intensive and unrealistic in busy hand therapy practices. Developing web-based, evidence-based psychological interventions specifically for hand therapy may be feasible in clinical practice and at home with reduced training and travel costs. Hand Therapy Online Coping Skills (HOCOS) is a program developed to supplement traditional hand therapy with therapist-assisted coping skills training based on principles from CBT and the Technology Acceptance Model.

Objective: This study aimed to describe the development and assess the usability of HOCOS to support hand therapists in the management of psychosocial problems.

Methods: The ADDIE model (Analysis, Design, Development, Implementation, and Evaluation) of system design was applied to create HOCOS. The usability testing of HOCOS involved a 2-stage process. In the first step, heuristic testing with information and communications technology (ICT) experts was completed using two sets of heuristics: Monkman heuristics and the Health Literacy Online (HLO) checklist. The second step involved user testing with hand therapists performing a series of online and face-to-face activities, completing 12 tasks on the website using the think-aloud protocol, completing the system usability scale (SUS) questionnaire, and a semistructured feedback interview in 2 iterative cycles. Descriptive statistics and content analyses were used to organize the data.

Results: In total, 4 ICT experts and 12 therapists completed usability testing. The heuristic evaluation revealed 15 of 35 violations on the HLO checklist and 5 of 11 violations on the Monkman heuristics. Initially, hand therapists found 5 tasks to be difficult but were able to complete all 12 tasks after the second cycle of testing. The cognitive interview findings were organized into 6 themes: task performance, navigation, design esthetics, content, functionality and features, and desire for future use. Usability issues identified were addressed in two iterative cycles. There was good agreement on all items of the SUS. Overall, therapists found that HOCOS was a detailed and helpful learning resource for therapists and patients.

Conclusions: We describe the development and usability testing of HOCOS; a new web-based psychosocial intervention for individuals with a hand and upper limb injuries. HOCOS targets psychosocial problems linked to prolonged pain and disability

XSL•FO
RenderX

by increasing access to therapist-guided coping skills training. We actively involved target users in the development and usability evaluation of the website. The final website was modified to meet the needs and preferences of the participants.

KEYWORDS

usability testing; upper extremities; psychosocial; internet; coping skills

Introduction

Background

Hand and upper limb injuries are some of the most common injuries in orthopedic settings [1,2], and approximately 11% to 20% of emergency department visits are because of hand and upper limb injuries [3,4]. In addition to pathophysiology, psychosocial factors can predict disability in individuals with hand and upper limb injuries [5,6]. These injuries have been shown to impact employment, body image [7], relationships [8], and functional abilities [9-11] negatively.

Most studies conducted in hand therapy have focused on maximizing physical recovery and adjustments with regard to medical or occupational therapy procedures [12-15]. Interventions such as joint protection [16], exercise therapy [17], mobilization [18], and modalities [19] in hand therapy have well-established benefits for pain and function. However, they do not directly target psychosocial factors that contribute to patient morbidity [20]. Several studies have established the mediating effect of psychological distress on hand and upper limb pain and disability [21-24] based on the far-reaching impact of psychosocial problems on pain and disability, and patient expectations after hand and upper limb injuries, a greater understanding of how to facilitate psychosocial adjustments is warranted [2,25]. Psychological interventions such as cognitive behavioral therapy (CBT) interventions have been shown to yield long-term [26] improvements in pain, daily function, quality of life, and overall mental health compared with active treatments alone for several musculoskeletal (MSK) problems [27,28] including knee pain [29], low back pain, [30,31] fibromyalgia, [32] preoperative spine [33] and post total joint surgery [34]. CBT is also cost-effective [35] and cost neutral when considering the overall health care sector and labor market perspective [33], with reduced health care utilization at the 5-year follow-up [26]. CBT has also been shown to be effective in improving adherence to exercise [36]. CBT techniques such as graded activity can be integrated into traditional physiotherapy [37,38].

In hand therapy, CBT may be efficient treatment to improve pain and distress by increasing adjustment to hand injury in relation to illness perception and coping strategies [2,39-41]. Unfortunately, traditional CBT is resource intensive and not feasible to implement in busy hand therapy practices because of prolonged face-to-face encounters and associated cost implications [42]. Web-based CBT is a potential emerging tool with modern interactive and communicative technologies for use in rural and urban areas, across languages and cultures, and on a global scale [43]. Web-based CBT has been shown to be effective for reducing catastrophization and improving the attitudes of patients with MSK conditions to exercise therapy

[44]. Current evidence supports the feasible and efficacious delivery of web-based CBT using nontraditional health professionals such as physiotherapists (PTs) and occupational therapists (OTs) [45], with reduced time commitment and treatment costs, and positive self-reported changes in the PTs' attitudes, confidence, and practice [46-48]. Therapist competence and therapeutic alliance are crucial factors influencing CBT [49]. Therapist competency can be developed online [50], and therapeutic alliance required for CBT to be effective does not diminish with the web-based delivery of CBT [51].

Hagemen et al [52] reported that almost 50% of outpatients presenting to hand surgery clinics investigated their symptoms online, which increases the potential to deliver evidence-based pain management and coping skills for HULI online. Further studies on the use of psychosocial interventions in HULI have the potential to convince payers to fund psychotherapy treatments, generate enthusiasm to include psychosocial treatments in educational curriculums, and advance incorporation of evidence-based psychosocial treatments in hand therapy recommendations for psychosocial problems [53]. In view of the evidence showing evidence-based CBT can be delivered via the internet and feasible to implement during wait times for hand therapy or in home-settings and reduce the costs associated with training providers and fewer hospital visits. From the foregoing, online evidence-based CBT is feasible to implement during wait times for hand therapy, is easy to use in home settings, and reduces costs associated with training providers and leads to fewer hospital visits. To meet the needs of patients with hand and upper limb injuries at risk of prolonged pain and disability because of psychosocial treatments, we decided to develop an intervention that incorporates evidence from CBT in orthopedic practice.

Hand Therapy Online Coping Skills Program

Hand Therapy Online Coping Skills (HOCOS) is an evidence-based and theory-based psychosocial coping skills program based on CBT principles [37] and the Technology Acceptance Model [54]. HOCOS was developed by Folarin Babatunde (PT) during his doctoral studies at McMaster University In collaboration with a team of PTs (JM and LM), OT (MS), hand therapists (JM and MS) and an orthopedic surgeon (RG). HOCOS involves five 'hand therapist-guided' modules. It is a multi-component, interactive online-based program consisting of hand and upper-limb specific information covering pain education and training in coping skills (activity-rest cycling, pleasant activity scheduling, problem-solving, identifying and challenging negative thoughts, relaxation response and their applications) to daily life for adults with hand and upper limb injuries (Table 1). Asynchronous learning was facilitated using PowerPoint (Microsoft Office)

presentations, audio files, workbooks, and downloadable PDF (Adobe) and Word (Microsoft Office) files. Links to evidence-based external educational resources were included to reinforce learning. The program was designed to supplement traditional hand therapy with coping skills training. The design and development of HOCOS were guided by the 5 steps of the Analysis, Design, Development, Implementation, and Evaluation from the ADDIE model [55,56]. The structure and specific

ingredients of HOCOS were based on the following recommendations suggested by Bennell et al [57]: (1) impact of psychosocial factors on pain and disability in hand injuries, (2) evidence base for the effects of CBT on MSK conditions, (3) the importance of incorporating the management of hand injuries into a biopsychosocial framework, and (4) practical issues related to the delivery of the intervention.

Table 1. Outline of hand therapy online coping skills session contents.

Session	Projected duration	Outline of content
1	1 week	• Logging in and account set up using provided password • Completing battery of questionnaires (demographic information, self-report of hand pain and function, psychosocial factors, and assessment) • Information provided about the psychosocial aspects of prolonged pain • Introduction to the module contents • Introduction to SMART (Specific, Measurable, Achievable, Realistic or Relevant and Timed) goals and using the program calendar to plan activities • Providing information on contacts for technical difficulties
2	2 weeks	• Module 1: introduction to pain (meaning, definition, and impact on recovery) • This module teaches concepts from therapeutic neuroscience education using stories and metaphors • Promote interest in exercise and physical activity
3	2-4 weeks	• Module 2: introduction to the cognitive model • Encourages users to identify and rate their moods • Encourages users to reflect on their thinking style and identify patterns • Encourages users to follow the guidelines for completing a thought record
4	2 weeks	• Module 3: introduction to activity management principles • Encourages users to pace activities to avoid boom and bust situations • Encourages users to review their day planner and spot patterns of overactivity and underactivity • Encourages users to focus on activities that have high mastery and pleasure value
5	2 weeks	• Module 4: introduction to problems limiting recovery and how to solve them • Encourages users to consider barriers to coping exercises and reflect on how to overcome those barriers • Discussion on challenges to exercise and physical activity adherence and steps to regain control
6	2 weeks	• Module 5: introduction to stress management, relaxation response, and sleep training • Encourages users to reflect on the cycle of stress, muscle tension, and pain • Encourages users to practice and adopt 1 or 2 relaxation techniques to their day plan • Encourages users to include downtime in their daily plan
7	Posttraining (single or multiple modules)	• This session focuses on how to continue recommended activities after completing the program • Users are encouraged to continue to access the resources on the website if necessary • Clinicians are provided with follow-up strategies to ensure patient success • Users complete a feedback form on their experience and a battery of questionnaires to measure their progress

Objectives

This paper aimed to provide a brief overview of the web-based system and to report on its usability from the perspectives of information and communications technology (ICT) experts and clinicians practicing in the field of hand therapy. Usability testing is a critical step in the development of online interventions and involves obtaining feedback to understand what is positive or negative about a system and identify existing gaps in content or functionality using iterative cycles of prototype alteration [58].

Methods

Design and Procedure

A mixed methods usability testing approach with semistructured interviews, observations, and questionnaires was undertaken, with iterative cycles to determine the usability of HOCOS and to further refine the prototype [59,60]. Participants were recruited using snowball sampling by asking key informants to suggest another participant who they believe is suitable for the study and introducing that person to the researcher [61]. This paper reported on step 3 of the ADDIE process (Figure 1).

ADDIE is commonly used in instructional development as a systematic way to achieve the desired results [62].

Figure 1. Diagram depicting the first 3 steps of Assessment or analysis, Design, Development, of the ADDIE Model. HOCOS: Hand Therapy Online Coping Skills; ICT: information and communications technology.
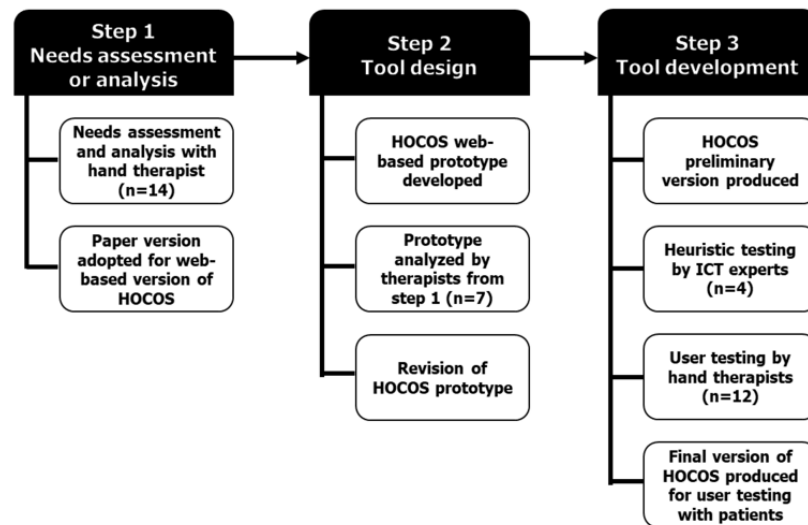
Participants

We recruited ICT experts online through the *Weebly support* (Weebly) portal to participate as heuristic evaluators in phase 1 of usability testing. In phase 2, PTs and OTs based in Ontario, Canada, and practicing in the clinical area of hand therapy were invited to participate in the study to enhance the development of HOCOS. Clinicians were messaged directly using contact details available to the public on the Canadian Society of Hand Therapy (CSHT) website. Interested participants contacted the research team directly by telephone or email and were provided with a letter of information and signed consent forms before data collection. Log-in access to the password-protected HOCOS website was provided free of charge.

Procedure

One of the researchers, FB, facilitated data collection by conducting interviews, taking notes, and observing participants' behavior. Appointments were made to meet with participants at the study site or at a desired destination within 2 hours of the study site. A brief description of the study was provided to each participant, with emphasis that the evaluation was about the content and functionality of the website. An explanation of cognitive interviews and information about privacy, and protection of the data collected were also provided. Before the interviews, demographic data, including age, gender, educational level, practice area, and use of technology, were collected. All participants were identified by pseudonyms to ensure anonymity [63]. According to Nielsen [56], 5 users are adequate to identify most usability problems. Current evidence shows that 80% of usability problems can be identified with 4 to 9 participants and 95% with 9 participants [64], thus we proposed a convenience sample of 12 participants for usability testing and to account for attrition. The usability testing protocol was approved by the Western University Health Sciences Research Ethics Board (no. 108064). Guided by steps 1 and 2 (Figure 1), we revised the prototype and developed a preliminary version of HOCOS that was tested by ICT experts (n=4) and therapists (n=12). No incentive was provided to participate in this study. Parking costs were covered for participants involved in face-to-face cognitive interviews.

Phase 1: Heuristic Testing

Heuristic testing is a usability inspection method completed by usability experts and involves evaluating an application to find usability problems, assigning them to a specific category of heuristics and ascribing a severity rating [58]. ICT experts were given a brief introduction to the background and rationale of the web portal under review and given instructions on how to conduct the heuristic testing. Between March and May 2018, the evaluators each separately conducted a heuristic evaluation of HOCOS through a page-by-page review of the website and noted *violations* where the interface did not conform to two sets of heuristics of predetermined criteria: the Monkman heuristics [65] and Health Literacy Online (HLO) checklist [66]. HLO was designed for the creation of usable online health content and comprises 35 separate criteria, categorized into 5 domains: write actionable content, display content clearly on the page, organize content and simplify navigation, engage users, and testing site with users with limited literacy skills [66]. The fifth domain was not factored in this study because this study focused on system design and development rather than implementation in practice. Monkman heuristics [65] comprises 11 checklist items and was designed for experienced heuristic evaluators by summarizing design guidelines from the HLO guide and incorporating research from electronic health (eHealth)/health literacy and usability literature [67]. The evaluators conferred using Skype clx and aggregated their results only after completing individual reviews. This phase resulted in the construction of a list of usability violations that were used to inform design changes before user testing commenced.

Phase 2: User Testing

User testing involved asking each participant to go through the website using the *thinking aloud* method [68], followed by a semistructured interview to elicit further feedback about user

interaction. Each session was completed during a 1.5- to 2-hour face-to-face visit between September 2018 and March 2019. This enabled the researcher to capture the ongoing thought processes of the participants while going over the program and any difficulties encountered [69]. First, participants were required to log on to the website, read an introductory script, and familiarize themselves with the online learning environment using hyperlinks to move between pages. Next, the participants completed the following tasks: (1) logging in, (2) reading the introductory page, (3) completing a set of psychosocial outcome measures, (4) listening to an audio recording, (5) reading a PowerPoint presentation, (6) downloading a PDF or Word document script, (7) completing one activity in the workbook, (8) setting up an activity plan for homework, (9) finding contents by browsing, (10) finding contents by searching, (11) completing a feedback form, and (12) contacting the web manager. These tasks tested the user's ability to follow the session plan and the amount of assistance required to use the online electronic tools. The facilitator did not offer any help during the tasks unless explicitly requested by the participants [68,70]. The facilitator encouraged the participants to talk about what they felt, saw, or thought while browsing through the website during the cognitive interviews. Verbal probes were also used to clarify the participant's answers [70].

The facilitator also asked the participants to explain or demonstrate the information in the video related to the module that was reviewed, such as metal practice and breathing exercises in module 1 and module 5, respectively. The participant's ability to follow the instructions correctly was observed, and any difficulties, doubts, and reports were documented using a 3-point scale (1=correctly demonstrated, 2=assistance required from an evaluator or replaying the video, and 3=difficulty demonstrating the activity correctly after being assisted) [56]. At the participant's request, whole or specific areas of content were revisited. On the basis of the benchmark by Rubin and Chisnell [71], a task was classified as a usability problem requiring attention to remedy if more than 70% of participants were unable to complete the task. The system usability scale (SUS) questionnaire [72,73] was used to evaluate satisfaction. SUS comprises 10 open-ended, polarity balance–based questions with a 5-point Likert scale for responses. The average scores were categorized based on a descriptor rating scale [74]. Finally, the facilitator interviewed each participant using a semistructured interview guide (Multimedia Appendix 1) adopted from the study by Stinson et al [75] to obtain feedback about navigation, content, and layout at the end of the second cycle of user testing.

Data Analysis

All interviews were audiotaped and transcribed verbatim in an anonymized format. The usability testing and interview data were analyzed together using triangulation [76]. Content analysis [77] of transcripts from the thinking aloud sessions, field notes, and feedback interviews was coded using predetermined codes related to usability issues (navigation, content, layout, learnability, errors, and satisfaction) after each iterative cycle. The interviews from the first cycle were analyzed and used to make minor modifications to the website before evaluation in the second cycle of testing. Very few modifications to the prototype were required after the second cycle of testing. To calculate the SUS score, the score contributions from each item are summed. Each item's score contribution ranges from 0 to 4. For items 1, 3, 5, 7, and 9, the score contribution is the scale position minus 1. For items 2, 4, 6, 8, and 10, the contribution is 5 minus the scale position. Quantitative data from SUS (10 questions, each scored from 0 to 4 points) were transformed by multiplying by 2.5 to convert scores to a 0 to 100 range and categorized using adjective ratings [74]. The descriptive analysis (means and SD) of the quantitative data was conducted using Stata 13 software for Microsoft Office.

Results

Participants' Characteristics

We enrolled four ICT experts to act as evaluators during heuristics evaluation, which meets the optimal requirement for detecting all usability problems [78]. During user testing, 26 clinicians agreed to participate in this study (14 for needs assessment, and 12 for usability testing). A total of 69% (18/26) of participants were female (Table 2). Most participants (17/26, 65%) had a background in occupational therapy, had at least 16 years of experience in hand therapy (10/26, 38%), and practiced in outpatient rehabilitation facility (10/26, 38%). Most therapists were very comfortable using a computer/tablet or internet. See Table 2 for participants' characteristics.

Table 2. Demographic and computer and internet use characteristics of therapists participating in needs assessment/analysis and usability testing of the study (N=26).

Demographics	Needs assessment (n=14)	Usability testing (n=12)
Age (years), n (%)		
21-30	3 (21)	2 (17)
30-40	5 (35)	2 (17)
40-50	2 (14)	3 (25)
>50	4 (29)	5 (42)
Gender, n (%)		
Male	6 (43)	2 (17)
Female	8 (57)	10 (83)
Prefer not to say	0 (0)	0 (0)
Profession, n (%)		
Occupational therapists	10 (71)	7 (58)
Physiotherapists	4 (29)	5 (42)
Education, n (%)		
Entry level (baccalaureate degree)	4 (29)	3 (25)
Entry level (master degree)	8 (57)	8 (57)
PhD	2 (15)	1 (8)
Work experience (years), n (%)		
<5	2 (15)	0 (0)
10-15	4 (29)	3 (25)
16-20	6 (43)	4 (33)
>20	2 (15)	5 (42)
Practice setting, n (%)		
Private practice	2 (15)	2 (17)
Acute care	3 (21)	3 (25)
Inpatient rehabilitation	2 (15)	3 (25)
Outpatient rehabilitation	6 (42.8)	4 (33)
Other (teaching)	1 (6)	0 (0)
Employment, n (%)		
Full time	9 (64)	9 (75)
Part time	3 (21)	3 (25)
Casual	1 (6)	0 (0)
Information about computer use, n (%)		
Computer/tablet use at home		
Yes	12 (85)	12 (100)
No	2 (15)	0 (0)
Computer/tablet use at work		
Yes	14 (100)	12 (100)
No	0 (0)	0 (0)
Hours spent on computer/tablet each week		
≤5	0 (0)	0 (0)
>5	14 (100)	12 (100)

XSL·FO
RenderX

Demographics	Needs assessment (n=14)	Usability testing (n=12)
Hours spent on the internet each week		
≤5	5 (35)	4 (33)
>5	9 (65)	8 (67)
Comfort level on computer/tablet		
Not at all comfortable	0 (0)	0 (0)
A little comfortable	0 (0)	0 (0)
Comfortable	4 (29)	4 (33)
Very comfortable	10 (71)	8 (67)
Comfort level on the internet		
Not at all comfortable	0 (0)	0 (0)
A little comfortable	2 (14)	0 (0)
Comfortable	4 (29)	4 (33)
Very comfortable	8 (57)	8 (67)

Phase 1: Heuristic Testing

The heuristic evaluation of HOCOS against the HLO checklist identified violations in 15 of the 35 criteria with violations seen across all domains (Multimedia Appendix 2). Domain 1 showed violations in 2 of the 7 criteria. There were 4 violations in the 13 criteria for domain 2. Most of the violations were represented in domain 3, with 6 of the 10 violations reported. Violations included (1) the home page image not representing the context of the website, (2) lack of a search function, and (3) links that are difficult to differentiate from the surrounding text or other graphic elements. Corrections were made and included adding a welcome image on the home page, adding a search function, and creating a box around link icons. Domain 4 revealed 3 violations in the 5 criteria because of heavy reliance on text-based information, lack of quizzes or forms, and lack of social media sharing options. We included more pictures and reduced the words per page, creating a separate link for the form. We decided against adding a social media link because of privacy concerns and the sensitive nature of psychosocial issues. Evaluation of HOCOS using Monkman heuristics identified violations in 5 of the 11 criteria (Multimedia Appendix 3) including lack of options for tailoring information to the user, poor use of plain language including medical jargon and Gunning Fog readability index greater than 8 [79,80], information in multiple languages, few succinct summaries versus more detailed information, need for scrolling to find important information, and poor communication of risks. The remaining violations were managed by adding activities that could be personalized, editing the content for therapists and patients using the Gunning Fog index, creating a summary of key points in the slides, adding an icon to important information, and adding a disclaimer to express inherent risks and benefits of the program.

Phase 2: User Testing

This included findings from the user task performance and cognitive interviews (thinking aloud) components of usability testing of the HOCOS.

Task Performance

We measured user performance based on ease of navigating through the site, assessing the ease of learning for a first-time user without familiarity with the interface, and the frequency and importance of errors. Errors observed during usability testing were reported in 3 categories: completed with ease, completed with help, and not completed [75]. The performance of the 10 tasks is presented in Table 3. In summary, seven tasks were completed easily by participants: logging in, browsing, reading the introductory pages, listening to audio files, reading PowerPoint presentations, filling a homework plan, contacting the researchers, and downloading a document. The remaining five tasks revealed difficulties with usability. Navigation errors were defined as failures to locate functions, excessive keystrokes, or failures to follow recommended screen flow [81]. Five participants were not able to find the *assessment* page to fill outcome measures. The page was accessible through the *resources* page, although the opening comments on the page highlighted contents on the resource page. There were 6 participants who did not realize that the workbook contained both educational information and homework despite text alongside the introduction highlighting different module assignments.

Control usage errors were defined as improper toolbars or entry field usage [75]. Five participants were unable to identify the icons for submitting answers to some activities on the modules. This error was corrected by typing *click on the link to write your answers* on the link to provide answers. Providing feedback using the website form was the most difficult task for participants. Users did not click on the next page at the end of every module where the feedback form was placed. We included a text highlighting where to find the feedback form on the module's introductory page and on the final page of every module. Presentation errors were defined as failures to locate and properly act upon desired information or selection errors because of labeling obscurities [75].

Table 3. Task performance findings during usability testing (N=12).

Task performance	Completed, n (%)		Not completed, n (%)
	With ease	With help	
Cycle 1			
Log in to the website	8 (66)	3 (25)	1 (8)
Read information on the home page and each module's introductory page	10 (83)	2 (17)	N/A[a]
Complete a questionnaire from the list of outcome measures	4 (33)	3 (25)	5 (42)
Listen to an audio recording	8 (66)	4 (33)	N/A
Read a PowerPoint slide	8 (66)	3 (25)	1 (8)
Download a PDF or Word document of a workbook or PowerPoint slide	10 (83)	2 (17)	N/A
Complete 1 activity in a workbook	4 (33)	2 (17)	6 (50)
Set up an activity plan for homework	7 (58)	3 (25)	2 (17)
Find content of interest by browsing	9 (42)	2 (25)	1 (8)
Find content of interest by searching	4 (33)	3 (25)	5 (42)
Complete a feedback form	2 (17)	3 (25)	7 (58)
Contact the website manager	8 (33)	4 (50)	N/A
Cycle 2			
Log in to the website	12 (100)	N/A	N/A
Read information on the home page and each module's introductory page	12 (100)	N/A	N/A
Complete a questionnaire from the list of outcome measures	10 (83)	2 (17)	N/A
Listen to an audio recording	8 (66)	4(33)	N/A
Read a PowerPoint slide	12 (100)	N/A	N/A
Download a PDF or Word document of a workbook or PowerPoint slide	12 (100)	N/A	N/A
Complete 1 activity in a workbook	8 (66)	4 (33)	N/A
Set up an activity plan for homework	10 (83)	2 (17)	N/A
Find content of interest by browsing	9 (75)	3 (25)	N/A
Find content of interest by searching	8 (66)	4 (33)	N/A
Complete a feedback form	9 (75)	3 (25)	N/A
Contact the website manager	12 (100)	N/A	N/A

[a]N/A: not applicable.

Searching was a bit of a challenge for 5 participants because they did not know what to search for, unsure of search terms to use, or struggled to come up with a health topic in the context of the website. Participants completed the 12 tasks in phase 2 at the end of the second cycle of testing.

Cognitive Interviews

The key usability findings from the thinking aloud interviews were organized into the following themes: design aesthetics, content, functionality and features, and desire for future use. Multimedia Appendix 4 shows participants' quotes from cognitive interviews.

Design Esthetics

Overall design aesthetics was critical to enhancing engagement and motivation to use the website and related to the layout, navigation, visual assets, and appeal. Participants liked the idea of different textures, colors, and cultures represented in the

graphics. It was suggested that the font sizes should be set at size 14 to 16, and a large amount of information should be grouped and broken up with visual assets (graphics and illustrations). In response, we divided the PowerPoint slides into parts A and B and/or C to reduce information overload and reduce the feeling of being overwhelmed. Part C was created as an addendum with the caption, *Please see part C for a deeper learning on this topic.* Users also recommended that the most important message on each page should be at the top of the page. As the modules were stand-alone content, the participants suggested that a decision tree or matrix would reduce the burden of prescribing the appropriate module to patients based on their presentation and treatment goals. In response, we created a matrix with information on key learning points, indications, and contraindications for each module. For example, patients with paradoxical responses to visualization avoid thinking about their hand injury, and those focused on the loss may find mental imagery distressing. Changes were also suggested to some

XSL•FO
RenderX

features to increase user interest and reduce negative responses. For example, we changed the titles *Mental Practice* to *Picturing My Movement*, *Thought Reprocessing* to *Healthy Thinking*, and *Board of Directors* to *Thinking Traps*.

Content

Program content was described in terms of completeness, understandability, quality, credibility, relevance, and interest. The comments on program content, such as texts, images, and multimedia components, were generally positive. The layout structure of presenting information in different formats and having a summary of key points after each lesson was valued. All participants judged that the site content was relevant and credible. Generally, participants were pleased with the completeness of the website, but additional content was suggested. Examples of additions included creating reflective pieces to improve engagement with the slideshows and linking activities under *Mental Practice* to portray the multisensory nature of hand movement. HOCOS was created with a focus on understandability, and all text developed to meet grade 6 to 7 reading levels. Most participants valued this consideration and commented that the information, language level, and medical term explanations were helpful in furthering understanding of topics that were unclear or new to them. However, some of the language used still had to be changed to conjure everyday talk and meet societal norms such as changing wife/husband to spousal partner and routine to day-to-day. Several language changes were made to clarify meaning, such as changing tissues to body, thought record to thought journal, healthy to uninjured, and food for thought to pause-stop-think.

Functionality and Features

These refer to the adaptive and interactive features on the website and included module 1 to 5 audio clips, printable PDF information forms for patient and clinician users, and videos of simulated patients completing module activities. It was agreed that these features allowed for an increased level of personalization of HOCOS content to meet the individual needs of the users. To further enhance participants' motivation and engagement, we added the following functions: interactive questions (quizzes after each PowerPoint presentation), an *Ask an Expert* link to allow users to send an email question to the web developers and a goal plan journal to keep track of goals and activities. Participants suggested having features that allowed the program to support social interactions among participants, such as a discussion board. However, because of budgetary and time constraints, we were unable to include these functions in HOCOS. Other features that were introduced to help patients incorporate the new information to their daily routine was the *How to Make It Work Guide*.

Desire for Future Use

Overall, participants received HOCOS very well and expressed a desire to use the program in the future. It was agreed that the website would be especially useful if available to patients from initial contact for presurgical screening with surgeons or immediately after surgery in acute care. The therapists commented that they valued the fact that the site content focused on supplementing current hand therapy practice for patients struggling with psychosocial issues. The accompanying navigation of the workflow would make it easy to prioritize programs for their patients. Most participants suggested that collaboration with the CSHT and hand programs in Ontario would help facilitate increased uptake in the hand therapy community.

System Usability Scale and User Satisfaction

The SUS scores from both cycles of usability testing are listed in Table 4. Scores above 68 (SD 12.5) indicate above-average usability [82]. The mean SUS score for this study improved from 62.5 (SD 8.5) to 84 (SD 8.2), indicating that the average participants were highly satisfied with the usability of this online learning tool on all items of the SUS questionnaire, in terms of learnability, efficiency, memorability, errors, and satisfaction [56]. After addressing cycle 2 usability issues, we made some revisions to the final version of HOCOS. Specifically, a do-it-yourself (DIY) guide was created to support each module, a *Go to homepage* tab was created as a signpost to the respective sessions after logging in, and a navigation tutorial video and informational videos on the clinical impact of psychosocial factors on hand and upper limb injuries were created. Finally, we included patient-friendly resources on chronic pain, problem solving, time management, and a sleep guide. Overall, therapists found that HOCOS was a detailed and helpful learning resource for therapists and patients. Participants liked the web layout, tabs for modules, and resource page. There were no reported harms or unintended effects on participants, privacy breaches, or technical problems during usability testing. Overall, HOCOS system usability improved from good to excellent based on adjective rating scale described by Bangor et al [74].

Table 4. System usability scale (N=12).

Questionnaire items	Cycle 1, mean (SD)[a]	Cycle 2, mean (SD)[a]
1. I think that I would like to use this website frequently (+)[b]	3 (0.8)	4 (0.5)
2. I found the website unnecessarily complex (−)[c]	2 (0.7)	3 (0.9)
3. I thought the website was easy to use (+)[b]	2 (0.6)	3 (0.0)
4. I think that I would need the support of a technical person to be able to use this website (−)[c]	3 (1.08)	4 (0.5)
5. I found the various functions in the product were well integrated (+)[b]	3 (0.5)	3 (0.5)
6. I thought there was too much inconsistency in this website (−)[c]	3 (0.4)	4 (0.5)
7. I imagine that most people would learn to use this product very quickly (+)[b]	2 (0.7)	3 (0.4)
8. I found the website very awkward to use (−)[c]	3 (0.5)	4 (0.5)
9. I felt very confident using the website (+)[b]	2 (0.4)	3 (0.4)
10. I needed to learn a lot of things before I could get going with this website (−)[c]	2 (0.9)	3 (0.4)
Total score of items 1, 3, 5, 7, and 9	2.4 (0.5)	3.2 (0.6)
Total score of items 2, 4, 6, 8, and 10	2.6 (0.7)	3.6 (0.5)
Total score	25 (3.4)	34 (3.2)
SUS[d] score[e]	62 (8.5)	84 (8.2)

[a]Rating scale, 1=strongly disagree and 5=strongly agree.

[b]For items 1, 3, 5, 7, and 9, the score contribution is the scale position minus 1.

[c]For items 2, 4, 6, 8, and 10, the contribution is 5 minus the scale position.

[d]SUS: system usability scale.

[e]SUS score=total score×2.5.

Final Version of Hand Therapy Online Coping Skills

The final version of HOCOS was built on the Weebly platform, customized and styled using platform add-ons and publicly available pictures on Creative Commons. The platform included a landing page, a resource library, tabs for each of the modules, a feedback page, an assessment page, a goals page, and therapist- or patient-specific resources (Figure 2). Multimedia Appendix 1 gives a brief overview of the HOCOS content. Sessions can be accessed by logging in and completed using a suggested timetable. Figure 3 shows a navigation pathway to complete the 5 modules. Each module can be completed as stand-alone materials based on patient presentation. However, we recommend that every patient complete the introductory and pain education sections. Completing all five modules is projected to take approximately 6 to 8 weeks based on the structure of similar coping skill programs [83]. HOCOS is designed to be beneficial for both acute and chronic hand injuries. A therapist manual was also developed based on participant feedback.

Figure 2. Navigation workflow of the Hand Therapy Online Coping Skills (HOCOS) training program. HCP: health care professional.
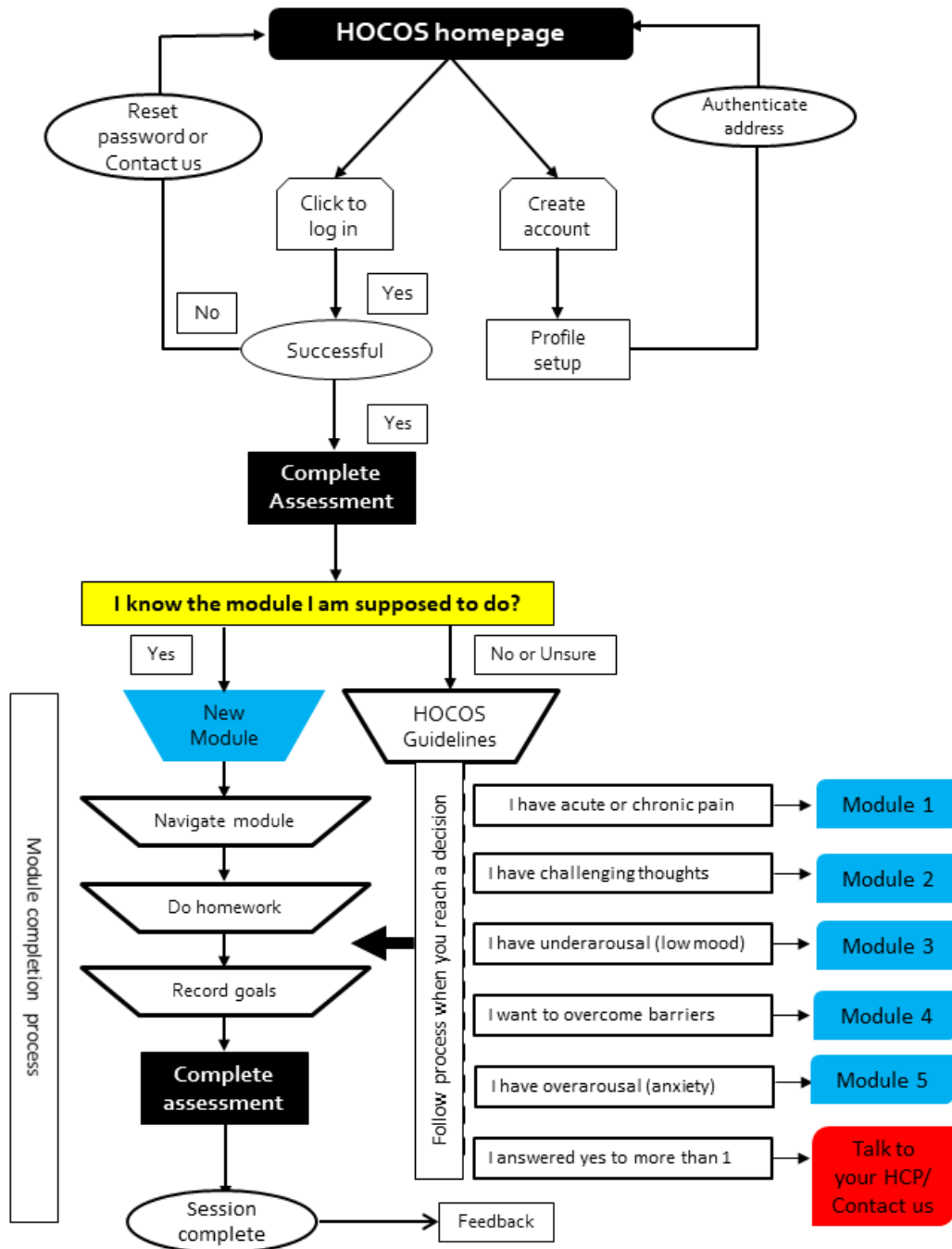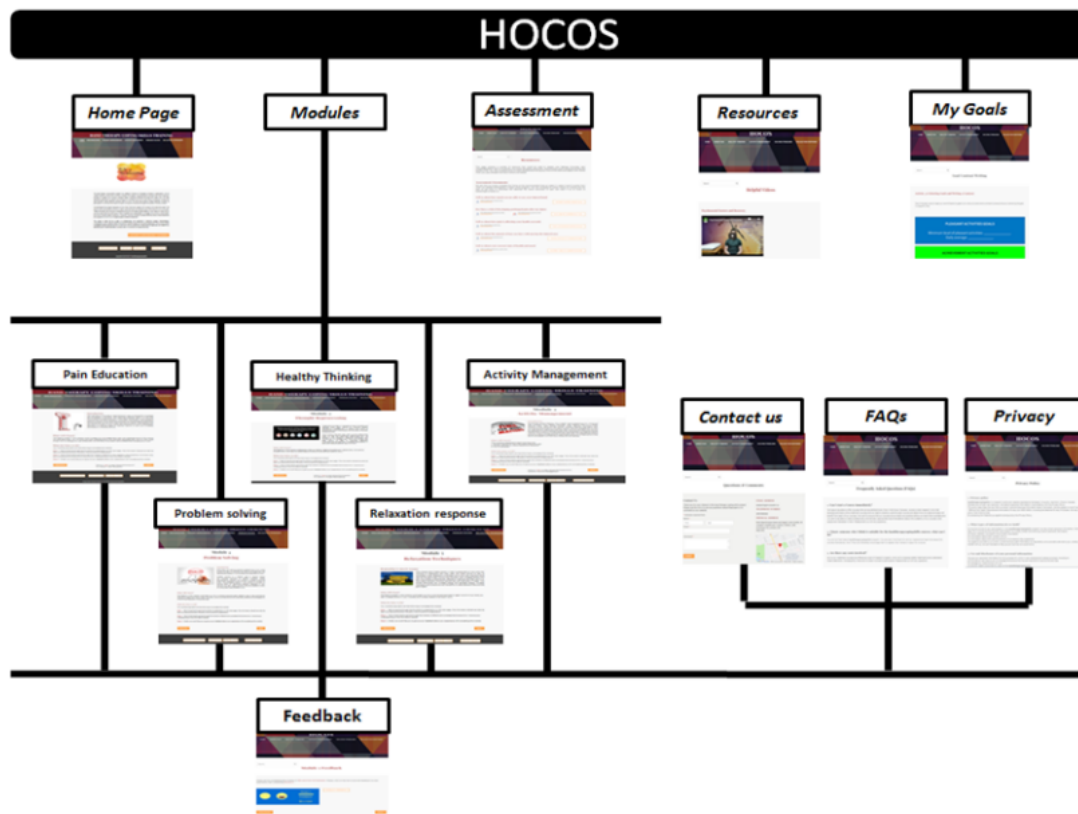
Figure 3. Screenshot of webpages showing the architecture of the website. FAQs: frequently answered questions; HOCOS: Hand Therapy Online Coping Skills.

Discussion

Principal Findings

The overall objectives of this paper were to provide an overview of HOCOS and report the findings from usability testing with ICT experts and clinicians practicing hand therapy. HOCOS is designed to help patients with hand and upper limb injuries learn how to better manage psychosocial issues. The uniqueness of HOCOS is an interface design that offers learning opportunities to both clinicians and patients. Overall, therapists were pleased with the objective and content of HOCOS and found it a useful resource for meeting patients' needs in hand therapy.

Usability Testing

Formal usability testing is a key process required to ensure the relevance of content and make the website easy to use, learn, efficient, and acceptable to users [75]. Usability testing uncovered several violations during heuristic testing with ICT experts. Furthermore, user performance errors and areas for enhancing user satisfaction were also identified by therapists during user testing. Therapists reported some positive features of the website including being simple, user-friendly, and engaging and having a functional design that was accessible on several browsers following usability testing. Several changes were made to the online portal that corrected the errors uncovered and improved overall user satisfaction. Although HOCOS was initially designed for all modules to be completed together, feedback from clinicians highlighted the benefits of having modules as stand-alone options to reduce potential

participant burden. Hand therapists expressed confidence that patients could execute the activities in the workbook, especially with the DIY guide. Testing also demonstrated that the primary condition of the patients determined the modules that therapists chose to introduce and apply in clinical practice. This process was enhanced by providing a guide on how to use the features, when to introduce the modules, and how the website may fit within the broader tool kits used in hand therapy.

This study contributes to the dearth of literature on the usability testing of web-based portals developed for managing psychosocial factors in orthopedic hand and upper extremity services. Chad-Friedman et al [84] reported the use of an online interface designed to deliver a brief 60-second mindfulness exercise for hand and upper limb pain with improvements in state anxiety, pain intensity, distress, anxiety, depression, and anger after watching the video. In another study by Westernberg et al [42], a free online mindfulness-based video exercise was targeted at individuals with upper extremity conditions and psychosocial problems. Study findings reported improvement in momentary pain, anxiety, depression, and anger in patients with low levels of pain and psychologic distress. Similarly, Vranceanu et al [85] described the Toolkit for Optimal Recovery (TOR), a 4-session, live video, and manualized program informed by the fear-avoidance model to prevent chronic pain in at-risk adults with orthopedic injuries. TOR combines relaxation response with CBT, Acceptance and Commitment therapy skills. HOCOS provides a larger platform offering multiple options to therapists and patients using concepts from pain education science, relaxation response and behavior change techniques. On the basis of postcognitive interview feedback,

therapists involved in this study preferred an online program that teaches patients how to change maladaptive cognitions and not simply accept such thoughts for long term effects. To close existing gaps in the literature, HOCOS was designed using CBT principles which teaches patients how to challenge automatic thoughts by holding them up to disproving evidence and then change them into different thoughts [86,87]. CBT begins by identifying a primary treatment goal and continuous striving to meet those goals [88]. HOCOS also provides modules that can be targeted at psychosocial problems associated with acute to subacute and chronic hand and upper limb conditions with therapist guidance. This is important because the untargeted use of psychological interventions in hand therapy and when self-directed by patients has been shown to demonstrate no benefits [89].

Dissemination of evidence-based therapies remains poor in routine practice [90]. Although allied health care professionals (HCPs) are aware of the benefits of incorporating psychological interventions within their practice, they feel insufficiently trained to optimize their use of such interventions [91]. Barriers to practicing the evidence-based therapies include a lack of access to resources that contain such evidence [92,93] and limited usable formats of the evidence [94]. Training hand therapists to manage the physical and psychological sequelae of hand and upper limb conditions using HOCOS would increase their knowledge of psychosocial interventions and build their capacity and confidence to deliver it in clinical practice.

Limitations

Our study should be viewed with consideration of certain limitations encountered. The study was conducted among hand therapists in Ontario, and most participants were comfortable using the computer and the internet, which limit the generalizability of the study results. This may not be representative of the end users, such as patients seen in most hand therapy clinics. In recruiting participants for this study, we chose snowball sampling, a form of convenience sampling. This increases the risk of compiling a nonrepresentative sample. We planned to create an online platform that is user friendly for a significant portion of patients with hand and upper limb injuries who are mostly elderly [95], low skilled [96], and with less education [97]. These groups of individuals tend to be less computer literate, and to this end, we did our best to incorporate recommendations to ensure accessibility and ease of use in the web design and simplify the user experience [66]. This included a larger font size, white space around texts, and a simple color scheme to enhance readability.

The presence of one of the researchers (FB) during the usability testing sessions may have affected the behavior of end users

conducting the testing. The participants may have felt reluctant to be critical despite encouragement to highlight both weak and strong features of the website. Furthermore, we were unable to test the HOCOS website in the context of the patient-user's experience to gain a comprehensive view of the system's functioning in a clinical setting because of financial and time constraints. This needs to be addressed in future research by examining the effectiveness of HOCOS in a randomized controlled trial to determine if the present system design can contribute to improved outcomes in practice.

User testing of an online intervention should include the ultimate end users, including patients, to allow for the examination of factors related to participants (age, gender, and education), disease (severity and duration of symptoms), and experience (access to and comfort with using the internet and computers) [75,98]. On the basis of ergonomic quality and safety principles, it has been recommended that prototypes of eHealth interventions should be fully inspected and walked through by HCPs before exposure to potentially vulnerable user groups such as individuals with significant psychosocial problems after a hand injury [99]. Financial and time constraints were significant barriers to testing HOCOS in patients with hand injuries. The next phase of the project is to evaluate the impact of HOCOS training on the actual implementation of the program on patients. We plan to carry out further testing in a proof-of-concept study to establish if individuals with hand and upper limb conditions and psychosocial problems are willing and able to complete the HOCOS program, complete the activities correctly, and adhere to the program principles.

Conclusions

This study provides initial support for the usability of HOCOS. Ensuring that therapists were involved in the design and development process of HOCOS enhanced the user-centeredness and user-friendliness of the website. Usability testing during the formative stage of eHealth intervention development is necessary to ensure that online interventions are effective and acceptable to potential users. HOCOS has the potential to increase access and acceptability of coping skills training programs for many individuals with hand and upper limb injuries who are not able to receive hospital- or clinic-based treatment psychotherapy. We plan to conduct a pilot study to determine the feasibility of the website for adults with hand and upper limb injuries and further refine the tool for a fully powered randomized controlled trial. If effective in improving outcomes, this program could be used as a template to develop more interventions targeting the psychosocial challenges confronting individuals with hand and upper limb injuries.

XSL·FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Semistructured interview guide for conducting cognitive interviews during usability testing.
[PDF File (Adobe PDF File), 33 KB - humanfactors_v7i2e17088_app1.pdf]

Multimedia Appendix 2
Heuristic evaluation of Hand Therapy Online Coping Skills by information and communications technology experts using general and readability guidelines of the Health Literacy Online checklist.
[PDF File (Adobe PDF File), 108 KB - humanfactors_v7i2e17088_app2.pdf]

Multimedia Appendix 3
Heuristic evaluation of Hand Therapy Online Coping Skills by information and communications technology experts using health-specific usability guidelines based on Monkman heuristics.
[PDF File (Adobe PDF File), 93 KB - humanfactors_v7i2e17088_app3.pdf]

Multimedia Appendix 4
Sample comments for each of the themes derived from analysis of the cognitive interview and feedback interview transcripts.
[PDF File (Adobe PDF File), 126 KB - humanfactors_v7i2e17088_app4.pdf]

References

1. Junqueira GD, Lima AL, Boni R, Almeida JC, Ribeiro RS, Figueiredo LA. Incidence of acute trauma on hand and wrist: a retrospective study. Acta Ortop Bras 2017;25(6):287-290 [FREE Full text] [doi: 10.1590/1413-785220172506169618] [Medline: 29375262]
2. Turkington C, Dempster M, Maguire J. Adjustment to hand injury: cross-sectional survey exploring adjustment in relation to illness perceptions and coping strategies. J Hand Ther 2018;31(4):502-510. [doi: 10.1016/j.jht.2017.05.021] [Medline: 28684197]
3. de Putter CE, Selles RW, Polinder S, Panneman MJ, Hovius SE, van Beeck EF. Economic impact of hand and wrist injuries: health-care costs and productivity costs in a population-based study. J Bone Joint Surg Am 2012 May 2;94(9):e56. [doi: 10.2106/JBJS.K.00561] [Medline: 22552678]
4. de Jong JP, Nguyen JT, Sonnema AJ, Nguyen EC, Amadio PC, Moran SL. The incidence of acute traumatic tendon injuries in the hand and wrist: a 10-year population-based study. Clin Orthop Surg 2014 Jun;6(2):196-202 [FREE Full text] [doi: 10.4055/cios.2014.6.2.196] [Medline: 24900902]
5. Vranceanu A, Bachoura A, Weening A, Vrahas M, Smith RM, Ring D. Psychological factors predict disability and pain intensity after skeletal trauma. J Bone Joint Surg Am 2014 Feb 5;96(3):e20. [doi: 10.2106/JBJS.L.00479] [Medline: 24500592]
6. Vranceanu A, Jupiter JB, Mudgal CS, Ring D. Predictors of pain intensity and disability after minor hand surgery. J Hand Surg Am 2010 Jun;35(6):956-960. [doi: 10.1016/j.jhsa.2010.02.001] [Medline: 20381981]
7. Rumsey N, Harcourt D. Body image and disfigurement: issues and interventions. Body Image 2004 Jan;1(1):83-97. [doi: 10.1016/S1740-1445(03)00005-6] [Medline: 18089143]
8. Hannah SD. Psychosocial issues after a traumatic hand injury: facilitating adjustment. J Hand Ther 2011;24(2):95-102; quiz 103. [doi: 10.1016/j.jht.2010.11.001] [Medline: 21236639]
9. Bot AG, Bossen JK, Mudgal CS, Jupiter JB, Ring D. Determinants of disability after fingertip injuries. Psychosomatics 2014;55(4):372-380. [doi: 10.1016/j.psym.2013.08.005] [Medline: 24360524]
10. Ring D. The role of science and psychology in optimizing care of hand illness. J Hand Ther 2011;24(2):82-83. [doi: 10.1016/j.jht.2010.11.004] [Medline: 21306872]
11. Chan J, Spencer J. Adaptation to hand injury: an evolving experience. Am J Occup Ther 2004;58(2):128-139. [doi: 10.5014/ajot.58.2.128] [Medline: 15068148]
12. Keir PJ, Zuniga AF, Mulla DM, Somasundram KG. Relationships and mechanisms between occupational risk factors and distal upper extremity disorders. Hum Factors 2019 Jul 17 epub ahead of print. [doi: 10.1177/0018720819860683] [Medline: 31314601]
13. Shi Q, Sinden K, MacDermid JC, Walton D, Grewal R. A systematic review of prognostic factors for return to work following work-related traumatic hand injury. J Hand Ther 2014;27(1):55-62; quiz 62. [doi: 10.1016/j.jht.2013.10.001] [Medline: 24268193]

14. Roesler ML, Glendon AI, O'Callaghan FV. Recovering from traumatic occupational hand injury following surgery: a biopsychosocial perspective. J Occup Rehabil 2013 Dec;23(4):536-546. [doi: 10.1007/s10926-013-9422-4] [Medline: 23400585]

15. Daud AZ, Yau MK, Barnett F, Judd J, Jones RE, Nawawi RF. Integration of occupation based intervention in hand injury rehabilitation: a randomized controlled trial. J Hand Ther 2016;29(1):30-40. [doi: 10.1016/j.jht.2015.09.004] [Medline: 26847318]

16. Bobos P, Nazari G, Szekeres M, Lalone EA, Ferreira L, MacDermid JC. The effectiveness of joint-protection programs on pain, hand function, and grip strength levels in patients with hand arthritis: a systematic review and meta-analysis. J Hand Ther 2019;32(2):194-211. [doi: 10.1016/j.jht.2018.09.012] [Medline: 30587434]

17. Menta R, Randhawa K, Côté P, Wong JJ, Yu H, Sutton D, et al. The effectiveness of exercise for the management of musculoskeletal disorders and injuries of the elbow, forearm, wrist, and hand: a systematic review by the Ontario protocol for traffic injury management (OPTIMa) collaboration. J Manipulative Physiol Ther 2015 Sep;38(7):507-520 [FREE Full text] [doi: 10.1016/j.jmpt.2015.06.002] [Medline: 26130104]

18. Sultana SS, MacDermid JC, Grewal R, Rath S. The effectiveness of early mobilization after tendon transfers in the hand: a systematic review. J Hand Ther 2013;26(1):1-20; quiz 21. [doi: 10.1016/j.jht.2012.06.006] [Medline: 23116645]

19. Szekeres M, MacDermid JC, Birmingham T, Grewal R, Lalone E. The effect of therapeutic whirlpool and hot packs on hand volume during rehabilitation after distal radius fracture: a blinded randomized controlled trial. Hand (N Y) 2017 May;12(3):265-271 [FREE Full text] [doi: 10.1177/1558944716661992] [Medline: 28453348]

20. Bennell KL, Ahamed Y, Jull G, Bryant C, Hunt MA, Forbes AB, et al. Physical therapist-delivered pain coping skills training and exercise for knee osteoarthritis: randomized controlled trial. Arthritis Care Res (Hoboken) 2016 May;68(5):590-602 [FREE Full text] [doi: 10.1002/acr.22744] [Medline: 26417720]

21. Das DS, Vranceanu A, Ring DC. Contribution of kinesophobia and catastrophic thinking to upper-extremity-specific disability. J Bone Joint Surg Am 2013 Jan 2;95(1):76-81. [doi: 10.2106/JBJS.L.00064] [Medline: 23283376]

22. Farzad M, Asgari A, Dashab F, Layeghi F, Karimlou M, Hosseini SA, et al. Does disability correlate with impairment after hand injury? Clin Orthop Relat Res 2015 Nov;473(11):3470-3476 [FREE Full text] [doi: 10.1007/s11999-015-4228-7] [Medline: 25739342]

23. Roh YH, Lee BK, Noh JH, Oh JH, Gong HS, Baek GH. Effect of anxiety and catastrophic pain ideation on early recovery after surgery for distal radius fractures. J Hand Surg Am 2014 Nov;39(11):2258-64.e2. [doi: 10.1016/j.jhsa.2014.08.007] [Medline: 25283489]

24. Roh YH, Noh JH, Oh JH, Gong HS, Baek GH. To what degree do pain-coping strategies affect joint stiffness and functional outcomes in patients with hand fractures? Clin Orthop Relat Res 2015 Nov;473(11):3484-3490 [FREE Full text] [doi: 10.1007/s11999-015-4269-y] [Medline: 25822453]

25. Briet JP, Hageman MG, Overbeek CL, Mudgal C, Ring DC, Vranceanu A. Factors associated with met expectations in patients with hand and upper extremity disorders: a pilot study. Psychosomatics 2016;57(4):401-408. [doi: 10.1016/j.psym.2016.02.004] [Medline: 27080458]

26. Groot D, van Hooff ML, Kroeze RJ, Monshouwer M, O'Dowd J, Horsting P, et al. Long-term results of an intensive cognitive behavioral pain management program for patients with chronic low back pain: a concise report of an extended cohort with a minimum of 5-year follow-up. Eur Spine J 2019 Jul;28(7):1579-1585. [doi: 10.1007/s00586-019-05967-6] [Medline: 31069525]

27. Richmond H, Hall AM, Copsey B, Hansen Z, Williamson E, Hoxey-Thomas N, et al. The effectiveness of cognitive behavioural treatment for non-specific low back pain: a systematic review and meta-analysis. PLoS One 2015;10(8):e0134192 [FREE Full text] [doi: 10.1371/journal.pone.0134192] [Medline: 26244668]

28. Tong F, Dannaway J, Enke O, Eslick G. Effect of preoperative psychological interventions on elective orthopaedic surgery outcomes: a systematic review and meta-analysis. ANZ J Surg 2020 Mar;90(3):230-236. [doi: 10.1111/ans.15332] [Medline: 31334592]

29. Ismail A, Moore C, Alshishani N, Yaseen K, Alshehri MA. Cognitive behavioural therapy and pain coping skills training for osteoarthritis knee pain management: a systematic review. J Phys Ther Sci 2017 Dec;29(12):2228-2235 [FREE Full text] [doi: 10.1589/jpts.29.2228] [Medline: 29643612]

30. Anheyer D, Haller H, Barth J, Lauche R, Dobos G, Cramer H. Mindfulness-based stress reduction for treating low back pain: a systematic review and meta-analysis. Ann Intern Med 2017 Jun 6;166(11):799-807. [doi: 10.7326/M16-1997] [Medline: 28437793]

31. Mariano TY, Urman RD, Hutchison CA, Jamison RN, Edwards RR. Cognitive behavioral therapy (CBT) for subacute low back pain: a systematic review. Curr Pain Headache Rep 2018 Feb 23;22(3):15. [doi: 10.1007/s11916-018-0669-5] [Medline: 29476270]

32. Lauche R, Cramer H, Dobos G, Langhorst J, Schmidt S. A systematic review and meta-analysis of mindfulness-based stress reduction for the fibromyalgia syndrome. J Psychosom Res 2013 Dec;75(6):500-510. [doi: 10.1016/j.jpsychores.2013.10.010] [Medline: 24290038]

33. Rolving N, Sogaard R, Nielsen CV, Christensen FB, Bünger C, Oestergaard LG. Preoperative cognitive-behavioral patient education versus standard care for lumbar spinal fusion patients: economic evaluation alongside a randomized controlled trial. Spine (Phila Pa 1976) 2016 Jan;41(1):18-25. [doi: 10.1097/BRS.0000000000001254] [Medline: 26536443]

34. das Nair R, Anderson P, Clarke S, Leighton P, Lincoln NB, Mhizha-Murira JR, et al. Home-administered pre-surgical psychological intervention for knee osteoarthritis (HAPPiKNEES): study protocol for a randomised controlled trial. Trials 2016 Jan 27;17:54 [FREE Full text] [doi: 10.1186/s13063-016-1165-z] [Medline: 26818407]

35. Smittenaar P, Erhart-Hledik JC, Kinsella R, Hunter S, Mecklenburg G, Perez D. Translating comprehensive conservative care for chronic knee pain into a digital care pathway: 12-week and 6-month outcomes for the hinge health program. JMIR Rehabil Assist Technol 2017 Apr 5;4(1):e4. [doi: 10.2196/rehab.7258] [Medline: 28582253]

36. Room J, Hannink E, Dawes H, Barker K. What interventions are used to improve exercise adherence in older people and what behavioural techniques are they based on? A systematic review. BMJ Open 2017 Dec 14;7(12):e019221 [FREE Full text] [doi: 10.1136/bmjopen-2017-019221] [Medline: 29247111]

37. Brunner E, de Herdt A, Minguet P, Baldew S, Probst M. Can cognitive behavioural therapy based strategies be integrated into physiotherapy for the prevention of chronic low back pain? A systematic review. Disabil Rehabil 2013 Jan;35(1):1-10. [doi: 10.3109/09638288.2012.683848] [Medline: 22607157]

38. Hall A, Richmond H, Copsey B, Hansen Z, Williamson E, Jones G, et al. Physiotherapist-delivered cognitive-behavioural interventions are effective for low back pain, but can they be replicated in clinical practice? A systematic review. Disabil Rehabil 2018 Jan;40(1):1-9. [doi: 10.1080/09638288.2016.1236155] [Medline: 27871193]

39. Bot AG, Bekkers S, Arnstein PM, Smith RM, Ring D. Opioid use after fracture surgery correlates with pain intensity and satisfaction with pain relief. Clin Orthop Relat Res 2014 Aug;472(8):2542-2549 [FREE Full text] [doi: 10.1007/s11999-014-3660-4] [Medline: 24777731]

40. Vranceanu A, Riklin E, Merker VL, Macklin EA, Park ER, Plotkin SR. Mind-body therapy via videoconferencing in patients with neurofibromatosis: an RCT. Neurology 2016 Aug 23;87(8):806-814. [doi: 10.1212/WNL.0000000000003005] [Medline: 27449066]

41. Beleckas CM, Wright M, Prather H, Chamberlain A, Guattery J, Calfee RP. Relative prevalence of anxiety and depression in patients with upper extremity conditions. J Hand Surg Am 2018 Jun;43(6):571.e1-571.e8 [FREE Full text] [doi: 10.1016/j.jhsa.2017.12.006] [Medline: 29395589]

42. Westenberg RF, Zale EL, Heinhuis TJ, Özkan S, Nazzal A, Lee S, et al. Does a brief mindfulness exercise improve outcomes in upper extremity patients? A randomized controlled trial. Clin Orthop Relat Res 2018 Apr;476(4):790-798 [FREE Full text] [doi: 10.1007/s11999.0000000000000086] [Medline: 29480886]

43. Kumar V, Sattar Y, Bseiso A, Khan S, Rutkofsky IH. The effectiveness of internet-based cognitive behavioral therapy in treatment of psychiatric disorders. Cureus 2017 Aug 29;9(8):e1626 [FREE Full text] [doi: 10.7759/cureus.1626] [Medline: 29098136]

44. Garg S, Garg D, Turin TC, Chowdhury MF. Web-based interventions for chronic back pain: a systematic review. J Med Internet Res 2016 Jul 26;18(7):e139 [FREE Full text] [doi: 10.2196/jmir.4932] [Medline: 27460413]

45. Terpstra JA, van der Vaart R, van Koulil SS, van Dam A, Rosmalen JG, Knoop H, et al. Becoming an eCoach: training therapists in online cognitive-behavioral therapy for chronic pain. Patient Educ Couns 2018 Sep;101(9):1702-1707. [doi: 10.1016/j.pec.2018.03.029] [Medline: 29661704]

46. Cowell I, O'Sullivan P, O'Sullivan K, Poyton R, McGregor A, Murtagh G. The perspectives of physiotherapists on managing nonspecific low back pain following a training programme in cognitive functional therapy: a qualitative study. Musculoskeletal Care 2019 Mar;17(1):79-90. [doi: 10.1002/msc.1370] [Medline: 30468555]

47. Stiles-Shields C, Kwasny MJ, Cai X, Mohr DC. Therapeutic alliance in face-to-face and telephone-administered cognitive behavioral therapy. J Consult Clin Psychol 2014 Apr;82(2):349-354 [FREE Full text] [doi: 10.1037/a0035554] [Medline: 24447003]

48. Pihlaja S, Stenberg J, Joutsenniemi K, Mehik H, Ritola V, Joffe G. Therapeutic alliance in guided internet therapy programs for depression and anxiety disorders-a systematic review. Internet Interv 2018 Mar;11:1-10 [FREE Full text] [doi: 10.1016/j.invent.2017.11.005] [Medline: 30135754]

49. Haug T, Nordgreen T, Öst LG, Tangen T, Kvale G, Hovland OJ, et al. Working alliance and competence as predictors of outcome in cognitive behavioral therapy for social anxiety and panic disorder in adults. Behav Res Ther 2016 Feb;77:40-51. [doi: 10.1016/j.brat.2015.12.004] [Medline: 26708332]

50. Cooper Z, Bailey-Straebler S, Morgan KE, O'Connor ME, Caddy C, Hamadi L, et al. Using the internet to train therapists: randomized comparison of two scalable methods. J Med Internet Res 2017 Oct 18;19(10):e355 [FREE Full text] [doi: 10.2196/jmir.8336] [Medline: 29046265]

51. Everitt HA, Landau S, O'Reilly G, Sibelli A, Hughes S, Windgassen S, ACTIB Trial Group. Assessing telephone-delivered cognitive-behavioural therapy (CBT) and web-delivered CBT versus treatment as usual in irritable bowel syndrome (ACTIB): a multicentre randomised trial. Gut 2019 Sep;68(9):1613-1623 [FREE Full text] [doi: 10.1136/gutjnl-2018-317805] [Medline: 30971419]

52. Hageman MG, Anderson J, Blok R, Bossen JK, Ring D. Internet self-diagnosis in hand surgery. Hand (N Y) 2015 Sep;10(3):565-569 [FREE Full text] [doi: 10.1007/s11552-014-9707-x] [Medline: 26330798]

53. Kelly JD. CORR insights: does a brief mindfulness exercise improve outcomes in upper extremity patients? a randomized controlled trial. Clin Orthop Relat Res 2018 Apr;476(4):799-800 [FREE Full text] [doi: 10.1007/s11999.0000000000000193] [Medline: 29406459]

54. Holden RJ, Karsh B. The technology acceptance model: its past and its future in health care. J Biomed Inform 2010 Feb;43(1):159-172 [FREE Full text] [doi: 10.1016/j.jbi.2009.07.002] [Medline: 19615467]

55. Levac D, Glegg SM, Camden C, Rivard LM, Missiuna C. Best practice recommendations for the development, implementation, and evaluation of online knowledge translation resources in rehabilitation. Phys Ther 2015 Apr;95(4):648-662. [doi: 10.2522/ptj.20130500] [Medline: 25301966]

56. Srikesavan CS, Williamson E, Eldridge L, Heine P, Adams J, Cranston T, et al. A web-based training resource for therapists to deliver an evidence-based exercise program for rheumatoid arthritis of the hand (iSARAH): design, development, and usability testing. J Med Internet Res 2017 Dec 13;19(12):e411 [FREE Full text] [doi: 10.2196/jmir.8424] [Medline: 29237581]

57. Bennell KL, Egerton T, Pua Y, Abbott JH, Sims K, Buchbinder R. Building the rationale and structure for a complex physical therapy intervention within the context of a clinical trial: a multimodal individualized treatment for patients with hip osteoarthritis. Phys Ther 2011 Oct;91(10):1525-1541. [doi: 10.2522/ptj.20100430] [Medline: 21817012]

58. Nielsen J. Heuristic evaluation. In: Nielsen J, Mack RL, editors. Usability Inspection Methods. New York, USA: Wiley; 1994:25-62.

59. Allen M, Currie LM, Bakken S, Patel VL, Cimino JJ. Heuristic evaluation of paper-based web pages: a simplified inspection usability methodology. J Biomed Inform 2006 Aug;39(4):412-423 [FREE Full text] [doi: 10.1016/j.jbi.2005.10.004] [Medline: 16321575]

60. Wichansky AM. Usability testing in 2000 and beyond. Ergonomics 2000 Jul;43(7):998-1006. [doi: 10.1080/001401300409170] [Medline: 10929833]

61. Sadler GR, Lee H, Lim RS, Fullerton J. Recruitment of hard-to-reach population subgroups via adaptations of the snowball sampling strategy. Nurs Health Sci 2010 Sep 1;12(3):369-374 [FREE Full text] [doi: 10.1111/j.1442-2018.2010.00541.x] [Medline: 20727089]

62. Khalil MK, Elkhider IA. Applying learning theories and instructional design models for effective instruction. Adv Physiol Educ 2016 Jun;40(2):147-156 [FREE Full text] [doi: 10.1152/advan.00138.2015] [Medline: 27068989]

63. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000 Aug;23(4):334-340. [doi: 10.1002/1098-240x(200008)23:4<334::aid-nur9>3.0.co;2-g] [Medline: 10940958]

64. Cazañas-Gordón A, de san Miguel A, Mora EP. Estimating sample size for usability testing. Enfoque 2017 Feb 24;8(1):172-185 [FREE Full text] [doi: 10.29019/enfoqueute.v8n1.126]

65. Monkman H, Kushniruk A. eHealth literacy issues, constructs, models, and methods for health information technology design and evaluation. Knowl Manag eLearn 2015;7:541-549. [doi: 10.34105/j.kmel.2015.07.036]

66. UC Berkeley Library. 2017. Health Literacy Online: A Guide for Simplifying the User Experience URL: https://update.lib.berkeley.edu/2017/05/01/health-literacy-online-a-guide-for-simplifying-the-user-experience/ [accessed 2018-01-08]

67. Monkman H, Griffith J, Kushniruk AW. Evidence-based heuristics for evaluating demands on ehealth literacy and usability in a mobile consumer health application. Stud Health Technol Inform 2015;216:358-362. [Medline: 26262071]

68. van Someren MW, Barnard YF, Sandberg JA. The Think Aloud Method: A Practical Guide to Modelling Cognitive Processes. London, UK: Academic Press; 1994.

69. Jaspers MW, Steen T, van den Bos C, Geenen M. The think aloud method: a guide to user interface design. Int J Med Inform 2004 Nov;73(11-12):781-795. [doi: 10.1016/j.ijmedinf.2004.08.003] [Medline: 15491929]

70. Irwin DE, Varni JW, Yeatts K, de Walt DA. Cognitive interviewing methodology in the development of a pediatric item bank: a patient reported outcomes measurement information system (PROMIS) study. Health Qual Life Outcomes 2009 Jan 23;7:3 [FREE Full text] [doi: 10.1186/1477-7525-7-3] [Medline: 19166601]

71. Rubin J, Chisnell D, Spool J. Analyze data and observations. In: Rubin J, editor. Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests. New York, USA: John Wiley and Sons; Jul 21, 2008:245-265.

72. Brooke J. System usability scale: a quick and dirty usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation In Industry. Philadelphia, Pennsylvania: Taylor and Francis; 1996.

73. Brooke J. System usability scale (SUS): a retrospective. J Usability Stud 2013;8(2):29-40 [FREE Full text] [doi: 10.5555/2817912.2817913]

74. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009;4(3):114-123 [FREE Full text] [doi: 10.5555/2835587.2835589]

75. Stinson J, McGrath P, Hodnett E, Feldman B, Duffy C, Huber A, et al. Usability testing of an online self-management program for adolescents with juvenile idiopathic arthritis. J Med Internet Res 2010 Jul 29;12(3):e30 [FREE Full text] [doi: 10.2196/jmir.1349] [Medline: 20675293]

76. Colorafi KJ, Evans B. Qualitative descriptive methods in health science research. Health Env Res Desig 2016 Jul;9(4):16-25. [doi: 10.1177/1937586715614171] [Medline: 26791375]

77. Neuendorf K. The Content Analysis Guidebook. Thousand Oaks, CA: Sage Publications; 2017.

78. Nielsen J, Landauer TK. A Mathematical Model of the Finding of Usability Problems. In: Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems. 1993 Presented at: CHI'93; April 24-29, 1993; Amsterdam, The Netherlands p. 24-29 URL: https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=A+Mathematical+Model+of+the+Finding+of+Usability+Problems.&btnG= [doi: 10.1145/169059.169166]

79. Stossel LM, Segar N, Gliatto P, Fallar R, Karani R. Readability of patient education materials available at the point of care. J Gen Intern Med 2012 Sep;27(9):1165-1170 [FREE Full text] [doi: 10.1007/s11606-012-2046-0] [Medline: 22528620]

80. Gunning R. The Technique of Clear Writing. New York, USA: McGraw-Hill; 1968.

81. 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 Mar;67:69-79 [FREE Full text] [doi: 10.1016/j.jbi.2017.01.005] [Medline: 28088527]

82. 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]

83. Dear BF, Gandy M, Karin E, Ricciardi T, Fogliati VJ, McDonald S, et al. The pain course: a randomised controlled trial comparing a remote-delivered chronic pain management program when provided in online and workbook formats. Pain 2017 Jul;158(7):1289-1301. [doi: 10.1097/j.pain.0000000000000916] [Medline: 28394850]

84. Chad-Friedman E, Talaei-Khoei M, Ring D, Vranceanu A. First use of a brief 60-second mindfulness exercise in an orthopedic surgical practice; results from a pilot study. Arch Bone Jt Surg 2017 Nov;5(6):400-405 [FREE Full text] [Medline: 29299495]

85. Vranceanu A, Jacobs C, Lin A, Greenberg J, Funes CJ, Harris MB, et al. Results of a feasibility randomized controlled trial (RCT) of the Toolkit for Optimal Recovery (TOR): a live video program to prevent chronic pain in at-risk adults with orthopedic injuries. Pilot Feasibility Stud 2019;5:30 [FREE Full text] [doi: 10.1186/s40814-019-0416-7] [Medline: 30820341]

86. Longmore RJ, Worrell M. Do we need to challenge thoughts in cognitive behavior therapy? Clin Psychol Rev 2007 Mar;27(2):173-187. [doi: 10.1016/j.cpr.2006.08.001] [Medline: 17157970]

87. Hofmann SG, Gómez AF. Mindfulness-based interventions for anxiety and depression. Psychiatr Clin North Am 2017 Dec;40(4):739-749 [FREE Full text] [doi: 10.1016/j.psc.2017.08.008] [Medline: 29080597]

88. Knoerl R, Smith EM, Weisberg J. Chronic pain and cognitive behavioral therapy: an integrative review. West J Nurs Res 2016 May;38(5):596-628. [doi: 10.1177/0193945915615869] [Medline: 26604219]

89. Goudie S, Dixon D, McMillan G, Ring D, McQueen M. Is use of a psychological workbook associated with improved disabilities of the arm, shoulder and hand scores in patients with distal radius fracture? Clin Orthop Relat Res 2018 Apr;476(4):832-845 [FREE Full text] [doi: 10.1007/s11999.0000000000000095] [Medline: 29406451]

90. Rapport F, Clay-Williams R, Churruca K, Shih P, Hogden A, Braithwaite J. The struggle of translating science into action: foundational concepts of implementation science. J Eval Clin Pract 2018 Feb;24(1):117-126 [FREE Full text] [doi: 10.1111/jep.12741] [Medline: 28371050]

91. Alexanders J, Anderson A, Henderson S. Musculoskeletal physiotherapists' use of psychological interventions: a systematic review of therapists' perceptions and practice. Physiotherapy 2015 Jun;101(2):95-102. [doi: 10.1016/j.physio.2014.03.008] [Medline: 25435081]

92. Zadro J, O'Keeffe M, Maher C. Do physical therapists follow evidence-based guidelines when managing musculoskeletal conditions? Systematic review. BMJ Open 2019 Oct 7;9(10):e032329 [FREE Full text] [doi: 10.1136/bmjopen-2019-032329] [Medline: 31591090]

93. Bennett S, Whitehead M, Eames S, Fleming J, Low S, Caldwell E. Building capacity for knowledge translation in occupational therapy: learning through participatory action research. BMC Med Educ 2016 Oct 1;16(1):257 [FREE Full text] [doi: 10.1186/s12909-016-0771-5] [Medline: 27716230]

94. Scurlock-Evans L, Upton P, Upton D. Evidence-based practice in physiotherapy: a systematic review of barriers, enablers and interventions. Physiotherapy 2014 Sep;100(3):208-219. [doi: 10.1016/j.physio.2014.03.001] [Medline: 24780633]

95. Mauck BM, Swigler CW. Evidence-based review of distal radius fractures. Orthop Clin North Am 2018 Apr;49(2):211-222. [doi: 10.1016/j.ocl.2017.12.001] [Medline: 29499822]

96. Grivna M, Eid HO, Abu-Zidan FM. Epidemiology of isolated hand injuries in the United Arab Emirates. World J Orthop 2016 Sep 18;7(9):570-576 [FREE Full text] [doi: 10.5312/wjo.v7.i9.570] [Medline: 27672570]

97. Heffernan KJ, Chang S, Maclean ST, Callegari ET, Garland SM, Reavley NJ, et al. Guidelines and recommendations for developing interactive ehealth apps for complex messaging in health promotion. JMIR Mhealth Uhealth 2016 Feb 9;4(1):e14 [FREE Full text] [doi: 10.2196/mhealth.4423] [Medline: 26860623]

98. Fu MR, Axelrod D, Guth AA, Wang Y, Scagliola J, Hiotis K, et al. Usability and feasibility of health IT interventions to enhance self-care for lymphedema symptom management in breast cancer survivors. Internet Interv 2016 Sep;5:56-64 [FREE Full text] [doi: 10.1016/j.invent.2016.08.001] [Medline: 28255542]

99. Harte R, Glynn L, Rodríguez-Molinero A, Baker PM, Scharf T, Quinlan LR, et al. A human-centered design methodology to enhance the usability, human factors, and user experience of connected health systems: a three-phase methodology. JMIR Hum Factors 2017 Mar 16;4(1):e8 [FREE Full text] [doi: 10.2196/humanfactors.5443] [Medline: 28302594]

Abbreviations

ADDIE: Analysis, Design, Development, Implementation, and Evaluation
CBT: cognitive behavioral therapy
CIHR: Canadian Institutes Health Research
CSHT: Canadian Society of Hand Therapy
DIY: do-it-yourself
eHealth: electronic health
HCP: health care professional
HLO: Health Literacy Online
HOCOS: Hand Therapy Online Coping Skills
ICT: information and communications technology
MSK: musculoskeletal
OT: occupational therapist
PT: physiotherapist
SUS: system usability scale
TOR: Toolkit for Optimal Recovery

XSL•FO
RenderX

Original Paper

The Mobile Alliance for Maternal Action Text Message–Based mHealth Intervention for Maternal Care in South Africa: Qualitative User Study

Jesse Coleman[1,2], BA, MSc, PhD; Jaran Eriksen[2,3], MD, PhD; Vivian Black[4], MBBCh, MD; Anna Thorson[2], MPH, MD, PhD; Abigail Hatcher[5], PhD

[1]Wits Reproductive Health & HIV Institute, School of Medicine, University of Witwatersrand, Johannesburg, South Africa

[2]Department of Global Public Health, Karolinska Institutet, Stockholm, Sweden

[3]Division of Clinical Pharmacology, Department of Laboratory Medicine, Karolinska Institutet at Karolinska University Hospital, Stockholm, Sweden

[4]Department of Clinical Microbiology and Infectious Diseases, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

[5]School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

Corresponding Author:
Jesse Coleman, BA, MSc, PhD
Wits Reproductive Health & HIV Institute
School of Medicine
University of Witwatersrand
22 Esselen Street
Hillbrow
Johannesburg, 2001
South Africa
Phone: 27 833991066
Email: denots@gmail.com

Abstract

Background: Using mobile technology to support health care (mobile health [mHealth]) has been shown to improve health outcomes across a multitude of health specialties and across the world. Exploring mHealth user experiences can aid in understanding how and why an intervention was successful. The Mobile Alliance for Maternal Action (MAMA) was a free maternal mHealth SMS text messaging service that was offered to pregnant women in Johannesburg, South Africa, with the goal of improving maternal, fetal, and infant health outcomes. We conducted focus group discussions with MAMA users to learn about their experiences with the program.

Objective: The aim of this qualitative study was to gather opinions of participants of the MAMA maternal mHealth service regarding health care atmosphere, intervention use, and intervention feedback.

Methods: Prenatal and postnatal women (N=15) from public antenatal and postnatal care sites in central Johannesburg who were receiving free maternal health text messages (MAMA) participated in 3 focus group discussions. Predefined discussion topics included personal background, health care system experiences, MAMA program recruitment, acceptability, participant experiences, and feedback.

Results: The feedback regarding experiences with the health system were comprised of a few reports of positive experiences and many more reports of negative experiences such as long wait times, understaffed facilities, and poor service. Overall acceptability for the maternal text message intervention was high. Participants reflected that the messages were timely, written clearly, and felt supportive. Participants also reported sharing messages with friends and family.

Conclusions: These findings suggest that maternal mHealth interventions delivered through text messages can provide timely, relevant, useful, and supportive information to pregnant women and new mothers especially in settings where there may be mistrust of the health care system.

XSL·FO
RenderX

Introduction

Attendance to antenatal care and postnatal follow-up care visits which provide professional maternal and infant health services during pregnancy is important for healthy maternal, neonatal, and child health outcomes [1,2]. Such visits allow health professionals to identify and treat maternal and neonatal health issues and have been found to decrease mortality and morbidity [3]. Together antenatal care, postnatal follow-up care visits, and infant vaccinations constitute the core of the maternal, neonatal, and child health continuum of care [4].

South Africa did not meet key child and maternal mortality United Nation Millennium Development Goals (goals 4 and 5) by 2015, largely due to the high prevalence of HIV [5]; however, these mortality statistics did show improvement. In 2007, at the height of the HIV epidemic, South Africa had 48.1 infant deaths per 1000 live births which halved to 23.6 per 1000 by 2013 [6]. Between 2004 and 2015, South African child (under 5 years of age) mortality decreased from 66.9 to 38.0 deaths per 1000 [6], and in 2013, the maternal mortality ratio was estimated to be 148 per 100 000 [7]. Despite these positive changes, more improvement is needed in order to achieve national and global maternal, neonatal, and child health goals.

Mobile technology, when used to support health care services, is often referred to as mHealth (mobile health) [8]. Previous systematic reviews of maternal mHealth interventions in low- or middle-income countries have highlighted the improvement in maternal and neonatal health outcomes, but have also recommended further research into the factors contributing to successful and unsuccessful mHealth interventions in practice [9-12]. A gap in research exists because most mHealth evaluations tend to use quantitative methodology. There is a need for rigorous and continued evaluation of mHealth interventions in order to understand how they worked and why they succeeded (or did not succeed) and to ensure future mHealth interventions are implemented successfully. Process evaluations of mHealth interventions have been used only a handful of times globally [13-15] and were used as part of pilot projects to identify participant need and interest rather than to look at large-scale implementation [16-18]. A qualitative analysis [19] of a nationally implemented (in South Africa) maternal mHealth intervention included only one clinic in the study. Another South African study [20] investigated the acceptance of ad hoc use of mobile technology by patients and providers, rather than that of a specific mHealth intervention. This study aims to address gaps in knowledge by exploring the experiences of participants of the Mobile Alliance for Maternal Action South Africa (MAMA) project, an SMS text message–based maternal mHealth intervention that was offered in Johannesburg between 2012 and 2014.

MAMA Intervention Overview

The MAMA intervention sent maternal health and infant care information by SMS text message to approximately 12,000 pregnant women and new mothers in Johannesburg, South Africa, throughout pregnancy and until their infant was one year of age [21]. At the time of recruitment, women were given the option to receive one of two types of text message content—general maternal health information or prevention of mother-to-child transmission of HIV maternal health information; however, due to a high rate of women who were pregnant and HIV-positive [22], both streams of messages contained some HIV content, such as regular HIV-testing reminders (see Multimedia Appendix 1). The difference between the prevention of mother-to-child transmission of HIV content stream and the general maternal health content stream was that approximately 20 general maternal health support–related messages were replaced with prevention of mother-to-child transmission of HIV–related messages. The intervention predated, however retrospectively, was in line with the World Health Organization Classification of Digital Health Interventions [23], whereby specific digital health interventions can be used to address health system challenges; in the case of the MAMA intervention, targeted health information was transmitted to a certain demographic (pregnancy) clients to provide health education and to decrease attrition rates [23].

Through routine operational research [22], MAMA SMS text message recipients had previously provided feedback regarding a number of contextual factors such as poverty, violence, alcohol, social support, the underresourced health care system, the high rate of HIV infection, and the high rate of miscarriage. Participant discussions of poverty and violence subthemes included topics of long-term unemployment and sharing living space with multiple other families, as well as perspectives on the effects of excessive alcohol use on both themselves and their community. Income-related concerns that were provided during feedback included not being able to afford high-quality medical services, witnessing verbal or physical fights between couples on the subject of finances, and being unable to regularly afford meals that included meat. Social support subthemes included social norms such as being able to turn to siblings and older generation members of the family for support after delivery to enable a safe and supportive environment for themselves and their newborns.

Previous nonrandomized quantitative studies [24,25] that investigated MAMA health outcomes looked at mother-infant pairs who received the MAMA SMS text messages compared to the mother-infant pairs who did not receive MAMA SMS text messages and showed that those in the intervention arm had a higher rate of antenatal care attendance, an increased likelihood of a vaginal birth, a reduced likelihood of emergency cesarean delivery, and were more likely to have attended all recommended postnatal follow-up care visits up to one year after birth. Furthermore, an analysis [26] found that the MAMA text message intervention would be a cost-effective strategy to improve antenatal care attendance and vaccination rates, even if only brought to scale in Gauteng, one of South Africa's provinces.

Methods

Study Design

This was a qualitative study with an inductive and descriptive design. An inductive approach involves drawing codes, categories, or themes directly from the data, and is useful when knowledge about a phenomenon is limited [27].

Study Setting and Participants

In late 2013 and early 2014, adult women (18 years of age or older) attending routine antenatal and postnatal follow-up care services at 3 sites were invited to participate in focus group discussions. Participants were purposively selected to identify women who were either at various stages of pregnancy or after delivery with infants less than one year of age on the day of recruitment. Potential participants were women who were already receiving MAMA messages and who were identified by asking; if the women responded affirmatively, they were invited to participate in the study. Women who agreed to participate in the study provided informed consent. Of the 21 women who were invited to participate, 15 women agreed to participate in the focus group discussions.

All three sites in the study were public health care facilities in Hillbrow, Johannesburg. Hillbrow has a high population density with high diversity, predominantly low-income households, and had an unemployment rate estimated at 23% in 2013 as well as high rates of behaviors such as alcohol use and gender-based violence [28]. In Hillbrow in 2013, 27% of women who were pregnant were HIV-positive [24].

Study Procedures

A focus group discussion guide was created prior to the study and was designed to elicit feedback from participants about their experiences related to the intervention text messages (Multimedia Appendix 2). The topics covered general questions about the message content, usefulness, the signup procedure, and sharing of the messages with others. Focus group discussions were timed so that participants had received at least four months of messages which allowed them to have sufficient experience to provide feedback but was early enough in the intervention life span to allow for optimization, change, and improvement, if necessary.

Focus group discussions were held in a private room in an antenatal and postnatal follow-up care site that offered the intervention. A total of 3 focus group discussions, each with 4-6 participants, were held. Each discussion group lasted between 60 and 90 minutes and was conducted in English. At each focus group discussion, 2 to 4 research staff were present, one of whom was experienced in qualitative research and who acted as the moderator. The other research staff observed and translated between local languages and English, when necessary.

Data Collection and Analysis

Audio recordings of each focus group discussion were transcribed verbatim, and managed in Dedoose [29], an online qualitative data analysis tool. Focus group discussions were separately read and coded by 3 members of the study team who then agreed on a hierarchical coding system. The hierarchical coding system was refined using an inductive-deductive approach based upon the focus group discussion interview guide and initial review of the transcripts. General categories were identified, reviewed, and then organized into major categories and subcategories to capture specific detail (Table 1). The 3 researchers then analyzed the text using latent content analysis inspired by Graneheim and Lundman [30]. When differences of opinion arose, coding was compared, reviewed, and discussed until there was consensus.

Table 1. Focus group discussion themes (categories and subcategories).

Category	Subcategory
Factors contextualizing the intervention	Poverty/employment/income
	Social support
	Experiences of the public health care system
	Positive
	Negative
Factors contributing to intervention success	Recruitment was facilitated by helpful staff members
	Privacy concerns were allayed
	Communication preferences
	Messages arrived regularly
	Text message content was accessible
Project feedback	Relevance of text message content
	Trust in the content
	Acceptability of the intervention

Ethics

The study was approved by the Human Research Ethics Committee (Medical; M120649) at the University of the Witwatersrand in Johannesburg. Participation in the study was voluntary and informed consent was given by each participant prior to the collection of any personal information. Participants were informed that they were not required to disclose their HIV status.

Results

Participants

The 15 participants ranged in age from 20 to 36 years (median 31, IQR 7). All women were black, African, and residents of Hillbrow, Johannesburg. Women who were pregnant (prenatal, n=8) ranged from being between 26 and 39 weeks pregnant at the time of their focus group discussion, the other women (postnatal, n=7) had given birth between one week and 52 weeks prior. Each focus group discussion included both post and prenatal participants as well as a women who ranged in age from their twenties to thirties (group 1: 20-36; group 2: 28-35; group 3: 21-36 years of age). All participants received MAMA SMS text messages sent twice a week for at least 16 weeks (ie, at least 32 SMS text messages).

Table 2. Overview of focus group participant characteristics.

Focus group	Participants, n	Age (years), range	Prenatal, n (weeks gestation)	Postnatal, n (weeks since)
Group 1	5	20-36	3 (30-34)	2 (4; 20)
Group 2	4	28-35	2 (26; 39)	2 (1; 52)
Group 3	6	21-36	4 (34-39)	2 (17; 34)

Experiences of the Public Health Care System

There were divergent opinions about the health care system. A few participants had positive feelings and experiences, but most expressed negative feelings. Those with positive experiences mentioned having trust in the medical procedures and the experts who work there. On the other hand, a number of participants described having poor opinions of both the health care system and staff. One participant was able to differentiate her opinion between the system and individuals who worked within it. Specific topics are explored in more detail below.

Feedback about the health care system related to HIV testing, care, and treatment was positive. There were comments about public HIV clinics being more trustworthy than private clinics that conducted HIV tests:

> Sometimes the tests from the [private] doctors they come wrong but normally at the clinic, if you know you are testing at the clinic I don't think your status would ever come wrong, if it's negative it will come back negative. [Postnatal woman, 32 years of age]

In addition, there was a perceived benefit in the antenatal HIV testing services:

> For me, I think the most important reason that you should book at the clinic is so that you may know your [HIV] status before you proceed with the pregnancy [and] so that your baby will be checked so that you can proceed with your pregnancy [knowing your baby is healthy]. [Postnatal woman, 32 years of age]

Notwithstanding HIV testing, care, and treatment services, health care–related feedback was less positive and participants were vocal about their negative health care experiences. Analyzing these experiences, 3 main barriers were identified: long wait times, poor treatment (by staff), and that staff seen to be overworked. One spoke about staff treating patients with disdain and disrespect, making them feel that they must "obey" and that they had made a mistake by becoming pregnant:

> I wouldn't suggest anyone to go to the clinic especially [clinic name] if she's pregnant, no I wouldn't suggest [it].... There [at the antenatal clinic] if you are pregnant you are being treated like you are stupid and if you don't obey that stupidity they won't help you, you don't get the dignity as a human being… You're just nothing just because you are pregnant which is not fair. [Prenatal woman, 33 years of age]

Others described the long wait times and the perception that staff chose not to treat all patients who arrived on a given day:

> People are coming 3 o'clock [in the morning]; imagine a pregnant person coming 3 o'clock to [wait until] 07:30. People have to come that early because [the staff] only see a small number [of patients a day] or you get turned away; no they shouldn't do that. [Unidentified participant, focus group 3]

Participants in focus group discussions recounted feeling scared of the treatment by staff, but explained that they continued to seek these services because of the importance of antenatal care for infant health:

> The problem is that even if you try to talk to them they will tell you that you challenging them...we scared but we just thinking of our babies...I always tell my friends just go there don't worry about how they treating us think about our babies. [Postnatal woman, 36 years of age]

Factors Contributing to Intervention Success

This section identifies aspects of the intervention that were mentioned by focus group discussion participants as being useful or enabling its use and integrating the information that they received into action. These included using a personalized and private recruitment technique, using an accepted and reliable

communication method, and receiving simple, relevant, and supportive message content.

Recruitment to the text message intervention was done with the support of study staff who identified themselves as working with the study as opposed to identifying themselves as working for the healthcare facility:

> *I think the approach was good because it was professional she was able to explain what is it that I'm expecting and the messages which will, the usefulness of the message and what will it be helping me with so I think she was professional enough.* [Prenatal woman, 36 years of age]

Participants also noted that privacy was an important issue. One participant appreciated the careful way that staff invited her to take part in the mHealth intervention, using a discreet conversation that did not disclose her HIV status to others:

> *What I was happy for is she asked me [if I wanted to receive HIV-related messages] in private she didn't just ask me in front of people so I just tell her that I know my status so she say those messages will help me so that I won't be having stress or to think too much about it so every Monday or Thursday I was always waiting for the messages so that if I'm not OK if I haven't gotten those messages I will feel better than before.* [Postnatal woman, 28 years of age]

Intervention Communication

The focus group discussion included a query about preferred communication methods which aimed to identify if text messages were a barrier in any way. The discussion covered text messages, radio, email, television, and print, but did not include smartphones due to their low utilization. Participants reported that text messages were their favorite method of communication as it felt more personal, was inexpensive ("free to receive!"), ubiquitous, and easy to use.

> *I also think like the cell phone is the easiest way because everyone is using a cell phone ...I think the cell phone SMS is the best.* [Postnatal woman, 28 years of age]

Most participants in focus group discussions were able to recall that the messages were sent twice a week, with all but one participant able to recite the precise time and days of the week.

> *Interviewer: So how often do you receive these SMS [messages]?*
>
> *Group response: Twice a week, Monday and Thursday, 9 o'clock exactly.*

Being able to count on the text messages to arrive on time, all the time, was mentioned by multiple participants. Two female participants in different focus group discussions mentioned that they waited for their Monday morning message before they decided to bring their infant to the clinic, to see if that day's message dealt with an issue that they had had over the weekend.

In South Africa, with 11 official languages, accessibility of the text message language was a concern among the implementers and was brought up at each focus group discussion. All focus group discussion participants claimed that English-language messages were not a barrier, but rather that English was the best option since it was easy to share with others who might not speak their mother tongue.

Intervention Feedback

Feedback regarding the intervention highlighted that it provided helpful, relevant information at appropriate times, was trustworthy, and was accepted.

Message content was brought up repeatedly throughout the focus group discussions as being an enabler of trust in the intervention and source of the messages. This trust is highlighted by discussions of the timeliness of messages and relevant to issues they were dealing with.

> *At a time my baby had a problem with her skin I received an SMS saying you can take that aqueous cream and just rub your baby and I did that I saw the skin of my baby change and I was so happy...I was so happy thank you so much...I am getting so much advice... Thank you.* [Postnatal woman, 20 years of age]

Other topics that were remembered by participants included nutrition during pregnancy, how to handle being pregnant and HIV-positive, learning about and preparing for delivery, understanding how to connect with a newborn, and dealing with teething. Participants also mentioned feeling more confident in caring for their infants as a result of the messages, and believed that the messages were trustworthy and were considered expert material.

Some participants identified their mothers as barriers to having healthy children. Two varying methods of dealing with this disconnect were shared; the first was explicitly telling their mother that the messages should be followed because it came from professionals and the second was telling their mother that their (the mother's) advice would be followed, but actually adhering to the message-based advice. The difference between which method was used was, in general, related to where their mothers were located. Women whose mothers were close by were more direct while women whose mothers were outside Johannesburg frequently used the second approach.

All participants showed an interest in continuing to receive SMS text messages after their baby reached one year of age. This was a recurring theme in all focus group discussion and was further evidenced by some focus group discussion participants claiming a willingness to pay for the messages, if necessary:

> *As now that I've received the SMS [text messages] I know about the SMS [text messages] already... if I have to pay I would pay because I know they are worthy, and SMS rate I would mind paying it.* [Prenatal woman, 33 years of age]

None of the focus group discussions brought up any negative comments about MAMA, nor did they have suggestions for improvement when probed for this.

Discussion

Principal Findings

In this qualitative study of user experiences of the MAMA maternal mHealth intervention, we found that, in contrast to poor experiences with the health system, MAMA maternal mHealth messages were considered to be reliable and useful. Despite mixed feelings regarding the quality of care provided by the health care system, participants were happy with the mHealth intervention and the content of the text messages. The text message intervention had an easy and discreet signup process, the use of text messages for communication was appropriate, and the content was accessible by the participants. The message content was reported to be relevant, trusted by the participants, and accepted by them.

Participants identified various barriers when trying to receive maternal health care and support. Each of these barriers may have been an indication of an underresourced health care system with high demand on staff. Participant feedback highlighted current health system barriers and disincentives for patients deciding how and when to attend health care facilities. Negative experiences at health care facilities, such as disrespectful staff and distrust of medical advice, among others, highlighted the need for a service that provides genuine, respectful, and trustworthy messaging for patients.

Focus group discussion participants generally had positive feedback and experiences of the SMS intervention because of the private recruitment, simple signup, and easy to understand messages that were relevant and timely. Most of the recruitment team were previously employed as HIV counselors which enabled them to be sensitive to the stigma that is related to HIV issues. This previous experience encouraged them to find recruitment methods that provided full patient confidentiality. Participants also reported having limited access to expert maternal health information outside of the intervention. Health care workers were seen as overworked and to only be able to provide limited support during maternal health care (antenatal care and postnatal follow-up care) visits. Thus, receiving timely and trustworthy maternal health information at no additional cost was seen as an enabler of good health, and it has, in fact, been shown to lead to better health [24]. Furthermore, given the national emphasis on HIV care, treatment, and support throughout South Africa over the last decade, the finding of a strong level of trust in the non-MAMA services provided by the public health care system was reassuring.

Comparison With Previous Work

Being outside the health care system enabled study staff to differentiate themselves from the health care system and its negative connotations. Additionally, the staff had the ability to focus entirely on the patient and provide as much information and support as was necessary to ensure the patient felt comfortable with the signup process. The combination of high HIV rates in the target population, significant HIV-associated stigma, and cramped waiting rooms where recruitment took place meant that recruiters had to be tactful. The depth of trust, relevance, and acceptability reported by participants was in line with another maternal mHealth study [15] that took place in rural northern Canada and which used focus group discussions and showed patient perception of the mHealth intervention to be highly acceptable and relevant. The same study [15] also found a high level of trust in the messaging that was provided; participants mentioned that they could believe the messages because of the source. This similarity shows the perceived pedigree of the message content is an important factor with regard to both the trust and the acceptability of an mHealth intervention.

The request for additional messages and the willingness to pay suggested that the text messages were not just accepted by focus group discussion participants but welcomed. Willingness to pay for mHealth services has not been studied extensively. While a willingness to pay was identified in this study, we feel this claim should be taken in context given the previously identified poverty-related issues. Participant acceptability of the intervention might be due to lack of stigma around pregnancy and infant care, and the relatively young age of participants, who were, by definition, of childbearing age. Watkins et al [20] suggested that acceptability of mHealth interventions could be based on age; older individuals reported having difficulty reading text on their phones and were less receptive to technology-based interventions. Additionally, maternal health is an area that has virtually no stigma associated with it, unlike other health conditions such as HIV. A qualitative study in Kenya [14] that looked at the acceptability of SMS text message–based HIV support reported that many individuals had concerns about the use of HIV-related terms and highlighted the potential for accidental disclosure of their HIV status. In contrast, MAMA was designed to support maternal and infant health. This might help reduce stigma, even though discussions of HIV were held privately, and this in turn may encourage women to get care for this critical health issue.

Participant feedback suggested that mHealth interventions may feel more compassionate than in-person visits. Lester et al [30] highlighted this same issue in their study which showed HIV-positive individuals who received SMS text messages from health care workers had improved clinical health outcomes compared to that of nonrecipients; many who received the messages reported that it seemed "like someone cares" (p 1843). Patient feelings that someone cares about them and their pregnancy could be a contributing factor to the positive effect demonstrated in previous MAMA research [24,25].

Limitations

This qualitative study included the feedback of only 15 participants in Johannesburg, South Africa, and might not be representative of the population as a whole. Additionally, there was potential for participants to be affected by social-desirability bias [32] as the focus group discussions were conducted close to the recruitment site. This could be a reason for the few critical statements about the intervention. Conversely, previous qualitative maternal mHealth studies [33,34] have shown that individuals tend to see mHealth based interventions in a positive light even before they are offered, which could partially explain the responses of participants in the current study. We are also aware that the thematic focus group discussion guide contains several closed questions which could have made the discussion

less free. Surprisingly, there were no negative comments regarding MAMA. This could be due to the participants perceiving the researchers as coming from MAMA, and therefore, not wanting to give any negative feedback; however, the participants were outspoken and seemed honest when discussing other topics. We believe the participants genuinely saw MAMA as an important tool that helped them during their pregnancy. Lastly, these results should not replace acceptability testing in other situations or among other populations.

Conclusions

Maternal mHealth interventions, delivered through text messages can provide timely, relevant, useful, and supportive information to pregnant women and new mothers, especially where mistrust in the health care system may exist. Maternal, neonatal, and child health is a field where this combination (timely, relevant, and supportive) is especially important and mHealth could be a tool used to attain maternal, neonatal, and child health goals, globally.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Examples of MAMA SMS.
[DOCX File , 20 KB - humanfactors_v7i2e14078_app1.docx]

Multimedia Appendix 2
Focus Group Discussion Guide.
[DOCX File , 21 KB - humanfactors_v7i2e14078_app2.docx]

References

1. Spietz A, Kelly J. The importance of maternal mental health during pregnancy: theory, practice, and intervention. Public Health Nurs 2002;19(3):153-155. [doi: 10.1046/j.0737-1209.2002.19301.x] [Medline: 11967099]
2. Byrom S, Edwards G, Bick D. Essential Midwivery Practice: Postnatal care. Chichester, United Kingdom: Wiley; 2009.
3. Standards for improving quality of maternal and newborn care in health facilities Internet. Geneva, Switzerland: World Health Organization; 2016.
4. Kerber KJ, de Graft-Johnson JE, Bhutta ZA, Okong P, Starrs A, Lawn JE. Continuum of care for maternal, newborn, and child health: from slogan to service delivery. Lancet 2007 Oct 13;370(9595):1358-1369. [doi: 10.1016/S0140-6736(07)61578-5] [Medline: 17933651]
5. Statistics SA. Millennium Development Goals: Country Report. Pretoria, South Africa: Statistics South Africa; 2015.
6. Millennium Development Goal 4:Reduce Child Mortality. Pretoria, South Africa: Statistics South Africa; 2015.
7. Pattinson R. Saving Mothers 2011-2013: Sixth report on the Confidential Enquiries into Maternal Deaths in South Africa. South Africa: National Committee on the Confidential Enquiries into Maternal Deaths, South Africa; 2014.
8. Kay M, Santos J, Takane M. mHealth: New horizons for health through mobile technologies. : World Health Organization Global Observatory for eHealth; 2011 Presented at: GSMA mHA Mobile Health Summit; 7 June 2011; Cape Town, South Africa p. 66-71 URL: http://www.who.int/entity/ehealth/mhealth_summit.pdf
9. Sondaal SFV, Browne JL, Amoakoh-Coleman M, Borgstein A, Miltenburg AS, Verwijs M, et al. Assessing the effect of mHealth interventions in improving maternal and neonatal care in low- and middle-income countries: a systematic review. PLoS One 2016 May;11(5):e0154664 [FREE Full text] [doi: 10.1371/journal.pone.0154664] [Medline: 27144393]
10. Lee SH, Nurmatov UB, Nwaru BI, Mukherjee M, Grant L, Pagliari C. Effectiveness of mHealth interventions for maternal, newborn and child health in low- and middle-income countries: Systematic review and meta-analysis. J Glob Health 2016 Jun;6(1):010401 [FREE Full text] [doi: 10.7189/jogh.06.010401] [Medline: 26649177]
11. Watterson JL, Walsh J, Madeka I. Using mHealth to improve usage of antenatal care, postnatal care, and immunization: a systematic review of the literature. Biomed Res Int 2015;2015:153402 [FREE Full text] [doi: 10.1155/2015/153402] [Medline: 26380263]

12. Poorman E, Gazmararian J, Parker RM, Yang B, Elon L. Use of text messaging for maternal and infant health: a systematic review of the literature. Matern Child Health J 2015 May;19(5):969-989. [doi: 10.1007/s10995-014-1595-8] [Medline: 25081242]

13. Saberi P, Siedle-Khan R, Sheon N, Lightfoot M. The use of mobile health applications among youth and young adults living with HIV: focus group findings. AIDS Patient Care STDS 2016 Jun;30(6):254-260 [FREE Full text] [doi: 10.1089/apc.2016.0044] [Medline: 27214751]

14. Ronen K, Unger JA, Drake AL, Perrier T, Akinyi P, Osborn L, et al. SMS messaging to improve ART adherence: perspectives of pregnant HIV-infected women in Kenya on HIV-related message content. AIDS Care 2018 Apr;30(4):500-505. [doi: 10.1080/09540121.2017.1417971] [Medline: 29254362]

15. Munro S, Hui A, Salmons V, Solomon C, Gemmell E, Torabi N, et al. SmartMom text messaging for prenatal education: a qualitative focus group study to explore Canadian women's perceptions. JMIR Public Health Surveill 2017 Feb 07;3(1):e7 [FREE Full text] [doi: 10.2196/publichealth.6949] [Medline: 28174149]

16. Huang S, Li M. Piloting a mHealth intervention to improve newborn care awareness among rural Cambodian mothers: a feasibility study. BMC Pregnancy Childbirth 2017 Oct 16;17(1):356 [FREE Full text] [doi: 10.1186/s12884-017-1541-z] [Medline: 29037179]

17. Labrique A, Vasudevan L, Chang LW, Mehl G. H_pe for mHealth: more "y" or "o" on the horizon? Int J Med Inform 2013 May;82(5):467-469 [FREE Full text] [doi: 10.1016/j.ijmedinf.2012.11.016] [Medline: 23279850]

18. Hao W, Hsu Y, Chen K, Li H, Iqbal U, Nguyen P, et al. LabPush: a pilot study of providing remote clinics with laboratory results via short message service (SMS) in Swaziland, Africa - a qualitative study. Comput Methods Programs Biomed 2015 Jan;118(1):77-83. [doi: 10.1016/j.cmpb.2014.10.005] [Medline: 25453385]

19. Wolff-Piggott B, Coleman J, Rivett U. The clinic-level perspective on mHealth implementation: a South African case study. Information Technology for Development 2017 Jan 06;24(3):532-553 [FREE Full text] [doi: 10.1080/02681102.2016.1233858]

20. Anstey Watkins JOT, Goudge J, Gómez-Olivé FX, Griffiths F. Mobile phone use among patients and health workers to enhance primary healthcare: A qualitative study in rural South Africa. Soc Sci Med 2018 Feb;198:139-147 [FREE Full text] [doi: 10.1016/j.socscimed.2018.01.011] [Medline: 29335160]

21. Den HW. Mama puts power in mom's hands. Brand South Africa Internet. URL: http://www.mediaclubsouthafrica.com/tech/3343-mama-puts-power-in-mom-s-hands [accessed 2017-06-05]

22. Wits RHI. Routine Monitoring and Evaluation Data. Wits Reproductive Health and HIV Institute. 2014. URL: http://www.wrhi.ac.za/uploads/files/Wits_RHI_Annual_Review_2013-2014.pdf [accessed 2020-06-09]

23. World Health Organization. Classification of digital health interventions 1.0. Geneva, Switzerland: World Health Organization; 2018. URL: https://apps.who.int/iris/bitstream/handle/10665/260480/WHO-RHR-18.06-eng.pdf;jsessionid=140F328DEF5031E678809731A3891D0C?sequence=1 [accessed 2020-06-09]

24. Coleman J, Bohlin KC, Thorson A, Black V, Mechael P, Mangxaba J, et al. Effectiveness of an SMS-based maternal mHealth intervention to improve clinical outcomes of HIV-positive pregnant women. AIDS Care 2017 Jul;29(7):890-897. [doi: 10.1080/09540121.2017.1280126] [Medline: 28107805]

25. Coleman J, Lau CV, Black V, Thorson A, Eriksen J. Improving maternal and infant health using mHealth text messages in a low-resource urban setting in South Africa. Unpublished 2021.

26. LeFevre A, Cabrera-Escobar MA, Mohan D, Eriksen J, Rogers D, Neo Parsons A, et al. Forecasting the value for money of mobile maternal health information messages on improving utilization of maternal and child health services in Gauteng, South Africa: cost-effectiveness analysis. JMIR Mhealth Uhealth 2018 Jul 27;6(7):e153 [FREE Full text] [doi: 10.2196/mhealth.8185] [Medline: 30054263]

27. 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]

28. Rees H, Delany-Moretlwe S, Scorgie F, Luchters S, Chersich MF. At the heart of the problem: health in Johannesburg's inner-city. BMC Public Health 2017 Jul 04;17(Suppl 3):554 [FREE Full text] [doi: 10.1186/s12889-017-4344-2] [Medline: 28832289]

29. Dedoose. Los Angeles, CA: Sociocultural Research Consultants LLC; 2017. URL: http://www.dedoose.com [accessed 2020-06-09]

30. 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]

31. Lester RT, Ritvo P, Mills EJ, Kariri A, Karanja S, Chung MH, et al. Effects of a mobile phone short message service on antiretroviral treatment adherence in Kenya (WelTel Kenya1): a randomised trial. Lancet 2010 Nov 27;376(9755):1838-1845. [doi: 10.1016/S0140-6736(10)61997-6] [Medline: 21071074]

32. Stewart D, Shamdasani P, Rook D. Focus Groups: Theory and Practice. Thousand Oaks, USA: Sage Publications; 2007:978-0761925828.

33. Cormick G, Kim NA, Rodgers A, Gibbons L, Buekens PM, Belizán JM, et al. Interest of pregnant women in the use of SMS (short message service) text messages for the improvement of perinatal and postnatal care. Reprod Health 2012;9:9. [doi: 10.1186/1742-4755-9-9] [Medline: 22866753]

XSL•FO
RenderX

34. Hmone MP, Dibley MJ, Li M, Alam A. A formative study to inform mHealth based randomized controlled trial intervention to promote exclusive breastfeeding practices in Myanmar: incorporating qualitative study findings. BMC Med Inform Decis Mak 2016 Jun 04;16:60 [FREE Full text] [doi: 10.1186/s12911-016-0301-8] [Medline: 27260252]

Abbreviations

MAMA: Mobile Alliance for Maternal Action
mHealth: mobile health

Original Paper

Comparison of the Effects of Automated and Manual Record Keeping on Anesthetists' Monitoring Performance: Randomized Controlled Simulation Study

Man-Kei Tse[1], BSocSci, MPhil; Simon Y W Li[1], BSc, MSc, PhD; Tsz Hin Chiu[2], MBBS; Chung Wai Lau[2], MBBS; Ka Man Lam[2], MBBS; Chun Pong Benny Cheng[2], MBBS

[1]Department of Applied Psychology, Lingnan University, Hong Kong, China (Hong Kong)
[2]Department of Anaesthesia and Intensive Care, Tuen Mun Hospital, Hong Kong, China (Hong Kong)

Corresponding Author:
Simon Y W Li, BSc, MSc, PhD
Department of Applied Psychology
Lingnan University
WYL-306, WYL Building
8 Castle Peak Road, Tuen Mun
Hong Kong,
China (Hong Kong)
Phone: 852 26167129
Email: simonli2@ln.edu.hk

Abstract

Background: Anesthesia information management systems (AIMSs) automatically import real-time vital signs from physiological monitors to anesthetic records, replacing part of anesthetists' traditional manual record keeping. However, only a handful of studies have examined the effects of AIMSs on anesthetists' monitoring performance.

Objective: This study aimed to compare the effects of AIMS use and manual record keeping on anesthetists' monitoring performance, using a full-scale high-fidelity simulation.

Methods: This simulation study was a randomized controlled trial with a parallel group design that compared the effects of two record-keeping methods (AIMS vs manual) on anesthetists' monitoring performance. Twenty anesthetists at a tertiary hospital in Hong Kong were randomly assigned to either the AIMS or manual condition, and they participated in a 45-minute scenario in a high-fidelity simulation environment. Participants took over a case involving general anesthesia for below-knee amputation surgery and performed record keeping. The three primary outcomes were participants' (1) vigilance detection accuracy (%), (2) situation awareness accuracy (%), and (3) subjective mental workload (0-100).

Results: With regard to the primary outcomes, there was no significant difference in participants' vigilance detection accuracy (AIMS, 56.7% vs manual, 56.7%; P=.50), and subjective mental workload was significantly lower in the AIMS condition than in the manual condition (AIMS, 34.2 vs manual, 46.7; P=.02). However, the result for situation awareness accuracy was inconclusive as the study did not have enough power to detect a difference between the two conditions.

Conclusions: Our findings suggest that it is promising for AIMS use to become a mainstay of anesthesia record keeping. AIMSs are effective in reducing anesthetists' workload and improving the quality of their anesthetic record keeping, without compromising vigilance.

KEYWORDS

Introduction

An anesthesia information management system (AIMS) is a computer-based system that automatically imports real-time vital signs from physiological monitors to replace traditional handwritten records [1] and is increasingly being adopted by hospitals [2]. Despite the increasing popularity of AIMSs, recent studies on AIMSs mainly addressed the completeness of

XSL•FO
RenderX

anesthetic records [3,4] but not the other attributes that are central to anesthetists' monitoring performance, such as situation awareness and mental workload. The purpose of this paper was to report a full-scale high-fidelity simulation that compared the effects of AIMS use and manual record keeping on anesthetists' monitoring performance.

Vigilance is the ability to maintain sustained attention over a long period of monitoring [5]. The most recent studies examining the effect of automated record keeping on vigilance were conducted 20 years ago [6,7]. Those studies focused on visual vigilance, which was operationalized as the time taken by participants to detect visual stimuli, including simulated abnormal values on a patient monitor [6] and flashing of an alarm light [7]. Anesthetists' vigilance was not affected when record keeping was carried out by machines or assistants [6].

Situation awareness refers to one's mental representation of the status of a dynamically changing environment. Situation awareness is measured at the following three levels: perception (level 1), comprehension (level 2), and projection (level 3) [8]. Situation awareness is critical to the administration of anesthesia because anesthetists need to monitor and be aware of numerous patient physiological variables (perception), detect unstable conditions and intervene appropriately (comprehension), and anticipate the effects of the intervention (projection) [9]. Situation awareness affects and is affected by mental workload, which is characterized as a subjective experience of the level of attentional demands imposed by performing tasks [10]. Noel suggested that anesthetists might become less attentive to the details of anesthetic events and patients' status when they do not have to scan patients' vital signs and write them down, as required in manual charting [11].

An AIMS would change the role of anesthetists from active processers of information to passive recipients [12,13]. As a result, anesthetists might be less attentive to the operating room (OR) surroundings and their patients' status during monitoring. However, an AIMS is expected to reduce anesthetists' subjective mental workload. Our three hypotheses specify that when compared with anesthetists who use manual record keeping, anesthetists who use AIMSs would have lower vigilance detection accuracy (H1), would have lower situation awareness accuracy (H2), and would experience lower subjective mental workload (H3).

Methods

Study Design and Approval

A parallel group experimental design was employed in this study. Ethical approval was obtained from Tuen Mun Hospital (TMH) (NTWC/CREC/17065) and Lingnan University (EC-063/1617). Written informed consent was obtained from all participants in advance and their data were deidentified.

Participants

Participants were recruited from among the members of the Anaesthesia and Intensive Care Unit, TMH between September 2017 and March 2018. Participants were eligible if they were resident trainees or specialists. Based on the limited availability of anesthetists, we included 10 participants in each of the two conditions (ie, AIMS and manual), with a total of 20 participants. To achieve simple randomization of group assignment, one experimenter (MKT) placed 10 red (representing the AIMS condition) and 10 green (representing the manual condition) stickers into an opaque envelope and then randomly drew a sticker to generate the allocation sequence. As soon as participants enrolled in the study, they were assigned to a condition according to the allocation sequence.

Simulation Design

A full-scale high-fidelity simulation was carried out in an OR at TMH. A clinical scenario specific for this study was designed by three anesthetists (THC, CPC, and KML). The scenario was designed to simulate uneventful monitoring with few critical incidents at intervals [14]. The scenario was set during the intraoperative portion of an emergency amputation below the right knee with general endotracheal anesthesia. It lasted for 45 minutes and comprised the following three phases: (1) preincident, (2) incident, and (3) postincident. The pre- and postincident phases were relatively uneventful, but the incident phase included the following three clinically relevant events: tourniquet pain, tourniquet deflation, and bleeding. The patient vital signs and progression were designed by an anesthetist (THC) and verified by a consultant anesthetist (CPC). When participants entered the simulation, they were asked to take over a case from a senior anesthetist (THC), who was a confederate in the study.

Apart from the participant, the simulation involved seven people, each with a specific role as follows: (1) senior anesthetist (THC); (2) runner nurse (a registered nurse colleague at TMH); (3) surgeon (CWL); (4) scrub nurse (KML); (5) patient simulator operator (CPC); and (6) two experimenters (MKT and SYWL). The confederates and the patient simulator operator were clinicians from TMH. The two experimenters were researchers from Lingnan University.

Each simulation session was recorded by two digital video recorders; one captured a general view of the OR (Figure 1A) and the other was head-mounted (GoPro Hero 5; GoPro, San Mateo, California, USA) to capture the participant's point of view (Figure 1B). A Fluke ProSim 8 Vital Signs Patient Monitor Simulator (Fluke Biomedical, Cleveland, Ohio, USA) was connected to a SimMan 3G (Laerdal Medical AS, Stavanger, Norway) patient simulator and a physiological monitor to display vital signs during the simulation.

Figure 1. Video capture from the perspective of the operating room (A) and participant (B) while the participant was entering data into the anesthesia information management system during the simulation scenario.

(A)   (B)
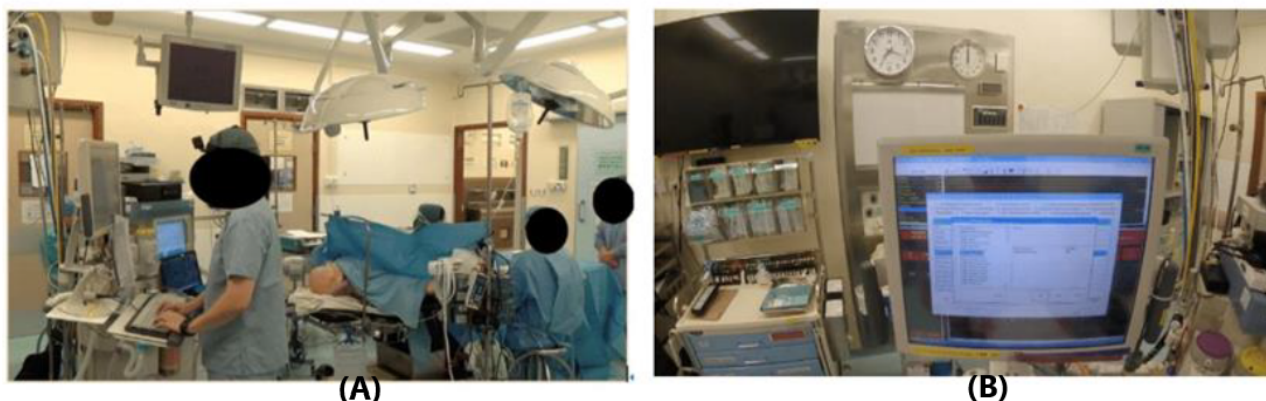
Before the simulation began, participants were given a briefing to introduce them to the purpose of the study. The participants were then informed about the role of each confederate and the function of the patient simulator. In a training session, participants were given instructions and demonstrations on how to respond to assessments of vigilance, situation awareness, and mental workload during the simulation. Participants in the manual condition were also trained on how to manually complete an anesthetic record, because resident anesthetists at the hospital use an AIMS in their usual work practice. The simulation began when the senior anesthetist completed the handover to the participant. The participants were debriefed when the simulation was completed.

Design of Situation Awareness Queries

The situation present assessment method (SPAM) [15] was used to measure participants' situation awareness. At predetermined moments of the simulation, the experimenter MKT called the participants' mobile phone to deliver situation awareness queries. The queries covered the three levels of situation awareness (perception, comprehension, and projection). For generating the situation awareness queries, we followed the process recommended by Endsley [16] to conduct a goal-directed task analysis (GDTA), which involved semistructured interviews, formulating a goal tree, and extracting and finally translating situation awareness requirements into scenario-specific queries. Details of the GDTA and situation awareness requirements are provided in Multimedia Appendix 1 and Multimedia Appendix 2, respectively. A total of nine situation awareness queries (Table 1) were generated with input from five anesthetists (CPC, KML, THC, an associate consultant, and a resident specialist).

Table 1. The nine situation awareness queries used in the scenario with their locations of information and their target answers.

Phase, Situation awareness queries	Location of the information	Target answer
Preincident		
Level 1: What is the level of hemoglobin of the patient?	Preoperative assessment	Approximately 11
Level 2: What is the most possible cause for the patient's hypertension?	• Physiological monitor (BP[a], baseline BP) • Understanding of the surgical procedure • Medical knowledge	Tourniquet pain
Level 3: If you do not provide any intervention, what would happen to the BP?	• Physiological monitor (BP, baseline BP) • Understanding of the surgical procedure • Medical knowledge	Increase
Incident		
Level 1: What is the patient's baseline BP?	• AIMS/manual record • Physiological monitor	125/80
Level 2: What is the most likely cause of the patient's hypotension?	• Physiological monitor (HR[b], BP) • Understanding of the surgical procedure • Medical knowledge	Bleeding/volume loss
Level 3: If you do not provide any intervention, what would happen to the end-tidal CO_2?	• Ventilator (CO_2, baseline CO_2, medical knowledge) • Understanding of the surgical procedure	Increase
Postincident		
Level 1: How much blood has the patient lost?	• Suction bottle (volume of blood) • Communication with nurses (volume of saline drip applied) • Blood gauze	500-700 mL (within ±5% is acceptable)
Level 2: Is the bleeding controlled? Why?	• Suction tubing sound • Suction bottle • Physiological monitor (BP, HR) • Surgical field (eg, blood gauze)	Yes, there is no more blood in suction tubing/HR and BP become normal
Level 3: If you do not provide any intervention, what would happen to the hemoglobin level?	• Medical knowledge • Understanding of the surgical procedure • Blood analysis	Increase. Not enough volume replacement, making the haemoglobin concentration higher. Or decrease. Due to severe blood loss

[a]BP: blood pressure.

[b]HR: heart rate.

Primary Outcomes

There were three primary outcomes as follows: (1) accuracy of detecting suction tubing sounds (ie, vigilance detection accuracy), which were sounds made from actual suction tubing controlled by the scrub nurse (KML); (2) accuracy of correctly answering scenario-specific situation awareness queries (ie, situation awareness accuracy); and (3) self-reported mental workload ratings on The National Aeronautics and Space Administration Task Load Index (NASA-TLX) [17]. Measurements of the primary outcomes were performed by the experimenters MKT and SWL at predetermined times during the 45-minute scenario. Figure 2 shows how the measures were distributed over the three phases of the scenario.

Figure 2. Design of the predetermined vital signs used in the clinical scenario of the simulation and the timeline of vigilance (V), situation awareness (S), and mental workload (W) assessments. BP: blood pressure; HR: heart rate.
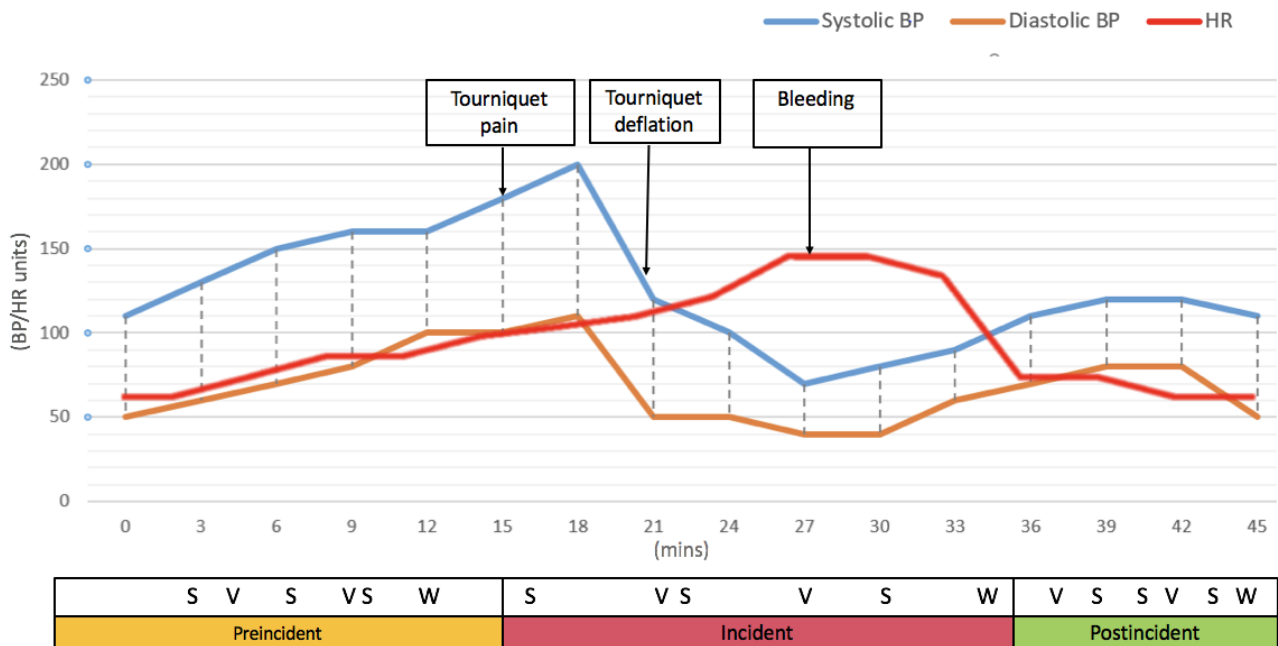
Secondary Outcomes

The secondary outcomes involved the distribution of the participants' time across different task activities (ie, task time distribution), the quality of their anesthesia record (ie, anesthesia record completeness), and their attitude toward the AIMS. We assessed participants' attitude toward the AIMS in terms of trust and acceptance, using a 45-item questionnaire (Multimedia Appendix 3) after the simulation was completed.

Statistical Analysis

Operationalization of the Primary Outcomes

Vigilance was operationalized as detection accuracy for each participant. The score was calculated as the proportion (%) of the six tubing sounds that a participant detected. Situation awareness was operationalized as response accuracy, which was calculated as the proportion (%) of the nine situation awareness queries that the participant answered correctly. Each participant's answers to the situation awareness queries were first evaluated against a predetermined marking scheme. When an answer did not match the target answer, an anesthetist (THC), who was blinded to the condition allocation, helped determine the accuracy of the answer according to expert judgement.

We performed the subjective mental workload measurement at the end of each simulation phase, in which participants rated each NASA-TLX dimension on a scale from 0 (lowest) to 100 (highest). The NASA-TLX comprises six dimensions (mental demand, physical demand, temporal demand, effort, frustration, and performance). The mean overall TLX score for each participant was calculated across the three simulation phases.

Operationalization of the Secondary Outcomes

Participants' task activities in the simulation were video recorded and were reviewed to extract data on the different task

activities. Task time distribution for each individual task category was computed as a percentage of the time spent on that category over the total time for all four tasks, including (1) entering record data, (2) monitoring the patient (eg, looking at the patient record, physiological monitor, anesthetic gas machine, or simulated patient), (3) performing patient care activities (eg, administering medication into patient's intravenous access), and (4) interacting with the surgical team (eg, talking to the surgeon, asking the runner nurse to order medication, etc). Data were not coded for tasks that did not fall into any of the four task categories (eg, tidying up equipment wires, walking around the OR, etc).

Two raters assessed the participants' anesthetic records for completeness using the 15-item checklist by Edwards et al [4], which was modified from the Australian and New Zealand College of Anesthetists' recommendations on anesthetic records [18]. The two raters were an anesthetist (THC) and a consultant anesthetist (CPC), and they scored each checklist item with 1 (present), 0.5 (partially present), or 0 (absent) for the anesthetic records. The scoring was carried out by the raters independent from each other. The scores of individual checklist items were summed to produce a total score for each anesthetic record.

The trust and acceptance questionnaire had the following two parts: "trust in the AIMS" (adapted from a scale on trust in automated systems [19]) and "acceptance of the AIMS" (adapted from a scale based on the technology acceptance model [20-22]). All items in the questionnaire were rated on a 5-point Likert scale, with 1 indicating *strongly disagree*, 2 indicating *disagree*, 3 indicating *neutral*, 4 indicating *agree*, and 5 indicating *strongly agree*. Separate mean scores for trust and acceptance were calculated for each participant.

Prior to analysis, the Shapiro-Wilk test and Levene test were performed to assess the normality and homogeneity of variance,

respectively, of the studentized residuals of the data. The independent sample *t* test was used to compare differences between the manual and AIMS conditions for normally distributed data. The Mann-Whitney *U* test was performed for non-normally distributed data.

According to the directions of the hypotheses, one-tailed significance tests were performed for the primary outcomes, whereas two-tailed tests were performed for the secondary outcomes. Task time distributions of the four tasks were compared between the two conditions with Bonferroni correction to obtain a more stringent alpha level of .0125 (.05/4).

Results

Response Rate

All 20 participants completed the trials without any dropout (Figure 3). Participants in the AIMS condition and those in the manual condition had comparable years of experience in anesthesia, with mean experience durations of 3.4 and 3.2 years, respectively.

Figure 3. CONSORT disgram for the simulation study.
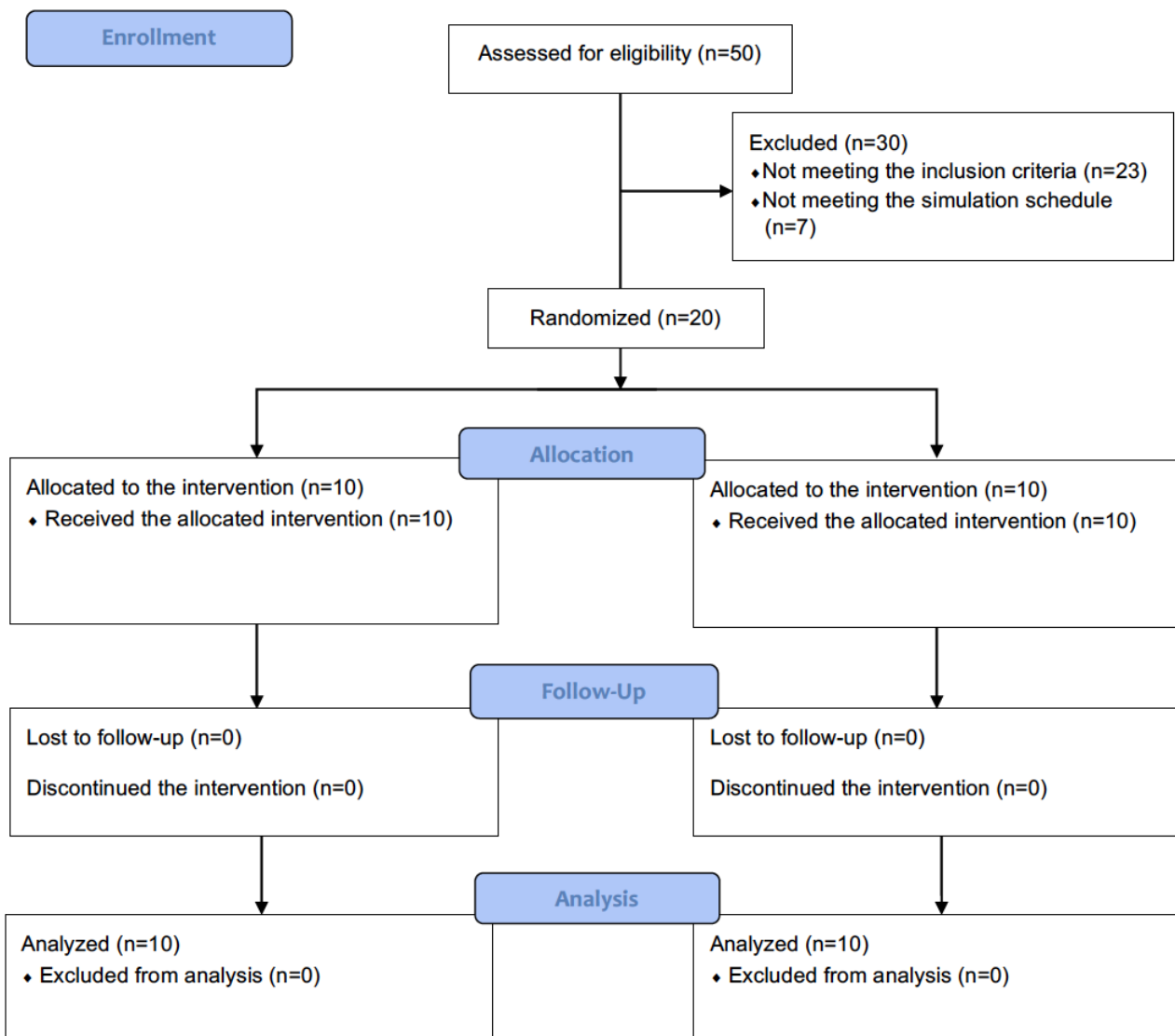
Primary Outcomes

There was no significant difference in vigilance accuracy between the AIMS (mean 56.7%, SD 32.6%) and manual conditions (mean 56.7%, SD 31.6%) (t_{18}=0.00, *P*=.50, one-tailed); therefore, H1 was not supported. Although there was no significant difference in situation awareness accuracy between the AIMS (median 88.9%, range 66.7%-100%) and manual conditions (median 88.9%, range 77.8%-100%) (U=40.5, *P*=.48), we carried out a post-hoc power analysis using G*Power

[23] on the basis of an emerging difference in trend between the two conditions. The achieved power (1 – β) calculated was 0.13, which was below the lowest conventionally acceptable level of 0.8. This suggests that the study did not have enough power to detect a difference in situation awareness accuracy between the AIMS and manual conditions. Therefore, H2 was inconclusive. However, we found that participants reported a significantly lower overall TLX score in the AIMS condition (mean 34.2, SD 12.5) than in the manual condition (mean 46.7,

SD 11.5) (t_{18}=−2.34, P=.02, one-tailed). Therefore, H3 was supported.

Secondary Outcomes

Some video data were not coded (30% in the AIMS condition and 26% in the manual condition), as they either could not be classified or involved tasks that did not fall into our predefined task categories. Of the data that were coded according to the four task categories, only the proportion of time spent on record data entry differed significantly between the AIMS (mean 26.0%, SD 4.9%) and manual conditions (mean 33.7%, SD 6.9%) (t_{18}=−2.87, P=.01, two-tailed). We also found that the level of completeness of anesthetic records was significantly higher in the AIMS condition (median 100%, range 93%-100%) than in the manual condition (median 75%, range 55%-87%) (U=0.000, P<.001, two-tailed). The two raters for record completeness had a high degree of reliability, with an average intraclass correlation coefficient of 0.893 and a 95% CI ranging from 0.68 to 0.96 ($F_{19,19}$=11.59, P<.001). Finally, data from the trust and acceptance survey indicated that 45% (9/20) of respondents showed a positive attitude (*agree* or *highly agree*) of trust toward the AIMS and the remaining 55% (11/20) showed a neutral attitude. In terms of acceptance, 90% (18/20) of respondents showed a positive attitude (*agree* or *highly agree*) and 10% (2/20) showed a neutral attitude.

Discussion

Overall Findings

Despite the increasing adoption rate of AIMSs in hospitals [2], their effect on the monitoring performance of anesthetists has not been thoroughly examined. This study compared the effects of AIMS use and manual record keeping in terms of anesthetists' levels of vigilance, situation awareness, and subjective mental workload with a randomized controlled trial in a high-fidelity simulation setting. The primary outcomes indicated that while there was no relevant difference in participants' vigilance between AIMS use and manual record keeping, subjective mental workload was much lower among participants using the AIMS than among those using the manual method. However, the effect on situation awareness accuracy was inconclusive because the study was under-powered to detect its difference between the two conditions.

AIMS use might have two advantages over manual record keeping with respect to mental workload. First, the lower subjective mental workload with AIMS use might be a product of reduced physical movements. Informal inspection of our GoPro video data revealed that participants in the manual condition exhibited extensive head movements owing to the shifting of attention between the physiological monitor and the paper anesthesia chart. These movements may imply that more cognitive and perceptual activities (eg, remembering, looking, and searching for information) are involved in manual record keeping, and thereby, they result in higher subjective mental workload. Second, manual record keeping might have placed a high demand on participants' prospective memory (remembering a future task) [24], because they needed to remind themselves to update vital signs on the paper chart regularly.

The secondary outcomes indicated further benefits of AIMS use. First, participants who used the AIMS spent about 8 percentage points less of their total time on record data entry than those who used manual record keeping. This result confirms previous findings that electronic record keeping allows anesthesia residents to spend less time on record keeping as compared to that with manual record keeping [7]. Second, AIMS use produced more complete anesthetic records than those produced by manual record keeping. This finding is consistent with the result of a previous study that retrospectively assessed 400 anesthetic records created by AIMS or manual record keeping methods [4] and reported more complete AIMS records than manual records. It is likely that AIMS use spares anesthetists from charting patients' vital signs and allows them to spend more time on including other required information in the anesthetic records. Third, the attitude survey of AIMS use indicated that participants had a positive attitude toward trusting and accepting AIMS use in their practice.

Compared with previous studies on AIMS use that only examined visual vigilance [6,7], our study tested auditory vigilance. In this study, vigilance was operationalized as participants' accuracy of detecting suction tubing sounds. This stimulus was chosen based on its clinical relevance, given that anesthetists often interpret it as a sign of patient blood loss during surgery. Although a direct comparison to visual vigilance might be impossible, our current results and those from previous studies suggest that AIMS use does not harmfully decrease anesthetists' vigilance level [6,7]. However, irrespective of the type of record keeping, participants in this study demonstrated only a fair vigilance level in that they only detected, on average, 3.2 out of all 6 suction sounds (54%) in the vigilance assessments. We had not anticipated this result, but given the clinical importance of detecting suction sounds, this should be further investigated in future studies.

Limitations

This study had six limitations. First, our simulated scenario only represented anesthetic cases that involve an uneventful period followed by critical incidents. Therefore, our findings can only be applied to the context of anesthesia with critical incidents. In anesthesia, many cases occur without any critical events. When the anesthetic procedure is uneventful, the effect of AIMS use on anesthetists' vigilance and situation awareness might be different because complacency might arise, and this warrants further investigation.

Second, our participants were more accustomed to AIMS use than manual record keeping in their usual practice because junior anesthetists at TMH are trained on the AIMS but not on manual record keeping. Therefore, participants in our simulation had to be retrained on manual record keeping for comparison. While this retraining might seem artificial, it was the aim of TMH's Department of Anaesthesia & ICU to investigate the tacit assumption of the effectiveness of AIMS use over manual record keeping. Retraining in the manual condition might have increased participants' perceived mental workload, degraded their vigilance, and decreased their record keeping efficiency. This possible confounding factor could be addressed in future

studies by sample screening or providing participants with prolonged training in manual record keeping.

Third, the findings of our study cannot be generalized to all models or brands of AIMSs. Other models of AIMSs might have different functions or interfaces and might interact with anesthetists differently.

Fourth, the participants, experimenters, and confederates were not blinded to the condition assigned to each participant owing to the nature of the manual and automated record keeping conditions.

Fifth, although our results suggest that AIMS use reduced the time spent on record data entry, it is unclear whether the time reduction led to an increase in time spent on monitoring patients or performing patient care activities. This could be addressed in future studies by examining how anesthetists reallocate the time saved with AIMS use to other tasks.

Sixth, we used a GoPro camera attached to each participant's head in an attempt to capture visual data. However, the GoPro camera, at its best, could only provide us with the participant's gaze direction. If accurate visual attention data are to be gathered, a mobile eye tracker should be used in future studies. Eye tracking data would allow for not only better inference of participants' visual attention in general, but also identification of what activities they focus on when not interacting with the AIMS.

Conclusions

Despite the increasing popularity of AIMSs in hospitals, no previous studies have analyzed their effects on comprehensive monitoring performance. The findings of this study provide support for the adoption of AIMSs in the OR by demonstrating a number of benefits of AIMS use, including reducing anesthetists' perceived mental workload, saving their time spent on data entry, and producing complete anesthetic records, without compromising vigilance. Moreover, the majority of our anesthetists expressed a positive attitude toward trusting and accepting AIMSs in the OR.

The level of automation in health care is likely to increase as medical technology advances. It is important to know the effects that automation will have on patient care, as it could affect clinicians' care quality and, ultimately, patients' well-being and safety.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The goal-directed task analysis.
[PDF File (Adobe PDF File), 556 KB - humanfactors_v7i2e16036_app1.pdf]

Multimedia Appendix 2
Pooled situation awareness requirements for the scenario.
[PDF File (Adobe PDF File), 417 KB - humanfactors_v7i2e16036_app2.pdf]

Multimedia Appendix 3
Questionnaire for trust and acceptance of anesthesia clinical information systems.
[PDF File (Adobe PDF File), 784 KB - humanfactors_v7i2e16036_app3.pdf]

Multimedia Appendix 4
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 2251 KB - humanfactors_v7i2e16036_app4.pdf]

References

1. Kadry B, Feaster W, Macario A, Ehrenfeld J. Anesthesia information management systems: past, present, and future of anesthesia records. Mt Sinai J Med 2012;79(1):154-165. [doi: 10.1002/msj.21281] [Medline: 22238048]

XSL•FO
RenderX

2.   Stol I, Ehrenfeld J, Epstein R. Technology Diffusion of Anesthesia Information Management Systems into Academic Anesthesia Departments in the United States. Anesthesia & Analgesia 2014;118(3):644-650 [FREE Full text] [doi: 10.1213/ane.0000000000000055]

3.   Driscoll WD, Columbia MA, Peterfreund RA. An observational study of anesthesia record completeness using an anesthesia information management system. Anesth Analg 2007 Jun;104(6):1454-1461. [doi: 10.1213/01.ane.0000264082.54561.d8] [Medline: 17513641]

4.   Edwards K, Hagen S, Hannam J, Kruger C, Yu R, Merry A. A randomized comparison between records made with an anesthesia information management system and by hand, and evaluation of the Hawthorne effect. Can J Anaesth 2013 Oct;60(10):990-997. [doi: 10.1007/s12630-013-0003-y] [Medline: 23934747]

5.   Davies DR, Parasuraman R. The Psychology Of Vigilance (Organizational and Occupational Psychology). London: Academic Press; 1982.

6.   Loeb R. Manual record keeping is not necessary for anesthesia vigilance. J Clin Monit 1995 Jan;11(1):9-13. [doi: 10.1007/bf01627413] [Medline: 7745463]

7.   Weinger M, Herndon O, Gaba D. The effect of electronic record keeping and transesophageal echocardiography on task distribution, workload, and vigilance during cardiac anesthesia. Anesthesiology 1997 Jul;87(1):144-55; discussion 29A. [doi: 10.1097/00000542-199707000-00019] [Medline: 9232145]

8.   Endsley M. Toward a Theory of Situation Awareness in Dynamic Systems. Hum Factors 2016 Nov 23;37(1):32-64 [FREE Full text] [doi: 10.1518/001872095779049543]

9.   Drews F, Westenskow D. The right picture is worth a thousand numbers: data displays in anesthesia. Hum Factors 2006;48(1):59-71. [doi: 10.1518/001872006776412270] [Medline: 16696257]

10. Tsang PS, Vidulich MA. Mental workload and situation awareness. In: Salvendy G, editor. Handbook of Human Factors and Ergonomics. Hoboken, NJ: Wiley; 2012:243-273.

11. Noel TA. Computerized anesthesia records may be dangerous. Anesthesiology 1986 Feb;64(2):300. [doi: 10.1097/00000542-198602000-00046] [Medline: 3946826]

12. Weinger MB, Smith NT. Vigilance, alarms, and integrated monitoring systems. In: Ehrenwerth J, Eisenkraft JB, editors. Anesthesia Equipment: Principles and Applications. St. Louis, MO: Mosby; 1993:350-384.

13. Endsley M, Kiris E. The Out-of-the-Loop Performance Problem and Level of Control in Automation. Hum Factors 2016 Nov 23;37(2):381-394 [FREE Full text] [doi: 10.1518/001872095779064555]

14. Slagle J, Weinger M. Effects of intraoperative reading on vigilance and workload during anesthesia care in an academic medical center. Anesthesiology 2009 Feb;110(2):275-283. [doi: 10.1097/ALN.0b013e318194b1fc] [Medline: 19194155]

15. Durso FT, Dattel AR, Banbury S, Tremblay S. SPAM: The real-time assessment of SA. In: Banbury S, Tremblay S, editors. A cognitive approach to situation awareness: Theory and application. Aldershot, England: Ashgate; 2004:137-154.

16. Endsley MR. A Survey of Situation Awareness Requirements in Air-to-Air Combat Fighters. The International Journal of Aviation Psychology 1993 Apr;3(2):157-168. [doi: 10.1207/s15327108ijap0302_5]

17. Hart S, Staveland L. Development of NASA-TLX (Task Load Index): Results of empiricaltheoretical research. In: Hancock PA, Meshkati N, editors. Advances in psychology. Amsterdam, The Netherlands: North-Holland; 1988:139-183.

18. Australian and New Zealand College of Anesthetists. 2006. The anesthesia record recommendations on the recording of an episode of anesthesia care URL: http://www.anzca.edu.au/documents/ps06-2006-the-anaesthesia-record-recommendations-o [accessed 2018-03-30]

19. Jian J, Bisantz AM, Drury CG. Foundations for an Empirically Determined Scale of Trust in Automated Systems. International Journal of Cognitive Ergonomics 2000 Mar;4(1):53-71. [doi: 10.1207/s15327566ijce0401_04]

20. Davis F. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly 1989 Sep;13(3):319 [FREE Full text] [doi: 10.2307/249008]

21. Venkatesh V, Davis F. A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. Management Science 2000 Feb;46(2):186-204 [FREE Full text] [doi: 10.1287/mnsc.46.2.186.11926]

22. Gagnon MP, Orruño E, Asua J, Abdeljelil AB, Emparanza J. Using a modified technology acceptance model to evaluate healthcare professionals' adoption of a new telemonitoring system. Telemed J E Health 2012;18(1):54-59 [FREE Full text] [doi: 10.1089/tmj.2011.0066] [Medline: 22082108]

23. Faul F, Erdfelder E, Lang A, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007 May;39(2):175-191. [doi: 10.3758/bf03193146] [Medline: 17695343]

24. McDaniel M, Einstein G. Prospective memory: An overview and synthesis of an emerging field. Thousand Oaks, CA: SAGE Publications; 2007.

Abbreviations

AIMS: anesthesia information management system
GDTA: goal-directed task analysis
NASA-TLX: National Aeronautics and Space Administration's Task Load Index
OR: operating room

SPAM: situation present assessment method
TMH: Tuen Mun Hospital

XSL·FO
RenderX

<u>Original Paper</u>

A Cyber-Physical System for Near Real-Time Monitoring of At-Home Orthopedic Rehabilitation and Mobile–Based Provider-Patient Communications to Improve Adherence: Development and Formative Evaluation

Timothy Stevens[1], MS; Ryan S McGinnis[2], PhD; Blake Hewgill[2], BS; Rebecca H Choquette[3], ATC; Timothy W Tourville[3,4], ATC, PhD; Jean Harvey[5], PhD; Richard Lachapelle[6], MS; Bruce D Beynnon[3], MS, PhD; Michael J Toth[3,6,7], PhD; Christian Skalka[1], PhD

[1]Department of Computer Science, University of Vermont, Burlington, VT, United States

[2]Department of Electrical and Biomedical Engineering, University of Vermont, Burlington, VT, United States

[3]Department of Orthopedics and Rehabilitation, University of Vermont, Burlington, VT, United States

[4]Department of Rehabilitation and Movement Sciences, University of Vermont, Burlington, VT, United States

[5]Department of Nutrition and Food Sciences, University of Vermont, Burlington, VT, United States

[6]Department of Molecular Physiology and Biophysics, University of Vermont, Burlington, VT, United States

[7]Department of Medicine, University of Vermont, Burlington, VT, United States

Corresponding Author:
Michael J Toth, PhD
Department of Medicine
University of Vermont
Health Science Research Facility Rm 126B
149 Beaumont Ave
Burlington, VT, 05405
United States
Phone: 1 (802) 656 7989
Fax: 1 (802) 656 0747
Email: michael.toth@med.uvm.edu

Abstract

Background: Knee extensor muscle performance is reduced after lower extremity trauma and orthopedic surgical interventions. At-home use of neuromuscular electrical stimulation (NMES) may improve functional recovery, but adherence to at-home interventions is low. Greater benefits from NMES may be realized with closer monitoring of adherence to at-home prescriptions and more frequent patient-provider interactions.

Objective: This study aimed to develop a cyber-physical system to monitor at-home adherence to NMES prescription and facilitate patient-provider communications to improve adherence in near real time.

Methods: The RehabTracker cyber-physical system was developed to accomplish this goal and comprises four components: (1) hardware modifications to a commercially available NMES therapy device to monitor device use and provide Bluetooth functionality; (2) an iPhone Operating System–based mobile health (mHealth) app that enables patient-provider communications in near real time; (3) a clinician portal to allow oversight of patient adherence with device use; and (4) a back-end server to store data, enable adherence analysis, and send automated push notifications to the patient. These four elements were designed to be fully compliant with the Health Insurance Portability and Accountability Act. The system underwent formative testing in a cohort of patients following anterior cruciate ligament rupture (n=7) to begin to assess face validity.

Results: Compared with the NMES device software–tracked device use, the RehabTracker system recorded 83% (40/48) of the rehabilitation sessions, with 100% (32/32) of all sessions logged by the system in 4 out of 7 patients. In patients for whom tracking of automated push notifications was enabled, 100% (29/29) of the push notifications sent by the back-end server were received by the patient. Process, hardware, and software issues contributing to these inaccuracies are detailed.

XSL•FO

RenderX

Conclusions: RehabTracker represents a promising mHealth app for tracking and improving adherence with at-home NMES rehabilitation programs and warrants further refinement and testing.

Introduction

Background

Traumatic injury to the knee joint, including rupture of the anterior cruciate ligament (ACL), is common and highly debilitating [1]. Despite surgical reconstruction and rehabilitation, many patients suffer muscle weakness following the index trauma and surgical intervention that persists for years after surgery [2,3] and are not satisfied with their knee functionality [4]. Current rehabilitation regimens are designed to restore muscle function to preinjury or presurgery levels but are only marginally effective [5]. This may be due, in part, to how pain, impaired neural activation, restricted range of knee motion, and risk for damaging the healing ACL graft limit the rehabilitation regimens available in the early, postinjury, and postsurgical periods. There is a need for improved rehabilitation modalities that can be used at these times to mitigate atrophy and weakness and improve long-term function.

Neuromuscular electrical stimulation (NMES) is an ideal candidate intervention for the early postinjury and postsurgical periods. This patient-directed therapy uses a portable, hand-held device to initiate muscle contraction by passing current through electrodes placed over the muscle of interest. NMES is effective at preventing skeletal muscle atrophy caused by experimentally induced muscle disuse in closely monitored research studies [6,7] and is approved by the Food and Drug Administration (FDA) for this indication following injury and surgery. In fact, studies show that it prevents quadriceps weakness and atrophy following ACL rupture and surgical reconstruction [8,9] and enhances long-term functional recovery [10]. However, use of NMES in orthopedic patients in the postinjury and early postsurgical periods may be limited by the need for associated costly outpatient clinic visits. Although NMES devices are amendable to home use, low adherence to home-based rehabilitation interventions [11-13] have tempered enthusiasm for its use in this setting. Some NMES devices allow for covert monitoring of adherence in the device software that provides the capacity to track at-home NMES use; however, as this oversight is retrospective, there is no opportunity to intervene and correct nonadherence as it occurs. New tools for administering and monitoring rehabilitation may improve adherence with at-home interventions such as NMES by allowing closer provider oversight and facilitating patient-provider interactions to address specific adherence issues.

Sensors and communication equipment can be coupled to medical devices to form cyber-physical systems that allow for near real-time monitoring of treatment adherence and clinical status as well as provide novel opportunities for patient-provider communication. As of 2018, over 95% of adults in the United States owned a cell phone [14], suggesting tremendous potential for cyber-physical systems to improve treatment adherence and monitor health outcomes. Such systems have been developed to monitor glucose levels in diabetics [15] and posture and joint loading in patients following hip surgery [16], to name a few. In patients recovering from musculoskeletal injury and surgery, efforts to develop cyber-physical systems to aid rehabilitation have been limited. Recent reports describe Web-based resources to assist patients with self-guided, at-home orthopedic rehabilitation [17]. To our knowledge, however, no reports have described the construction of a cyber-physical system to support adherence monitoring of at-home rehabilitation with NMES.

Objectives

To address this technological gap as well as the clinical needs of patients and rehabilitation professionals, we sought to develop a cyber-physical system, comprising an instrumented medical device, a mobile health (mHealth) app, and back-end server architecture, to monitor and improve adherence with at-home NMES therapy. Along with the basic functionality, a crucial feature of our system is compliance with the Health Insurance Portability and Accountability Act (HIPAA), which is a law that requires strict protections for the storage of and access to protected health information. In addition, we performed initial formative testing of the system in patients following ACL injury to assess its functionality in a real-world setting and discern whether the system would be suitable for further development and testing.

Methods

System Components

The RehabTracker cyber-physical system comprises four main components (Figure 1): (1) the modified NMES device, (2) the iPhone Operating System (iOS) mobile app (mHealth app), (3) the back-end server, and (4) the clinician portal Web interface. As patients perform NMES with the modified device, rehabilitation session data are recorded. After each session, patients use the RehabTracker app to transfer the session data from the embedded device to the database. In turn, the Web-based data are viewed by the clinician and used for automated adherence tracking and push notifications. The following sections detail each component of the system.

Figure 1. Simplified overview of the components of the RehabTracker cyber-physical system and user interactions. The RehabTracker mobile health (mHealth) app receives and transmits neuromuscular electrical stimulation (NMES) use data and serves as a conduit for patient-provider interactions. A secure, Health Insurance Portability and Accountability Act–compliant, back-end server receives the device use data, displays the adherence data for care provider review, and sends automated push notifications to the mHealth app, with the goal of improving adherence to the NMES prescription. NMES: neuromuscular electrical stimulation.
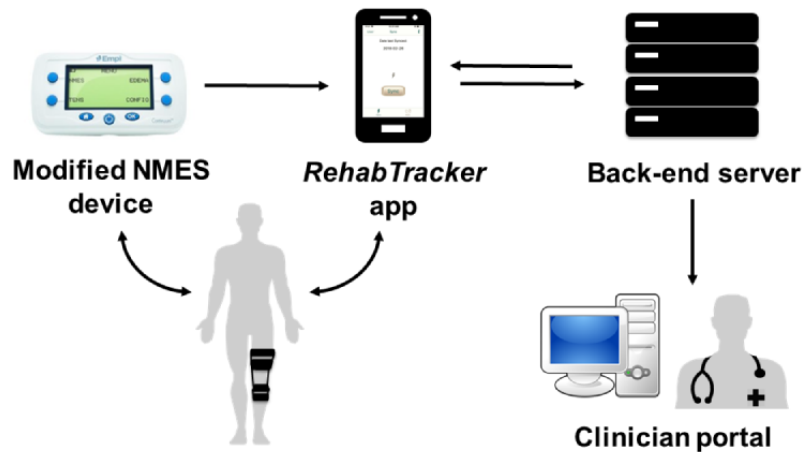
Neuromuscular Electrical Stimulation Hardware and Software Development

The core of this system is the Empi Continuum (Figure 2), an FDA-cleared multifunctional electrotherapy device that offers adjunctive electrophysical rehabilitation therapies, including the NMES therapy considered herein. The main goals of the modifications to this system were (1) to render it capable of tracking device use and (2) communicate data to a companion iOS app via Bluetooth 4.0.

The instrumentation to achieve both these goals is built upon the RedBearLab's Blend BLE (Figure 2), a small development board that includes an integrated microcontroller and Bluetooth 4.0 module. The four output leads of the EMPI device are passed through a custom rectifier and voltage divider circuit (Figure 2) and sampled through two analog inputs of the Blend. The resulting signal provides a direct measure of device activity that is logged quantitatively using custom firmware on the device. The Blend was also integrated with a real-time clock to provide absolute time stamps that were used to characterize the duration of each rehabilitation session. All this information is communicated to a mobile phone via Bluetooth low energy.

The combined Blend-EMPI system is powered by two AA rechargeable nickel metal hydride batteries and is controlled by a master switch (Figure 2), ensuring that a rehabilitation session cannot be completed without also being tracked by the monitoring hardware. A step-up regulator (Figure 2) is used to provide the requisite 5 V to the Blend. The EMPI's internal low-voltage cutoff is retained, ensuring that the system cannot be used if the batteries are no longer capable of providing

sufficient power to enable a standardized rehabilitation session. If this state is reached, an indicator light-emitting diode (Figure 2) will not illuminate when the master power switch is engaged. The Blend system is enclosed within a 3D-printed housing that is secured to the back of the EMPI (Figure 2). The batteries are housed in an external enclosure that allows for easy replacement by the user and acts as a *kickstand* for the device, when in use (Figure 2).

The Blend firmware serves two purposes: (1) processes the voltages from the NMES device into session data and (2) sends the data to the RehabTracker mHealth app via Bluetooth. During a therapy session, the NMES device outputs alternating waves of high and low voltage, which trigger muscle contraction and relaxation, respectively. When enabled, the Blend constantly monitors the NMES device voltage and automatically identifies the beginning and the end of a therapy session as recorded by an onboard real-time clock based on a simple threshold-based state machine. As a measure of the session intensity, an average maximum voltage variable is calculated by determining the peak voltage achieved during each muscle contraction cycle and averaging across the session. When a session is completed, the data for that session (start time, end time, and average peak voltage) are stored in the Blend's local storage until a sync is initiated by the user. Data storage space is not an issue as the Blend's 256 KB local storage can hold the data of far more sessions (approximately 24 bytes per session) than a patient would complete. Once a sync is initiated by the user, stored session data are sent to the RehabTracker mHealth app and deleted from the Blend's storage. Data transmission from the modified NMES device to the mobile app is enabled by the Blend firmware and is transferred in a custom format.

Figure 2. Modified neuromuscular electrical stimulation system can track the duration, intensity, and timing of rehabilitation sessions (labeled as "a"). Exploded view of the system, comprising a RedBearLab's Blend BLE (labeled as "c-i") and custom circuitry for quantifying device usage (labeled as "c-ii") and integrated within a 3D-printed enclosure secured to the back of the EMPI (labeled as "b-iii"). Power is provided by two AA batteries secured in an external housing (labeled as "b-iv") and is controlled by a master switch (labeled as "b-i") and step-up regulator (labeled as "c-iii"). An external light-emitting diode (labeled as "b-ii") indicates when the device is powered on.
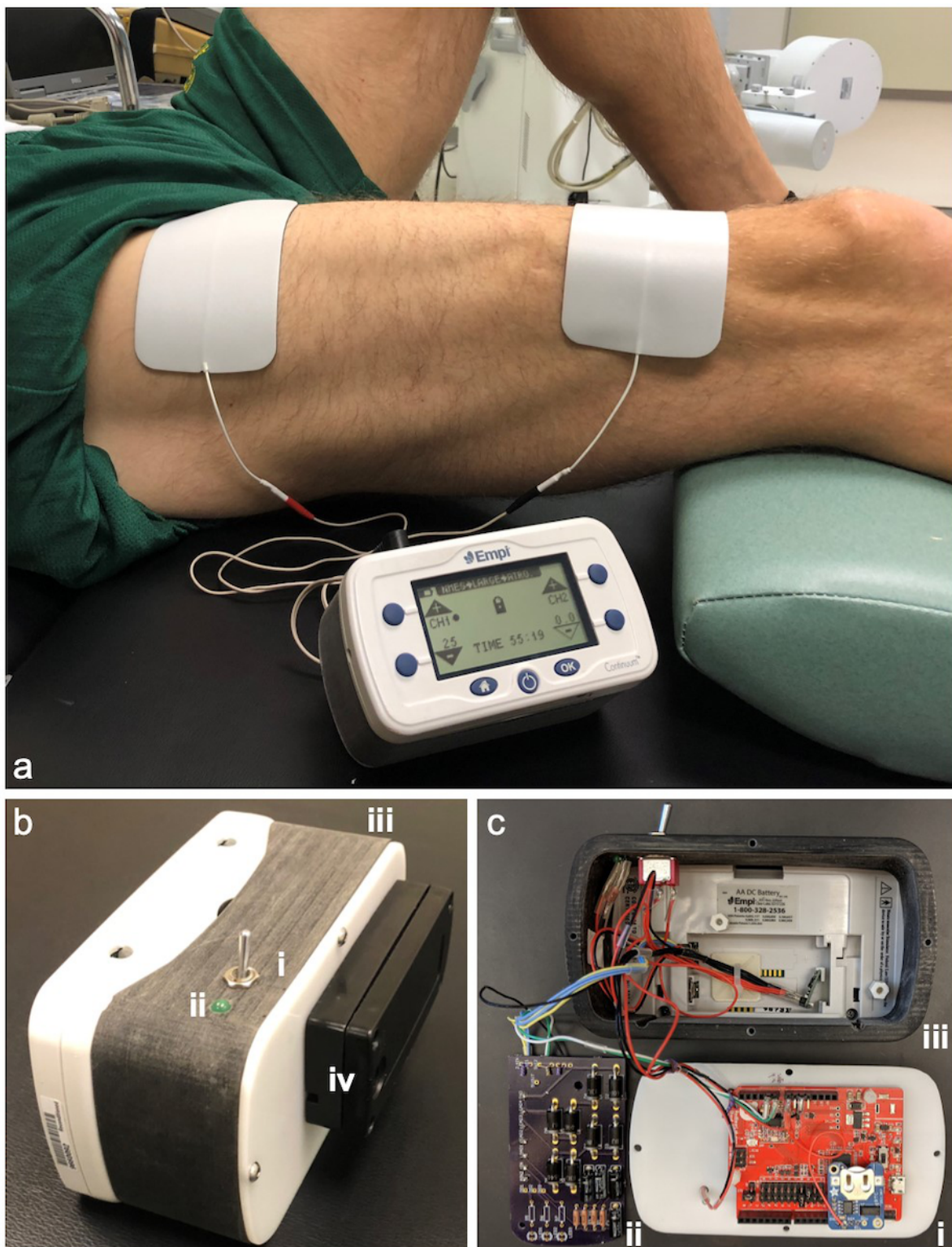
Mobile Health App Development

Patients predominantly interact with the system through the mHealth app. The app allows patients to view and sync their session data (Figure 3) to the database. Communication with the patient via push notifications is also supported by the app (Figure 3). The app was developed for iOS because of its high preference in the age demographic most likely to sustain traumatic knee injuries [18]. However, all supported functionality could be replicated on the Android platform without technical barriers.

Figure 3. Sample screenshots from the RehabTracker mobile health app, which includes functionality to sync session data with the secure database (shown in screenshot "a") and encourages patients' adherence to the push notifications (shown in screenshot "b").
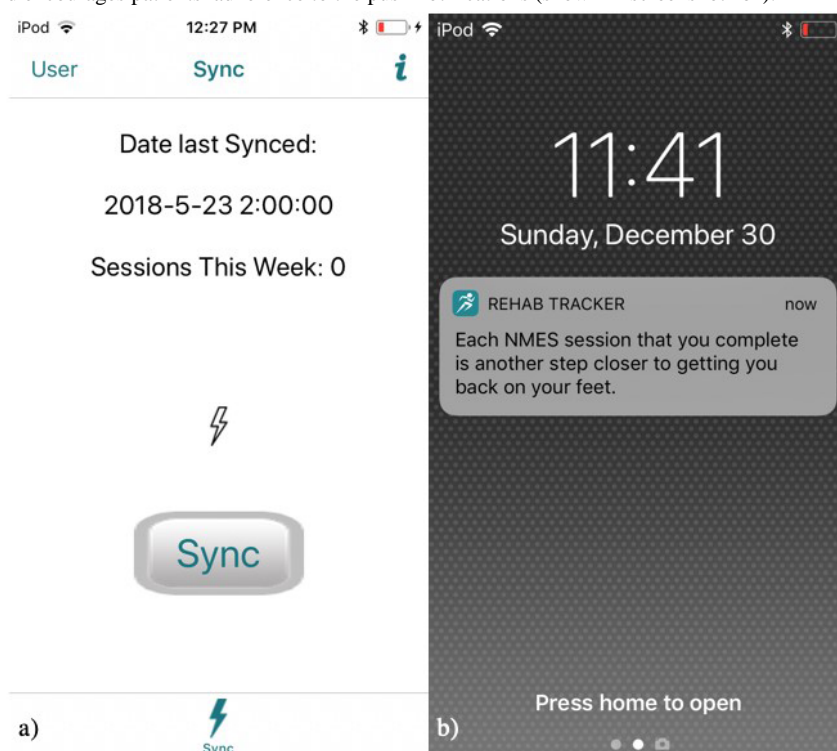
App Interface

The most significant feature of the mobile app is the sync feature (Figure 3), which sends patient data from the modified NMES device, through the app, to the database via our representational state transfer (REST) application programming interface (API). When a patient presses the sync button on the app, it initiates the Bluetooth scanning procedure in an attempt to establish a connection with any nearby modified NMES device. The process occurs from within the app, and patients do not need to pair their phone with their NMES device. After establishing a connection, the app reads session data from the device's Bluetooth characteristic until it reaches the end-of-message symbol or the connection times out after 30 seconds. Parenthetically, we did not include a layer of security that required the modified NMES device to be paired with a specific user mobile device as the close oversight of the software and device disbursement in this study meant that the likelihood of the two devices and/or apps being in close proximity was extremely remote. Nonetheless, future versions will incorporate device pairing.

Data collected from the NMES device are parsed into JavaScript Object Notation objects and stored locally on the patient's phone. The app also attempts to upload these locally stored sessions to the database via our REST Web API. On-phone data storage is only used to prevent data loss from network errors or when internet service is unavailable. These copies are deleted after the data are successfully uploaded to the database. In

addition to the one-button sync feature, the iOS app also provides user authentication for patients. This introduces additional security to the app and ensures that the data synced through the app will only be associated with the logged-in user.

Push Notifications

RehabTracker uses automated push notifications to communicate positive and motivational messages to patients on an almost daily basis, an approach designed to reduce overhead time for clinicians in the daily process of communicating with patients regarding adherence to their rehabilitation prescription. The notification method, based on social cognitive theory [19] and outlined in Figure 4, reminds and encourages patients to complete NMES sessions, without contacting them so frequently that they ignore notifications. There are three events that initiate a push notification, including (1) a completed session, (2) a week starting with no sessions, and (3) the end of a week. For each type of event, a message, randomly chosen from an extensive set of messages associated with that event and the patient's level of adherence, is sent to the patient via an app notification. Mixing up the messages and assuring that there is sufficient number of messages make the experience less automated. The examples of message content for each event type are listed in Textbox 1, and a screenshot of how the notifications appear to patients is presented in Figure 3. Generally, the messages are designed to enhance motivation and encourage the patients' confidence in their ability to comply with rehabilitation.

Figure 4. Patient adherence evaluation algorithm. Adherence is checked twice weekly. Positive and remedial push notifications are sent based on the number of neuromuscular electrical stimulation sessions completed.
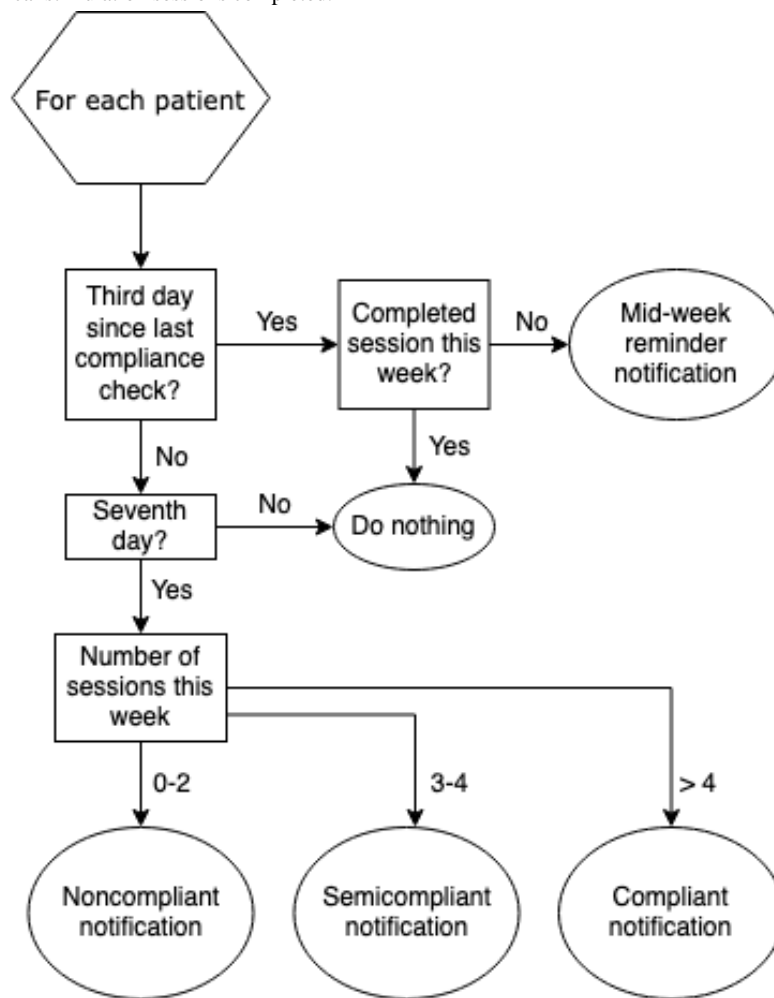
Textbox 1. Examples of text used in different types of notifications sent to patients.

After session

- "Great job finishing your NMES session!"

Compliant

- "Fantastic week! You met your goal for NMES sessions. Keep up the good work!"

Semicompliant

- "We would like to see you complete 5 NMES sessions per week. You did great last week but didn't quite sync 5 sessions. Is there anything we can do to help?"

Noncompliant

- "Tough week? We saw that you did not sync 5 sessions of NMES. Let us know if there is anything we can do to help."

Midweek reminder

- "We noticed you have not completed an NMES session in the last 3 days. Is there a problem? Can we help? Give us a call"

After a patient syncs session data with the system, they receive a positive reinforcement notification from the system. As patients should be completing five sessions every week (each 7-day period starting with the first day of RehabTracker system use), these are the most frequent notifications. If a patient begins their week with three consecutive days without syncing a session, then a reminder notification is sent inquiring if they forgot to sync their sessions or if they are having problems with the device or app. Finally, at the end of a patient's week, the patient receives a notification reporting their total adherence for the week. We have created three subcategories for patient adherence notifications for these weekly communications: (1)

compliant (>4 sessions completed), (2) almost compliant (3-4 sessions completed), and (3) noncompliant (<3 sessions completed). This classification scheme ensures that the patients who are completing some, but not all, of the required sessions receive positive reinforcement while also being nudged to improve adherence during the next week. In noncompliant patients, care providers are alerted to the patients' noncompliance so that additional patient-provider communication can be undertaken to resolve issues contributing to the nonadherence.

Clinician Portal Web Interface Development

The clinician portal provides near real-time access to patient adherence and session data. The site's front end is written in hypertext preprocessor (PHP) language and directly interfaces with the MySQL database. Queries are sanitized by our structured query language (SQL) database API. The site is also hosted with hypertext transfer protocol secure (HTTPS) to ensure HIPAA compliance. Data are immediately updated after a patient's data sync. Upon logging in with a username and password, clinician users receive various views of their patients' data (Figure 5). The home screen sorts patients by their adherence status to alert clinicians of noncompliant patients who may need more attention. Clinicians can also see device use data on a session-specific level. The portal also allows clinicians to enroll patients into RehabTracker and allows admin clinicians to add other clinicians.

Figure 5. Sample screenshot of the clinician portal, which identifies patients who are nonadherent; provides summary information on the current patients; and offers functionality to view all patients, individual patient sessions, and to add clinicians (administrator only) and patients.

Back-End Server

The back-end server that enables the RehabTracker cyber-physical system includes the following components/features: (1) REST API, (2) MySQL database, (3) adherence assessment, (4) push notifications, and (5) HIPAA adherence and security. Each component is described in more detail below.

Representational State Transfer Application Programming Interface

The RehabTracker iOS mobile app follows a 3-tier client-server architecture when interacting with the database. For the middle tier, we implemented a REST API in PHP. The primary functions of the RehabTracker cyber-physical system are enabled by this API. When a patient presses the sync button in the iOS app, all session data are transferred through the API to the database. It also creates a push notification for the patients when they sync, to provide immediate positive feedback. Additional features of the Web API include authentication of users.

MySQL Database

A MySQL database was used to store data on a HIPAA-compliant server. A total of five tables store and relate clinician, patient, session, push, and notification data. The patient and clinician tables store information about patients that will remain constant throughout the study. When a patient completes a session, a new row is added to the session table to log it, and whenever a new push notification is sent, it is added to the push table. The notification table contains the text of the push notification ordered by the push notification category.

Adherence Assessment

RehabTracker automatically checks a patient's adherence on the third and the seventh day of each week relative to their start

date (Figure 4). A scheduled server process performs these checks. According to the NMES prescription for these patients, adherence is defined as the completion of a 1-hour session each day for 5 days every week. The push notification regime associated with this adherence checking is discussed in detail above.

Push Notifications Infrastructures

The RehabTracker server uses three components to generate, store, and send Apple push notifications (APNs). The adherence script and API generate push notifications based on the user data, the database stores the notifications, and a scheduled server job sends the notifications with APN. The data sync API endpoint and adherence script described above upload push notifications to the database. Each notification is stored in the database with the ID of the patient who is receiving the notification and the index of the notification body.

Every hour, the push notification script queries the database for all unsent notifications, builds notification objects with the notification bodies and the universally unique identifier of the patients' phones, and sends these notification objects to users with APN. Notifications are stored in the database agnostic of purpose; that is, with regards to adherence or completed session. Accordingly, the same notification-sending script is used for both. The push notification script is written in Python and takes advantage of the Python APNs library for sending APNs. The script also uses the REST API to receive all information from the database. This minimizes exposed database credentials and ensures that the messages are encrypted when sent from the database to the script.

Health Insurance Portability and Accountability Act Compliance and Security

All patient data in RehabTracker are anonymous. When a patient is signed up to use the RehabTracker, they receive an anonymous user ID from the clinician with no correlation to their personal information. The patient uses this ID as their RehabTracker username, and all data in the database for that patient is related to this ID. This approach ensures that data stored in the database are anonymous, and, thus, are not protected health information, as defined by HIPAA [20]. The app itself does not use passwords. Instead, users log in with their anonymous ID. In addition, the use of the app requires the customized NMES device. All these elements further reinforce data security.

The database, Web API, and provider portal are hosted on a HIPAA-compliant server. The database uses Research Electronic Data Capture to secure permanent data storage. To access the data, one must use the Web API or provider portal. Each of these points of access actively limits the SQL injection to minimize the data one may access. The API and portal are also implemented using HTTPS, taking advantage of the transport layer security encryption. In addition, the provider portal is username and password protected. Only developers have root privileges. These security and privacy measures have allowed us to test RehabTracker in the clinical setting.

Pilot Clinical Study

We undertook a formative evaluation of the RehabTracker system on a convenience sample of patients who were taking part in a clinical trial of NMES use in patients who had suffered ACL injury and undergone surgical reconstruction (NCT02945553).

Participants

A total of 7 patients (3 women and 4 men) were selected to use RehabTracker. All patients were aged between 18 and 50 years; had BMI <35 kg/m^2; had suffered an acute, unilateral, first-time ACL rupture; and were scheduled to undergo reconstructive surgery. Patients were excluded based on the following criteria: (1) history of knee injury/surgery of either leg or nonsurgical intervention; (2) abnormal laxity in any other lower extremity besides the injured ACL; (3) signs or symptoms of arthritis, autoimmune or inflammatory disease, or diabetes; (4) grade IIIb or greater articular cartilage lesions; and (5) women who were pregnant or planning on becoming pregnant. Written informed consent was obtained from all volunteers before their participation, and all protocols and procedures were approved by the Committee on Human Research in the Medical Sciences at the University of Vermont.

Procedures

Patients (age: mean 22 years, SE 1; BMI: mean 26 kg/m^2, SE 1) used RehabTracker for 1 to 2 weeks, performing between 5 and 10 NMES rehabilitation sessions. The RehabTracker system was used after a patient had experience using NMES therapy so that the errors in use of the device would not affect the usability of the RehabTracker system. Each patient was enrolled in the RehabTracker study by a single participating clinician. Following enrollment, the RehabTracker app was downloaded to their personal iOS device using Apple's TestFlight service. The download links were sent to an email address created as a part of the enrollment procedure. While using the RehabTracker system, patients simultaneously logged their NMES use on paper-based log sheets. Patients' self-report logs were kept to document the NMES device use and were compared with the use data recorded by RehabTracker to verify the functional correctness of the system for recording NMES sessions.

Results

Functional Correctness

We conducted an initial assessment of the functional correctness of the system, specifically, its ability to perform the two main functions for which it was designed: (1) to sense and convey information about the NMES device use to the back-end server/clinician portal and (2) to function as a platform for provider-patient communications about device use adherence. Numerical results are provided as these are the basis for the assessment of functionality; however, we acknowledge that these numbers reflect the prototype functionality rather than the true adherence to the NMES prescription. Nonetheless, our efforts to define functionality uncovered issues that will be addressed in future prototypes.

With regard to the first design goal of our system, before we could define the aspects of the system that functioned properly and the ones that did not, we had to define whether the sessions that the patients completed were tracked by our system. Of the total number of sessions that the patients self-reported using the modified NMES device, 75% (55/73) were recorded by the RehabTracker system. In 2 patients, 100% (26/26) of the self-reported sessions were tracked by the system, whereas 2 patients had ~90% (20/22), one patient had 75% (6/8) and one patient had 23% (3/13) of the self-reported sessions tracked by the system. Finally, in another patient, none of the sessions (n=4) were recorded as the modified NMES device was not synced with the RehabTracker app. This failure is likely related to the patient not using the device, as no sessions were recorded by the device use tracking feature of the NMES device stock software. Thus, excluding this last patient, 80% of the patient-reported sessions were recorded by the system.

We initially chose patient-reported NMES device use as our comparator for RehabTracker system functionality. However, in the home environment, patient-reported adherence data may be less reliable [11]. To explore this possibility, we used the covert monitoring feature built into the EMPI Continuum software that tracks device use. Using this device use monitoring feature, results from our broader trial with nonmodified EMPI Continuum devices showed that patients overreported device use by approximately 12%. Overreporting of device use by patients could cause bias in our assessment of the functionality of the RehabTracker system, specifically toward the RehabTracker recording less sessions compared with patient-reported device use.

Owing to this potential bias in self-report, we also examined the ability of the RehabTracker system to monitor home NMES device use by comparing the sessions recorded by the system with those logged by the device use monitoring feature of the software. This approach uncovered an issue with our device modifications on the NMES device's internal software. In two patients, the number of device-reported sessions were spuriously high (34 and 135 sessions), suggesting that, in some cases, the NMES device modifications made for the RehabTracker system interfered with the covert monitoring feature such that it did not accurately record the number of sessions completed. Not considering these two patients, the device software reported 48 total sessions completed by the remaining patients. For these remaining patients, the self-report records showed 55 sessions completed, which represents a 15% (55/48) overestimate. This estimate is in accord with the data from the unmodified devices used by other participants throughout the remainder of the trial showing overreporting bias. The RehabTracker system reported 40 sessions for these same patients, which corresponds to 83% (40/48) of the device-reported sessions.

With regard to the second goal of providing a platform for communications aimed at improving device use adherence, patients for whom the tracking of automated push notifications was enabled (n=3), received 100% (29/29) of their expected push notifications. An additional 2 patients reported receiving push notifications; however, we did not have records that they received these notifications as the patients' device ID became dissociated from our system log because of log-ins from different devices. In addition, two patients did not receive the push notifications as the notification-sending script was not scheduled to run during their system participation. The system functioned properly for patients with complete notification logs.

Other Software and Process Issues

We discovered software issues with the mHealth app during formative testing. In the first version, a bug in the authentication logic prevented users from logging in. This was fixed, but it delayed the patients' use of the system. In addition, the patient enrollment and app installation procedure required coordination among the developers and study coordinators and several distributed steps. The added complexity of this process delayed the patients' participation.

Hardware Issues

Several hardware issues were uncovered with testing. First, the peak NMES device voltage logged by the RehabTracker did not match with the patients' self-reports of intensity. One volunteer did not self-report the device intensity and another did not sync sessions, leaving 5 volunteers for comparison. Among these patients, the RehabTracker's intensity matched the patient-reported intensity on average (36, SD 13 vs 36, SD 2 arbitrary units, respectively), but did not match on an individual basis ($r=-0.20$; $P=.75$). This individual inaccuracy is likely related to the high degree of variability in the device intensity recorded by the RehabTracker device. Second, the modifications to the stock NMES device to enable the RehabTracker system to monitor device use created problems with the NMES use tracking feature available as part of the stock device software. As reported above, in two patients, the number of device use sessions were spuriously high (34 and 135). The reason for this error was not readily apparent and will require further testing. Third, one patient stopped using the modified NMES device when the batteries ran out as the patient felt that the battery case was too cumbersome to open. Finally, there was a problem with the real-time clock that caused clock drift. We addressed this by time-stamping data in the app as they were received from the device, using time on the device hosting the app instead.

Discussion

Principal Findings

Knee extensor muscle performance is profoundly reduced in the postinjury and early postsurgical period following ACL reconstruction [3,21] because of a combination of neural, biomechanical, and pain limitations. Although orthopedic rehabilitation aims to prevent or remediate these maladaptations, most fall short of this goal, as evidenced by persistent atrophy and weakness in the years following surgery [3,22,23]. This loss of strength could have consequences for the development of future joint pathology [3], and interventions to prevent its development may contribute to better health outcomes and restoration of normal lower extremity function.

One potential reason why many rehabilitation programs do not remediate the loss of strength and function to preinjury levels is that a large proportion of the prescribed therapeutic exercises and activities are performed at home, particularly during the

XSL·FO
RenderX

early postsurgical period. In this environment, there is little or no oversight by the rehabilitation practitioners. Greater functional improvements may be realized with closer oversight by clinicians and more frequent provider-patient communications. In particular, approaches that facilitate patient-provider communications and seek to improve adherence could prevent strength deficits from developing. To this end, we described the construction of a cyber-physical system that enables the monitoring of home-based NMES use and an mHealth app that facilitates communication of device use data to clinicians and provides a platform for automated positive and remedial messages via push notifications to patients who are geared to improve device use adherence.

This study builds on our prior work to develop early rehabilitation programs for patients suffering ACL injury who have undergone surgical reconstruction [24,25]. In the prior work, rehabilitation was performed during clinic visits and was supervised by study personnel. Although this is a rigorous approach to test the safety and efficacy of such interventions, it is impractical in a real-world clinical setting. NMES has shown promise in preserving muscle size and function [8,9] and enhancing long-term functional recovery [10] in orthopedic surgical patients. The use of NMES over extended periods of time after injury and surgery may be problematic, however, as much of this rehabilitation would need to be performed at home, where adherence is generally low. Moreover, adherence with rehabilitation interventions decreases over time [25]. The cyber-physical system described here seeks to address these issues and improve adherence.

A goal of the RehabTracker system is to enhance the patient-provider communication to improve adherence with NMES device use prescription. Accordingly, our approach for these communications deserves some discussion. The patients' psychological responses to knee injury, surgery, and/or rehabilitation may be important for their ability to return to prior levels of activity and function [26] and, in turn, their satisfaction with surgical outcomes. Self-efficacy, defined as an individual's perceived ability to successfully engage in targeted behaviors, is associated with rehabilitation adherence and functional outcomes in patients who experience ACL trauma and undergo surgical reconstruction [27]. With this in mind, the design of our communication system was grounded in social cognitive theory [19], which emphasizes improving patients' self-efficacy toward adherence with rehabilitation prescriptions. Evidence shows that this approach supports adherence to the targeted behaviors for weight loss [28], diabetes management adherence [29], cancer screening [30], and smoking cessation programs [31].

We used a hybrid communications approach that included both automated messaging as well as weekly provider phone calls to patients. The admixture of manual and automated communications could be modified according to the number of outpatient rehabilitation visits, patient needs, and/or insurance coverage for telehealth services. The automated messaging is the most frequent contact and is designed to alleviate burdens on clinicians to monitor patient adherence on a daily basis, while also providing early positive feedback to patients for their adherence to the NMES use prescription. Short messaging

service, or texting, has been used more widely than push notifications in health care, including preventative care [32], disease management [33,34], and patient education [35] applications. However, the effectiveness of these interventions has been difficult to judge because of the poor quality of the evidence [36,37]. Although they have been less studied, we used push notifications as they allow more control over the appearance of notifications and support integrating functionality into the notifications (eg, to launch the app). Push notifications also do not require the patients' cell phone numbers, reducing the possibility of identifying patients using the data in the database, which raises privacy concerns that may influence the patients' attitude toward using the app.

One goal of this study was to conduct an initial assessment of the functionality of the system. With regard to the goal of tracking at-home NMES use, RehabTracker functioned as designed for most of the patients. Most patients were able to successfully use the RehabTracker cyber-physical system to sync data to the database and received push notifications from the automated communication system. In this context, our initial prototype of the system shows promise and warrants further development and testing. However, we noted process-oriented, hardware, and software shortcomings that are areas for improvement.

The two major system usage problems were largely because of process-oriented issues. In the first instance, problems were encountered by patients setting up the app. A total of 5 patients set up RehabTracker in a knowledgeable clinician's presence. All of these patients were able to use the RehabTracker device for the duration of their trial period without problem. The first two patients who set up RehabTracker independently had difficulties. We initially allowed patients to enroll independently through the Apple beta testing program, TestFlight, to simulate real-world usage. However, this process was more involved than downloading a traditional iOS app and proved too difficult for these patients to perform independently. Following these difficulties, one patient did not feel confident that the sync procedure downloaded data from the NMES device to the app and the back-end server, which affected their use of the system. In the second issue, when one patient's modified NMES device ran out of batteries, they switched to a backup device rather than replacing the batteries. Both incidents represent usability concerns for the system, albeit minor ones that can be corrected with hardware (ie, easier access to battery compartment) and software (ie, patient setting up the app with clinician/study personnel and better user notification to confirm data sync) adjustments.

Our ability to identify the functionality of software, firmware, and hardware components of the RehabTracker system was influenced by the fact that we used a convenience sample from an on-going clinical trial to assess the functionality. Imprecision in tracking at-home use of the NMES device hampered assigning specific functionality defects as user- or system-specific. Despite this problem, our study identified several issues with the initial prototype of this system that provide valuable information for future modification.

As data from our broader trial showed that patients overreport NMES use, we used the device use tracking feature built into the NMES device software for comparison with the device use recorded by our system. However, in two patients, we discovered that this was affected by the hardware modifications made to the device to enable the RehabTracker system to track device use. The nature of this interference is unclear and will necessitate further testing and likely redesign of the sensor system for tracking NMES current outflow. More broadly, this initial functionality assessment reveals that our study, similar to many in the field [11], lacks an accurate and objective criterion for assessing at-home exercise/rehabilitation participation and, in turn, functionality and usability of mHealth systems designed to track at-home interventions. Accordingly, future efforts to evaluate the system's functional correctness will need to incorporate initial laboratory-based assessments to test the usability and functionality of the hardware, firmware, and software before assessments in the home environment.

Our sensor system also did not accurately track the NMES device's current delivery on an individual basis. Parity in these measures was not a primary design goal as the NMES device intensity varies substantially from day-to-day following injury and in the early postsurgical period, with changes in the level of fluid infiltration in the surrounding tissue (ie, edema). Nonetheless, some measurable NMES device intensity serves as a verification of the device use, and the sensor system can be further refined to improve accuracy in future iterations. In addition, we noted problems with the real-time clock that were likely because of a design flaw in the battery connection of the real-time clock utilized in our prototype. Both of these issues can be addressed in the next prototype iteration by using higher fidelity components to improve upon the accuracy and precision of recording device use.

With regard to the goal of the system to communicate with patients via push notifications, our system functioned largely as designed. That is, sessions that were successfully monitored by the system initiated the automated push notification communications protocol designed to improve adherence to the NMES device use. We did, however, note several process-oriented issues that prevented us from documenting and confirming the receipt of push notifications, which will be fixed in future testing.

Conclusions

The cyber-physical system that we describe provides a system to collect data on home-based NMES use and communicates NMES device use data from patients to the clinical environment in a HIPAA-compliant, noninvasive manner in near real time. These collected data can be reviewed directly by care providers in the clinician portal and are also available to automated subsystems for actively tracking and improving patients' adherence through positive and remedial push notifications. Our system differs from prior work in this field that focused on telehealth systems to aid in the performance of classical home rehabilitation [17,27,38]. In contrast, RehabTracker seeks to provide an mHealth-assisted adjunctive rehabilitation modality for patients recovering from orthopedic surgery to supplement existing at-home programs along with an automated messaging intervention grounded in social cognitive theory and designed to improve adherence. With further advancements and testing, the RehabTracker system has the potential to improve provider monitoring of patients' adherence with at-home NMES prescription.

Acknowledgments

Conflicts of Interest

None declared.

References

1. Griffin LY, Albohm MJ, Arendt EA, Bahr R, Beynnon BD, Demaio M, et al. Understanding and preventing noncontact anterior cruciate ligament injuries: a review of the Hunt Valley II meeting, January 2005. Am J Sports Med 2006 Sep;34(9):1512-1532. [doi: 10.1177/0363546506286866] [Medline: 16905673]
2. Palmieri-Smith RM, Thomas AC, Wojtys EM. Maximizing quadriceps strength after ACL reconstruction. Clin Sports Med 2008 Jul;27(3):405-24, vii. [doi: 10.1016/j.csm.2008.02.001] [Medline: 18503875]
3. Tourville T, Jarrell K, Naud S, Slauterbeck JR, Johnson RJ, Beynnon BD. Relationship between isokinetic strength and tibiofemoral joint space width changes after anterior cruciate ligament reconstruction. Am J Sports Med 2014 Feb;42(2):302-311 [FREE Full text] [doi: 10.1177/0363546513510672] [Medline: 24275860]
4. Ingelsrud LH, Granan L, Terwee CB, Engebretsen L, Roos EM. Proportion of patients reporting acceptable symptoms or treatment failure and their associated KOOS values at 6 to 24 months after anterior cruciate ligament reconstruction: a study from the Norwegian Knee Ligament Registry. Am J Sports Med 2015 Aug;43(8):1902-1907. [doi: 10.1177/0363546515584041] [Medline: 25977523]

5. Thomas AC, Villwock M, Wojtys EM, Palmieri-Smith RM. Lower extremity muscle strength after anterior cruciate ligament injury and reconstruction. J Athl Train 2013;48(5):610-620 [FREE Full text] [doi: 10.4085/1062-6050-48.3.23] [Medline: 24067150]

6. Dirks ML, Wall BT, Snijders T, Ottenbros CL, Verdijk LB, van Loon LJ. Neuromuscular electrical stimulation prevents muscle disuse atrophy during leg immobilization in humans. Acta Physiol (Oxf) 2014 Mar;210(3):628-641. [doi: 10.1111/apha.12200] [Medline: 24251881]

7. Gibson JN, Smith K, Rennie MJ. Prevention of disuse muscle atrophy by means of electrical stimulation: maintenance of protein synthesis. Lancet 1988 Oct 1;2(8614):767-770. [doi: 10.1016/s0140-6736(88)92417-8] [Medline: 2901612]

8. Hasegawa S, Kobayashi M, Arai R, Tamaki A, Nakamura T, Moritani T. Effect of early implementation of electrical muscle stimulation to prevent muscle atrophy and weakness in patients after anterior cruciate ligament reconstruction. J Electromyogr Kinesiol 2011 Aug;21(4):622-630. [doi: 10.1016/j.jelekin.2011.01.005] [Medline: 21334221]

9. Hauger AV, Reiman MP, Bjordal JM, Sheets C, Ledbetter L, Goode AP. Neuromuscular electrical stimulation is effective in strengthening the quadriceps muscle after anterior cruciate ligament surgery. Knee Surg Sports Traumatol Arthrosc 2018 Feb;26(2):399-410. [doi: 10.1007/s00167-017-4669-5] [Medline: 28819679]

10. Stevens-Lapsley JE, Balter JE, Wolfe P, Eckhoff DG, Kohrt WM. Early neuromuscular electrical stimulation to improve quadriceps muscle strength after total knee arthroplasty: a randomized controlled trial. Phys Ther 2012 Feb;92(2):210-226 [FREE Full text] [doi: 10.2522/ptj.20110124] [Medline: 22095207]

11. Bollen JC, Dean SG, Siegert RJ, Howe TE, Goodwin VA. A systematic review of measures of self-reported adherence to unsupervised home-based rehabilitation exercise programmes, and their psychometric properties. BMJ Open 2014 Jun 27;4(6):e005044 [FREE Full text] [doi: 10.1136/bmjopen-2014-005044] [Medline: 24972606]

12. Hall AM, Kamper SJ, Hernon M, Hughes K, Kelly G, Lonsdale C, et al. Measurement tools for adherence to non-pharmacologic self-management treatment for chronic musculoskeletal conditions: a systematic review. Arch Phys Med Rehabil 2015 Mar;96(3):552-562. [doi: 10.1016/j.apmr.2014.07.405] [Medline: 25108098]

13. Vuorenmaa M, Ylinen J, Piitulainen K, Salo P, Kautiainen H, Pesola M, et al. Efficacy of a 12-month, monitored home exercise programme compared with normal care commencing 2 months after total knee arthroplasty: a randomized controlled trial. J Rehabil Med 2014 Feb;46(2):166-172 [FREE Full text] [doi: 10.2340/16501977-1242] [Medline: 24241606]

14. Pew Research Center. 2019 Jun 12. Mobile Fact Sheet URL: http://www.pewinternet.org/fact-sheet/mobile/ [accessed 2019-05-27] [WebCite Cache ID 78KSNoLpw]

15. Quinn CC, Clough SS, Minor JM, Lender D, Okafor MC, Gruber-Baldini A. WellDoc mobile diabetes management randomized controlled trial: change in clinical and behavioral outcomes and patient and physician satisfaction. Diabetes Technol Ther 2008 Jun;10(3):160-168. [doi: 10.1089/dia.2008.0283] [Medline: 18473689]

16. Iso-Ketola P, Karinsalo T, Vanhala J. HipGuard: A Wearable Measurement System for Patients Recovering From a Hip Operation. In: Proceedings of the 2008 Second International Conference on Pervasive Computing Technologies for Healthcare. 2008 Presented at: PCTHEALTH'08; January 30-February 1, 2008; Tampere, Finland p. 196-199. [doi: 10.1109/pcthealth.2008.4571068]

17. Dunphy E, Hamilton FL, Spasić I, Button K. Acceptability of a digital health intervention alongside physiotherapy to support patients following anterior cruciate ligament reconstruction. BMC Musculoskelet Disord 2017 Nov 21;18(1):471 [FREE Full text] [doi: 10.1186/s12891-017-1846-0] [Medline: 29162071]

18. Piper Sandler. Minneapolis, MN: Piper Jaffray; 2019. Taking Stock with Teens. Spring 2019 Survey URL: http://www.pipersandler.com/3col.aspx?id=5552 [accessed 2020-03-19]

19. Bandura A. Social Learning Theory. Englewood Cliffs, NH: Prentice-Hall; 1977.

20. Malin B. US Department of Health and Human Services. 2012. Guidance Regarding Methods for De-Identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule URL: https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html [accessed 2019-05-13] [WebCite Cache ID 78KSWpPS5]

21. Lieber RL, Silva PD, Daniel DM. Equal effectiveness of electrical and volitional strength training for quadriceps femoris muscles after anterior cruciate ligament surgery. J Orthop Res 1996 Jan;14(1):131-138. [doi: 10.1002/jor.1100140121] [Medline: 8618155]

22. Järvelä T, Kannus P, Latvala K, Järvinen M. Simple measurements in assessing muscle performance after an ACL reconstruction. Int J Sports Med 2002 Apr;23(3):196-201. [doi: 10.1055/s-2002-23171] [Medline: 11914983]

23. Lautamies R, Harilainen A, Kettunen J, Sandelin J, Kujala UM. Isokinetic quadriceps and hamstring muscle strength and knee function 5 years after anterior cruciate ligament reconstruction: comparison between bone-patellar tendon-bone and hamstring tendon autografts. Knee Surg Sports Traumatol Arthrosc 2008 Nov;16(11):1009-1016. [doi: 10.1007/s00167-008-0598-7] [Medline: 18712355]

24. Beynnon BD, Johnson RJ, Naud S, Fleming BC, Abate JA, Brattbakk B, et al. Accelerated versus nonaccelerated rehabilitation after anterior cruciate ligament reconstruction: a prospective, randomized, double-blind investigation evaluating knee joint laxity using roentgen stereophotogrammetric analysis. Am J Sports Med 2011 Dec;39(12):2536-2548. [doi: 10.1177/0363546511422349] [Medline: 21952714]

XSL•FO
RenderX

25. Beynnon BD, Uh BS, Johnson RJ, Abate JA, Nichols CE, Fleming BC, et al. Rehabilitation after anterior cruciate ligament reconstruction: a prospective, randomized, double-blind comparison of programs administered over 2 different time intervals. Am J Sports Med 2005 Mar;33(3):347-359. [doi: 10.1177/0363546504268406] [Medline: 15716250]

26. Ardern CL, Taylor NF, Feller JA, Whitehead TS, Webster KE. Psychological responses matter in returning to preinjury level of sport after anterior cruciate ligament reconstruction surgery. Am J Sports Med 2013 Jul;41(7):1549-1558. [doi: 10.1177/0363546513489284] [Medline: 23733635]

27. Levinger P, Hallam K, Fraser D, Pile R, Ardern C, Moreira B, et al. A novel web-support intervention to promote recovery following Anterior Cruciate Ligament reconstruction: A pilot randomised controlled trial. Phys Ther Sport 2017 Sep;27:29-37. [doi: 10.1016/j.ptsp.2017.06.001] [Medline: 28822956]

28. Schippers M, Adam PC, Smolenski DJ, Wong HT, de Wit JB. A meta-analysis of overall effects of weight loss interventions delivered via mobile phones and effect size differences according to delivery mode, personal contact, and intervention intensity and duration. Obes Rev 2017 Apr;18(4):450-459. [doi: 10.1111/obr.12492] [Medline: 28187246]

29. Haider R, Sudini L, Chow CK, Cheung NW. Mobile phone text messaging in improving glycaemic control for patients with type 2 diabetes mellitus: A systematic review and meta-analysis. Diabetes Res Clin Pract 2019 Apr;150:27-37. [doi: 10.1016/j.diabres.2019.02.022] [Medline: 30822496]

30. Uy C, Lopez J, Trinh-Shevrin C, Kwon S, Sherman S, Liang P. Text messaging interventions on cancer screening rates: a systematic review. J Med Internet Res 2017 Aug 24;19(8):e296 [FREE Full text] [doi: 10.2196/jmir.7893] [Medline: 28838885]

31. Whittaker R, McRobbie H, Bullen C, Rodgers A, Gu Y. Mobile phone-based interventions for smoking cessation. Cochrane Database Syst Rev 2016 Apr 10;4:CD006611 [FREE Full text] [doi: 10.1002/14651858.CD006611.pub4] [Medline: 27060875]

32. Armstrong AW, Watson AJ, Makredes M, Frangos JE, Kimball AB, Kvedar JC. Text-message reminders to improve sunscreen use: a randomized, controlled trial using electronic monitoring. Arch Dermatol 2009 Nov;145(11):1230-1236. [doi: 10.1001/archdermatol.2009.269] [Medline: 19917951]

33. Fischer HH, Moore SL, Ginosar D, Davidson AJ, Rice-Peterson CM, Durfee MJ, et al. Care by cell phone: text messaging for chronic disease management. Am J Manag Care 2012 Feb 1;18(2):e42-e47 [FREE Full text] [Medline: 22435883]

34. Horvath T, Azman H, Kennedy G, Rutherford G. Mobile phone text messaging for promoting adherence to antiretroviral therapy in patients with HIV infection. Cochrane Database Syst Rev 2012 Mar 14(3):CD009756 [FREE Full text] [doi: 10.1002/14651858.CD009756] [Medline: 22419345]

35. de Tolly K, Skinner D, Nembaware V, Benjamin P. Investigation into the use of short message services to expand uptake of human immunodeficiency virus testing, and whether content and dosage have impact. Telemed J E Health 2012;18(1):18-23. [doi: 10.1089/tmj.2011.0058] [Medline: 22150712]

36. Palmer M, Barnard S, Perel P, Free C. Mobile phone-based interventions for improving adherence to medication prescribed for the primary prevention of cardiovascular disease in adults. Cochrane Database Syst Rev 2018 Jun 22;6:CD012675 [FREE Full text] [doi: 10.1002/14651858.CD012675.pub2] [Medline: 29932455]

37. Vodopivec-Jamsek V, de Jongh T, Gurol-Urganci I, Atun R, Car J. Mobile phone messaging for preventive health care. Cochrane Database Syst Rev 2012 Dec 12;12:CD007457 [FREE Full text] [doi: 10.1002/14651858.CD007457.pub2] [Medline: 23235643]

38. Ayoade M, Baillie L. A Novel Knee Rehabilitation System for the Home. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Toronto, Ontario, Canada: Association for Computing Machinery; 2014 Presented at: CHI'14; April 26 - May 1, 2014; Ontario, Toronto, Canada p. 2521-2530. [doi: 10.1145/2556288.2557353]

39. GitHub. Rehab-Tracker URL: https://github.com/timstevens1/Rehab-Tracker [accessed 2020-03-19]

40. The University of Vermont. Rehab Tracker URL: http://www.uvm.edu/~cchao/html_doxygen/index.html [accessed 2020-03-19]

Abbreviations

ACL: anterior cruciate ligament
API: application programming interface
APN: Apple Push Notification
FDA: Food and Drug Administration
HIPAA: Health Insurance Portability and Accountability Act
HTTPS: hypertext transfer protocol secure
iOS: iPhone Operating System
mHealth: mobile health
NMES: neuromuscular electrical stimulation
PHP: hypertext preprocessor
REST: representational state transfer
SQL: structured query language

XSL•FO
RenderX

Original Paper

Recommendations for Developing Support Tools With People Suffering From Chronic Obstructive Pulmonary Disease: Co-Design and Pilot Testing of a Mobile Health Prototype

Alan Davies[1,2*], PhD; Julia Mueller[1,2*], PhD; Jean Hennings[1*], PhD; Ann-Louise Caress[3*], PhD; Caroline Jay[4*], PhD

[1]School of Health Sciences, University of Manchester, Manchester, United Kingdom

[2]Manchester Academic Health Science Centre, Manchester, United Kingdom

[3]School of Human and Health Sciences, University of Huddersfield, Huddersfield, United Kingdom

[4]Department of Computer Science, University of Manchester, Manchester, United Kingdom

[*]all authors contributed equally

Corresponding Author:
Alan Davies, PhD
School of Health Sciences
University of Manchester
Room 1002
Vaughan House, Portsmouth St
Manchester, M13 9GB
United Kingdom
Phone: 44 7843112378
Email: alan.davies-2@manchester.ac.uk

Abstract

Background: Gaps exist between developers, commissioners, and end users in terms of the perceived desirability of different features and functionalities of mobile apps.

Objective: The objective of this study was to co-design a prototype mobile app for people with chronic obstructive pulmonary disease (COPD). We present lessons learned and recommendations from working on a large project with various stakeholders to develop a mobile app for patients with COPD.

Methods: We adopted a user-centered, participatory approach to app development. Following a series of focus groups and interviews to capture requirements, we developed a prototype app designed to enable daily symptom recording (experience sampling). The prototype was tested in a usability study applying the *think aloud* protocol with people with COPD. It was then released via the Android app store, and experience sampling data and event data were captured to gather further usability data.

Results: A total of 5 people with COPD participated in the pilot study. Identified themes include familiarity with technology, appropriate levels for feeding back information, and usability issues such as manual dexterity. Moreover, 37 participants used the app over a 4-month period (median age 47 years). The symptoms most correlated to perceived well-being were *tiredness* ($r=0.61$; $P<.001$) and *breathlessness* ($r=0.59$; $P<.001$).

Conclusions: Design implications for COPD apps include the need for clearly labeled features (rather than relying on colors or symbols that require experience using smartphones), providing weather information, and using the same terminology as health care professionals (rather than simply lay terms). Target users, researchers, and developers should be involved at every stage of app development, using an iterative approach to build a prototype app, which should then be tested in controlled settings as well as *in the wild* (ie, when deployed and used in real-world settings) over longer periods.

KEYWORDS

XSL·FO
RenderX

Introduction

Context

An estimated 65 million people worldwide and 1.2 million people in the United Kingdom alone [1] have chronic obstructive pulmonary disease (COPD). COPD is projected to become the third leading cause of death by 2030 [2,3]. COPD entails a significant personal, economic, and societal burden [4].

COPD patients are predominately older adults, (ex)smokers, and from lower socioeconomic backgrounds [5]. Strategies are needed to help patients manage and monitor their condition over time and between medical assessments to ensure effective long-term management [6]. Importantly, COPD tends to occur alongside various comorbidities (such as coronary heart disease, lung cancer, anxiety, depression, and osteoporosis) due to shared risk factors (eg, aging and smoking) and shared underlying pathophysiological mechanisms [7]. This increases the disease burden, worsens patients' prognosis, and leads to increased health care costs [7].

Background

The Department of Health in the United Kingdom has issued recommendations for the prescription of apps as part of the care strategy for various long-term conditions, such as COPD [8]. As COPD patients may not recall the small daily fluctuations in their lung symptoms due to the highly symptomatic nature of the disease, telemonitoring via mobile health (mHealth) technologies may facilitate early intervention through daily monitoring of symptoms [2].

mHealth technologies have been shown to reduce costs associated with long-term COPD management [2]. Qualitative insights show that using mHealth to complement regular care is acceptable to both COPD patients and their health care professionals [9]. The advantages of Web-based health interventions include cost-effectiveness, round-the-clock availability, customizability to personal preferences, and anonymity (when compared with face-to-face interactions) [10]. However, mHealth apps also entail issues and risks, such as a lack of quality control and lack of evaluations of their effectiveness, and privacy and security risks [10-12]. A systematic review of mobile apps used for self-management of chronic conditions concluded that apps can potentially improve health outcomes in long-term conditions through improved symptom management [13]. A systematic review and meta-analysis on randomized controlled trials (RCTs) assessing mobile apps for COPD self-management found evidence for a lower risk of hospital admissions among app users as compared with usual care [14]. A further systematic review and

meta-analysis of RCTs found significant improvements in health-related quality of life across 557 COPD patients who used smart technology compared with face-to-face or written support [15]. However, the evidence stemmed from only three studies and was deemed to be of poor quality [15]. Moreover, a review of relevant literature on apps as well as apps available in app stores for people with COPD identified a scarcity of published literature regarding the effectiveness of the apps [2]. There is a clear need for quality-controlled, effective, and acceptable health apps if mHealth is to play a significant role in efforts to prevent and manage diseases.

Objectives

This study aimed to provide in-depth insights into the views of people with COPD and their caregivers regarding the use of apps in COPD, highlighting key topics and issues around usability that need to be taken into consideration during app development.

Methods

Study Context

This study formed part of the large, multistakeholder project *CityVerve*. The CityVerve initiative resulted in the city of Manchester receiving Innovate UK funding for a 2-year period to work on a range of initiatives to apply technology to four separate use cases in partnership with the National Health Service, industry, and universities. The use cases comprised health and social care, transport and travel, culture, and energy and environment. One of the focuses of the health and social care case was COPD. The use of mHealth with this demographic was considered, and a mobile app was developed using a co-design approach and prototyping with participants who were diagnosed with COPD. Project members included academic researchers, software engineers, and health care professionals.

Study Design

The approach taken in this study is similar to the *iterative convergent mixed-methods design* proposed by Alwashmi et al [16]. The authors propose the use of a mixed methods framework in which both qualitative and quantitative data collection and analysis are used in iterations to develop mHealth interventions and enhance the usability of such interventions [16].

Ethical Approval

Ethical approval for the study was granted by the University of Manchester (2017-2941-4477). An overview of the main phases can be seen in Textbox 1.

XSL·FO

RenderX

Textbox 1. Overview of the project phases.

Phase 1: Patient and public involvement work and prototyping

- Aim: To co-design a prototype app with chronic obstructive pulmonary disease (COPD) patients and caregivers

- Methods: Co-design, paper prototyping, focus groups, and direct observation

- Deliverables: A first version digital of the digital prototype on a mobile device

- Participants: approximately 48

Phase 2: Usability study

- Aim: To see how people would actually use the app and identify any usability issues

- Methods: Think Aloud, interviews, System Usability Scale (survey), and observations

- Deliverables: Data on user feedback, survey results, and update prototype with results

- Participants: 5

Phase 3: Wider deployment and usability testing

- Aim: To test the app with a wider number of COPD patients not directly involved in its design

- Methods: Experience sampling, event data collection, and analysis

- Deliverables: User event data

- Participants: 37

Procedure

Development and evaluation of the app took place during the three main phases. The *first phase* consisted of patient and public involvement (PPI) work, leading to the development of *mock-up* app designs and subsequently a prototype app. PPI work was carried out by attending and hosting a number of events for COPD patients and their families. We also invited members of a COPD self-help group to attend several sessions to discuss their needs and preferences. Together with participants, we developed paper prototypes for the app. The main stakeholders involved in the project were patients, families, caregivers, and members of the CityVerve project team. As this phase constituted PPI work rather than research, sessions were not audio recorded. Notes were taken following discussion and observation of the participants.

The *second phase* involved a usability study with the prototype app, using the "think aloud" paradigm. Participants were instructed to set up the app, browse its features, and verbalize their thoughts about the app as they proceeded. Participants then completed the System Usability Scale (SUS) questionnaire [17]. The "think aloud" sessions were audio recorded and transcribed. In addition, 2 researchers attended each session to observe and record notes. Transcripts and observation notes were subsequently analyzed thematically using framework analysis [18]. This is a qualitative method where themes are represented in columns and participants in rows, with participants' responses summarized in the associated theme's column (Multimedia Appendix 1). This facilitates comparison across participants for the respective themes. All data were coded by 1 researcher, and a random subselection of quotations was coded by a second independent researcher to test for coding agreement between the reviewers (interrater reliability). Cohen kappa indicated substantial agreement (84.6% agreement; $\kappa=0.75$).

The *third phase* involved deploying the app on the Google Play store for a 4-month period to collect the self-reported symptom data and event data to evaluate the app over a longer period with a wider group of the target population. The app was advertised by disseminating a 150-word summary of the study (or in the case of Twitter, a 220-character summary) to relevant groups on social media, various mailing lists, websites, and newsletters, with a link to the app's playstore page. For mailing lists, newsletters, websites, and Twitter, we approached the relevant admin who sent the information out or posted it to their website and Twitter account. For Facebook groups, we first approached the admin to obtain permission before posting. Targeted groups were either COPD-related (eg, COPD support groups) or likely to include an older demographic (eg, aging-related mailing lists) because COPD mostly affects older adults [19].

The App

The app (Figure 1) was developed as a hybrid Web app using Apache Cordova and the Ionic framework and was designed to exploit the experience sampling method (ESM), which allows for the daily submission of data rather than retrospective completion, leading to potential memory bias [20]. Once downloaded, the app takes users through several setup screens. These screens collect some one-time data about the user (Textbox 2).

The main menu page displayed the current weather data, including temperature, humidity, and wind speed, as seen in the gray box in the first screen (home page) in Figure 1. The *My symptoms* page allowed users to record their general well-being, measured on a 3-point scale (*great*, *so-so,* and *bad*). Following this, they entered information about five key symptoms of COPD (breathlessness, coughing, mucus, tiredness, and sleep quality on a 4-point scale) and their medication use (Figure 2).

XSL•FO

RenderX

The scales used to assess symptoms and well-being were adapted from the Britain Breathing app [21].

Participants could view their self-reported ratings for each symptom in the form of graphs presented on the *My data* page (Figure 1, center). This is accessed by tapping the *View my symptoms* menu option on the home page (Figure 1, left).

Educational material about COPD was presented on the *what is COPD?* page. The *About* page contained information about the study and the participant information sheet. Each page had a help button displayed in the top right-hand corner, which included a video with an audio description of how to use that page.

Figure 1. Screenshots of app: (left) home/landing page with the main menu and weather widget (center) graph showing ratings for breathlessness symptom and (right) self-reported impact of symptoms.
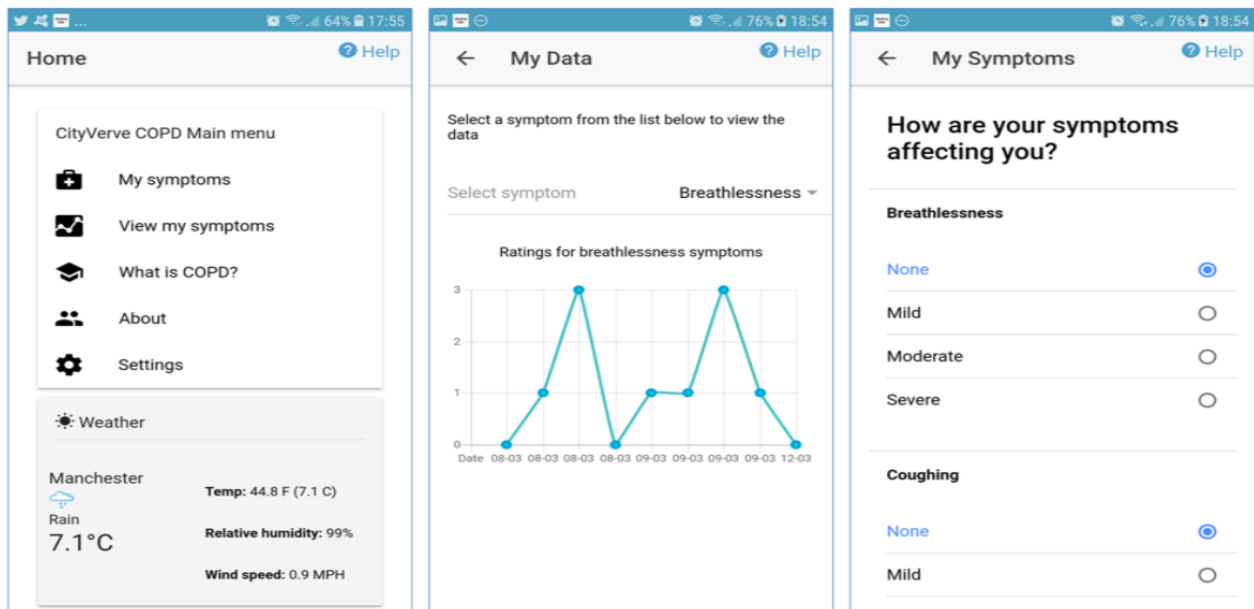
Textbox 2. One-time setup data captured by the app (modifiable subsequently via the settings screen).

- Gender (male or female)
- Year of birth
- Chronic obstructive pulmonary disease status (yes or no)
- Smoking status (yes daily, yes occasionally, no never, or no but I used to)
- Number of days they exercise for at least 30 min (0-14 or >14)
- Daily reminder for recording symptoms and medication (if relevant)

Figure 2. Screenshots of the app pages for medication questions and flare-up medications.
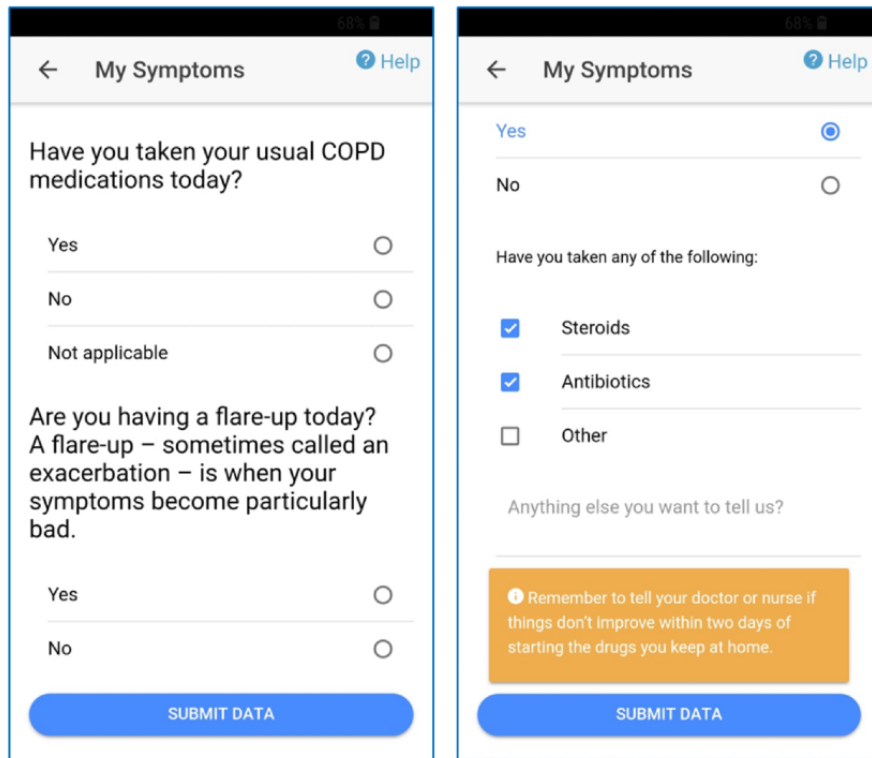
Results

Phase 1: User Engagement

Overall, users were enthusiastic about the concept of using an app to manage and monitor their condition. They welcomed the idea of being able to easily share information with family members, caregivers, and health care professionals, especially as several participants described difficulties communicating with doctors. Participants saw a value in the app for COPD patients in terms of recording, viewing, and monitoring symptoms, monitoring medication, and providing information about weather and characteristics of different locations to enable planning of activities. Figure 3 shows an example paper prototype developed with users.

Figure 3. Homescreen mock-up.
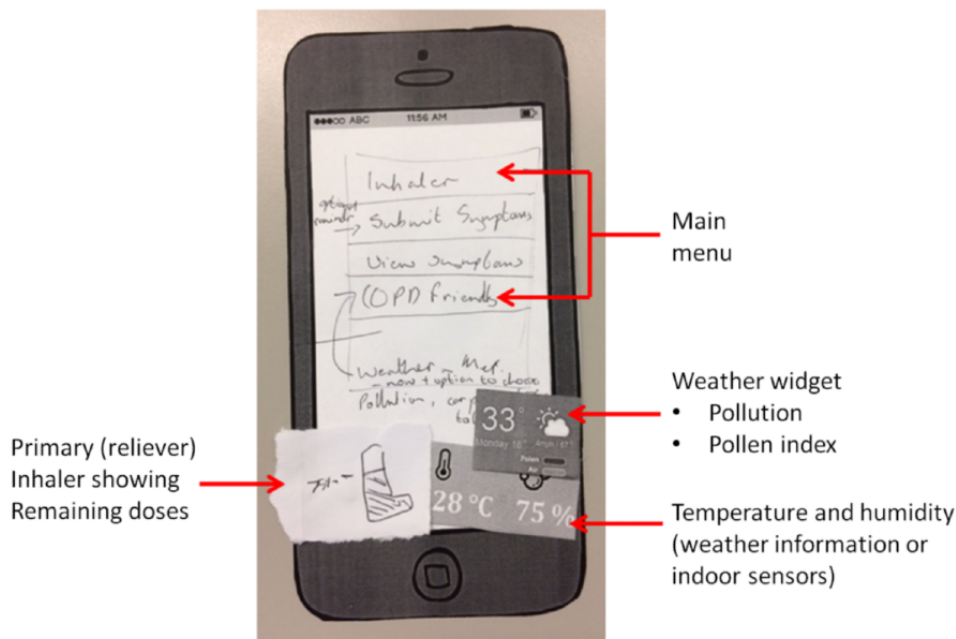
Chronic Obstructive Pulmonary Disease–Friendly Map

Participants suggested the inclusion of a *COPD friendly map*, which could be utilized to display locations that were *friendly* to COPD patients. This would include being able to view COPD-relevant details, such as the presence of elevators, stairs, gradients, weather, air pollution, pollen, and air quality. They were interested in using this information to help them plan journeys to places that would cater to their individual health needs. It was not possible to implement this feature for the prototype app due to limited timescales. This was not deemed to be a core requirement of the app by the participants but rather a desirable but optional, additional feature. Given the technical demands of providing such a feature that would involve allowing users to mark features on maps and rate areas and sharing this information with each other, this was not implemented in the prototype app.

Smart Inhaler

Participants commented that a smart inhaler that links to the app would be useful for monitoring the remaining doses. They suggested that an inhaler icon should appear on the app home screen, which would indicate remaining levels of medication. Participants were also keen to obtain feedback on the effectiveness of their inhaler technique. Smart inhaler integration was not implemented in the prototype app owing to project time constraints.

Symptom and Medication Monitoring

Participants emphasized the utility of symptom monitoring due to the fluctuations in the condition leading to participants being unsure about how best to use their medication for symptom control. They described difficulties recalling these details during their infrequent medical reviews. They indicated a preference for a visual depiction of this information. On the basis of the patients' recommendations, we selected five symptoms to be monitored within the app (breathlessness, coughing, mucus, tiredness, and sleep quality).

Phase 2: Usability Study

On the basis of phase 1, we developed a prototype that was tested with 5 people (4 females and 1 male) in the usability study. This included 4 people, who self-identified as having COPD, and 1 caregiver. The participants comprised 5 volunteers (4 females, all aged ≥55 years), and all of them had clinician-diagnosed COPD for more than 5 years. Participants were recruited through a respiratory patient support group and a general practice patient forum, and all of them had prior experience of being members of research advisory or stakeholder groups; hence, they were familiar with medical terminology. All but 1 participant used a smartphone at least occasionally. Moreover, 1 participant stated that she did not usually use a smartphone. Participants gave an average score of 32.4 out of 55 (SD 8.47) on the SUS, suggesting that they rated its usability as poor. The results of the qualitative analysis of the audio transcriptions and observational notes reveal the main issues faced by users. The main themes identified are shown in Figure 4. The number in brackets indicates the number of times that theme was mentioned.

Figure 4. Treemap depicting the number of references made to each coded theme.
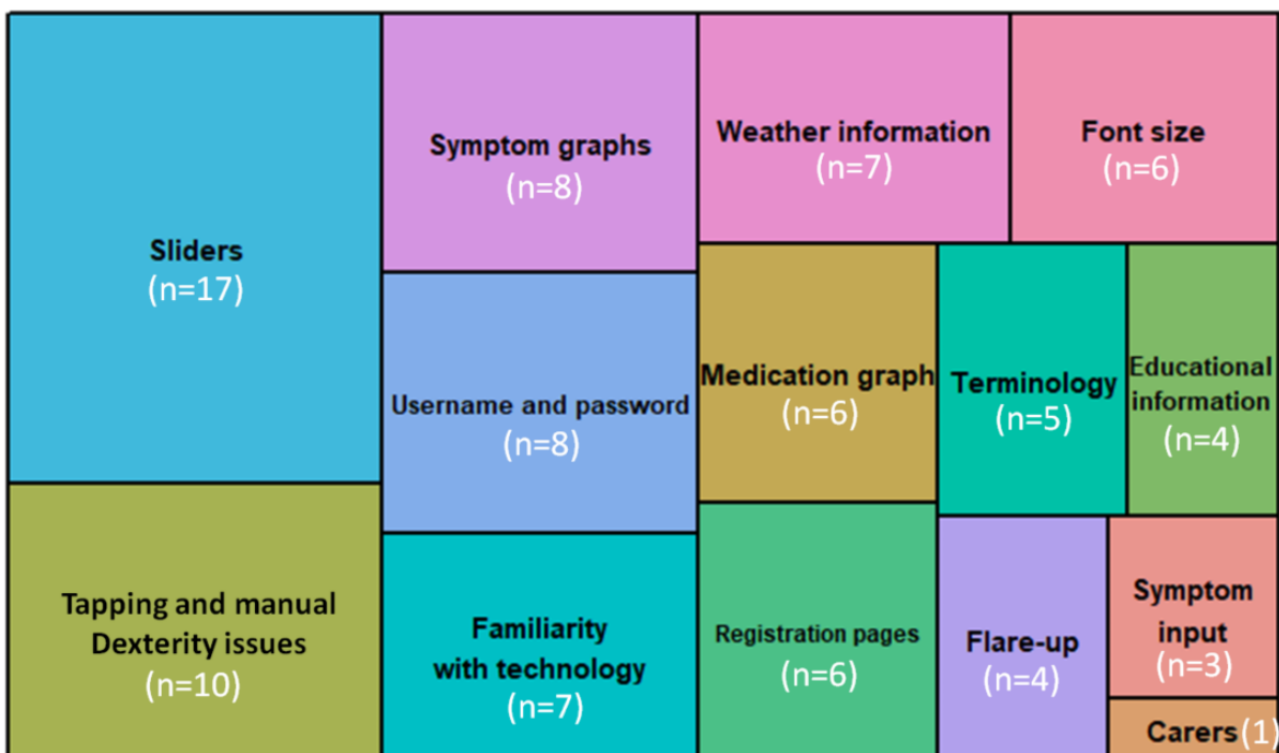
App Information Displayed to Users

Weather Information

Participants were specifically interested in information about humidity, wind, and temperature changes. They also indicated that they would prefer weather information for a longer period to plan activities in the future:

> it might be helpful if it could give you more than one day, the anticipated, so for instance you could think, oh, I don't think I better go shopping today I'll go tomorrow. [P3, COPD]

This information was also considered useful for caregivers:

> It's interesting for me to know that because if I look at it and see that it's anything like 90 or above I know he's not going to be in a fit state to do anything today, and I don't suggest going out or do anything, you know. So that's one of the places where a carer needs to know. [P4, caregiver]

Educational Information About Chronic Obstructive Pulmonary Disease

When reflecting on experiences of their initial diagnosis, participants suggested that future COPD patients would want to know more about the nature of the condition and the prognosis. Participants particularly liked diagrams and felt that these were most helpful for understanding airway constrictions and how they cause symptoms:

> I think the diagrams are very clear and good. They are helpful, and definitely show the changes that take place, and why, and it helps you think why you might get some of the symptoms that you do." [P1, COPD]

Information Feedback Level

A main feature of our experience sampling app involved feeding back information to users to enable them to track their symptoms and disease progression. We were interested in assessing which level of information would best suit participants. Participants were able to interpret the symptom graphs (Figure 1) displayed by the app accurately. This was elicited by directly questioning participants about their understanding of the graphs and what they represented:

> Interviewer 2: "Okay, so can I just ask, on that graph, so what do you reckon it shows, when it goes up and when it goes down? What does that show?"
>
> Participant 1: "When it goes up and down, like I've just said, the further it goes up, the more breathless I am, and for me, that's how it would look for me, and when it's down to zero, I'm not breathless at all."
>
> Interviewer 2: "Yeah, okay. And on there, on the horizontal axis?"
>
> Participant 1: "The days, that's fine."

Participants preferred graphs to be simple, clean, and minimal. They made several suggestions to improve the appearance and interpretability of graphs, including avoiding the use of color alone, and ensuring graphs can be presented in a sufficiently large format on the screen.

One participant was unable to interpret the graph and expressed little interest in understanding symptom fluctuation with medication use:

> Participant 5: "Well, that's your meds, isn't it, and that's your...Okay. [...] Yeah. But I don't know whether that would make any difference to me, I don't know whether it's relevant."
>
> Interviewer 2: "Mmm. What would you think? Looking at this graph, what would you think does it tell us about the medication in this case?"
>
> Participant 5: "Well, the green one is medication not taken. No, that's the red one. Yeah. Obviously, the green one is taken. I don't know what it really means to me; it doesn't really mean anything."

Terminology

Participants recommended using terminology commonly used by health care practitioners. They suggested that people with the condition would be familiar with *medical* terminology due to its frequent use by health care practitioners; they did, however, suggest including other commonly used synonyms for mucus, such as *phlegm* and *sputum*.

Usability Issues

Certain design features posed a challenge to participants due to manual dexterity issues, visual impairments, and unfamiliarity with smartphone usage. Participants frequently accidentally triggered functions within the app:

> It didn't give me a chance to rate how I was feeling, just clicked straight over into how are your symptoms affecting you? [P1, COPD]

Participants struggled, particularly, with typing due to the small size of the fields assigned to each letter within their smartphone's default virtual keyboard. This meant that security features such as inputting a unique username and password were challenging:

> Aren't they small these? [tries to type username] Well that was supposed to say [NAME]. [...] But the size of your finger doesn't cater for these does it? [P3, COPD]

These issues were exacerbated when coupled with visual impairments:

> Yeah. Okay. Invalid password. My passwords don't match, that's because I'm not very good at hitting keys either. I have a stigmatism, which means I tend to go one to the right or one to the left, or that bit [P4, caregiver]

Although most participants reported no issues with the font size, 1 participant reported that the font was *just about [her] limit* and suggested that being able to zoom in on the text would improve readability.

Manual and visual issues as well as lack of familiarity meant that certain design features common to many modern smartphone apps, such as sliders (Figure 5), were particularly challenging to participants:

It's hard for me because [...] I'm not used to sliding things, that's the bit I forget to do. [P4, caregiver]

The toggle switch (Figure 5) proved particularly challenging for users. Users were unsure how to engage with this function (tap or slide), struggled with the small size, and struggled to interpret its meaning (the slider turned grey when switched off and green when switched on, as is common in many smartphone apps). Participants preferred large buttons that they could *hit properly*:

Interviewer 2: "So, can I just ask, does that mean you prefer not to have a daily reminder?"

Participant 1: "No."

Interviewer 2: "Or would you like to have one?"

Participant 1: "Yes, the daily reminder would be good."

Interviewer 2: "Okay."

Participant 1: "So, the green means that it's set to remind?"

Participants also appeared to prefer clearly labeled buttons, rather than implicit symbols or color coding:

Participant 5: "Yeah. Enable a daily reminder to complete your symptoms and information. We did that, but it was only when you asked me if I didn't want a reminder. [...] The fact that I wouldn't have done if I hadn't have asked you what you wanted."

Interviewer 1: "Right."

Interviewer 2: "So would it maybe be more useful if it said something like, do you want a daily reminder, and then there would be yes or no, and you'd click yes or no?"

Participant 5: "Yeah, exactly."

Figure 5. Screenshots of the app pages using slider controls (left of the vertical bar) before they were replaced following user feedback by a more explicit list structure (right of the vertical bar).
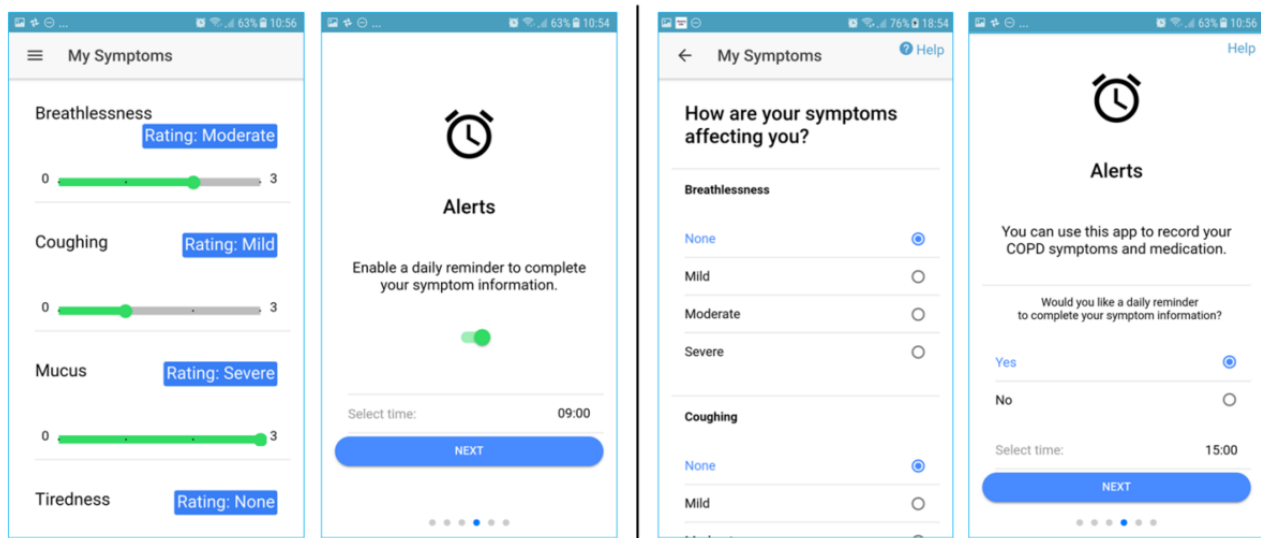
Technology Support and Help Requirements

Our observations revealed that participants may require help and support with basic aspects of smartphone usage, such as instructions on how to tap on items using the fingertip, as the following interaction highlights:

Participant 3: [tries selecting an item on screen using fingernail]

Interviewer 1: "Can I suggest that if you use your finger rather than your nail because that's..."

Participant 3: "It's also put the wrong date of birth in. It doesn't like me." [tries to select item on screen by pressing the fingertip down and holding]

Interviewer 2: "I think if you try..."

Interviewer 1: "Shorter. Yeah, just a bit shorter, yeah, there we go."

Several users appeared to search for help functions when they were unsure how to proceed, and they expressed a need for additional support and instruction:

They don't tell you where any of these things, it's not just you, they don't tell you where to tap. Do you have to tap on female or do you want it tapped on the dot at the end? That might make a difference. [P3, COPD]

Familiarity With Mobile Technology

All participants reported low familiarity and confidence in smartphone usage:

You're dealing with a complete technophobe here. [P1, COPD]

I presume that's a phone is it? [P3, COPD]

Validity of Participant Feedback

In assessing participants' feedback regarding the app against the researchers' observation notes, some discrepancies became evident. For example, observation notes indicate that participant 1 struggled with the slider controls, tending to tap rather than drag the control, and switching functions off when they intended to switch them on. When asked their opinion regarding the sliders, the participant nevertheless replied:

It's nice, it's not bad to use.

Following feedback from the usability study, design changes were made to mediate the issues identified. These included the following:

- Adding a help feature (help videos)

- Replacing the sliders with clearly labeled option lists (Figure 5)
- Improving the size and readability of text and input fields.

Figure 6 highlights the entire iterative design and development phases, including the key changes made in each iteration of the app leading to the final changes in iteration 3 suggested for the final *deployment* version of the app.

Figure 6. Iterative design and development phases. COPD: chronic obstructive pulmonary disease.
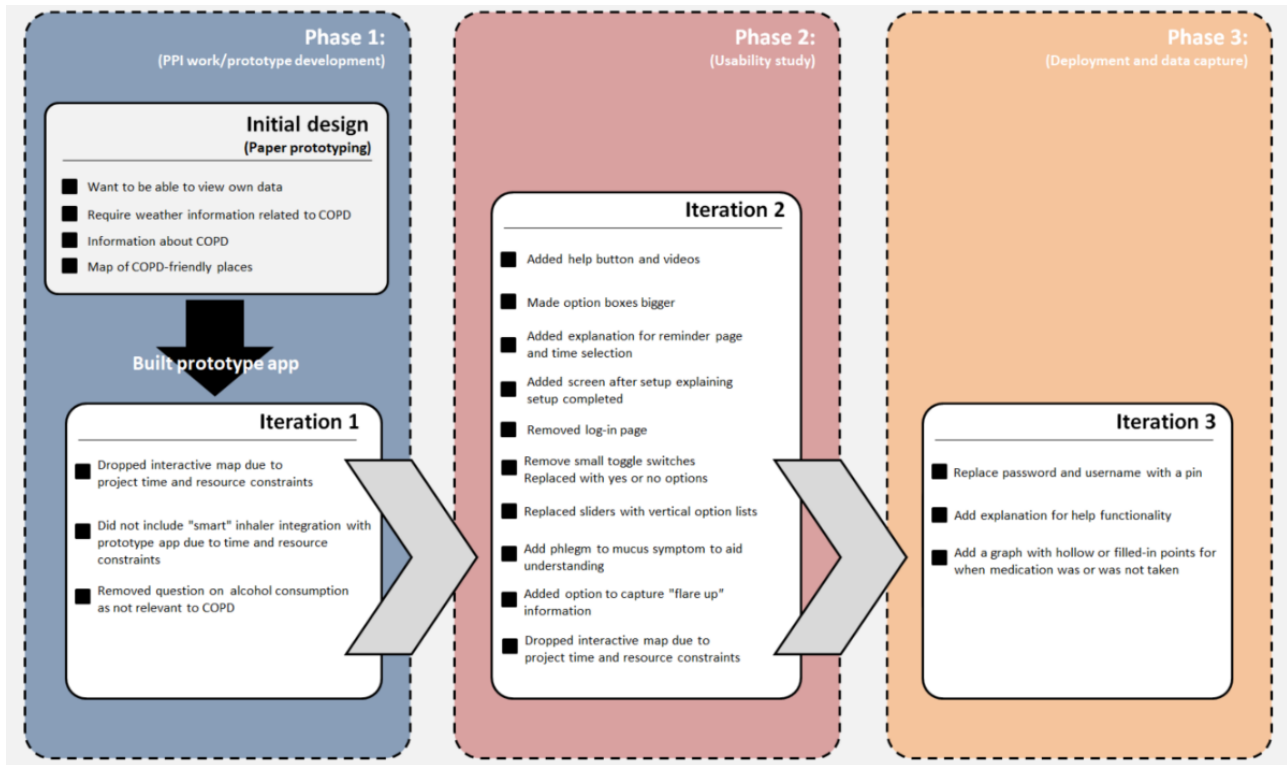
Phase 3: Pilot Testing Phase

Experience Sampling

In addition to the two devices the app was directly installed on during the UE, a further 41 users downloaded the app from the Google Play store. Moreover, 48% of the downloads came from the United Kingdom and 36.59% from the United States. A total of 37 participants used the app between May 2018 and August 2018; median age of the participants was 47 years (mean 45.2 years, SD 23.8 years). Table 1 shows participants' demographics. In 83% (31/37) of cases, participants identified as having a diagnosis of COPD and 16% (6/37) identified as not having a COPD diagnosis. A total of 25 users entered the data only once.

Figure 7 shows participants' symptom ratings over the 4-month period. The symptoms most correlated to perceived wellness coded with the variable *howFeeling* were *tiredness* (r=0.61;

P<.001; 95% CI 0.51-0.68) and *breathlessness* (r=0.59; P<.001; 95% CI 0.49-0.66). The average symptom ratings over the 4-month period (Table 2) show that the reported impact was highest for tiredness and breathlessness. Due to the small sample size and variation in symptom reporting (discussed in the Event Data section) in the pilot data, we would need to collect more data points for a longer period to carry out a more well-informed subsequent analysis. This should also account for seasonal variations in symptom reporting. Currently, these data are insufficient to draw inferences from, as the increased spread of symptoms over time may be related to the decrease in data points over time or other unknown factors. The subsequent analysis of the number of users reporting symptoms and the frequency of symptom reporting was important as it helped to provide some context to the data presented in Figure 7 (eg, being able to see that the majority of consistent reporting data were derived from only 3 participants, despite 37 people having recorded their symptoms at least once).

Table 1. Demographic data self-reported by participants (N=37).

Demographic	Value, n (%)
Gender	
Male	23 (62)
Female	14 (37)
Self-identified COPD[a] status	
Has COPD diagnosis	31 (83)
Has no COPD diagnosis	6 (16)
Smoking status	
Yes, daily	4 (10)
Yes, occasionally	8 (21)
No, never	4 (10)
No, but I used to	21 (56)
Exercise frequency	
1 day per week	10 (27)
2 days per week	9 (24)
3 days per week	9 (24)
4 days per week	4 (10)
5 days per week	3 (8)
6 days per week	2 (5)

[a]COPD: chronic obstructive pulmonary disease.

Figure 7. Average score for all participants (N=37) per month for the impact of symptoms. Note that "how feeling" is rated on a 3-point scale (0=great, 1=so-so, and 3=bad), and the other symptoms are rated on a 4-point scale (0=none, 1=mild, 2=moderate, and 3=severe).
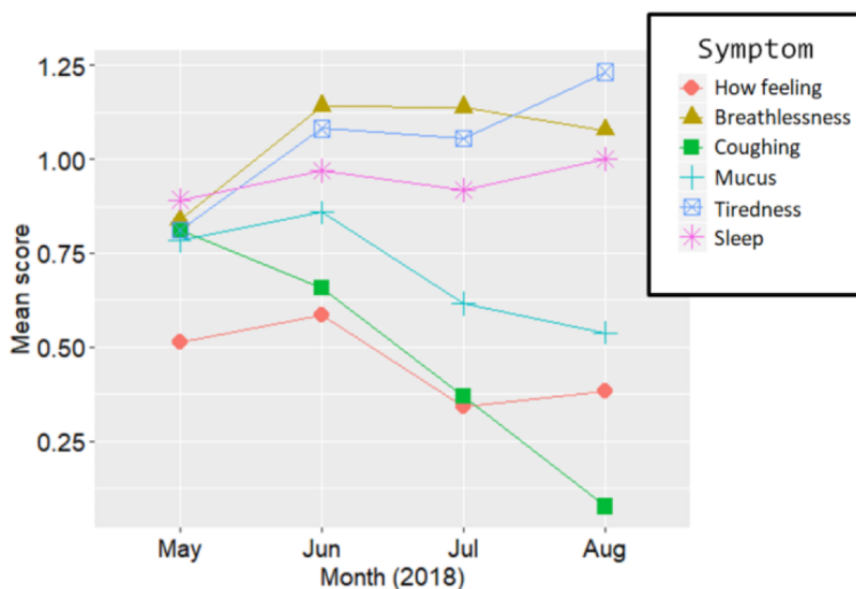
Table 2. Average rating for each symptom for all participants (N=37) over a 4-month period.

Symptom	Value, mean (SD)
Perceived wellness (How feeling)	0.481 (0.59)
Breathlessness	1.085 (0.76)
Coughing	0.554 (0.79)
Mucus	0.747 (0.78)
Tiredness	1.036 (0.99)
Sleep quality	0.941 (0.75)

Event Data

To gain a deeper understanding of how users were using the app outside of a controlled setting, we deployed an event capturing system into the app [22]. Event data were then extracted and analyzed using a combination of the R [23] package *bupaR v0.4.0* and the *WevQuery* tool [24] to apply pattern mining to the event data.

Figure 8 shows the number of times data were submitted by each participant over the 4-month period. This suggests a high dropout and low usage curve, with the exception of 3 participants who entered their symptom data on a regular basis (P36, P9, and P5). A total of 21 participants entered their data only once. Moreover, 2 participants (P18 and P37) exhibited unexpected behavior and entered their data 10 and 7 times in a single day, respectively, and did not enter any subsequent data.

The help button that appears on each page was selected by 4 users a total of 8 times. Figure 9 and Textbox 3 highlight the most common sequences of events between pages. This shows that, as intended, the symptom logging and viewing of data are among the most common sequences, suggesting that users viewed their data regularly after submitting it.

Figure 8. The number of times each participant submitted his or her self-reported symptom data per day per month.

Figure 9. Most common sequence of events within an episode between pages.

Textbox 3. Common mouse down event sequence patterns.

Mouse down event patterns

- Menu (symptoms page button) —> symptom page —>medication page (n=173)

- Menu (view my data button; n=86)
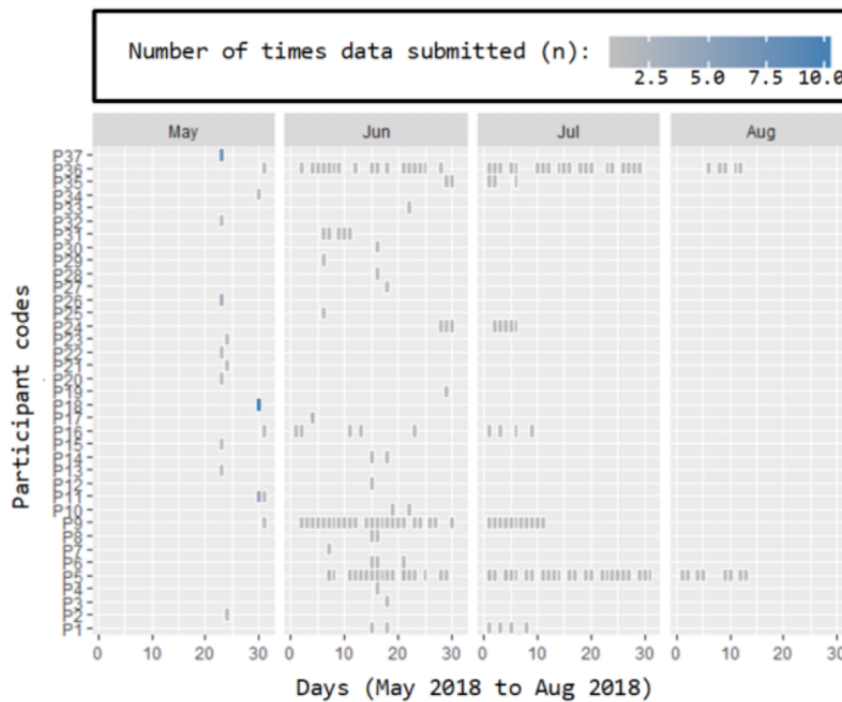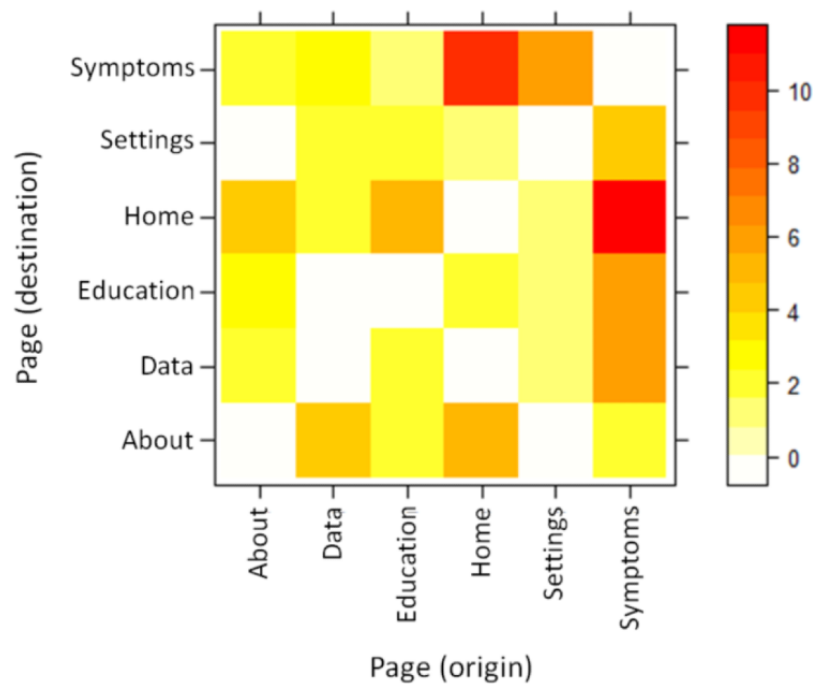
- Symptom page —> menu (view my data button; n=51)

- Medication page —> menu (view my data button; n=51)

Discussion

Principal Findings

We found that many of the issues affecting our work with participants were similar to those faced by O'Connor et al [25], who attempted to co-design an app with dementia patients and their caregivers, including unfamiliarity with technology and incorrect perceptions about how users would interact with the technology. This led to the initial design ideas being significantly revised. So far, very little qualitative research regarding the perception of patients and their caregivers and how they experience technology has been carried out [26]. The findings of our study reinforce the importance of considering the unique users of the technology we are proposing. When working with people with chronic diseases, their age, education, and Information Technology literacy should be taken into account [2]. This is especially relevant for the COPD population.

Many design features commonly used in mobile apps are hitherto unknown to the COPD population, such as the sliders that were initially included in the app. Our research indicates that such design elements are not optimal for a COPD population. Previous research corroborates this by showing that older users require larger touch targets (minimum 15×15 mm) with sufficiently large gaps between touch targets (minimum 6

mm), ideally complemented through support functions such as speech input [27].

Besides the physical impairments, our study highlights the importance of considering lack of familiarity and confidence in technology use. Our observations as well as participants' verbalizations while navigating the app suggest that older users may require more help functions within apps that provide support regarding app usage as well as general smartphone usage (eg, how to tap items). Research shows that among those older than 65 years, the majority do not feel confident in using computers, smartphones, and other electronic devices [28].

We used ESM to capture patients' symptom information, as it has been shown to provide an adaptive and personalized system for the monitoring and adaptation of treatment strategies as well as increasing ecological validity and reducing memory biases [20,29]. ESM has also proven to be a useful method in other conditions where the impact of symptoms varies over time [21]. Information about COPD was also included in the app. This was introduced following input from patients attending lung events, before app construction, as several patients highlighted a lack of information from their general practitioner about the condition.

Our study highlights that physical impairments and lack of familiarity can particularly affect initial registration processes,

such as creating usernames and passwords. This step is crucial for health-related apps, as password protection is needed to protect sensitive data. On the basis of our findings, we recommend that this process be kept brief and simple, for example, by using a simple 4-digit pin. Registration pages should also be supported through a feedback system to identify errors immediately and reduce user frustration [27].

Given the demographic qualities of the COPD population, the development of apps for this target population may be reticent to present information in the form of graphs; however, our research suggests that graphs can be interpreted correctly and perceived as useful if they are presented in a simple, clear format. Previous research confirms that graphics and multimedia should be used sparingly and purposefully when targeting older users, and text alternatives should be provided where possible [27].

Another important insight from our work reflects on the validity of user feedback. Features that users appeared to struggle with according to observation notes were nevertheless described as *fine* by users, suggesting a social desirability bias [30]. This underscores the importance of developing apps not only based on users' verbalizations and responses to specific questions (such as standardized usability scales) but on observations of user behavior while engaging with the app. This limitation has been noted in previous usability evaluations [31,32] and requires due consideration during app development.

Attrition rates are a known issue in digital health interventions [33]. The high level of attrition seen in our study is, therefore, not entirely unexpected. An understanding of why people discontinue use is important and worthy of further research given that retention is key for the management of chronic conditions over time [33]. This is especially challenging as chronically ill people may experience *diary fatigue* and be unlikely to keep accurate records of their condition, especially when unwell [34,35]. Another factor that may have contributed to the lack of use is that the reminder notification did not work on all devices due to an issue with the local notifications plugin path in a package of the Ionic framework.

The usage data indicate that the app was used as intended with most of the activity surrounding the recording and viewing of symptom data, which was the app's principal purpose. The low usage of the *help* feature may suggest that the app was also fairly intuitive for most people. The biggest issue surrounds the lack of continued use of the app over time. More research is needed to identify the intrinsic and extrinsic reasons behind the attrition in this context. Longer-term analysis of app usage before dropout using *in the wild* (deployed in the real world) event capture may help to shed light on some of these reasons.

Design Implications

As a result of the usability work carried out with participants, we can infer several considerations for designing COPD support apps, including the following:

- Large font size
- Large clearly labeled input (eg, buttons)
- Avoid items requiring greater tactile manipulation than tapping (ie, sliders)
- Provide easy to access help feature on each screen
- Label items clearly; do not rely on intuitive features
- Include the ability to zoom into content
- Passwords can be difficult; consider reducing the required length and inclusion of special characters.

Chronic Obstructive Pulmonary Disease–Specific Implications

The implications specific to those with COPD include:

- Information about weather was especially welcomed by participants, given its effect on the condition. Participants were keen to have everything in one place (app) rather than using multiple different sources of information.
- Future apps should consider implementing COPD-friendly maps and linking to smart inhalers to show remaining medication levels and feedback regarding usage techniques.
- Participants valued having information about the condition in the app, as many highlighted during PPI events that they often received little or no information following diagnosis. They valued easily accessible (not too technical) and reliable information.
- When developing disease support apps, discussing terminology with patients and their health care providers to determine which terminology they usually use is advisable. It may be acceptable to use medical terminology, as long as this is commonly used. In some cases, this may be more beneficial than trying to use lay terms. However, this may be restricted to those with a longer history of COPD who have had regular interactions with health care professionals regarding their COPD.

Recommendations

The principal recommendation based on our experience of working with a challenging target population on a large project with multiple stakeholders is that the population of interest, researchers, and developers should all be involved at every stage of the project and that an iterative approach be used to build a prototype. This prototype should then be tested in a wider environment with a larger group of the target population, where their interaction with the app is evaluated over a longer period to determine further issues and acceptability of the final intervention. The main stages are summarized in Figure 10.

Figure 10. Recommended steps.
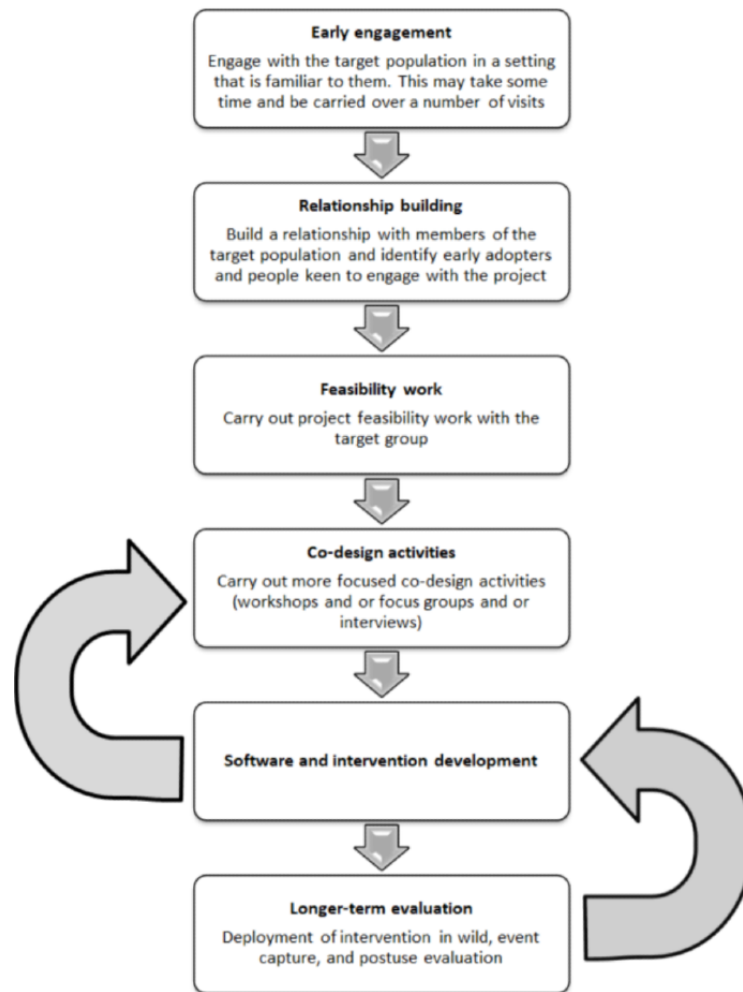
Limitations

Our PPI work involved participants from COPD support groups who expressed an interest in building a smartphone app. Such participants are more likely to be educated about COPD and may have more familiarity with technology than the standard COPD user. We attempted to mitigate this by capturing patterns of event behavior (ie button presses and transitions) from a wider group of users post deployment. There is an argument that such interventions in general do not adequately address the *digital divide*, as only those with access to such interventions in the first place may benefit. This rules out many older and low-income users [36]. Moreover, as our usability study was conducted with individuals who had a COPD diagnoses of more than 5 years' duration, findings may not adequately represent those who were recently diagnosed. In addition, the time of year is likely to have had an impact on the reported symptom frequency and severity. Finally, it should be noted that the 3-point scale for well-being used in our app is not a validated measure. It was developed based on user engagement and has

high face validity, but it is not clear how well it correlates with actual well-being.

Conclusions

We found that working with members of the target population at all stages of the project was a useful strategy; stakeholder engagement aids the development of research interventions that are both adaptive to the needs of the patient and the preferences of the provider [37]. This is different from a traditional researcher-led approach or a pure software engineering approach, such as *agile*. Placing the target population at the center of the work and iteratively building the intervention with the target users allows for the creation of a more acceptable final product. The combination of qualitative data analysis and data collected from open-source event capture tools also served to offer a further insight into app usage and dropout. We must try where possible to mitigate bias when carrying out such work, including social desirability bias and biases associated with participants who display more familiarity with technology than the typical target end user.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Example of framework analysis.
[DOCX File , 91 KB - humanfactors_v7i2e16289_app1.docx]

References

1. Lung disease in the UK | British Lung Foundation. 2018. Chronic Obstructive Pulmonary Disease (COPD) Statistics URL: https://statistics.blf.org.uk/copd [accessed 2018-03-20] [WebCite Cache ID 6y3yXrUnY]

2. Sobnath DD, Philip N, Kayyali R, Nabhani-Gebara S, Pierscionek B, Vaes AW, et al. Features of a mobile support app for patients with chronic obstructive pulmonary disease: literature review and current applications. JMIR Mhealth Uhealth 2017 Mar 20;5(2):e17 [FREE Full text] [doi: 10.2196/mhealth.4951] [Medline: 28219878]

3. World Health Organization. 2017 Dec 1. Chronic Obstructive Pulmonary Disease (COPD) URL: https://www.who.int/news-room/fact-sheets/detail/chronic-obstructive-pulmonary-disease-(copd) [accessed 2020-03-30]

4. Fletcher MJ, Upton J, Taylor-Fishwick J, Buist SA, Jenkins C, Hutton J, et al. COPD uncovered: an international survey on the impact of chronic obstructive pulmonary disease [COPD] on a working age population. BMC Public Health 2011 Aug 1;11:612 [FREE Full text] [doi: 10.1186/1471-2458-11-612] [Medline: 21806798]

5. Gershon AS, Dolmage TE, Stephenson A, Jackson B. Chronic obstructive pulmonary disease and socioeconomic status: a systematic review. Chron Obstruct Pulmon Dis 2012 Jun;9(3):216-226. [doi: 10.3109/15412555.2011.648030] [Medline: 22497534]

6. Lopez-Campos J, Calero C, Quintana-Gallego E. Symptom variability in COPD: a narrative review. Int J Chron Obstruct Pulmon Dis 2013;8:231-238 [FREE Full text] [doi: 10.2147/COPD.S42866] [Medline: 23687444]

7. Cavaillès A, Brinchault-Rabin G, Dixmier A, Goupil F, Gut-Gobert C, Marchand-Adam S, et al. Comorbidities of COPD. Eur Respir Rev 2013 Dec;22(130):454-475 [FREE Full text] [doi: 10.1183/09059180.00008612] [Medline: 24293462]

8. Department of Health and Social Care, The Rt Hon Andrew Lansley CBE. The Government of UK. 2012. GPs to 'Prescribe' Apps for Patients URL: https://www.gov.uk/government/news/gps-to-prescribe-apps-for-patients [accessed 2020-03-30]

9. Korpershoek YJ, Vervoort SC, Trappenburg JC, Schuurmans MJ. Perceptions of patients with chronic obstructive pulmonary disease and their health care providers towards using mHealth for self-management of exacerbations: a qualitative study. BMC Health Serv Res 2018 Oct 4;18(1):757 [FREE Full text] [doi: 10.1186/s12913-018-3545-4] [Medline: 30286761]

10. Boulos MN, Brewer AC, Karimkhani C, Buller DB, Dellavalle RP. Mobile medical and health apps: state of the art, concerns, regulatory control and certification. Online J Public Health Inform 2014;5(3):229 [FREE Full text] [doi: 10.5210/ojphi.v5i3.4814] [Medline: 24683442]

11. Martínez-Pérez B, de la Torre-Díez I, López-Coronado M. Privacy and security in mobile health apps: a review and recommendations. J Med Syst 2015 Jan;39(1):181. [doi: 10.1007/s10916-014-0181-3] [Medline: 25486895]

12. Vorrink S, Huisman C, Kort H, Troosters T, Lammers J. Perceptions of patients with chronic obstructive pulmonary disease and their physiotherapists regarding the use of an ehealth intervention. JMIR Hum Factors 2017 Sep 19;4(3):e20 [FREE Full text] [doi: 10.2196/humanfactors.7196] [Medline: 28928110]

13. Whitehead L, Seaton P. The effectiveness of self-management mobile phone and tablet apps in long-term condition management: a systematic review. J Med Internet Res 2016 May 16;18(5):e97 [FREE Full text] [doi: 10.2196/jmir.4883] [Medline: 27185295]

14. Yang F, Wang Y, Yang C, Hu H, Xiong Z. Mobile health applications in self-management of patients with chronic obstructive pulmonary disease: a systematic review and meta-analysis of their efficacy. BMC Pulm Med 2018 Sep 4;18(1):147 [FREE Full text] [doi: 10.1186/s12890-018-0671-z] [Medline: 30180835]

15. McCabe C, McCann M, Brady AM. Computer and mobile technology interventions for self-management in chronic obstructive pulmonary disease. Cochrane Database Syst Rev 2017 May 23;5:CD011425 [FREE Full text] [doi: 10.1002/14651858.CD011425.pub2] [Medline: 28535331]

16. Alwashmi MF, Hawboldt J, Davis E, Fetters MD. The iterative convergent design for mobile health usability testing: mixed methods approach. JMIR Mhealth Uhealth 2019 Apr 26;7(4):e11656 [FREE Full text] [doi: 10.2196/11656] [Medline: 31025951]

17. Brooke J. Hell - Jens Oliver Meiert. 1996. SUS - A quick and dirty usability scale URL: https://hell.meiert.org/core/pdf/sus.pdf [accessed 2020-03-30]

18. Spencer L, Ritchie J, O'Connor W, Morrell G, Ormston R. Analysis in practice. In: Ritchie J, editor. Qualitative Research Practice: A Guide for Social Science Students and Researchers. London: Sage Publications; 2013:295-343.

XSL•FO
RenderX

19. Feenstra TL, van Genugten ML, Hoogenveen RT, Wouters EF, Rutten-van Mölken MP. The impact of aging and smoking on the future burden of chronic obstructive pulmonary disease: a model analysis in the Netherlands. Am J Respir Crit Care Med 2001 Aug 15;164(4):590-596. [doi: 10.1164/ajrccm.164.4.2003167] [Medline: 11520721]

20. Scollon CN, Kim-Prieto C, Diener E. Experience sampling: promises and pitfalls, strengths and weaknesses. J Happiness Stud 2003;4(1):5-34. [doi: 10.1023/A:1023605205115]

21. Vigo M, Hassan L, Vance W, Jay C, Brass A, Cruickshank S. Britain Breathing: using the experience sampling method to collect the seasonal allergy symptoms of a country. J Am Med Inform Assoc 2018 Jan 1;25(1):88-92 [FREE Full text] [doi: 10.1093/jamia/ocx148] [Medline: 29240930]

22. Apaolaza A, Harper S, Jay C. Understanding Users in the Wild. In: Proceedings of the 10th International Cross-Disciplinary Conference on Web Accessibility. 2013 Presented at: W4A'13; May 13-15, 2013; Rio de Janeiro, Brazil p. 1-4. [doi: 10.1145/2461121.2461133]

23. R Core Team. The R Project for Statistical Computing. 2017. URL: https://www.r-project.org/ [accessed 2020-03-30]

24. Apaolaza A, Vigo M. WevQuery: testing hypotheses about web interaction patterns. Proc ACM Hum-Comput Interact 2017 Jun 30;1(1):1-17. [doi: 10.1145/3095806]

25. O'Connor S, Bouamrane MM, O'Donnell CA, Mair FS. Barriers to co-designing mobile technology with persons with dementia and their carers. Stud Health Technol Inform 2016;225:1028-1029. [Medline: 27332466]

26. Brunton L, Bower P, Sanders C. The contradictions of telehealth user experience in chronic obstructive pulmonary disease (COPD): A qualitative meta-synthesis. PLoS One 2015;10(10):e0139561 [FREE Full text] [doi: 10.1371/journal.pone.0139561] [Medline: 26465333]

27. Darvishy A, Hutter HP. Recommendations for age-appropriate mobile application design. Stud Health Technol Inform 2017;242:676-686. [Medline: 28873871]

28. Anderson M, Perrin A. Pew Research Center. 2017 May 17. Tech Adoption Climbs Among Older Adults URL: https://www.pewresearch.org/internet/2017/05/17/tech-adoption-climbs-among-older-adults/ [accessed 2020-03-30]

29. Verhagen SJ, Berben JA, Leue C, Marsman A, Delespaul PA, van Os J, et al. Demonstrating the reliability of transdiagnostic mHealth Routine Outcome Monitoring in mental health services using experience sampling technology. PLoS One 2017;12(10):e0186294 [FREE Full text] [doi: 10.1371/journal.pone.0186294] [Medline: 29023568]

30. Holtgraves T. Social desirability and self-reports: testing models of socially desirable responding. Pers Soc Psychol Bull 2004 Mar;30(2):161-172. [doi: 10.1177/0146167203259930] [Medline: 15030631]

31. Starling R, Nodulman JA, Kong AS, Wheeler CM, Buller DB, Woodall WG. Usability testing of an HPV information website for parents and adolescents. Online J Commun Media Technol 2015 Oct;5(4):184-203 [FREE Full text] [doi: 10.29333/ojcmt/2532] [Medline: 26594313]

32. Nelson LA, Mayberry LS, Wallston K, Kripalani S, Bergner EM, Osborn CY. Development and usability of REACH: a tailored theory-based text messaging intervention for disadvantaged adults with type 2 diabetes. JMIR Hum Factors 2016 Sep 8;3(2):e23 [FREE Full text] [doi: 10.2196/humanfactors.6029] [Medline: 27609738]

33. Bauer AM, Iles-Shih M, Ghomi RH, Rue T, Grover T, Kincler N, et al. Acceptability of mHealth augmentation of Collaborative Care: A mixed methods pilot study. Gen Hosp Psychiatry 2018;51:22-29 [FREE Full text] [doi: 10.1016/j.genhosppsych.2017.11.010] [Medline: 29272712]

34. Caraceni A, Galbiati A, Brunelli C, Gorni G, Martini C, Zecca E, et al. Cancer patient compliance in the self-administration of a pain assessment tool. J Pain Symptom Manage 2004 May;27(5):417-424 [FREE Full text] [doi: 10.1016/j.jpainsymman.2004.01.002] [Medline: 15120770]

35. Papapetropoulos SS. Patient diaries as a clinical endpoint in Parkinson's disease clinical trials. CNS Neurosci Ther 2012 May;18(5):380-387 [FREE Full text] [doi: 10.1111/j.1755-5949.2011.00253.x] [Medline: 22070400]

36. Choi NG, Dinitto DM. The digital divide among low-income homebound older adults: internet use patterns, eHealth literacy, and attitudes toward computer/internet use. J Med Internet Res 2013 May 2;15(5):e93 [FREE Full text] [doi: 10.2196/jmir.2645] [Medline: 23639979]

37. Shelef DQ, Rand C, Streisand R, Horn IB, Yadav K, Stewart L, et al. Using stakeholder engagement to develop a patient-centered pediatric asthma intervention. J Allergy Clin Immunol 2016 Dec;138(6):1512-1517. [doi: 10.1016/j.jaci.2016.10.001] [Medline: 27744029]

Abbreviations

COPD: chronic obstructive pulmonary disease
ESM: experience sampling method
mHealth: mobile health
PPI: patient and public involvement
RCTs: randomized controlled trials
SUS: System Usability Scale

XSL·FO
RenderX

Original Paper

Teamwork and Safety Attitudes in Complex Aortic Surgery at a Dutch Hospital: Cross-Sectional Survey Study

Alexander D Hilt[1], MD; Ad A Kaptein[2], PhD; Martin J Schalij[1], MD, PhD; Jan van Schaik[3], MD

[1]Department of Cardiology, Leiden University Medical Center, Leiden, Netherlands

[2]Department of Medical Psychology, Leiden University Medical Center, Leiden, Netherlands

[3]Department of Vascular Surgery, Leiden University Medical Center, Leiden, Netherlands

Corresponding Author:
Jan van Schaik, MD
Department of Vascular Surgery
Leiden University Medical Center
Albinusdreef 2
Leiden, 2333 ZA
Netherlands
Phone: 31 715299407
Email: J.van_Schaik@lumc.nl

Abstract

Background: Improving teamwork in surgery is a complex goal and difficult to achieve. Human factors questionnaires, such as the Safety Attitudes Questionnaire (SAQ), can help us understand medical teamwork and may assist in achieving this goal.

Objective: This paper aimed to assess local team and safety culture in a cardiovascular surgery setting to understand how purposeful teamwork improvements can be reached.

Methods: Two cardiovascular surgical teams performing complex aortic treatments were assessed: an endovascular-treatment team (ETT) and an open-treatment team (OTT). Both teams answered an online version of the SAQ Dutch Edition (SAQ-NL) consisting of 30 questions related to six different domains of safety: teamwork climate, safety climate, job satisfaction, stress recognition, perceptions of management, and working conditions. In addition, one open-ended question was posed to gain more insight into the completed questionnaires.

Results: The SAQ-NL was completed by all 23 ETT members and all 13 OTT members. Team composition was comparable for both teams: 57% and 62% males, respectively, and 48% and 54% physicians, respectively. All participants worked for 10 years or more in health care. SAQ-NL mean scores were comparable between both teams, with important differences found between the physicians and nonphysicians of the ETT. Nonphysicians were less positive about the safety climate, job satisfaction, and working climate domains than were the physicians ($P<.05$). Additional education on performed procedures, more conjoined team training, as well as a hybrid operating room were suggested by participants as important areas of improvement.

Conclusions: Nonphysicians of a local team performing complex endovascular aortic aneurysm surgery perceived safety climate, job satisfaction, and working conditions less positively than did physicians from the same team. Open-ended questions suggested that this is related to a lack of adequate conjoined training, lack of adequate education, and lack of an adequate operating room. With added open-ended questions, the SAQ-NL appears to be an assessment tool that allows for developing strategies that are instrumental in improving quality of care.

Introduction

The World Health Organization (WHO) has stated that knowledge on human factors (HF), especially nontechnical skills, is crucial in developing safe environments for patients [1]. A 2017 analysis of the Dutch health care system showed that nontechnical aspects of work were understudied in professional training [2,3]. Nontechnical dimensions of teamwork, such as communication, stress awareness, and shared decision making, all contribute to the effectiveness of teamwork. Importantly,

failing to invest in these issues may have negative effects on patient safety and clinical outcomes [4-6]. The challenge lies in how to identify, analyze, and improve these nontechnical skills.

In aviation and offshore industries, for example, awareness of nontechnical skills is crucial in daily work. Training and improving nontechnical skills are often part of corporate policies, with proven effects on safety [7,8]. Similarly, positive results have been observed in health care, although the number of studies is scarce [9,10]. Understanding the safety culture and climate within a team is central to improving nontechnical skills. This can be assessed through questionnaires such as the Safety Attitudes Questionnaire (SAQ), which is a medical HF questionnaire that has been validated in different medical domains. In 2016, the SAQ Dutch Edition (SAQ-NL) was the questionnaire validated in the Dutch language [11,12]. Although often used to assess an *ex ante* baseline and the *ex post* effect of team trainings, the SAQ-NL as a diagnostic tool is not commonly used to identify what exactly needs changing within a team nor to adjust subsequent training accordingly.

The outcome of complex aortic aneurysm surgery is highly dependent on team dynamics. Aortic aneurysms are defined as *complex* when important side branches are included in the aneurysm. This necessitates inclusion of these side branches in the vascular reconstruction, making the procedure high risk. Open, as well as endovascular complex aortic, reconstructions are associated with high mortality and morbidity rates. Both treatments are conducted by multidisciplinary teams.

In this study, the SAQ-NL was used as a diagnostic tool to examine teamwork and safety climate in two types of teams: an open-treatment team (OTT) and an endovascular-treatment team (ETT). The aim of this study was to understand, and ultimately help improve, teamwork conditions and safety climate in this high-risk setting. Primarily, it was hypothesized that (1) the SAQ-NL will provide insight into how teamwork and safety is perceived by different team members and (2) this knowledge may help guide future teamwork improvement strategies.

Methods

Terminology

Pinpointing safety culture and safety climate within a medical department is difficult, especially because they are not mutually exclusive. The safety culture of an organization is the product of individual and group values, traditions, perceptions, and competences that determine the commitment to, and the style and proficiency of, an organization's health and safety management [13]. An organization's safety culture is the context in which personal safety attitudes develop, persist, and are promoted [8]. It is like a "script" that is taught to every employee that is continuously formed, shaped, and reshaped not only by themselves, but also by their fellow "actors" in the work setting. This concept has been used widely since the 1980s in aviation, as well as industrial settings, such as power plants and offshore environments.

The safety climate is the manifestation of that safety culture in the behaviors and attitudes of professionals, for instance, during surgical procedures. When one would take a "snapshot" of such an environment, certain behavioral cues would be seen; for example, a surgeon being focused on the patient and on his or her tools, the scrub nurse seeing a drop in blood pressure, and the anesthetist reacting accordingly. This "snapshot" with all the interactions between professionals can be seen as the climate people are working in. This climate (ie, the "play" or the day-to-day atmosphere when working) is directly influenced by the department's culture (ie, the "script" which consists of perceptions, beliefs, and traditions). For example, when convention holds that nurses do not speak up when things go wrong, this negatively impacts the safety climate and often leads to errors and eventually diminished patient safety [14].

Measuring perceptions of safety and teamwork in a specific setting at a certain point in time (ie, during a surgical procedure) provides insight into the safety climate as well as the safety culture. Put differently, it allows for the assessment of how every "actor" plays their role and, while doing so, to what extent they are influenced by others and the "script" used. Figure 1 gives an overview of the terminology used.

Figure 1. Safety culture and safety climate (source: AD Hilt).
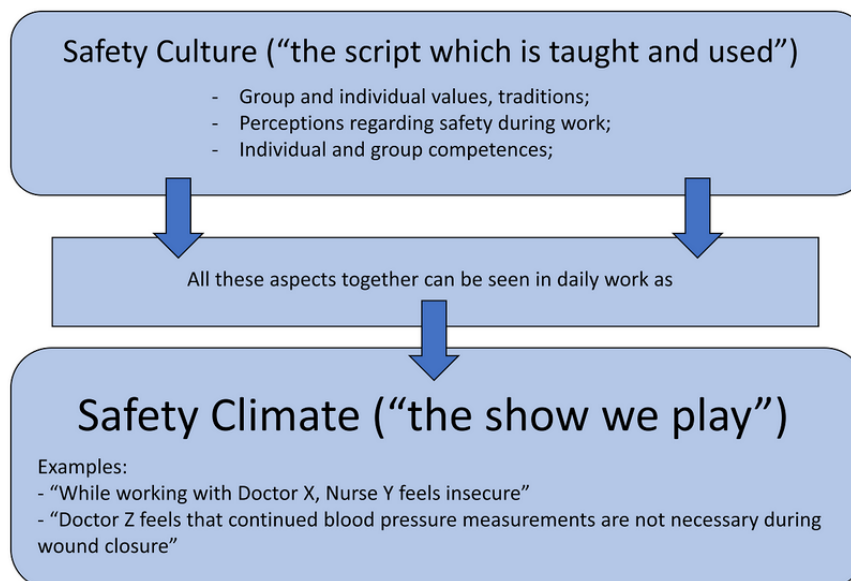
Design and Study Setting

This study followed a cross-sectional survey design. The Leiden University Medical Centre is one of eight university hospitals in the Netherlands. For this study, two complex aortic aneurysm treatment teams were evaluated: the ETT and the standard OTT.

There were two reasons for the assessment of the two teams. Firstly, the endovascular treatment is relatively new to this hospital, having been performed starting in 2013. Refinement of nontechnical skills is of great interest in this setting, since it has been shown that this improves patient safety and outcomes [10,15]. Secondly, the introduction of the endovascular treatment demanded a shift in work environment for part of the team.

The OTT continued to work in the familiar environment of their operating theater, whereas the ETT had to perform their procedures in an angiography suite, an environment where many team members were not used to working. For daily workflow of the ETT, it was important to understand how it was influenced by this shift in environment. An outline of routine ETT and OTT procedures is shown in Multimedia Appendix 1.

Study Population

The ETT consisted of 23 team members with a large diversity of radiology personnel, surgical staff, and the addition of a supplier specialist. The OTT consisted of 13 team members with predominantly surgical staff and perfusionists, the latter not being included in the ETT. Noticeably, a supplier specialist was present in the ETT but not the OTT. The specific role of the supplier specialist lies in participating in the discussion of stent type and design, as well as on-site product advice during the procedure. The supplier specialist is a standard, crucial team member of the ETT. Additionally, it should be noted that 2 vascular surgeons, 1 neurologist, and 1 clinical neurophysiology technician were part of both teams. The partial overlap of members of different teams is common in medical settings. All 4 interviewees with dual team membership were able to clearly distinguish between the two teams when answering our questions. In all further analyses, vascular surgeons, thoracic surgeons, radiologists, anesthetists, and neurologists are referred to as *physicians*, whereas scrub nurses, nurse anesthetists, clinical neurophysiology technicians, radiology technicians, supplier specialists, and perfusionists are referred to as *nonphysicians*. Table 1 summarizes the physician and nonphysician composition of both teams, as well as health care tenure and team tenure.

Table 1. Overview of team composition in the endovascular-treatment team (ETT) versus the open-treatment team (OTT).

Team and members	N (%)	Average health care tenure, years	Average team tenure, years
ETT (N=23)			
Radiologist	2 (9)	≥10	≥5
Thoracic surgeon	1 (4)	≥10	4
Anesthetist	3 (13)	≥10	≥5
Vascular surgeon	4 (17)	≥10	4
Neurologist	1 (4)	≥10	3
Radiology technician	5 (22)	≥10	≥5
Scrub nurse	3 (13)	8	≥5
Nurse anesthetist	1 (4)	≥10	≥5
Clinical neurophysiology technician	2 (9)	≥10	4
Supplier specialist	1 (4)	8	≥5
OTT (N=13)			
Thoracic surgeon	1 (8)	≥10	3
Anesthetist	2 (15)	≥10	1
Vascular surgeon	3 (23)	≥10	≥5
Neurologist	1 (8)	≥10	4
Scrub nurse	2 (15)	9	4
Nurse anesthetist	1 (8)	≥10	4
Clinical neurophysiology technician	1 (8)	≥10	4
Perfusionist	2 (15)	≥10	≥5

Human Factors and the Safety Attitudes Questionnaire

Research into HF aims to understand how humans function in different environments, in order to improve human performance and safety within these environments [16]. HF research has become a core part of major industries, such as aviation and the offshore industry, mainly because of the high dependence on human performance and its effect on safety. Teamwork safety has been extensively evaluated in aviation through HF questionnaires, originally through the Cockpit Management Attitudes Questionnaire (CMAQ) [7,17]. This questionnaire assessed the perceptions concerning safety climate and teamwork among personnel working on an aircraft. This was later refined into the Flight Management Attitudes Questionnaire (FMAQ) [7]. In the medical domain, intensive care units were the first to adopt a medical version of the FMAQ: the Intensive Care Unit Management Attitudes Questionnaire (ICUMAQ) [17]. Developed by Sexton et al, the SAQ is a refinement of the ICUMAQ for a health care setting. It has proven its psychometric and clinical quality in different clinical settings, as well as in the Dutch setting (ie, the SAQ-NL) [11,17,18]. The SAQ assesses 30 items in six domains: safety climate (SC), teamwork climate (TC), job satisfaction (JS), stress recognition (SR), perceptions of management (PoM), and working conditions (WC). The 30 items are each assessed on a 5-point Likert scale: disagree strongly (1), disagree slightly (2), neutral (3), agree slightly (4), and agree strongly (5). The WHO indicates that the SAQ is a valuable HF instrument for assessing medical teamwork dynamics in a standardized fashion [1]. For this study, the strong methodological foundation of the SAQ and its usability in the field were the main reasons to use it.

Additionally, to gain insight into teamwork, safety attitudes, and the meaning of the SAQ-NL outcomes, respondents were asked to answer the following open-ended question: "What are your top three recommendations for improving patient safety in this clinical area?" A Web-based survey of the SAQ-NL via Google Forms (Google) was sent to all ETT and OTT members (see Multimedia Appendix 2).

Statistics

Frequency tables for gender, professional positions, team tenure, and general health care tenure were generated to give an overview of both teams. Response patterns are shown as percentages. For normally distributed categorical data, a chi-square test was used to calculate statistical differences. For each SAQ dimension, mean scores and standard deviations were calculated per team (ie, ETT and OTT), per professional group (ie, physicians and nonphysicians), and per department. An unpaired t test was used to calculate differences between the SAQ-NL mean scores for the ETT and the OTT. A univariate analysis of variance (ANOVA) test was performed to evaluate whether there was a significant difference between average SAQ-NL scores among professional groups, the ETT and OTT, as well as the departments. Data from the open-ended questions were displayed in a descriptive manner; content analysis was used to analyze these. Two authors (ADH and JvS) labelled responses according to major themes that emerged from the

data. Cronbach alpha was calculated for all SAQ dimensions of our sample. For analysis, SPSS Statistics for Windows, version 23.0 (IBM Corp), was used. A *P* value of less than .05 was considered significant.

Biases

Teamwork and safety are delicate subjects, leading to a risk of response bias. Examples of response bias are *question order bias* and *social desirability bias*. The use of a self-administered questionnaire via an online survey is known to minimize the latter effect [19]. All questionnaire data were available only to the main researcher (ADH), who has no professional position in the ETT or the OTT.

Ethical Considerations

By Dutch law, no ethical approval was needed to conduct this study. All participants gave informed consent for participating in the study and the use of their pseudoanonymized data.

Results

Demographics

The ETT consisted of 23 members of which 13 (57%) were male and 11 (48%) were physicians. The OTT consisted of 13 members of which 8 (62%) were male and 7 (54%) were physicians. The composition of the teams regarding number of males and physicians was not significantly different (*P*=.60 and

P=.50, respectively; see Table 2). Team tenure of 5 years or more was more prevalent among the ETT (12/23, 52%) than among the OTT (3/13, 23%), but this difference was not statistically significant (*P*=.16; see Table 2). Both teams had a large proportion of members working 10 years or more in health care (ETT 19/23, 83%, vs OTT 12/13, 92%, *P*=.30). Long working weeks (ie, ≥50 hours) were more prevalent among the OTT than among the ETT; however, this difference was not significant (OTT 6/13, 46%, vs ETT 5/23, 22%, *P*=.50).

Mean Scores From the Dutch Safety Attitudes Questionnaire: Endovascular-Treatment Team Versus Open-Treatment Team

An overview of mean SAQ-NL scores with standard deviations per domain is shown in Table 3. Higher means were observed for the OTT; however, an independent-samples *t* test showed that for all SAQ-NL domains, no statistically significant differences existed between the ETT and OTT.

Mean scores for the SAQ dimensions for the ETT and OTT, respectively, were as follows: TC 3.7 (SD 0.37) vs 3.9 (SD 0.31), *P*=.40; SC 3.6 (SD 0.43) vs 3.7 (SD 0.31), *P*=.65; JS 4.1 (SD 0.50) vs 4.2 (SD 0.46), *P*=.39; SR 3.0 (SD 0.73) vs 3.1 (SD 0.92), *P*=.84; PoM 2.9 (SD 0.66) vs 3.1 (SD 0.51), *P*=.44; and WC 3.5 (SD 0.64) vs 3.6 (SD 0.70), *P*=.69. For our sample, all SAQ domains had an acceptable level of reliability (alpha≥.70), with the exception of the TC domain, which had poor reliability (alpha=.58).

Table 2. Demographics of the endovascular-treatment team (ETT) and the open-treatment team (OTT).

Demographic	ETT (N=23), N (%)	OTT (N=13), N (%)	*P* value
Male	13 (57)	8 (62)	.60
Physician	11 (48)	7 (54)	.50
Team tenure of ≥5 years	12 (52)	3 (23)	.16
Health care tenure of ≥10 years	19 (83)	12 (92)	.30
Weekly work time of ≥50 hours	5 (22)	6 (46)	.50
Response	23 (100)	13 (100)	N/A[a]

[a]N/A: not applicable.

Table 3. Scores from the Safety Attitudes Questionnaire Dutch Edition (SAQ-NL) per domain.

Respondents	Scores for each domain, mean (SD)					
	Teamwork climate	Safety climate	Job satisfaction	Stress recognition	Perceptions of management	Working conditions
Team						
Endovascular-treatment team (ETT) (N=23)	3.7 (0.37)	3.6 (0.43)	4.1 (0.50)	3.0 (0.73)	2.9 (0.66)	3.5 (0.64)
Open-treatment team (OTT) (N=13)	3.9 (0.31)	3.7 (0.31)	4.2 (0.46)	3.1 (0.92)	3.1 (0.51)	3.6 (0.70)
Positions within each team						
ETT						
Nonphysician[a] (n=12)	3.6 (0.43)	*3.4 (0.35)[e]*	*3.8 (0.41)[e]*	2.9 (0.61)	2.7 (0.67)	*3.2 (0.68)[e]*
Physician (n=11)	3.9 (0.31)	*3.9 (0.34)[e]*	*4.4 (0.33)[e]*	3.1 (0.86)	3.1 (0.64)	*3.9 (0.37)[e]*
OTT						
Nonphysician[a] (n=6)	3.8 (0.40)	3.7 (0.33)	4.0 (0.47)	3.0 (0.93)	2.9 (0.43)	3.5 (0.54)
Physician (n=7)	3.9 (0.23)	3.7 (0.33)	4.4 (0.39)	3.1 (0.98)	3.2 (0.52)	3.7 (0.83)
Department within each team						
ETT						
Surgery (n=8)	3.8 (0.35)	3.7 (0.39)	4.0 (0.56)	3.1 (0.62)	2.9 (0.86)	3.3 (0.56)
Anesthesiology (n=4)	3.9 (0.26)	4.0 (0.32)	4.4 (0.51)	2.6 (1.12)	3.0 (0.00)	4.1 (0.17)
Radiology (n=7)	3.7 (0.45)	3.4 (0.41)	4.1 (0.46)	3.2 (0.49)	2.5 (0.39)	3.2 (0.79)
Neurology (n=3)	3.4 (0.20)	3.5 (0.59)	4.0 (0.40)	3.4 (0.76)	3.6 (0.53)	4.0 (0.33)
Industry (n=1)[b]	4.4	4.1	4.2	2.0	3.6	4.0
OTT						
Surgery (n=8)	3.9 (0.30)	3.7 (0.38)	4.3 (0.46)	3.0 (1.01)	3.0 (0.51)	3.5 (0.39)
Anesthesiology (n=3)	3.6 (0.34)	3.7 (0.10)	4.3 (0.61)	2.8 (0.90)	2.8 (0.00)	3.4 (0.96)
Radiology (n=0)[c]	N/A[d]	N/A	N/A	N/A	N/A	N/A
Neurology (n=2)	4.0 (0.00)	3.7 (0.40)	4.1 (0.42)	3.8 (0.35)	3.7 (0.42)	4.7 (0.47)
Industry (n=0)[c]	N/A	N/A	N/A	N/A	N/A	N/A
Overlapping members within each team						
ETT (n=1 of each)[b]						
Vascular surgeon W	4.2	4.2	4.6	2.3	2.4	3.7
Vascular surgeon X	3.4	3.5	4.2	3.8	2.4	3.4
Neurologist Y	3.6	4.1	4.4	4.3	3.8	4.3
Clinical neurophysiology technician Z	3.4	3.3	4.0	3.3	3.0	3.7
OTT (n=1 of each)[b]						
Vascular surgeon W	4.2	4.2	5.0	1.8	3.4	4.0
Vascular surgeon X	4.2	3.2	4.4	3.7	2.6	3.3
Neurologist Y	4.0	4.0	4.4	4.0	4.0	5.0
Clinical neurophysiology technician Z	4.0	3.4	3.8	3.5	3.4	4.3

[a]Nonphysicians include scrub nurses, nurse anesthetists, clinical neurophysiology technicians, radiology technicians, supplier specialist, and perfusionists.

[b]Because there is only 1 member within this group (or within each group), SDs were not calculated.

[c]Because there are no members in this group, scores were not collected.

[d]N/A: not applicable.

[e]Statistical difference, $P<.05$.

XSL·FO
RenderX

Mean Scores From the Dutch Safety Attitudes Questionnaire: Physicians Versus Nonphysicians

Univariate ANOVA showed that for the ETT, there were significant differences between physicians and nonphysicians on mean scores for the SC, JS, and WC domains; physicians were significantly more positive about SC, JS, and WC compared to nonphysicians. Mean scores for these domains for physicians versus nonphysicians, respectively, were as follows: SC 3.9 (SD 0.34) vs 3.4 (SD 0.35), P=.002; JS 4.4 (SD 0.33) vs 3.8 (SD 0.41), P=.001; and WC 3.9 (SD 0.37) vs 3.2 (SD 0.68), P=.008. For the ETT, the supplier specialist did not have significantly different scores from the other nonphysicians (see Table 3); there was a slight trend toward higher TC (P=.08) and SC (P=.07) scores. For the OTT, besides a slight trend toward higher mean scores among physicians for the JS domain—3.7 (SD 0.83) vs 3.5 (SD 0.54), P=.12—no significant differences were found between scores from physicians and nonphysicians for all domains.

Mean Scores From the Dutch Safety Attitudes Questionnaire: Departmental Differences

Univariate ANOVA and independent t tests showed no statistical differences between members of different departments (ie, radiology, surgery, neurology, industry, and anesthesiology) among the ETT and OTT.

Subanalysis of Mean Scores From the Dutch Safety Attitudes Questionnaire: Overlapping Team Members

A total of 3 physicians and 1 technician filled out both the ETT and OTT questionnaires; the mean SAQ-NL scores are also shown in Table 3. An independent t test showed no significant differences between the ETT and OTT for any of the SAQ-NL domains in this group. Despite a slight trend toward lower JS among nonphysicians (P=.18), no significant differences were found for any of the domains when comparing physicians and nonphysicians in the ETT and OTT, both through univariate ANOVA.

When eliminating these 4 participants from the total analysis of physicians versus nonphysicians in the ETT and OTT, univariate ANOVA showed identical results for the ETT; mean scores for SC (P=.002), JS (P<.001), and WC (P=.008) were significantly lower among nonphysicians compared to physicians in the ETT but not in the OTT.

Open-Ended Questions

Out of 23 members in the ETT, 21 (91%) respondents together provided 50 comments. Of the 13 members in the OTT, 7 (54%) respondents together provided 14 comments. For the ETT, five themes were identified through content analysis. Comments were related to periprocedural planning; dynamics during procedures, both technical and nontechnical aspects; facilities present in the operating room (OR); and patient privacy (see Multimedia Appendix 3). In total, 23 out of 50 comments (46%) were related to teamwork between nonphysicians and physicians. Nonphysicians expressed their desire to be more involved in the surgical process (12/23 comments, 52%); individual example quotes were as follows: "... more open communication about the patients' status during surgery," "...

more clarification of the surgical steps taken," and "... more debriefing after performed surgery." Physicians found the education of nonphysicians to be an important issue (10/23 comments, 43%); individual example quotes were as follows: "... more time for extra training," "... more team members should attend the conjoined presurgery meetings," "... there should be more postsurgery evaluations together," and "... more open communication at different stages in surgery should be applied toward all." Additionally, the need for a hybrid OR (ie, fit for both open and endovascular treatment) was stressed (11/50 comments, 22%): "... a hybrid OR where all the radiology and surgery devices are available is a must."

For the OTT, two major themes were identified; comments were related to periprocedural planning and dynamics during procedures (ie, nontechnical aspects). In total, 6 out of 14 comments (43%) were education related. Nonphysicians wanted to be educated more (4/6 comments, 67%); individual example quotes were as follows: "... there should be more clinical classes about this procedure done by the anesthetist and surgeons" and "... there should be more dedicated trainings and preparation." Physicians also expressed a desire for more education of nonphysicians in the different phases of surgery (2/6 comments, 33%); individual example quotes were as follows: "... if there are lessons learned during procedures, we should conjointly evaluate them" and "... clinical evaluations after surgery should be evaluated with the whole team." An overview of relevant themes for both the ETT and OTT with example remarks is included in Multimedia Appendix 3.

Discussion

Principal Findings

The results of this study can be summarized as follows: (1) physicians from the ETT were more positive about SC, JS, and WC than were nonphysicians; (2) conjoined training sessions, education, postprocedural evaluation, and a hybrid OR are important topics for future improvements for both physicians and nonphysicians from the ETT; and (3) using the SAQ-NL with the addition of open-ended questions was an instrumental way of assessing the safety culture and climate of two surgical teams and to propose strategies to improve this further.

The findings of our local study suggest that there is room for improvement in teamwork within the ETT. Regarding SC, JS, and WC domains, physicians were more positive than nonphysicians, which was not observed in the OTT. These outcomes were specified by the answers to the open-ended questions. In particular, the remarks regarding more conjoined education on procedures and the desire for a hybrid OR provide a good explanation for the lower scores on the JS and WC domains, and possibly the SC domain, within the nonphysician group. Higher SC, JS, and WC scores reflected aspects of overall perceptions regarding commitment to safety, the work experience, and the quality of the work environment (ie, equipment and staffing), respectively. It is striking that this was different from the OTT. A reasonable explanation for lower JS and WC scores in the ETT may be that nonphysicians need to operate outside of their own habitat, in an environment (ie, the angiography suite) they are not familiar with and do not know

as well as the OR. This setup is due to the absence of adequate radiological facilities in the OR. This condition results in nonphysicians having to move large amounts of instruments and materials from the OR to the angiography suite. Having to work outside of their familiar environment and having to move surgical equipment is not necessary for OTT members, who operate in the OR where all materials are close at hand. Qualitative results suggest that building a hybrid OR must be prioritized to raise ETT scores to the level of OTT scores. A hybrid OR is a fully functional surgical theater that is equipped with advanced medical imaging devices, such as fixed C-arms, computed tomography scanners, or magnetic resonance imaging scanners. These imaging devices enable complex, minimally invasive surgery as well as *hybrid* procedures where minimally invasive techniques are combined with conventional *open* surgery.

The perceived need for more education and adequate working conditions could also explain the lower SC score among nonphysicians of the ETT. For future improvements, some suggestions would be cross-functional teaching between radiology technicians and scrub nurses, a more explicit definition of roles and use of equipment, and instruction for team members by physicians. SAQ-NL outcomes can be used after these improvements to measure the effect of these changes in working circumstances on teamwork.

Implications for Surgical Procedures

Previous studies have shown the effectiveness of using the SAQ as a measure to assess teamwork in different medical settings, largely focusing on measuring the effect of team trainings on daily work [20,21]. The SAQ-NL has not been solely used as a diagnostic tool.

Although no overall differences were found in our study between the ETT and OTT as a whole, there were important differences within the ETT. Physicians were more positive than nonphysicians. Through open-ended questions, important themes for improvement of daily procedures were found. Differences between physicians and nonphysicians are not new [10,22]. However, this is still an important finding, especially for a large tertiary referral hospital. Our findings are not only useful for patient-facing employees, but also for team managers. These findings stress not only the need for facilitating conjoined training and education, but also to direct this more specifically toward the needs of the employees. An example of the latter is *slowing down during surgery*, which enables team members to ask questions at certain key points during the surgical process [23].

Outcomes of the Dutch Safety Attitudes Questionnaire

Improving health care team culture and teamwork safety is not straightforward, and thorough assessments of workflow and interactions between different professionals are time-consuming. While improvements are necessary, trying to change the entire health care system at once is doomed to fail because of the complex nature of this working environment. For instance, it is questionable what the relevance of a national teamwork assessment would be, essentially assessing teamwork among thousands of people having no direct interaction with each other.

Therefore, as proposed by Sexton et al, it is especially important to put effort into the analysis of the working environment of patient-facing employees and focus on local settings [18].

Attitudinal surveys on a local team level can be a valuable addition to this. This study shows that small teams can be fruitfully assessed using the SAQ-NL. Firstly, the strength of using the SAQ-NL among small teams is that a complete response rate is more easily obtained. Secondly, the clinical implications of the study outcomes can be used immediately. For example, regarding the education-related remarks, a focus on more education during procedures can be started during the next surgery. The SAQ-NL could subsequently be used to monitor how such changes would influence a team's safety attitudes.

Lastly, the SAQ-NL is a useful tool in a cross-professional setting. Due to the intertwinement of work, the supplier specialist, for example, cannot be left out of the ETT analysis. The SAQ-NL in this sense is not restricted to particular professions.

Future Perspectives: Human Factors and Team Analysis

Assessing team processes such as SC through the SAQ-NL is a valuable addition to team analysis. A recent meta-analysis by Schmutz et al assessed the impact of team process analysis on team performance [24]. It showed that teams who are aware of processes during daily work were almost three times more likely to achieve high performance than teams who were not. In line with this meta-analysis, and as we hypothesized, we recognize the SAQ-NL as a valuable diagnostic tool for team process analysis, mainly to assess and create awareness of processes among team members that define their daily work.

With the knowledge of what needs attention during daily teamwork, a next step could be HF trainings, such as Crew Resource Management (CRM) or Team Strategies and Tools to Enhance Performance and Patient Safety (TeamSTEPPS) [25]. Both are proven to be effective in altering team performance through HF principles. They teach participants that people have certain strengths and weaknesses that can impact daily work in a good or bad way [16,26-28]. The SAQ is often used to monitor the effects of these HF trainings. O'Dea et al proposed in their meta-analysis that, while plausible, it is difficult to unambiguously link changes in team behavior or SAQ outcomes to a particular training [29]. However, regarding the SAQ, starting with a diagnostic approach of what needs attention in a team before commencing training, the effect of CRM or TeamSTEPPS could be better understood during the course of training. For our sample, a CRM or TeamSTEPPS training could aim at improving communication during crucial steps of the ETT procedures, in order to assure shared understanding between physicians and nonphysicians and hereby increase the SC.

Limitations

Our study has several limitations. Firstly, it is debatable what the clinical meaning or implication is of the difference between sections of the Likert scale in daily work. When looking at the ETT outcomes between nonphysicians and physicians, for

XSL•FO
RenderX

example, the difference for the JS domain is 0.6 and for the WC domain is 0.7. What this statistically significant difference implies, solely from the questionnaire's outcome, is not directly clear. However, using open-ended questions helps us understand this difference. Secondly, we are well aware that there is overlap in respondents filling out the SAQ-NL for both ETT and OTT. In this small group, no differences were found between physicians and nonphysicians for both the ETT and OTT. Correcting all data for this group did not alter the main outcomes. Thirdly, the original SAQ and the SAQ-NL showed good psychometric properties and good reliability (average Cronbach alpha of .76). In our study, the reliability was generally acceptable (alpha≥.70), with the exception of the TC domain, which had rather poor internal reliability (alpha=.58).

However, this is highly dependent on the number of subjects participating in the study and the number of items per dimension. Further use of the SAQ-NL and research in this setting should be stressed to evaluate the psychometric properties of the SAQ-NL.

Conclusions

Nonphysicians of a local team performing endovascular aortic aneurysm surgery perceived SC, JS, and WC less positively than physicians on the same team. Open-ended questions specified this to be related to a lack of adequate conjoined training, lack of adequate education, and lack of an adequate OR. The SAQ-NL can be a first step in developing strategies to improve quality of care.

Authors' Contributions

ADH and JvS conceived of the presented idea. JvS provided input from an earlier study of the ETT [15]. ADH, AAK, and JvS developed the theory. ADH performed the data collection and analysis. AAK, MJS, and JvS supervised the research and critically reviewed the findings. ADH and JvS drafted the manuscript. All authors discussed the results and contributed to the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Typical endovascular-treatment team (ETT) and open-treatment team (OTT) procedure days.
[DOCX File , 16 KB - humanfactors_v7i2e17131_app1.docx]

Multimedia Appendix 2
Safety Attitudes Questionnaire Dutch Edition (SAQ-NL).
[DOCX File , 15 KB - humanfactors_v7i2e17131_app2.docx]

Multimedia Appendix 3
Themes and example excerpts from analysis of the Safety Attitudes Questionnaire Dutch Edition (SAQ-NL) for the endovascular-treatment team (ETT) versus the open-treatment team (OTT).
[DOCX File , 16 KB - humanfactors_v7i2e17131_app3.docx]

References

1.   Flin R, Winter J, Sarac C, Raduma Tomas MA. Human Factors in Patient Safety: Review of Topics and Tools. Geneva, Switzerland: World Health Organization; 2009 Apr. URL: https://www.who.int/patientsafety/research/methods_measures/human_factors/human_factors_review.pdf [accessed 2020-03-04]

2.   Klopotowska J, Schutijser B, de Bruijne M, Wagner C. Tweede Evaluatie van het VMS Veiligheidsprogramma [Second Evaluation of the VMS Safety Program]. Utrecht, the Netherlands: NIVEL; 2016 Nov. URL: https://nivel.nl/sites/default/files/bestanden/Rapport_Tweede_Evaluatie_VMS.pdf [accessed 2020-03-04]

3.   Langelaan M, Broekens M, de Bruijne M, Groot JF, Moesker MJ, Porte PJ, et al. Monitor Zorggerelateerde Schade 2015/2016: Dossieronderzoek bij Overladen Patienten in Nederlandse Ziekenhuizen [Monitoring Health Care-Related Damage 2015/2016: Case Research of Deceased Patients in Dutch Hospitals]. Utrecht, the Netherlands: NIVEL; 2017. URL: https://nivel.nl/sites/default/files/bestanden/Rapport_Monitor_Zorggerelateerde_Schade_2017.pdf [accessed 2020-03-04]

4.   Kim SE, Kim CW, Lee SJ, Oh JH, Lee DH, Lim TH, et al. A questionnaire survey exploring healthcare professionals' attitudes towards teamwork and safety in acute care areas in South Korea. BMJ Open 2015 Jul 24;5(7):e007881 [FREE Full text] [doi: 10.1136/bmjopen-2015-007881] [Medline: 26209120]

5.   Committee on Quality of Health Care in America, Institute of Medicine, Kohn LT, Corrigan JM, Donaldson ME. To Err is Human: Building a Safer Health System. Washington, DC: National Academy Press; 2000. URL: https://www.ncbi.nlm.nih.gov/books/NBK225182/pdf/Bookshelf_NBK225182.pdf [accessed 2020-03-05]

6. Warnack E, Pachter HL, Choi B, DiMaggio C, Frangos S, Klein M, et al. Postinjury complications: Retrospective study of causative factors . JMIR Hum Factors 2019 Sep 26;6(3):e14819 [FREE Full text] [doi: 10.2196/14819] [Medline: 31573897]

7. Helmreich R, Merritt A, Wilhelm J. The evolution of Crew Resource Management training in commercial aviation. Int J Aviat Psychol 1999;9(1):19-32. [doi: 10.1207/s15327108ijap0901_2] [Medline: 11541445]

8. Mearns K, Whitaker SM, Flin R. Safety climate, safety management practice and safety performance in offshore environments. Saf Sci 2003 Oct;41(8):641-680. [doi: 10.1016/s0925-7535(02)00011-5]

9. Pal A, Lal R, Frizelle F. Aviation-based teamwork skills work for surgeons: Time for an 'aviation bundle'? ANZ J Surg 2018 Dec;88(12):1231-1235. [doi: 10.1111/ans.14892] [Medline: 30306705]

10. Haerkens MH, Kox M, Lemson J, Houterman S, van der Hoeven JG, Pickkers P. Crew Resource Management in the intensive care unit: A prospective 3-year cohort study. Acta Anaesthesiol Scand 2015 Nov;59(10):1319-1329 [FREE Full text] [doi: 10.1111/aas.12573] [Medline: 26079640]

11. Haerkens MH, van Leeuwen W, Sexton JB, Pickkers P, van der Hoeven JG. Validation of the Dutch language version of the Safety Attitudes Questionnaire (SAQ-NL). BMC Health Serv Res 2016 Aug 15;16(a):385 [FREE Full text] [doi: 10.1186/s12913-016-1648-3] [Medline: 27528393]

12. Ajlan AM, Harsh GR. The human factor and safety attitudes in neurosurgical operating rooms. World Neurosurg 2015 Jan;83(1):46-48. [doi: 10.1016/j.wneu.2013.08.039] [Medline: 24012553]

13. Mardon RE, Khanna K, Sorra J, Dyer N, Famolaro T. Exploring relationships between hospital patient safety culture and adverse events. J Patient Saf 2010 Dec;6(4):226-232. [doi: 10.1097/PTS.0b013e3181fd1a00] [Medline: 21099551]

14. Christian CK, Gustafson ML, Roth EM, Sheridan TB, Gandhi TK, Dwyer K, et al. A prospective study of patient safety in the operating room. Surgery 2006 Feb;139(2):159-173. [doi: 10.1016/j.surg.2005.07.037] [Medline: 16455323]

15. Stevens M, Schaik J. Implementing new technologies for complex care: The role of embeddedness factors in team learning. J Oper Manag 2019 Jun 10;66(1-2):112-134. [doi: 10.1002/joom.1034]

16. Russ AL, Fairbanks RJ, Karsh B, Militello LG, Saleem JJ, Wears RL. The science of human factors: Separating fact from fiction. BMJ Qual Saf 2013 Oct;22(10):802-808 [FREE Full text] [doi: 10.1136/bmjqs-2012-001450] [Medline: 23592760]

17. Sexton J, Thomas E, Helmreich R. Error, stress, and teamwork in medicine and aviation: Cross sectional surveys. BMJ 2000 Mar 18;320(7237):745-749 [FREE Full text] [doi: 10.1136/bmj.320.7237.745] [Medline: 10720356]

18. Sexton JB, Helmreich RL, Neilands TB, Rowan K, Vella K, Boyden J, et al. The Safety Attitudes Questionnaire: Psychometric properties, benchmarking data, and emerging research. BMC Health Serv Res 2006 Apr 03;6:44 [FREE Full text] [doi: 10.1186/1472-6963-6-44] [Medline: 16584553]

19. Nederhof AJ. Methods of coping with social desirability bias: A review. Eur J Soc Psychol 1985 Jul;15(3):263-280. [doi: 10.1002/ejsp.2420150303]

20. Savage C, Gaffney F, Hussain-Alkhateeb L, Olsson Ackheim P, Henricson G, Antoniadou I, et al. Safer paediatric surgical teams: A 5-year evaluation of Crew Resource Management implementation and outcomes. Int J Qual Health Care 2017 Oct 01;29(6):853-860. [doi: 10.1093/intqhc/mzx113] [Medline: 29024977]

21. Smiley K, Ofori L, Spangler C, Acquaah-Arhin R, Deh D, Enos J, et al. Safety culture and perioperative quality at the Volta River Authority Hospital in Akosombo, Ghana. World J Surg 2019 Jan;43(1):16-23. [doi: 10.1007/s00268-018-4763-y] [Medline: 30109388]

22. Makary MA, Sexton JB, Freischlag JA, Holzmueller CG, Millman EA, Rowen L, et al. Operating room teamwork among physicians and nurses: Teamwork in the eye of the beholder. J Am Coll Surg 2006 May;202(5):746-752. [doi: 10.1016/j.jamcollsurg.2006.01.017] [Medline: 16648014]

23. Moulton CE, Regehr G, Mylopoulos M, MacRae HM. Slowing down when you should: A new model of expert judgment. Acad Med 2007;82(Suppl):S109-S116. [doi: 10.1097/acm.0b013e3181405a76]

24. Schmutz JB, Meier LL, Manser T. How effective is teamwork really? The relationship between teamwork and performance in healthcare teams: A systematic review and meta-analysis. BMJ Open 2019 Sep 12;9(9):e028280 [FREE Full text] [doi: 10.1136/bmjopen-2018-028280] [Medline: 31515415]

25. Buljac-Samardzic M, Doekhie KD, van Wijngaarden JD. Interventions to improve team effectiveness within health care: A systematic review of the past decade. Hum Resour Health 2020 Jan 08;18(1):2 [FREE Full text] [doi: 10.1186/s12960-019-0411-3] [Medline: 31915007]

26. O'Connor P, Campbell J, Newon J, Melton J, Salas E, Wilson KA. Crew Resource Management training effectiveness: A meta-analysis and some critical needs. Int J Aviat Psychol 2008 Oct 16;18(4):353-368. [doi: 10.1080/10508410802347044]

27. Haerkens MH, Jenkins DH, van der Hoeven JG. Crew Resource Management in the ICU: The need for culture change. Ann Intensive Care 2012 Aug 22;2(1):39 [FREE Full text] [doi: 10.1186/2110-5820-2-39] [Medline: 22913855]

28. Sorra J, Khanna K, Dyer N, Mardon R, Famolaro T. Exploring relationships between patient safety culture and patients' assessments of hospital care. J Patient Saf 2012 Sep;8(3):131-139. [doi: 10.1097/PTS.0b013e318258ca46] [Medline: 22785344]

29. O'Dea A, O'Connor P, Keogh I. A meta-analysis of the effectiveness of Crew Resource Management training in acute care domains. Postgrad Med J 2014 Dec;90(1070):699-708. [doi: 10.1136/postgradmedj-2014-132800] [Medline: 25370080]

Abbreviations

ANOVA: analysis of variance
CMAQ: Cockpit Management Attitudes Questionnaire
CRM: Crew Resource Management
ETT: endovascular-treatment team
FMAQ: Flight Management Attitudes Questionnaire
HF: human factors
ICUMAQ: Intensive Care Unit Management Attitudes Questionnaire
JS: job satisfaction
OR: operating room
OTT: open-treatment team
PoM: perceptions of management
SAQ: Safety Attitudes Questionnaire
SAQ-NL: Safety Attitudes Questionnaire Dutch Edition
SC: safety climate
SR: stress recognition
TC: teamwork climate
TeamSTEPPS: Team Strategies and Tools to Enhance Performance and Patient Safety
WC: working conditions
WHO: World Health Organization

XSL•FO
RenderX

Original Paper

Value of Eye-Tracking Data for Classification of Information Processing–Intensive Handling Tasks: Quasi-Experimental Study on Cognition and User Interface Design

Stephan Wegner[1], MSc; Quentin Lohmeyer[1], Dr-Ing; Dimitri Wahlen[1], MSc; Sandra Neumann[2], PhD; Jean-Claude Groebli[2], PhD; Mirko Meboldt[1], Dr-Ing

[1]Product Development Group Zurich, Institute of Design, Materials and Fabrication, Department of Mechanical and Process Engineering, Swiss Federal Institute of Technology in Zurich, Zürich, Switzerland
[2]Peripal AG, Zürich, Switzerland

Corresponding Author:
Stephan Wegner, MSc
Product Development Group Zurich, Institute of Design, Materials and Fabrication
Department of Mechanical and Process Engineering
Swiss Federal Institute of Technology in Zurich
Leonhardstrasse 21
Zürich, 8092
Switzerland
Phone: 41 446324862
Email: stehess@ethz.ch

Abstract

Background: In order to give a wide range of people the opportunity to ensure and support home care, one approach is to develop medical devices that are as user-friendly as possible. This allows nonexperts to use medical devices that were originally too complicated to use. For a user-centric development of such medical devices, it is essential to understand which user interface design best supports patients, caregivers, and health care professionals.

Objective: Using the benefits of mobile eye tracking, this work aims to gain a deeper understanding of the challenges of user cognition. As a consequence, its goal is to identify the obstacles to the usability of the features of two different designs of a single medical device user interface. The medical device is a patient assistance device for home use in peritoneal dialysis therapy.

Methods: A total of 16 participants, with a subset of seniors (8/16, mean age 73.7 years) and young adults (8/16, mean age 25.0 years), were recruited and participated in this study. The handling cycle consisted of seven main tasks. Data analysis started with the analysis of task effectiveness for searching for error-related tasks. Subsequently, the in-depth gaze data analysis focused on these identified critical tasks. In order to understand the challenges of user cognition in critical tasks, gaze data were analyzed with respect to individual user interface features of the medical device system. Therefore, it focused on the two dimensions of dwell time and fixation duration of the gaze.

Results: In total, 97% of the handling steps for design 1 and 96% for design 2 were performed correctly, with the main challenges being task 1 insert, task 2 connect, and task 6 disconnect for both designs. In order to understand the two analyzed dimensions of the physiological measurements simultaneously, the authors propose a new graphical representation. It distinguishes four different patterns to compare the eye movements associated with the two designs. The patterns identified for the critical tasks are consistent with the results of the task performance.

Conclusions: This study showed that mobile eye tracking provides insights into information processing in intensive handling tasks related to individual user interface features. The evaluation of each feature of the user interface promises an optimal design by combining the best found features. In this way, manufacturers are able to develop products that can be used by untrained people without prior knowledge. This would allow home care to be provided not only by highly qualified nurses and caregivers, but also by patients themselves, partners, children, or neighbors.

XSL·FO
RenderX

KEYWORDS

human factors engineering; mobile eye tracking; benchmarking; home care; usability; self-management; quantitative research; quantitative evaluation

Introduction

Chronically ill patients cared for at home experience a higher health-related quality of life and a normalization of everyday life that is less dominated by the disease [1-3]. Therefore, 82% of end-stage renal disease patients and their families, if fully informed about their treatment options, would choose a home modality [4]. However, only 14% of dialysis patients in Europe are treated at home [5]. The main obstacle to home care is the availability of caregivers such as community nurses, neighbors, or relatives [6,7]. In order to allow a broad range of people the opportunity to ensure and support home care, one approach is to design medical devices with greater ease of use. This allows nonexperts to use medical devices that were originally too complicated to use. For user-centric development of such medical devices, it is essential to understand which user interface (UI) designs best support patients, caregivers, and health care professionals [8,9].

Human factors engineering drives user-oriented design and must test customized product UIs with intended users to determine the ideal level of mental workload. According to Kantowitz [10], mental workload is a subset of attention and the link between the demands of the environment and the capacity of the organism; it cannot be directly assessed. In a usability evaluation, the abstract term demand of the environment means fulfilling a task correctly. Consequently, when use errors occur, demand has not been met, and mental workload may have been too high or too low. This may be evaluated by analyzing the distribution and characteristic of attention in use error–related tasks.

Methods such as observations, questionnaires, and interviews are used to gain insight into the usability of an interface, but the focus is mainly on the graphical UI on a screen [9,11-15]. However, users gain most information through visual perception [16], and the short-term memory has only a limited capacity [17,18]. Therefore, it is difficult to understand the causes of use errors using traditional methods only.

Eye tracking provides a first-person perspective of the user and continuous localization of the gaze point. According to Hoang Duc et al [19], "tracking eye movements has the potential to provide a more direct measure of where attention is deployed since the direction of gaze is generally considered to be tightly coupled to the orienting of attention." Furthermore, "when people attend to a particular spatial location, there is greater neural processing in portions of the visual cortex corresponding to that location" [20]. Eye tracking thus allows objective feedback to find perception problems [21,22] and gain valuable insights into hotspots in attention distribution on the UI. This information can be used for both qualitative and quantitative evaluation of the usability of the UI. As a result, in recent years eye tracking has increasingly become a method for testing attention and improving or evaluating the features of UIs. Examples are web and print advertisements [23,24] and graphical representations like x-ray images of patients [25]. More complex subjects of the investigations include graphical UIs such as computer tomography interfaces [26] or spacecraft displays [27,28]. Further, there are single studies where eye tracking is used to evaluate highly interactive UIs of tangible products like smart TVs [29], smartwatches [12,21], or medical devices [30,31].

Most studies used a remote eye-tracking system where the stimulus is presented on a screen and participants are asked to sit still in front of a desk. Aside from this setup, mobile eye tracking with minimally invasive head-mounted systems provides a degree of freedom in movability. This promises natural user behavior in the testing of tangible medical devices [30].

In a first step of the eye-tracking data analysis, the raw gaze point data are classified into three events: fixation (nearly no eye movement), saccade (fast eye movement), and blink (closed eye). Since classified gaze data contain no semantic information on the looked-at objects or features, a second step of areas of interest mapping is needed. In this step, the single fixation events are manually assigned to the specific looked-at objects or UI features. As a result, data can subsequently be analyzed object-related in terms of durations of single fixations or cumulative dwell times (DTs) on an object or feature for a particular task. Fixation duration (FD), describing a property of visual attention per unit, is associated with the processing depth, which when increased leads to longer fixations [32-35], and with the rate of information extraction [23,35,36]. DT, describing the sum of visual attention related to specific objects or features, is associated with the length of the information extraction [28,37]. Thus, these measurements represent attention and, in the context of handling tasks, mental workload as a subset of attention in two dimensions.

Using the benefits of mobile eye tracking, we aimed to gain a deeper understanding of the challenges and differences in user cognition and thus identify obstacles to the user-friendliness of single UI features of a patient assistance device intended for home use in peritoneal dialysis (PD) therapy. This paper describes, to our knowledge, the first benchmark tests of two different UI designs based on physiological measurements using mobile eye tracking. The underlying research questions of this work are as follows:

- RQ1: Do slight differences in the design of the UI of a patient assistant device lead to differences in the effectiveness of use?
- RQ2: What are the differences in visual perception between two UI designs of a patient assistant device related to single task-relevant UI features in use error–related handling tasks?

Methods

Summary

The aim of this work was to gain a deeper understanding of the challenges in user cognition and thus the obstacles to user-friendliness of single UI features of a medical device. Therefore, a quasi-experimental study was conducted for data collection with the medical device with the intention of being as realistic as possible and representing the intended use. As a result, naïve representatives of the user group of patients (young adults and seniors) were recruited, and the study was conducted in the intended environment.

Stimuli

The stimuli of the study were prototypes with two different UI designs (D1 and D2, see Figure 1) of a medical device system.

The system consists of medical device, inlet for guiding and manipulating a bag system with dialysis fluid, and catheter, which is connected to the patient in the real therapy application. The most important interface features of the medical device are the buttons for manipulating the bag system and the lever for moving the inlet inside the device. The inlet has functions for fixing, clamping, and opening a predetermined breaking point feature inside the bag lines. The medical device system supports PD handling and is aimed at adults aged 18 and older. The stimuli provide acoustic (click sounds), haptic (positioning by stops), and visual (clear states and observation windows) feedback. Both prototypes support the same functionalities and require the same handling steps. At the top level, appearance of the UI designs was neutral in a monochrome design, as shown in Figure 1, to eliminate the effects of different coloring as an additional variable.

Figure 1. Illustration of user interface designs D1 (left) and D2 (right) including features lever and buttons 1-3. Additional parts for the therapy handling with the medical device are bag lines and catheter (standard parts used in therapy) and inlets D1 and D2, compatible with their respective designs.
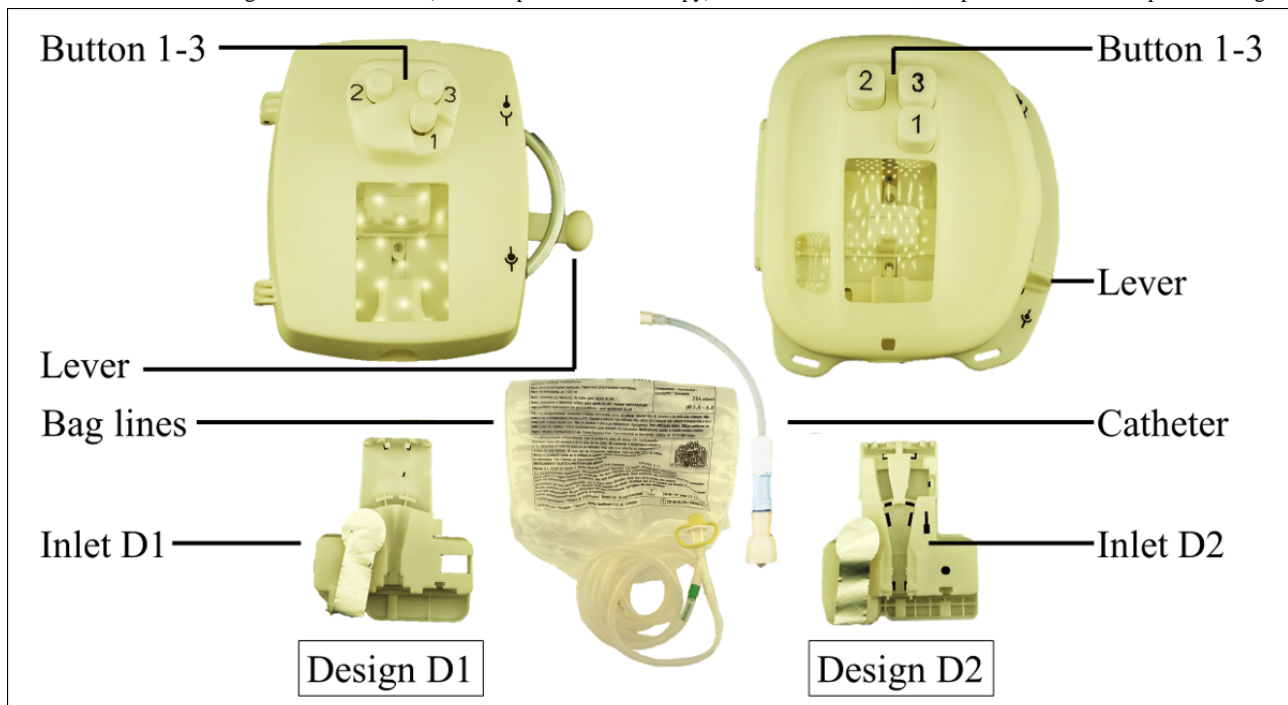
Recruitment and Data Exclusion

A total of 25 participants (18 men and 7 women, average 50.2 years, range 24 to 90 years) were recruited and participated in this study. The sample was recruited from a retirement home (10 men and 5 women, average 74.0 years, range 67 to 90 years) and from university (8 men and 2 women, average 25.1 years, range 24 to 26 years). In the PD patient population in Europe, 52% are younger than age 65 years [5]. Due to potential technical challenges with the eye-tracking technology related to the physiology of the eye area, which is especially relevant for seniors as reported by Bojko [38], more participants were invited than analyzed in the final analysis. All participants were in good physical and mental condition and assessed the suitability of study participation themselves. No participant was familiar with PD therapy or mobile eye tracking. All participants

had normal or corrected vision with contact or corrective lenses that could be connected to the mobile eye-tracking system.

One senior left the study prematurely after the first handling cycle and was therefore excluded from the analysis. For five seniors and one young adult, data quality was insufficient due to measure errors by the eye tracker resulting from drooping eyelids, watery eyes, or long eyelashes. In order to achieve a counterbalance in terms of the order of use of the two designs and represent the target population characteristic of PD patients in age, the data sets of a randomly selected senior and young adult were not included in the data analysis. Thus, a total of 16 data sets with 8 data sets from each group of young adults (25.0 years on average) and seniors (73.7 years on average) could be analyzed. Four participants in each group started with D1 and four with D2, achieving a complete counterbalance. As a result, this within-subject design mitigated the effect of individuality. Consequently, measures that naturally differ from participant

to participant, such as FDs, could be compared with this balanced design of the study.

Study Procedure

When participants arrived in the test environment, they were welcomed and thanked for their participation. Before the study began, participants were asked to read information on the goal of the study, data safety, and data management. If they agreed to participate in the study, they were asked to sign the consent form. Subsequently, participants put on the mobile eye-tracking system, and the moderator conducted a 3-point calibration. Since all participants were beginners in PD therapy and in the use of the device, the moderator briefly described the disease and associated PD therapy. Next, the moderator demonstrated the handling procedure with a low-level representation of the UI, designed and built for this purpose, and the devices. After the

introduction, participants performed the handling cycle of tasks 1 through 7 in a simulated PD therapy (see Figure 2), starting either with D1 or D2 and guided by written instructions. Each instruction was printed in a neutral design on an individual sheet to test the usability of the medical device and not the instruction. There was no time limit for the fulfillment of tasks, and the moderator assisted only in cases where the study would otherwise have had to be terminated due to the use error. Subsequent to the first completed handling cycle, participants were asked to give their feedback on usability in a semistructured interview with predefined high-level questions asking for general feedback on tasks related to use errors, guiding to the root causes of handling difficulties and use errors. Starting with the handling cycle, this process was repeated for the remaining prototype of the UI design.

Figure 2. Seven tasks in medical device handling cycle. User interacts manually with inlet, bag lines, catheter, and user interface features lever and buttons 1-3.
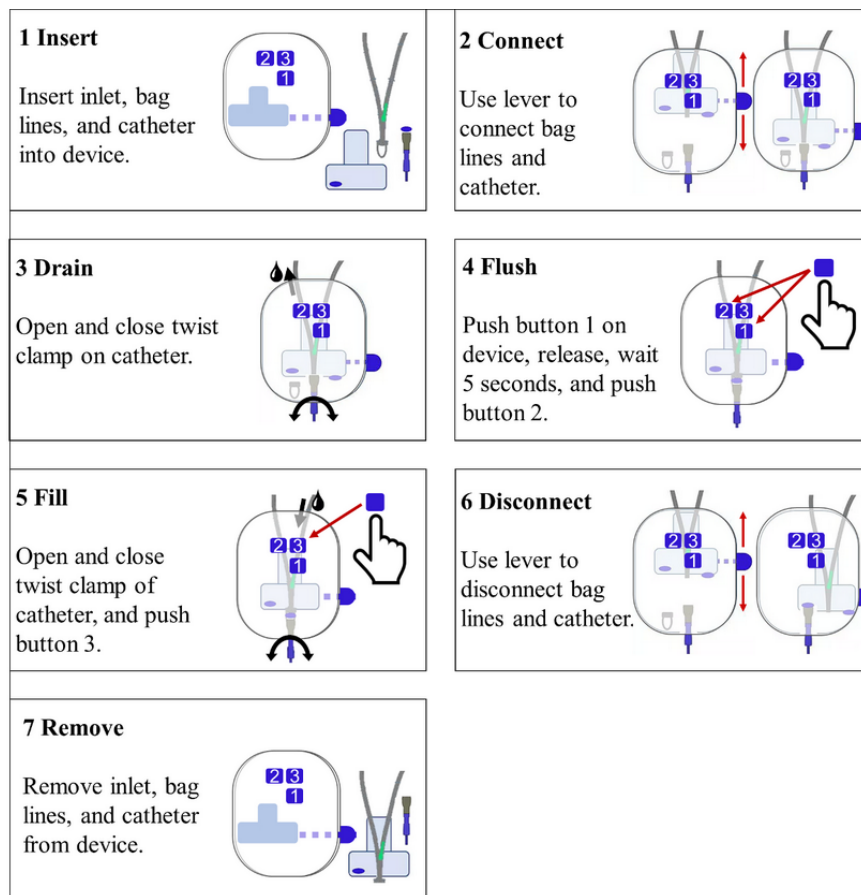
Data Analysis

In the data analysis, a 2-step approach was used. It started with the analysis of task effectivity searching for use error–related tasks. Subsequently, the in-depth gaze data analysis focused on these identified critical tasks.

For analysis of the task effectivity, the handling process of participants was observed via a live recording from the first person's perspective from the eye-tracking system. The performance in each task was evaluated by an observer. In the evaluation, two categories were distinguished according to the international standard IEC 62366-1 (2015). The first category,

safe use, is defined as "normal use without use error" [39]. The second category, use error, is defined as "user action or lack of user action while using the medical device that leads to a different result than that intended by the manufacturer or expected by the user" [39].

Gaze data were recorded with the mobile eye-tracking system, SMI Eye Tracking Glasses 2 (SensoMotoric Instruments GmbH), with a scene resolution of 1280×960 pixels (viewing angle: 60° horizontal, 46° vertical) of the front camera offering a sampling frequency of 24 Hz with the gaze point measurement having an accuracy of 0.5° over all distances. The raw gaze point data were classified into the events of fixations, saccades,

and blinks by SMI BeGaze version 3.7 (SensoMotoric Instruments GmbH). Subsequent to the areas of interest analysis, information on FD for depth and rate and DT for length of information extraction related to particular objects and features of the medical device system (Figure 1) were calculated. This information was used to understand the challenges of user cognition in use error–related tasks. Blinks were not considered in this work.

Combining the information on FD and DT, the data were analyzed with a multivariate analysis of variance (MANOVA) using SPSS Statistics 24 (IBM Corp). The MANOVA had one independent variable with two levels, D1 and D2 (see Figure 1), two dependent variables, FD and DT, both measured on a ordinal level and representing the rank of the mean

measurements for every participant for the UI features in error-related tasks.

For a better understanding of the two analyzed dimensions, length and depth of visual perception, Figure 3 combines information on the two measured parameters. Evaluating user perception of all participants as a whole, it shows the relationship between mean FD and mean DT for individual UI features of D1 compared with D2. Based on the two analyzed dimensions assigned to D1 in the coordinate origin, the mean FD and mean DT of D2 can be longer or shorter. Consequently, four different categories or patterns can be distinguished. A suggested interpretation of these patterns in terms of workload or gaze behavior is shown in Figure 3. Equations for calculating the values of the shift in both dimensions (ΔDT and ΔFD) from D1 to D2 for the diagram can be seen in Figure 4.

Figure 3. Visualization of shifts in two dimensions of the physiological gaze data measurements fixation duration and dwell time. The displayed shifts are from a Design D1 in the coordinate origin to a Design D2, presented in the middle column. In total, a distinction is made between the four categories. The right column explains the four patterns.

Category	Visual representation	Explanation
1	ΔFD [%], D2, ΔDT [%], D1	Combination of *longer dwell time* and *longer fixation duration*. This indicates *more scrutinizing* and thus a more focused and longer attention e.g. in handling tasks.
2	ΔFD [%], D2, ΔDT [%], D1	Combination of *shorter dwell time* and *longer fixation duration*. This indicates *less skimming* and thus a more focused and shorter attention e.g. while searching or checking.
3	ΔFD [%], D1, ΔDT [%], D2	Combination of *shorter dwell time* and *shorter fixation duration*. This indicates *less scrutinizing* and thus a more spread and shorter attention e.g. in handling tasks.
4	ΔFD [%], D1, ΔDT [%], D2	Combination of *longer dwell time* and *shorter fixation duration*. This indicates *more skimming* and thus a more spread and longer attention e.g. while searching or checking.

Figure 4. Equations for calculating the values of the shift in both dimensions of dwell time and fixation duration (ΔDT and ΔFD) from UI design D1 to D2.

$$\Delta \text{DT} = \frac{DT_{D2} - DT_{D1}}{DT_{D1}} \quad (1)$$

$$\Delta \text{FD} = \frac{FD_{D2} - FD_{D1}}{FD_{D1}} \quad (2)$$

Results

Each of the 16 participants performed 30 handling steps in the 7 tasks with both UI designs, resulting in 480 evaluated handling steps for each UI design. The results of the task performance are shown in Figure 5. Overall, 97% of the handling steps were performed correctly for D1 and 96% for D2.

According to the results, the main challenges were in task 1 (insert), task 2 (connect), and task 6 (disconnect) for both UI designs. The remaining four handling tasks were performed without errors, except for one missing catheter closure in task 5 (fill) with D2. Observed use errors in the first task were mainly

incorrectly inserted bag lines in the inlet. Further use errors were forgetting to attach the cap of the bag lines to a safety feature on the device and folding the protective film of the inlet outwards. All use errors were discovered and corrected by participants at a later stage of the handling cycle. In task 2 (connect), use errors occurred when the lever should have been used to connect bag lines and catheter. In task 6 (disconnect), some participants forgot to operate the lever for disconnecting catheter from bag lines and for placing a new cap onto catheter. In the semistructured interview, participants mentioned difficulties positioning bag lines and catheter in the inlet, oblivion of some details in the handling from the presentation, hesitation because of fear of breaking something, and misleading

XSL•FO

RenderX

wording in the instructions for tasks 2 (connect) and 6 (disconnect). In addition, participants gave positive and negative feedback on the overall impression and experience with the device.

Multimedia Appendix 1 focuses on handling tasks with observed use errors and shows the results of the data analysis of the physiological gaze data in both dimensions. The mean values for FD are given in milliseconds and for DT in seconds. The mean FD for single UI features was between 149 and 405 milliseconds. The mean DT for single UI features was between 0.3 and 28 seconds. At task 1 (insert), there were large shifts from the UI features bag lines, inlet, and catheter to the UI features levers and buttons. While the first group had average DTs between 7 and 28 seconds, the second group had average DTs between 0.3 and 3 seconds. At task 2 (connect) and task 6 (disconnect), the DT varied from less than 1 second for the buttons to 3 seconds for bag lines and catheter. For the mean FD, clustering was not found in any of the three tasks.

A MANOVA revealed statistically significant differences between D1 and D2 for catheter (Pillai trace=.216, $F_{2,29}$=3.985,

P=.03) and for lever (Pillai trace=.348, $F_{2,25}$=6.674, P=.005) in task 1 (insert) and for inlet (Pillai trace=.22, $F_{2,25}$=3.534, P=.045) in task 6 (disconnect). All other UI features showed no statistically significant differences in the three error-related tasks.

For better understanding, Figure 6 visualizes the data presented in Multimedia Appendix 1. As shown in Figure 3, this visualization combines FD and DT as two dimensions of the gaze data. In task 1 (insert; Figure 6A), the mean DT for all task-relevant UI features is longer for D2. The bag lines show a strong category 1 pattern, while the other two UI features show little to no shift for the mean FD. For task 2 (connect; Figure 6B), three UI features show a strong category 4 pattern, while the bag lines show mainly shorter mean DTs and only a slightly longer mean FD, thus showing a weak category 2 pattern. For task 6 (disconnect; Figure 6C), the UI elements located inside the device in this task show a strong category 2 pattern, while the lever on the outside of the device shows a strong category 1 pattern.

Figure 5. Comparison of task performance between D1 and D2 for all seven tasks. Evaluation in two categories, safe use and use error, according to International Electrotechnical Commission 62366-1 (2015).
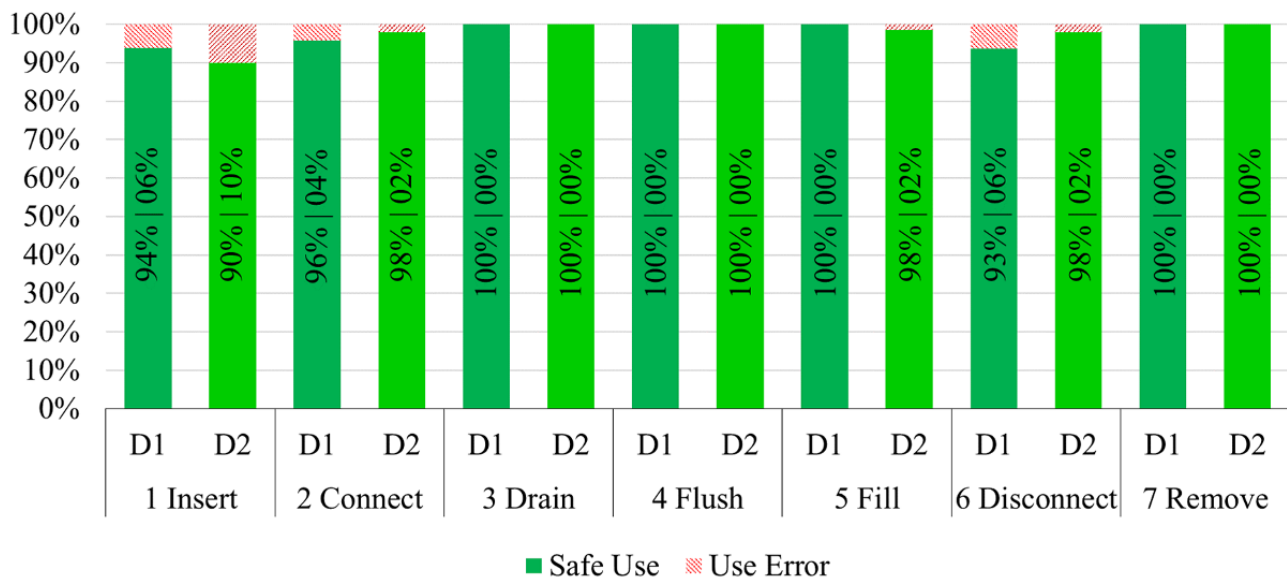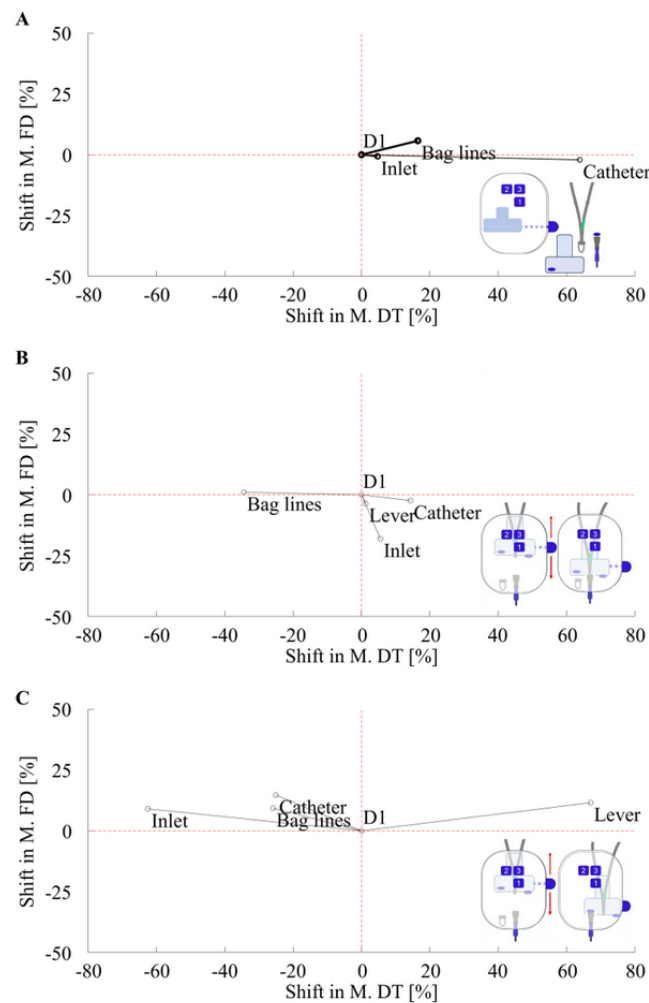
Figure 6. Shifts from D1 (in the coordinate origin) to D2 in terms of mean fixation duration (ordinate) and mean dwell time (abscissa) for task 1 insert (A), task 2 connect (B), and task 6 disconnect (C). The relevant user interface features in these three tasks are bag lines, inlet, catheter, and lever.

Discussion

Principal Findings

Task performance analysis generally showed little or no use errors in the various handling tasks for both UI designs (see Figure 5). The tasks with observed use errors were the insertion of material and connection and disconnection of bag lines and catheter. In line with the observations, participants described in semistructured interviews difficulties in the execution and in remembering of the correct handling step details in the observed use error–related tasks. Further, they reported misleading wording in the instructions as the explanation for their use errors in task 2 (connect) and task 6 (disconnect), thus providing additional information for the development of the supplementary material. In the first task, most use errors occurred when inserting the bag lines into the inlet. For this task, Figure 6A shows a category 1 pattern with longer mean DT and longer mean FD for the bag lines. Therefore, the results of gaze data analysis are consistent with the results of task performance. Gaze data shows more scrutinizing for D2 compared with D1 in order to insert bag lines and catheter into inlet and device. The longer and higher depth in visual perception indicates a higher mental workload for this task using D2.

When connecting and disconnecting the catheter, some participants missed pulling down the lever to connect bag lines and catheter again and putting a new cap on the catheter. For connecting and disconnecting bag lines and catheter, the most important interface features show category 4 patterns (task 2, Figure 6B) and category 2 patterns (task 6, Figure 6C). This visual pattern indicates more skimming behavior for task 2 and less skimming behavior for task 6. This in turn indicates more visual controls when connecting bag lines and catheter in task 2 for D2. Compared with task performance, this seems to result in slightly fewer use errors for D2 (2% vs 4%). For task 6, results indicate less visual searching associated with the relevant features inlet and catheter for D2 when a new cap is placed on the catheter. In a comparison of the two UI designs, the main difference between D1 and D2 is the position of the top window. With D2, the user can better see the inlet. This may help finding the important features while a new cap is placed on the catheter. Furthermore, the lever in task 6 shows a category 1 pattern associated with a longer and higher depth in visual perception for D2. Although the results show fewer use errors, handling the lever with D2 appears to be mentally more difficult than with D1.

When evaluating the total mental workload of the medical device system, the analyzed UI features of the medical device showed

shifts in both the mean FD and mean DT. The mean FD varied from 149 to 405 milliseconds in the critical tasks across all features (Multimedia Appendix 1). In order to be able to interpret these values, the results of three different task examples as described in the literature are compared. In a case study of a driving situation described by Velichkovsky et al [34], the values for the mean FD were between 499 and 543 milliseconds. Bojko et al [33] reported in an evaluation of drug label designs that the FD varied between 260 and 392 milliseconds. Just and Carpenter [35] observed a mean FD of 477 milliseconds observing the task of reading a scientific text. Compared with these studies, the mean FD of the handling cycle is in the same range as reading a drug label. The mean DT in the critical tasks varied in a range from 0.3 to 28.3 seconds (Multimedia Appendix 1). Especially in the first task, the insertion of the material in both UI designs required longer DT for bag lines and catheter compared with other tasks. This shows that this task requires special attention from the user. This is supported by significant differences in a MANOVA for the catheter in the considered task. The statistical analysis showed only in two other cases significant differences in the gaze data. The reason for merely three significant differences is probably because of the low level of variation in the design.

Based on the results of this study, benchmarking D1 and D2 showed the following. Inserting the material seemed to be challenging for both UI designs in general. Therefore, the guiding material (manual and quick starting guide) and training should focus on this task. The lever of D1 seemed to result in lower mental workload. It has a more dominant appearance compared with D2, where the lever is integrated into the housing for protection in case of a fall. The UI design D2 of the inlet seems to be easier to perceive visually. The higher position of the top window in D2 shows a positive impact on the task connecting and disconnecting bag lines and catheter.

Analysis of two dimensions of visual perception using eye tracking provided a detailed picture of the length and depth of the visual perception and therefore the challenges in user cognition and ease of use. Results highlighted the differences in information extraction for different UI features in single tasks. This information helped human factors engineering to focus the development on the critical UI features. Following this work, a summative study evaluated the final UI of the device. This final design and the instructions incorporated the results of this study, such as the detailed description of the insertion of the material and the coloring of the main UI features to guide the user's gaze. The summative study included patients, relatives, nurses, and physicians. They represented the later user population in the characteristic in age, preknowledge, and comorbidities. Patients had two types of comorbidities, such as arthritis and Reynaud syndrome, in addition to the renal disease with its own accompanying symptoms. The summative study confirmed the safety, efficiency, and effectiveness of use [40].

Limitations

Due to the novelty of the medical device presented in this study, there are several limitations regarding the results. First, participants were not patients in the real therapy. They were beginners who had no experience in this specific therapy or associated tasks. Furthermore, the device was not used in the real therapy application but in a simulation. These factors provide information on how forgetfulness or even dementia would influence use of the medical device in the later use by patients. Second, when the final product is used, individual training of the user is mandatory and labeling material supports the user. This support was not provided in this study. Instead, a presentation with an additional low-level representation of the UI and a neutral text of the seven tasks guided participants through the handling cycle. Consequently, the focus was on intuitive task performance and perception of information depending on the different UI designs. Third, the design of the two different top-level designs was similar due to a unicolored representation. This is not a strong contrast between the main UI functions and the rest of the medical device. As stated in Methods, this was chosen to eliminate influences of different coloring as an additional influencing factor. At the level of gaze data analysis, representation of the combination of mean FD and mean DT is the first published. Further research is needed to assess whether identified patterns apply to different usability studies with different tasks and stimuli.

Conclusion

The prototypes of the medical device system as stimuli of the study had only little differences in the single UI features. Consequently, results in the effectiveness of use revealed only marginal differences, with a maximum of 6% versus 10% use errors in task 1 (insert). Based on the two dimensions of the physiological gaze data measurements FD and DT, four distinct patterns could be distinguished between the two UI designs. A MANOVA revealed statistically significant differences in these patterns for three UI features.

Studying the impact on the usability of alternatives of different UI designs is crucial to understand which best supports the user. Traditional methods such as observation, interviews, or questionnaires tend to give feedback only at the level of the UI as a whole. Furthermore, when it comes to reporting usability issues or first impressions of the medical device during interviews or questionnaires, several challenges arise. Test participants may forget to report their impressions or adapt their answers to social expectations [30,41]. This makes it difficult to identify the root causes of usability problems and thus the necessary changes in UI design. In alignment with Lohmeyer et al [31] and Koester et al [30], this study showed that mobile eye tracking provides objective quantitative results based on physiological measurements related to individual UI features. These results can be used to evaluate usability in much more detail compared with traditional methods.

This information is crucial to be able to adapt the design of a product to the needs of the users. Therefore, results of usability testing must be more detailed than just a yes-or-no result of use errors. On the contrary, evaluation of each feature of the UI promises to achieve the best possible UI design by combining the best features found. This combined solution would therefore offer the highest level of usability. In this way, manufacturers can develop products that can be used even by untrained people without prior knowledge. This would allow home care to be provided not only by highly qualified nurses and caregivers,

but also by patients themselves, partners, children, or neighbors. This would contribute to removing barriers to home care and thus to a higher quality of life and normalization of everyday life, which is less dominated by illness for patients.

Multimedia Appendix 1
Analysis of eye tracking metrics for user interface features bag lines, inlet, catheter, lever, and buttons. Mean fixation duration (FD) in milliseconds and mean dwell time (DT) of the gaze in seconds for user interface designs D1 and D2. Evaluated tasks are task 1 (insert), task 2 (connect), and task 6 (disconnect). Multivariate analysis of variance analyzed the combination of FD and DT for significant differences according to the Pillai trace (p) between D1 and D2.
[DOCX File , 14 KB - humanfactors_v7i2e15581_app1.docx]

References

1. Rygh E, Arild E, Johnsen E, Rumpsfeld M. Choosing to live with home dialysis-patients' experiences and potential for telemedicine support: a qualitative study. BMC Nephrol 2012 Mar 19;13:13 [FREE Full text] [doi: 10.1186/1471-2369-13-13] [Medline: 22429705]
2. Brown EA, Johansson L, Farrington K, Gallagher H, Sensky T, Gordon F, et al. Broadening Options for Long-term Dialysis in the Elderly (BOLDE): differences in quality of life on peritoneal dialysis compared to haemodialysis for older patients. Nephrol Dial Transplant 2010 Nov;25(11):3755-3763 [FREE Full text] [doi: 10.1093/ndt/gfq212] [Medline: 20400451]
3. Alloatti S, Manes M, Paternoster G, Gaiter AM, Molino A, Rosati C. Peritoneal dialysis compared with hemodialysis in the treatment of end-stage renal disease. J Nephrol 2000;13(5):331-342. [Medline: 11063136]
4. Devoe DJ, Wong B, James MT, Ravani P, Oliver MJ, Barnieh L, et al. Patient education and peritoneal dialysis modality selection: a systematic review and meta-analysis. Am J Kidney Dis 2016 Sep;68(3):422-433. [doi: 10.1053/j.ajkd.2016.02.053] [Medline: 27125246]
5. ERA-EDTA Registry Annual Report 2016. Amsterdam: European Renal Association–European Dialysis and Transplant Association URL: https://www.era-edta-reg.org/files/annualreports/pdf/AnnRep2016.pdf [accessed 2020-05-05]
6. Wilson SE, Ouédraogo CT, Prince L, Ouédraogo A, Hess SY, Rouamba N, et al. Caregiver recognition of childhood diarrhea, care seeking behaviors and home treatment practices in rural Burkina Faso: a cross-sectional survey. PLoS One 2012;7(3):e33273 [FREE Full text] [doi: 10.1371/journal.pone.0033273] [Medline: 22428006]
7. Jones R, Jordan S. The implementation of crisis resolution home treatment teams in wales: results of the national survey 2007-2008. Open Nurs J 2010 Feb 18;4:9-19 [FREE Full text] [doi: 10.2174/1874434601004010009] [Medline: 20502646]
8. Dijkstra NE, Sino CGM, Heerdink ER, Schuurmans MJ. Development of eHOME, a mobile instrument for reporting, monitoring, and consulting drug-related problems in home care: human-centered design study. JMIR Hum Factors 2018 Mar 07;5(1):e10 [FREE Full text] [doi: 10.2196/humanfactors.8319] [Medline: 29514771]
9. Alnosayan N, Chatterjee S, Alluhaidan A, Lee E, Houston FL. Design and usability of a heart failure mhealth system: a pilot study. JMIR Hum Factors 2017 Mar 24;4(1):e9 [FREE Full text] [doi: 10.2196/humanfactors.6481] [Medline: 28341615]
10. Kantowitz BH. Attention and mental workload. 2000 Presented at: Proceedings of the 15th Triennial Congress IEA 2000/Human Factors and Ergonomics Society Annual Meeting; 2000; San Diego. [doi: 10.1177/154193120004402121]
11. Beatty AL, Magnusson SL, Fortney JC, Sayre GG, Whooley MA. VA FitHeart, a mobile app for cardiac rehabilitation: usability study. JMIR Hum Factors 2018 Jan 15;5(1):e3 [FREE Full text] [doi: 10.2196/humanfactors.8017] [Medline: 29335235]
12. Liang J, Xian D, Liu X, Fu J, Zhang X, Tang B, et al. Usability study of mainstream wearable fitness devices: feature analysis and system usability scale evaluation. JMIR Mhealth Uhealth 2018 Nov 08;6(11):e11066 [FREE Full text] [doi: 10.2196/11066] [Medline: 30409767]
13. George D, Hassali MA, Hss A. Usability testing of a mobile app to report medication errors anonymously: mixed-methods approach. JMIR Hum Factors 2018 Dec 21;5(4):e12232 [FREE Full text] [doi: 10.2196/12232] [Medline: 30578216]
14. Rotondi AJ, Spring MR, Hanusa BH, Eack SM, Haas GL. Designing eHealth applications to reduce cognitive effort for persons with severe mental illness: page complexity, navigation simplicity, and comprehensibility. JMIR Hum Factors 2017 Jan 05;4(1):e1 [FREE Full text] [doi: 10.2196/humanfactors.6221] [Medline: 28057610]

15.   Medicines and Healthcare Products Regulatory Agency. 2017. Human factors and usability engineering—guidance for medical devices including drug-device combination products URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/645862/HumanFactors_Medical-Devices_v1.0.pdf [accessed 2020-05-04]

16.   Rosenblum LD. See What I'm Saying: The Extraordinary Powers of our Five Senses. New York: WW Norton; 2010.

17.   Miller GA. The magical number seven plus or minus two: some limits on our capacity for processing information. Psychol Rev 1956 Mar;63(2):81-97. [Medline: 13310704]

18.   Peterson LR, Peterson MJ. Short-term retention of individual verbal items. J Exp Psychol 1959 Sep;58:193-198. [doi: 10.1037/h0049234] [Medline: 14432252]

19.   Duc AH, Bays P, Husain M. Eye movements as a probe of attention. Prog Brain Res 2008;171:403-411. [doi: 10.1016/S0079-6123(08)00659-6] [Medline: 18718333]

20.   Anderson J. Cognitive Psychology and its Implications. 6th Edition. New York: Worth; 2004.

21.   Wu Y, Cheng J, Kang X. Study of smart watch interface usability evaluation based on eye-tracking. In: Marcus A, editor. Design, User Experience, and Usability: Technological Contexts, Vol. 9748. Springer International Publishing: Cham; 2016:98-109.

22.   Groen M, Noyes J. Using eye tracking to evaluate usability of user interfaces: is it warranted? IFAC Proceedings Volumes 2010;43(13):489-493. [doi: 10.3182/20100831-4-FR-2021.00086]

23.   Wedel M, Pieters R. A review of eye-tracking research in marketing. In: Review of Marketing Research, Vol. 4. Bingley: Emerald Group Publishing Limited; 2008:123-147.

24.   Chynal P, Sobecki J, Szymanski J. Remote usability evaluation using eye tracking enhanced with intelligent data analysis. In: Hutchison D, Kanade T, Kittler J, Kleinberg JM, Mattern F, editors. Design, User Experience, and Usability. Design Philosophy, Methods, and Tools. Vol. 8012. Springer: Berlin; 2013:212-221.

25.   Harezlak K, Kasprowski P. Application of eye tracking in medicine: a survey, research issues and challenges. Comput Med Imaging Graph 2018 Apr;65:176-190. [doi: 10.1016/j.compmedimag.2017.04.006] [Medline: 28606763]

26.   Erol Barkana D, Açık A, Duru DG, Duru AD. Improvement of design of a surgical interface using an eye tracking device. Theor Biol Med Model 2014 May 7;11 Suppl 1:S4 [FREE Full text] [doi: 10.1186/1742-4682-11-S1-S4] [Medline: 25080176]

27.   Guo Q, Xue C, Lin Y, Niu Y, Chen M. A study for human-machine interface design of spacecraft display & control device based on eye-tracking experiments. In: Harris D, editor. Engineering Psychology and Cognitive Ergonomics: Cognition and Design. Vol. 10276. Springer International Publishing: Cham; 2017:211-221.

28.   Dill E, Young S. Analysis of eye-tracking data with regards to the complexity of flight deck information automation and management—inattentional blindness, system state awareness, and EFB usage. 2015 Presented at: 15th AIAA Aviation Technology, Integration, and Operations Conference; 2015; Dallas. [doi: 10.2514/6.2015-2901]

29.   Obrist M, Bernhaupt R, Beck E, Tscheligi M. Focusing on elderly: an iTV usability evaluation study with eye-tracking. In: Cesar P, Chorianopoulos K, Jensen JF, editors. Interactive TV: A Shared Experience. Vol. 4471. Springer: Berlin; 2007:66-75.

30.   Koester T, Brøsted JE, Jakobsen JJ, Malmros HP, Andreasen NK. The use of eye-tracking in usability testing of medical devices. Proc Int Symp Human Factors Ergonom Health Care 2017 May 15;6(1):192-199. [doi: 10.1177/2327857917061042]

31.   Lohmeyer Q, Schneider A, Jordi C, Lange J, Meboldt M. Toward a new age of patient centricity? The application of eye-tracking to the development of connected self-injection systems. Expert Opin Drug Deliv 2019 Feb;16(2):163-175. [doi: 10.1080/17425247.2019.1563070] [Medline: 30577710]

32.   Godfroid A, Spino LA. Reconceptualizing reactivity of think-alouds and eye tracking: absence of evidence is not evidence of absence. Language Learning 2015 Oct 08;65(4):896-928. [doi: 10.1111/lang.12136]

33.   Bojko A, Gaddy C, Lew G, Quinn A, Israelski E. Evaluation of drug label designs using eye tracking. Proc Human Factors Ergonom Soc 2016 Nov 05;49(11):1033-1037. [doi: 10.1177/154193120504901107]

34.   Velichkovsky B, Dornhoefer S, Pannasch S, Unema P. Visual fixations and level of attentional processing. 2000 Presented at: Proceedings of the symposium on eye tracking research & applications - ETRA '00; 2000; Palm Beach Gardens p. 79-85. [doi: 10.1145/355017.355029]

35.   Just MA, Carpenter PA. A theory of reading: from eye fixations to comprehension. Psychol Rev 1980 Jul;87(4):329-354. [Medline: 7413885]

36.   Kim JH, Zhao X, Du W. Assessing the performance of visual identification tasks using time window-based eye inter-fixation duration. Int J Industr Ergonom 2018 Mar;64:15-22. [doi: 10.1016/j.ergon.2017.09.002]

37.   Sarter NB, Mumaw RJ, Wickens CD. Pilots' monitoring strategies and performance on automated flight decks: an empirical study combining behavioral and eye-tracking data. Hum Factors 2007 Jun;49(3):347-357. [doi: 10.1518/001872007X196685] [Medline: 17552302]

38.   Bojko A. Eye Tracking the User Experience: A Practical Guide to Research. New York: Rosenfeld Media; 2013.

39.   International Electrotechnical Commission. IEC 62366-1:2015. Medical devices—Part 1: application of usability engineering to medical devices. 2015. URL: https://www.iso.org/standard/63179.html [accessed 2020-05-04]

40.   Hess S, Dubach M, Meboldt M, Foggensteiner L. Evaluating patient safety and ease of use of a novel connection-assist device for peritoneal dialysis. Patient Prefer Adherence 2019;13:1785-1790 [FREE Full text] [doi: 10.2147/PPA.S218663] [Medline: 31754299]

41.   Wiklund M, Kendler J, Strochlic A. Usability Testing of Medical Devices. Boca Raton: CRC Press; 2016.

Abbreviations

DT: dwell time
FD: fixation duration
MANOVA: multivariate analysis of variance
PD: peritoneal dialysis
UI: user interface

XSL•FO
RenderX

<u>Original Paper</u>

Usability of a Mobile App for Improving Literacy in Children With Hearing Impairment: Focus Group Study

Shelly DeForte[1*], PhD; Emre Sezgin[1*], PhD; Janelle Huefner[2], MA; Shana Lucius[2], MA; John Luna[1], MA; Anand A Satyapriya[3], MD; Prashant Malhotra[4], MD

[1]Research Information Solutions and Innovation, The Abigail Wexner Research Institute, Nationwide Children's Hospital, Columbus, OH, United States

[2]Clinical Therapies Department, Nationwide Children's Hospital, Columbus, OH, United States

[3]OhioHealth Riverside Methodist Hospital, Columbus, OH, United States

[4]The Hearing Program in the Pediatric Otolaryngology Department, Nationwide Children's Hospital, Columbus, OH, United States

[*]these authors contributed equally

Corresponding Author:
Emre Sezgin, PhD
Research Information Solutions and Innovation
The Abigail Wexner Research Institute
Nationwide Children's Hospital
700 Children's Drive
Columbus, OH, 43205
United States
Phone: 1 6143556814
Email: emre.sezgin@nationwidechildrens.org

Abstract

Background: Children with hearing loss, even those identified early and who use hearing aids or cochlear implants, may face challenges in developing spoken language and literacy. This can lead to academic, behavioral, and social difficulties. There are apps for healthy children to improve their spoken language and literacy and apps that focus on sign language proficiency for children with hearing loss, but these apps are limited for children with hearing loss. Therefore, we have developed an app called Hear Me Read, which uses enhanced digital stories as therapy tools for speech, language, and literacy for children with hearing loss. The platform has therapist and parent/child modes that allow (1) the selection of high-quality, illustrated digital stories by a speech-language pathologist, parent, or child; (2) the modification of digital stories for a multitude of speech and language targets; and (3) the assignment of stories by a therapist to facilitate individualized speech and language goals. In addition, Hear Me Read makes the caregiver a core partner in engagement through functionality, whereby the caregiver can record video and audio of themselves to be played back by the child.

Objective: This study aimed to evaluate the user experience of the Hear Me Read app through a focus group study with caregivers and their children.

Methods: We recruited 16 participants (8 children with and without hearing loss and 8 caregivers) to participate in 1-hour focus groups. Caregivers and children interacted with the app and discussed their experiences through a semistructured group interview. We employed thematic analysis methods and analyzed the data. We used feedback from the focus group to improve the elements of the app for a larger clinical trial assessing the impact of the app on outcomes.

Results: We identified three themes: default needs, specific needs, and family needs. Participants found the app to be esthetically pleasing and easy to use. The findings of this study helped us to identify usability attributes and to amend app functionalities to best fit user needs. Caregivers and children appreciated the enhancements, such as highlighting of parts of speech and caregiver reading of video playback, which were made possible by the digital format. Participants expressed that the app could be used to enhance family reading sessions and family interaction.

Conclusions: The findings from this focus group study are promising for the use of educational apps designed specifically for those with hearing loss who are pursuing listening and spoken language as a communication outcome. Further investigation is needed with larger sample sizes to understand the clinical impact on relevant language and literacy outcomes in this population.

XSL•FO
RenderX

Introduction

Background

Nearly 2 to 3 per 1000 newborns are born with hearing loss, making hearing loss one of the most common birth anomalies [1]. Approximately 1 to 3 million children in the United States, and 34 million children worldwide suffer from disabling hearing loss [2]. Children with hearing loss face challenges in developing spoken language and literacy. Historically, the reading skills of deaf children have been poor, with graduating teenager reading scores comparable with first- to fourth-grade reading level [3-5]. Infants with hearing loss can now be identified at birth and fit with hearing aids and cochlear implants early in life, allowing them to have greater access to sound and improved oral language abilities [6-9]. However, even with modern hearing technology, many children who are deaf or have hearing loss may continue to read at significantly lower levels than typically hearing peers [10-12]. Young children who do not attain early literacy skills are at a higher risk for academic and social problems [13-16].

To address these challenges, we have developed a mobile app called the *Hear Me Read*. The National Association of Education of Young Children and other educational institutions emphasize the importance of reading storybooks to young children to enhance literacy [17-20]. In line with this recommendation, our intention with Hear Me Read is to use digital stories as therapy tools for speech, language, and literacy and to develop a platform for delivery that enhances family engagement for children with hearing loss. The Hear Me Read app is not meant to replace traditional storybook reading, but rather to enhance the user experience by providing the additional content individuals with hearing loss are expected to benefit from (eg, lip reading and audio-visual combination).

Current Practice and Technology Use in the Education of Children With Hearing Impairment

Mobile phone and tablet technologies with digital electronic books (eBooks) and storybook reading are now commonplace in modern homes and schools. The impact of this technology on emerging literacy can be positive. However, digital eBooks can also be distracting when compared with traditional print storybook reading [21]. The impact of digital eBooks on the shared-book reading experience could be harmful if it increases distractions and reduces face-to-face interaction [22]. Thus, the design of *educational* apps should be done thoughtfully [23], especially in vulnerable populations such as children with hearing loss, with a focus on minimizing distractions and design flaws to promote engagement.

Although many apps deal specifically with certain parts of speech such as articulation, phonics, vocabulary, grammar, and comprehension, only a few are designed for children with hearing loss (Multimedia Appendix 1 shows a list of these apps). Furthermore, although sign language apps exist for children who are deaf or have hearing loss, there are no apps that are designed to develop spoken language and literacy in children with hearing loss pursuing a listening and spoken language outcome. The current *gold standard* for children with hearing loss to develop spoken language and literacy is through in-person one-on-one therapy sessions with a pediatric hearing loss expert (speech and language therapist or auditory-verbal therapist) [24]. Hear Me Read is developed to work specifically in the direction of a speech-language therapist, using digital stories as therapy extenders. Our app fits the principle of auditory-verbal intervention, where the caregivers are coached to be the *primary language facilitators* of their child's language and literacy skills. When coaching is carried over into the home setting, we observe the most significant progress. Our app is an extension of this philosophy by enabling caregivers and their children to complete therapy activities outside of therapy sessions.

Children with hearing loss vary widely in their auditory access and how they acquire language or literacy skills [25]. They may benefit from multiple presentation modalities, including auditory, visual, or a combination of these approaches [26]. However, data regarding the efficacy of digital storybook interventions targeting children with hearing loss or how these children use existing digital reading technology are scarce. In this study, our objective was to understand user needs and expectations and the usability of Hear Me Read. The findings from this study were used to inform the app's design process and improve the interface and functionality. Our next step will be to test Hear Me Read in a prospective study of children with hearing loss by measuring the impact of app usage on language and literacy outcomes.

Methods

Hear Me Read App

Hear Me Read is an interactive mobile app for improving language and literacy, which is targeted for auditory-based learning for children with hearing loss. Hear Me Read is composed of general features for improving literacy, including interactive storybook reading and syntax highlighting, audio-visual features including video recording and playback functionality, and a therapist mode, for the creation of individual therapeutic language and literacy goals. Some of the features, such as the ability to highlight and interact with vocabulary words that show a related image in Hear Me Read, are broadly applicable to children with and without hearing loss, whereas others, such as the playback of caregivers' recorded video narration or the ability to highlight particular auditory training words, are specific to children with hearing loss.

Hear Me Read provides an interactive digital environment for caregivers and children, using high-quality children's stories (provided by Highlights for Children, Inc) in multiple formats and modalities (Figure 1). Hear Me Read allows the same digital story to be read in multiple ways: (1) text alone, (2) with

illustrations, (3) with highlighted text targets, and (4) with audio and/or video recording of the parent reading (customized to display the text for the narrator to read and place it in a position relative to the camera that would produce a video where the narrator seems to be looking directly at the user).

Furthermore, Hear Me Read can help caregivers and therapists track reading progress and prescribe new reading assignments. In-app metrics can also measure the time spent in the book, the number of times read, and progress within the book. With these fundamental features, Hear Me Read is a one-of-a-kind app for children with hearing loss that leverages child-caregiver engagement. The development of the app used an iterative and user-centered approach. The layout of the user interface went through a few iterations during development and internal testing, as we narrowed down the scope of the project. Through observation and anecdotal feedback, we positioned and reshaped buttons to match the natural hand positions and interaction instincts of the users. A video introduction for the app is available in Multimedia Appendix 2.

Figure 1. Screenshots from the Me Read app. The app provides multiple user menu options based on the therapeutic objectives (upper left). The app allows caregivers to record video and audio of themselves reading sentences and words (upper right), and caregivers and children can play back this video and view other associated media (lower right) during reading sessions. The app gives users the option to highlight parts of speech for targeted learning (lower middle) and a view for therapists to select custom learning objectives (lower right).
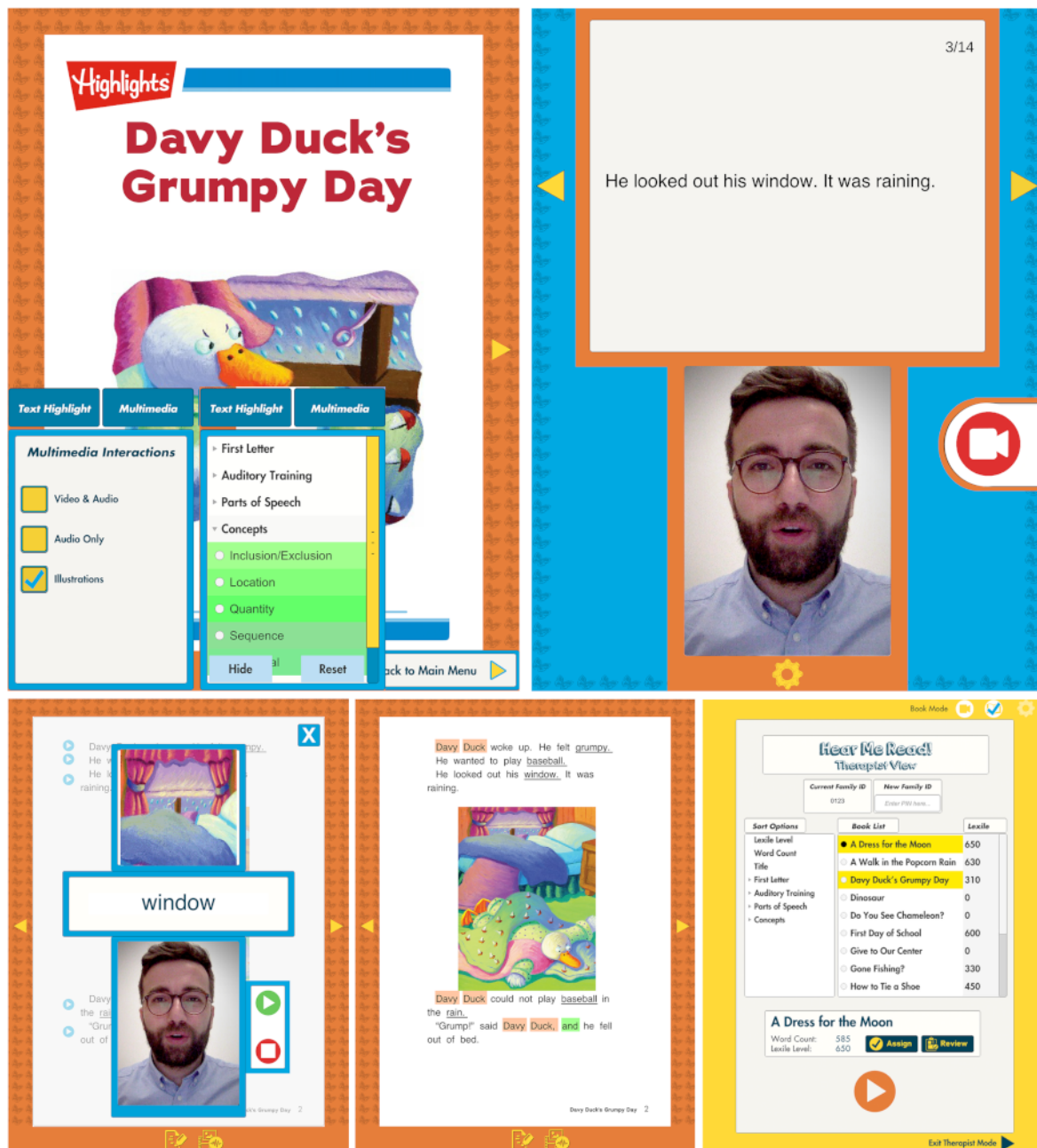
The app was developed for the iPhone operating system using Unity. The app is intended to be used with the built-in speakers of the mobile device it is played on, and no additional calibration is provided for use with headphones. The digital content for the children's book uses stories created by and retrieved with permission from Highlights Inc. We made a parser to translate the HTML exports of the PDFs of the Highlights books to a format that we could push into Unity. All content interactions are logged in the background and can be exported in comma-separated values (.csv) document format. The log files include user interaction event type, time of occurrence, and relative event state (eg, true/false, number, string). The event types include book open/close, book completion, narration start/stop, word click, app launch/close, video enable/disable, images enable/disable, audio enable/disable, highlight enable/disable, book assigned/unassigned, and Ling 6 sound [27]. Hear Me Read is currently not available for consumer use; however, this app will undergo further study in children with hearing loss to assess how it impacts language and literacy outcomes.

Focus Group Study Design

After approval of the Nationwide Children's Hospital's (NCH) institutional review board, we recruited children with and without hearing loss and their caregivers. We included children without hearing loss because it allows us to explore usability across all literacy and hearing levels. Hearing loss can go from near-normal to completely deaf, and literacy capabilities can go from far below or far above peers. In clinical assessment, we would have to separate groups, but we wanted to make sure we had the entire spectrum of abilities and ages for the usability part. Following a demonstration and a short trial period of the Hear Me Read app, caregiver-child groups participated in a focus group discussion. Semistructured interviews were held among caregiver-child groups. This method is suitable for collecting rich information to understand the needs and expectations of families toward a technology-based solution to improve language and literacy skills for children with hearing loss. Observing caregiver-child relationships and engagement with the technology is also helpful. The session objectives were to identify obstacles that may interfere with the regular use of the app; to verify that participants can interface with and find value in the reading, recording, and language tasks; and to understand how the app may influence users and fit within the daily life of caregiver-child groups. The coauthors, SD and ES, designed and moderated the interviews. PM and JH helped recruit participants.

Participants

Participants were recruited from a tertiary children's hospital system. They were drawn from the Hearing and Cochlear Implant program or speech-language therapy department or were healthy volunteers. Participants—children and their caregivers—were invited via phone calls or during clinic visits. Participants were informed about the study aims, voluntary participation, and financial compensation. Gift cards were provided to families who participated in the study.

Data Collection

A total of 2 focus group sessions were held in a pediatric therapy room for approximately 1 hour each. Before the focus group sessions, JH provided a written description of the study and collected consent forms from caregivers. Caregivers were then asked to complete a questionnaire (Multimedia Appendix 3) that provided information about child and caregiver hearing status, reading habits, and use of technology. The focus groups began with a brief introduction of the study and the team, study aims, participant's rights, and the agenda. Participants watched a short video, which demonstrated the app (Multimedia Appendix 2). Then, an assigned speech therapist assisted every 2 to 3 caregiver-child groups to use and test the app together. Families were encouraged to use the Hear Me Read app without additional guidance, and speech therapists primarily observed and answered questions as needed. Participants were provided iPads with the Hear Me Read app installed, and participants engaged with the app using the iPad speakers for playback. For 20 min, caregiver-child groups explored the app and completed the tasks of (1) recording a video, (2) selecting highlights for words and letters, and (3) reading a book. Following app use, the study administrators facilitated a semistructured focus group interview. The focus group participants (both children and caregivers) were first asked usability questions concerning the app's design and layout, functionality, ease of use, learnability, satisfaction, future use, and system reliability [28]. Usability questions were followed by questions intended to help understand the caregiver-child relationship with technology and use (eg, What technologies do you use at home? Which apps do you use mostly on your phone? Have you used any apps related to hearing before?). At the end of the session, caregiver-child groups filled out reaction cards together [29], where they highlighted the words from a list that express their feelings and opinions toward the app. The focus group sessions were audio recorded and transcribed.

Analysis

We employed a thematic analysis to analyze the data collected during the study. Thematic analysis is a common approach in qualitative research to identify, assess, and analyze the patterns in the data [30] and is commonly used to evaluate usability for mobile apps [31,32]. The recorded audio of the focus group sessions was transcribed, and meeting notes and observational notes were curated into a single document. Coauthors ES and SD implemented inductive thematic coding on Microsoft Excel software (Microsoft, Inc) following Braun and Clarke's [30] thematic analysis guideline. The following steps were used in the analysis process: (1) familiarizing with the data, (2) generating initial codes, (3) searching for themes, (4) reviewing and refining themes, (5) defining and naming themes, and (6) reporting the findings [30]. ES and SD went through multiple readings to extract codes and themes. Memos and observational notes were used to elaborate on the analysis. After independently developing themes and subthemes, ES and SD compared their independent coding schemes and agreed upon the 3 main themes used in the analysis and the division of individual comments. ES and SD then independently recoded 74 participant comments into 3 consensus themes (default needs: 26 items; specific needs: 34 items; and family need: 14 items).

XSL•FO

RenderX

Cohen kappa inter-rater reliability testing was employed to ensure rater agreement for the themes (Multimedia Appendix 4) [33]. The scores were in the range of 0.81 to 1.00, which counted as an *almost perfect* agreement for each theme [33] with a highly significant *P* value (95% CI 0.57-1.0; *P*<.001), which means the agreement is significantly different from what would be achieved by chance. We used RStudio 1.2 as statistical software (RStudio, Inc). The authors employed 2 additional sessions to discuss and build a consensus upon themes and codes with the coauthors, where discrepancies were present.

Results

Demographics

In total, 8 caregiver-child groups participated in the study (caregiver: n=8; children: n=8) divided into 2 focus groups (group 1 and group 2), with the younger children (aged 2-5 years) in group 1 and the older children (aged 7-13 years) in group 2. One child attended a session with 2 caregivers, and another single caregiver attended a session with 2 children. Demographic and hearing loss information for the child participants were collected (Table 1). The children's average age was 7.3 years (SD 3.5; range 2.4-12.4 years; median 7.2). Four (4/8, 50%) children were female. Five children (5/8, 63%) had moderate-to-severe sensorineural hearing loss, and all these children managed their hearing loss with bilateral cochlear implants. The caregivers of the children all used spoken language as the primary mode of communication at home, and 13% (2/15) of caregivers at home (1 child had a single caregiver at home) had hearing loss.

Table 1. Participant demographics.

ID	Group	Age (years)	Gender	Age ID[a] (months)	Hearing age[b] (months)	Management		Sensorineural hearing loss severity		CG[c] HL[d]	
						R[e]	L[f]	R	L	1	2
1	1	2.4	F[g]	4	25	CI[h]	CI	Sev-Prof[i]	Sev-Prof	No	No
2	1	4.8	F	36	36	CI	CI	Mod-Sev[j]	Mod-Sev	No	No
5	1	3.7	M[k]	0	4	CI	CI	Profound[l]	Profound	No	No
3	2	11.3	F	0	6	CI	CI	Profound	Profound	No	No
4	2	9.3	M	0	4	CI	CI	Profound	Profound	Yes	No
6	2	12.4	F	N/A[m]	N/A	None	None	NH[n]	NH	No	No
7	2	7.3	M	N/A	N/A	None	None	NH	NH	No	No
8	2	7.0	M	N/A	N/A	None	None	NH	NH	Yes	N/A

[a]Age ID: age of diagnosis of hearing loss.

[b]Hearing age: age of cochlear implantation.

[c]CG: caregiver.

[d]HL: hearing loss.

[e]R: right.

[f]L: left.

[g]F: female.

[h]CI: cochlear implant.

[i]Sev-Prof: severe to profound hearing loss (>70 dB to 91 dB HL).

[j]Mod-Sev: moderate-to-severe hearing loss (>40 dB to 70 dB HL).

[k]M: male.

[l]Profound: Profound hearing loss (>91 dB HL).

[m]N/A: not applicable.

[n]NH: normal hearing.

Technology Use

Questionnaire information regarding digital technology use in children in this study is mainly descriptive, given the small numbers (Table 2). All but one family used digital devices at home, although these devices were usually not used for digital reading.

Participants were given a survey before they interacted with Hear Me Read that asked about their current experiences with apps for hearing loss and features that they would ideally like in a reading app (Table 3). Participants are using training apps that provide visual and audio educational components. All caregivers of children with hearing loss knew what Ling 6 sounds were, and no children with normal hearing did (Table 3).

Table 2. Digital technology use.

ID	Device use at home		Apps commonly used		Digital reading device		
	Devices	Primary use at home	Caregiver app use	Child app use	Own?	Device	Reading done digitally
1	N/A[a]	N/A	N/A	N/A	Yes	Smartphone	None
2	Television[b], smartphone	Videos	Social media, news	ABC Mouse, YouTube Kids	No	N/A	N/A
3	Television smartphone[b]	Work	Text (Telegram)	YouTube Kids	No	N/A	N/A
4	Television, computer, iPad, smartphone[b]	Web, communication, music	Navigation (Mapquest, Waze)	Music, games, videos	No	N/A	N/A
5	iPad[b], smartphone[b]	Apps for speech, Netflix, work	Social media	Games	Yes	Kindle	1 in every 20 reading session
6	Television, computer, Kindle, smartphone[b]	News, communication, social media, music, videos	Web, email, text, Facebook, running tracking, music	Games, camera, Kindle	Yes	Kindle	6 in every 10 reading sessions
7	Television, computer, Kindle[b], smartphone	News, communication, social media, music, videos	Web, email, text, Facebook, running tracking, music	Games, Kindle	Yes	Kindle	2 in every 10 reading session
8	Television, iPad[b]	Learning	Facebook	iConnection	No	N/A	N/A

[a]N/A: not applicable.

[b]Devices most frequently used.

Table 3. Hearing loss app use.

ID	Apps used for hearing loss	Read differently because of hearing loss? In what way?	What would you like to have in a reading app?	Know Ling 6 sounds?
1	None	Yes—"We have done signs quite a bit with reading; Using our fingers to point out each word as we read."	"Show the sign after I reads it. That way, when a child knows the sign it can make the link from the sign to the audio part of the object; repeat on the page."	Yes
2	None	Yes—"I break down every part to make sure she understands what is happening."	"It would need to grab their attention. Also make them feel like they could understand and work the app."	Yes
3	My Signing Time, Sign and Sing	No	"An app that needs a passcode to exit the app while using it. Videos that demonstrate the task that is being teached. Vivid colors."	Yes
4	Lexia, RazKids	N/A[a]	"At this age, I would say to help (name) pronounce more challenging words and provide definitions of their meaning as she is reading."	Yes
5	Speech Stickers, Hope Words, Kids Vocab-Read Comp 1	No	"Have it have precise speech comprehension questions at the end."	Yes
6	None	N/A	N/A	N/A
7	None	N/A	N/A	N/A
8	None	N/A	"To slowly pronounce words as they are said and seen on screen."	N/A

[a]N/A: not applicable.

Focus Group Discussion Themes

Focus group analysis identified 3 major themes. The first theme, *default needs*, represents the generic needs from an app. The subthemes are ease of use, the intuitiveness of navigation, layout, and workflow, with comparisons to generic apps. The second theme, *specific needs*, represents the needs regarding hearing loss. The subthemes are the reading and language comprehension functionality, user engagement, and preferences, including attention, learning, and design. The third theme, *family needs*, includes family suggestions about how the app could be used, family interactions, and how Hear Me Read might fit within the daily lives of the participants. The subthemes are family relationships and daily life.

Theme 1: Default Needs (Ease of Use Compared With Generic Apps and Identifying Common Needs)

Ease of Use

Participants generally found the app to be easy to use with a pleasing design. Most participants agreed that the app compared favorably with other apps that they used frequently. Despite the difficulty in use because of the new components and complexity of functions, users confirmed that Hear Me Read is easy to learn and a fulfilling app at the first use. Caregivers and children were able to overcome difficulties and navigate the app after watching the short introductory video and briefly exploring the app:

> ..., did you find doing the tasks difficult? Was it easy and intuitive to learn it? Was there anything you got stuck on? [facilitator, group 1]

> No, just after I clicked around a little. [caregiver, group 1]

> Yeah, once you're exploring, you kind of get the hang of it. [caregiver, group 1]

Navigation

However, 2 functions are presented for modifying the usability of the app. The first was the page-turning functionality. Initially, the app required a long swiping motion across the bottom of the screen to navigate to a new page. However, participants, especially young children, found this method of navigation to be complicated:

> Children I think are going to really struggle. I think with flipping pages, if you have to have your finger in the small little corner to get it over, instead of just hitting a button to go to the next page or anywhere on the screen. [caregiver, group 1]

The chief complaint seemed to be that the swiping motion was too specific and not forgiving enough when general swiping movements or single taps were used. The children who were using the app tended to instinctively press a single point to navigate to the next page:

> Yeah, at first I just tapped it, and I thought it was like, if you tap it, it'll go. [child, group 2]

However, caregivers mentioned that they tended to see children using a swiping motion more often for other apps:

> I think kids in general, when they think of any kind of device they just kinda go like this [slides a finger from bottom to top of iPad] and it's just what they're used to doing. [caregiver, group 2]

Layout and Workflow

The second function with noted concerns was the recording function, which required the user to click record on every page. This action was not intuitive to most caregivers and caused them not to record once they progressed past the first page:

> I didn't realize you had to hit the record button after every page. We just kept right on going. [caregiver, group 2]

The app allows the ability to playback either video or audio. Although some participants only engaged with the audio, others

did not realize that audio only was an option. In addition, some participants did not comprehend that the recording could be deleted and replaced.

Theme 2: Specific Needs (Core Functionality, Engagement, and User Preferences)

Reading and Language Functionality

Participants mentioned multiple times the ability to highlight parts of speech as a favorite feature:

> I liked the feature of highlighting the word, vocabulary, stuff like that. That was a good idea. [caregiver, group 1]

> I loved how you could highlight pronouns and verbs and adjectives, and I see where that could be helpful with both of my children, both hearing and non-hearing. I think that's a real plus of the app. [caregiver, group 2]

User Engagement

However, groups 1 and 2 engaged in the game differently. For younger children (group 1), the video of a caregiver was especially engaging:

> When she did see me talking though, she did get really excited at first, like "I see mommy." [caregiver, group 1]

In addition, caregivers of younger children expressed concern that the text was too dense, and visual interest was insufficient to maintain engagement. Participants suggested increasing visual interest for younger children through pictures, colors, and font choices:

> I think if the book was longer and had some pictures... for each sentence have like a picture of what...the story is trying to tell. Cause he [referring to son] got excited when he saw the photos. [caregiver, group 1]

> I feel like since it's for a kid, it should have more color to attract kids. [caregiver, group 1]

> He's still reading books that are kind of fun to look at. Even chapter books that he reads, the font is a more playful font, and it's just more, I think, eye-catching for them. [caregiver, group 2]

User Preferences

Older caregiver-child groups wanted more advanced features, chapter books, and quizzes for comprehension:

> I think maybe you should add like a little quiz at the end or like something, just kinda refresh your memory. [child, group 2]

...so to have a couple questions at the end...that would make him think about what he just read. [caregiver, group 2]

Although the app is intended for caregivers to record themselves reading, some caregivers also suggested that the children could record themselves:

> Or you could even work with them, and they could even learn a sentence and record it themselves doing it; so I thought that was cool, too. Once they learn

the word, you could have them do the book and then play it back to them. [caregiver, group 1]

However, a caregiver also noticed that it was very distracting for her child to record himself:

I like the fact that I can record myself, but that it's extremely distracting for him to record himself. All he wants to do is look at himself...very distracting for him. [caregiver, group 2]

Theme 3: Family Needs (Integration of the App in Family Relationships and Daily Life)

Family Relationships

Participants emphasized the value of family engagement and joint use of the app for education and training of the children. In that regard, they agreed that they liked the use of custom videos of family. Participants also suggested that they could use the app to include extended family members (eg, grandmother) for video recording, which can help to bond family members through this media:

And that my face is on there, that he recognized. It's not like some random voice or a stranger's face or something, I like that. [caregiver, group 2]

But to your point, at least it could be a family member this way. [caregiver, group 2]

Yeah! Like a grandparent, that is a way they could help them. [caregiver, group 2]

So, you recognize the voice, it's not like some random person. [child, group 2]

Daily Life

Caregivers cited after school and bedtime as the most likely time for interaction with the app and estimated that they would spend 15 to 30 min interacting with their child and the app each day:

Well he started reading in school, so we would probably use it for after homework or something. [caregiver, group 2]

If you had a parent who traveled, I think it would be wonderful to re-record bedtime stories, and they see you. I mean, that is a really nice thing! [caregiver, group 2]

Some caregivers referenced current reading times for their children and suggested that they could use the app during these times:

He has required nightly reading so that would be nice. [caregiver, group 2]

I feel like before bedtime would be easiest when he does his reading. [caregiver, group 1]

User Reactions

At the end of the interview session, participants were given reaction cards with a series of words and asked to mark each word that they felt applied to the app. This method was used to quickly capture their thoughts and feelings about the app. In Figure 2, the reactions are arranged from the highest number of marks in the upper left corner to the lowest at the bottom right of the figure. The highest number of marks was found for those words most related to ease of use (usable, easy to use, and straight forward) and value (valuable, useful, and motivating). In addition, words related to positive design aspects had high and medium representation (desirable, attractive, appealing, and inviting). The words *intimidating* and *slow* each had one mark, whereas other negative words received no marks.

Figure 2. User reactions. Users were asked to mark words that they associated with the app. Displayed are the total counts for each word that was marked. Words were rearranged by count for display.

valuable 8	usable 8	attractive 5	appealing 3	sophisticated 0
useful 8	straight forward 7	desirable 6	accessible 3	overbearing 0
organized 6	fun 7	customizable 6	inviting 3	overwhelming 0
easy to use 5	high quality 5	motivating 4	exciting 3	patronizing 0
flexible 3	time-saving 3	reliable 3	empowering 3	predictable 0
efficient 2	collaborative 2	trustworthy 2	comprehensive 2	confusing 0
stimulating 2	simplistic 2	personal 2	fast 2	rigid 0
relevant 2	connected 1	fresh 1	intimidating 1	inconsistent 0
familiar 1	slow 1	stressful 0	busy 0	not valuable 0
unpredictable 0	unconventional 0	uncontrollable 0	frustrating 0	hard to use 0
too technical 0	complex 0	time-consuming 0	consistent 0	gets in the way 0

Discussion

Overview

Hear Me Read is an app built to expand upon the *gold standard* practice for developing literacy in children with hearing loss, which consists primarily of in-person therapy sessions with a pediatric hearing loss expert. Hear Me Read allows caregivers to work with therapists to develop lesson plans within the app, thereby facilitating at-home interactions between caregivers and their children regularly. Furthermore, the recording functionality allows the caregiver to prerecord reading sessions that the child can use alone. Children can, therefore, interact with the text in multiple ways. Caregiver and child engagement while using this technology is the core objective of this app. Joint caregiver-child engagement with media facilitates learning [34] and can enhance family relationships [35]. Few commercially available apps promote the caregiver's involvement while the child is engaging with the app. Hear Me Read is developed with a user-centered perspective and makes the caregiver a core partner in engagement through functionality, whereby the caregiver can record video and audio of themselves to be played back by the child, assign customized reading objectives, and read together with their child. Hear Me Read's utility is designed to supplement the joint book reading in which caregivers are already engaging their children. Hear Me Read is not designed to replace traditional printed storybooks. Caregivers can apply the knowledge from the Hear Me Read app to their shared storybook reading experiences using a variety of platforms, including both printed text and digital text.

The objective of this study was to use focus groups to investigate the usability of the Hear Me Read app within caregiver-child groups. Through participant interaction with the app followed by a focus group discussion, we sought to identify areas requiring modifications and improvements during the use of the app, to understand how caregivers and children interacted with the core functionality of the app, and to determine how they might use the app in their daily life.

Principal Findings

We conducted two focus groups, one with children aged 2 to 5 years and their caregivers and the other with children aged 7 to 12 years and their caregivers. All children with hearing loss had bilateral cochlear implants. The majority of family groups used either smartphones or Kindle as their primary digital devices, although not all participants used these devices for digital reading. When considering the generic attributes of Hear Me Read, participants agreed that the app was visually pleasing, generally easy to use, and compared favorably to other apps they use (both general apps and apps related to literacy). We identified two obstacles (default needs) for ease of use in the app, namely the page-turning functionality and the recording functionality. Without a clear consensus on the preferred way to turn the page (swiping vs pressing), we changed the navigation to accept either a general swiping motion or a button press on the right or left side of the display. In addition, our participants tended to assume that recording continued as they changed the pages, so we also updated our app to allow for continuous recording rather than requiring it to be halted and then reinitiated at each page. Despite these two usability issues,

XSL•FO

RenderX

participants marked the app high for ease of use in the reaction cards. The aforementioned favorable attributes and obstacles may be common issues for any user group. We have not identified any significant correlation with those who have hearing loss and these user preferences.

In terms of user-specific needs, a key objective was to determine if functionality deviates from non-digital standard practices. In that regard, participants found that customized options such as recording and playback and dynamic word highlighting were accessible and valuable. The participants also expressed appreciation for both the video recording and highlighting capabilities, suggesting that these features would be useful for improving literacy. Such user-level customization is a significant driving factor in the digital health app literature as well [36]. Most caregiver participants already included daily reading activities with their children, either after school or before bed, and were able to see using the app as fitting within their daily schedules for 15 to 30 min. Most of the participants expressed appreciation for the recording functionality and the ability to record a family member rather than to have a video of a stranger. They suggested that this feature may increase engagement because of familiarity and may also increase family bonding, even when caregivers and children are not reading together. This finding supports the importance of including coviewing design elements in development through joint media engagement with family members. This aspect will be investigated further in a clinical trial.

Limitations

Participants in this study interacted with the app for only a limited time (40 min), which may have limited their experiences to be able to interact with all the features of the app. We observed that different participants interacted with different features of the app. However, none of the participants were able to interface with the full suite of features. Testing time constraints may have limited further investigation of usability issues. The two focus groups represented a relatively small sample size (n=8) and may not have been representative of the full spectrum of the patient and caregiver population. Furthermore, heterogeneity in terms of the children's age and gender may limit the generalization of age- or gender-specific implications. The presence of children both with and without hearing loss means that our feedback was skewed toward more general features that are accessible to children with and without hearing loss. However, we found that the feedback we received from both groups was congruent.

The NCH clinic recruited participants with a snowball approach, and therefore, familiarity with some of the staff, including the speech therapists who were present during the focus group, may have had an inhibitory effect on negative feedback. Finally, the scope of this study was limited to an evaluation of the usability of core functions. We did not ask participants to consider evaluating all features (eg, the therapist tools) of the app in depth. Future studies should expand to evaluate all features with an extended user time, and therapeutic tasks should be assigned and personalized for each participant.

Conclusions

We conducted a focus group study for usability testing of the Hear Me Read app. Participants primarily found the app to be easy to use, esthetically pleasing, and valuable. Feedback from this study was used to improve the app and contribute to the literature by reporting user needs and expectations from children with hearing loss and caregiver population for a mobile app. The findings are promising for the use of educational apps designed specifically for the hearing loss population. Further investigation is needed with larger sample sizes and the actual impact on relevant language and literacy outcomes in this population.

Authors' Contributions

SD and ES designed and conducted the focus group, performed thematic analysis on the transcripts, and performed the primary preparation of the manuscript. JH was responsible for participant recruitment, assisted with focus group facilitation, and contributed to the app design. JL implemented the app and contributed to the app design methodology. AS contributed to the app conception and design process, and PM was the primary contributor to the app conception and design and conducted the background research.

Conflicts of Interest

None declared.

Multimedia Appendix 1
A table compiling existing apps for help with speech-language pathology, literacy, and hearing loss. Apps were found through a search on the Apple iTunes app store and a Google keyword search. The keywords for the search were related to speech therapy apps for children with hearing loss along with related terms, such as phonics, phonetics, and auditory memory.
[XLSX File (Microsoft Excel File), 10 KB - humanfactors_v7i2e16310_app1.xlsx]

Multimedia Appendix 2
A video explaining how to use the Hear Me Read app.
[MP4 File (MP4 Video), 12075 KB - humanfactors_v7i2e16310_app2.mp4]

Multimedia Appendix 3
Questionnaire for parent participants in the focus groups.
[DOCX File , 20 KB - humanfactors_v7i2e16310_app3.docx]

Multimedia Appendix 4
Cohen kappa inter-rater reliability testing results.
[DOCX File , 15 KB - humanfactors_v7i2e16310_app4.docx]

References

1. Centers for Disease Control and Prevention (CDC). Identifying infants with hearing loss-United States, 1999-2007. MMWR Morb Mortal Wkly Rep 2010 Mar 5;59(8):220-223 [FREE Full text] [Medline: 20203554]
2. World Health Organization. 2020 Mar 1. Deafness and Hearing Loss URL: https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss [accessed 2020-03-31]
3. Holt J. Classroom attributes and achievement test scores for deaf and hard of hearing students. Am Ann Deaf 1994 Oct;139(4):430-437. [doi: 10.1353/aad.2012.0274] [Medline: 7810461]
4. Traxler CB. The Stanford Achievement Test, 9th edition: national norming and performance standards for deaf and hard-of-hearing students. J Deaf Stud Deaf Educ 2000;5(4):337-348. [doi: 10.1093/deafed/5.4.337] [Medline: 15454499]
5. Wauters LN, van Bon WH, Tellings AE. Reading comprehension of Dutch deaf children. Read Writ 2006 Feb;19(1):49-76. [doi: 10.1007/s11145-004-5894-0]
6. Geers AE. Predictors of reading skill development in children with early cochlear implantation. Ear Hear 2003 Mar;24(1 Suppl):59S-68S. [doi: 10.1097/01.AUD.0000051690.43989.5D] [Medline: 12612481]
7. Marschark M, Rhoten C, Fabich M. Effects of cochlear implants on children's reading and academic achievement. J Deaf Stud Deaf Educ 2007;12(3):269-282. [doi: 10.1093/deafed/enm013] [Medline: 17526867]
8. Nittrouer S, Caldwell A, Lowenstein JH, Tarr E, Holloman C. Emergent literacy in kindergartners with cochlear implants. Ear Hear 2012;33(6):683-697 [FREE Full text] [doi: 10.1097/AUD.0b013e318258c98e] [Medline: 22572795]
9. Perfetti C, Sandak R. Reading optimally builds on spoken language: implications for deaf readers. J Deaf Stud Deaf Educ 2000;5(1):32-50. [doi: 10.1093/deafed/5.1.32] [Medline: 15454516]
10. Spencer LJ, Barker BA, Tomblin JB. Exploring the language and literacy outcomes of pediatric cochlear implant users. Ear Hear 2003 Jun;24(3):236-247 [FREE Full text] [doi: 10.1097/01.AUD.0000069231.72244.94] [Medline: 12799546]
11. Connor CM, Zwolan TA. Examining multiple sources of influence on the reading comprehension skills of children who use cochlear implants. J Speech Lang Hear Res 2004 Jun;47(3):509-526. [doi: 10.1044/1092-4388(2004/040)] [Medline: 15212565]
12. Musselman C. How do children who can't hear learn to read an alphabetic script? A review of the literature on reading and deafness. J Deaf Stud Deaf Educ 2000;5(1):9-31. [doi: 10.1093/deafed/5.1.9] [Medline: 15454515]
13. Dennis LR, Horn E. Strategies for supporting early literacy development. Young Except Child 2011 Sep 16;14(3):29-40. [doi: 10.1177/1096250611420553]
14. Hay I, Fielding-Barnsley R. Competencies that underpin children's transition into early literacy. Aust J Lang Lit 2009;32(2):148-162 [FREE Full text]
15. Massetti GM, Bracken SS. Classroom academic and social context: relationships among emergent literacy, behavioural functioning and teacher curriculum goals in kindergarten. Early Child Dev Care 2010 Apr;180(3):359-375. [doi: 10.1080/03004430801917401]
16. Missall KN, Mcconnell SR, Cadigan K. Early literacy development: skill growth and relations between classroom variables for preschool children. J Early Interv 2016 Jul 25;29(1):1-21. [doi: 10.1177/105381510602900101]
17. Bus AG, van Ijzendoorn MH, Pellegrini AD. Joint book reading makes for success in learning to read: a meta-analysis on intergenerational transmission of literacy. Rev Educ Res 1995;65(1):1. [doi: 10.2307/1170476]
18. Mol SE, Bus AG, de Jong MT. Interactive book reading in early education: a tool to stimulate print knowledge as well as oral language. Rev Educ Res 2009;79(2):979-1007. [doi: 10.3102/0034654309332561]
19. van Kleeck A, Woude JV. Book-sharing with preschoolers with language delays. In: van Kleeck A, Stahl SA, Bauer EB, editors. On Reading Books to Children: Parents and Teachers. Abingdon, United Kingdom: Routledge; 2003:58-92.
20. NAEYC, NAECS/SDE. National Association for the Education of Young Children. Washington DC; 2002. Early Learning Standards: Creating the Conditions for Success URL: https://www.naeyc.org/sites/default/files/globally-shared/downloads/PDFs/resources/position-statements/position_statement.pdf [accessed 2020-03-20]
21. Radesky JS, Schumacher J, Zuckerman B. Mobile and interactive media use by young children: the good, the bad, and the unknown. Pediatrics 2015 Jan;135(1):1-3. [doi: 10.1542/peds.2014-2251] [Medline: 25548323]

XSL•FO
RenderX

22. Parish-Morris J, Mahajan N, Hirsh-Pasek K, Golinkoff RM, Collins MF. Once upon a time: parent-child dialogue and storybook reading in the electronic era. Mind Brain Educ 2013 Aug 16;7(3):200-211. [doi: 10.1111/mbe.12028]

23. Hirsh-Pasek K, Zosh JM, Golinkoff RM, Gray JH, Robb MB, Kaufman J. Putting education in 'educational' apps: lessons from the science of learning. Psychol Sci Public Interest 2015 May;16(1):3-34. [doi: 10.1177/1529100615569721] [Medline: 25985468]

24. Stacey L. Children with hearing loss: developing listening and talking, birth to six, 3rd edition. Int J Audiol 2017 Jan 13;56(5):358-359. [doi: 10.1080/14992027.2016.1270470]

25. Runnion E, Gray S. What clinicians need to know about early literacy development in children with hearing loss. Lang Speech Hear Serv Sch 2019 Jan 28;50(1):16-33. [doi: 10.1044/2018_LSHSS-18-0015] [Medline: 30950774]

26. Alexander Graham Bell Association for the Deaf and Hard of Hearing. Washington DC: AG Bell; 2019. Evaluating Communications Options for Your Child URL: https://www.agbell.org/Families/Communication-Options [accessed 2020-03-20]

27. Agung K, Purdy S, Kitamura C. The ling sound test revisited. Aust New Zeal J Audiol 2005 May;27(1):33-41. [doi: 10.1375/audi.2005.27.1.33]

28. Zins AH, Bauernfeind U, del Missier F, Venturini A, Rumetshofer H, Frew AJ. An experimental usability test for different destination recommender systems. In: Frew AJ, editor. Information and Communication Technologies in Tourism 2004. NY: Springer; 2004:228-238.

29. Benedek J, Miner T. Measuring Desirability: New Methods for Evaluating Desirability in a Usability Lab Setting. In: Proceedings of the 2002 Usability Professionals Association. 2002 Presented at: UPA'09; July 8–12, 2002; Orlando, FL URL: https://www.researchgate.net/publication/228721563_Measuring_Desirability_New_methods_for_evaluating_desirability_in_a_usability_lab_setting

30. Braun A, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]

31. Hull P, Emerson JS, Quirk ME, Canedo JR, Jones JL, Vylegzhanina V, et al. A smartphone app for families with preschool-aged children in a public nutrition program: prototype development and beta-testing. JMIR Mhealth Uhealth 2017 Aug 2;5(8):e102 [FREE Full text] [doi: 10.2196/mhealth.7477] [Medline: 28768611]

32. Martin CA, Upvall MJ. A mobile phone HIV medication adherence intervention: acceptability and feasibility study. J Assoc Nurses AIDS Care 2016;27(6):804-816. [doi: 10.1016/j.jana.2016.07.002] [Medline: 27497514]

33. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012;22(3):276-282 [FREE Full text] [Medline: 23092060]

34. Kabali HK, Irigoyen MM, Nunez-Davis R, Budacki JG, Mohanty SH, Leister KP, et al. Exposure and use of mobile media devices by young children. Pediatrics 2015 Dec;136(6):1044-1050 [FREE Full text] [doi: 10.1542/peds.2015-2151] [Medline: 26527548]

35. Takeuchi L, Reed SR. The Joan Ganz Cooney Center at Sesame Workshop. NY: The Joan Ganz Cooney Center; 2011 Dec 8. The New Coviewing: Designing for Learning through Joint Media Engagement URL: https://joanganzcooneycenter.org/publication/the-new-coviewing-designing-for-learning-through-joint-media-engagement/ [accessed 2020-03-20]

36. Sama PR, Eapen ZJ, Weinfurt KP, Shah BR, Schulman KA. An evaluation of mobile health application tools. JMIR Mhealth Uhealth 2014 May 1;2(2):e19 [FREE Full text] [doi: 10.2196/mhealth.3088] [Medline: 25099179]

Abbreviations

eBook: electronic book
NCH: Nationwide Children's Hospital

XSL•FO

RenderX

XSL•FO

RenderX