JMIR Formative Research

Electronic, mobile, digital health approaches in cardiology and for cardiovascular health. Official partner journal of the European Congress on eCardiology and eHealth
Volume 5 (2021), Issue 8 ISSN: 2561-326X Editor in Chief: Nico Bruining, PhD

Contents

Original Papers

Patients' Experiences of Using a Self-help App for Posttraumatic Stress Disorder: Qualitative Study (e26852)
Lisa Riisager, Anne Christensen, Frederik Scharff, Ida-Marie Arendt, Israa Ismail, Marianne Lau, Stine Moeller. 4

Physiologic Response to the Pfizer-BioNTech COVID-19 Vaccine Measured Using Wearable Devices: Prospective Observational Study (e28568)
Alexander Hajduczok, Kara DiJoseph, Brinnae Bent, Audrey Thorp, Jon Mullholand, Stuart MacKay, Sabrina Barik, Jamie Coleman, Catharine Paules, Andrew Tinsley. 16

Mobile Health for Smoking Cessation Among Disadvantaged Young Women During and After Pregnancy: User-Centered Design and Usability Study (e24112)
Marloes Derksen, Monique Jaspers, Sander van Strijp, Mirjam Fransen. 39

Mobile App (WHEELS) to Promote a Healthy Lifestyle in Wheelchair Users With Spinal Cord Injury or Lower Limb Amputation: Usability and Feasibility Study (e24909)
Dirk Hoevenaars, Jasmijn Holla, Leonie te Loo, Johan Koedijker, Sarah Dankers, Han Houdijk, Bart Visser, Thomas Janssen, Sonja de Groot, Marije Deutekom, WHEELS Study Group. 61

Design of the Maternal Website EMAeHealth That Supports Decision-Making During Pregnancy and in the Postpartum Period: Collaborative Action Research Study (e28855)
Isabel Artieta-Pinedo, Carmen Paz-Pascual, Paola Bully, Maite Espinosa, EmaQ Group. 81

Artificial Intelligence–Based Chatbot for Anxiety and Depression in University Students: Pilot Randomized Controlled Trial (e20678)
Maria Klos, Milagros Escoredo, Angela Joerin, Viviana Lemos, Michiel Rauws, Eduardo Bunge. 94

The Experience of Key Stakeholders During the Implementation and Use of Trauma Therapy via Digital Health for Military, Veteran, and Public Safety Personnel: Qualitative Thematic Analysis (e26369)
Lorraine Smith-MacDonald, Chelsea Jones, Phillip Sevigny, Allison White, Alexa Laidlaw, Melissa Voth, Cynthia Mikolas, Alexandra Heber, Andrew Greenshaw, Suzette Brémault-Phillips. 103

Digital Patient-Reported Outcome Measures for Monitoring of Patients on Cancer Treatment: Cross-sectional Questionnaire Study (e18502)
Mayuran Sivanandan, Catherine Sharma, Pippa Bullard, Judith Christian. 121

Development of a Mobile Health App (TOGETHERCare) to Reduce Cancer Care Partner Burden: Product Design Study (e22608)
Ingrid Oakley-Girvan, Sharon Davis, Allison Kurian, Lisa Rosas, Jena Daniels, Oxana Palesh, Rachel Mesia, Arif Kamal, Michelle Longmire, Vasu Divi. 133

XSL•FO
RenderX

Understanding Physicians' Preferences for Telemedicine During the COVID-19 Pandemic: Cross-sectional Study (e26565)
Sarah Nies, Shae Patel, Melissa Shafer, Laura Longman, Iman Sharif, Paulo Pina. 143

Copyright as a Barrier to Music Therapy Telehealth Interventions: Qualitative Interview Study (e28383)
Amanda Reid, Alex Kresovich. 150

Efficiency, Usability, and Outcomes of Proctored Next-Level Exams for Proficiency Testing in Primary Care Education: Observational Study (e23834)
Birgitte Schoenmakers, Johan Wens. 163

Development and Evaluation of Acceptability and Feasibility of a Web-Based Intervention for Patients With Bipolar Disorder in Iran: Implementation Study (e23360)
Agaah Ashrafi, Maryam Tabatabaee, Vandad Sharifi. 171

Development of a Mobile App to Support Self-management of Anxiety and Depression in African American Women: Usability Study (e24393)
Terika McCall, Muhammad Ali, Fei Yu, Paul Fontelo, Saif Khairat. 181

Utilization of a Directly Supervised Telehealth-Based Exercise Training Program in Patients With Nonalcoholic Steatohepatitis: Feasibility Study (e30239)
Victoria Motz, Alison Faust, Jessica Dahmus, Benjamin Stern, Christopher Soriano, Jonathan Stine. 197

Satisfaction and Usability of an Information and Communications Technology–Based System by Clinically Healthy Patients With COVID-19 and Medical Professionals: Cross-sectional Survey and Focus Group Interview Study (e26227)
Ye Bae, Kyung Kim, Sae Choi, Taehoon Ko, Jun Lim, Meihua Piao. 202

Cyberbullying Prevention for Adolescents: Iterative Qualitative Methods for Mobile Intervention Design (e25900)
Megan Ranney, Sarah Pittman, Isabelle Moseley, Kristen Morgan, Alison Riese, Michele Ybarra, Rebecca Cunningham, Rochelle Rosen. . .
2 1 6

Dietary Intake and Health Status of Elderly Patients With Type 2 Diabetes Mellitus: Cross-sectional Study Using a Mobile App in Primary Care (e27454)
Joane Coleone, Ericles Bellei, Mateus Roman, Vanessa Kirsten, Ana De Marchi. 229

Evaluation and Refinement of a Bank of SMS Text Messages to Promote Behavior Change Adherence Following a Diabetes Prevention Program: Survey Study (e28163)
Megan MacPherson, Kaela Cranston, Cara Johnston, Sean Locke, Mary Jung. 243

Scoping the Need for a Tailored mHealth App to Improve Health and Well-being Behavioral Transformation in the Police: Exploring the Views of UK Police Workers via Web-Based Surveys and Client Meetings (e28075)
Emma Swanston, Andy Pulman, Huseyin Dogan, Jane Murphy, Fiona Bitters. 256

A Mobile Health App (WYZ) for Engagement in Care and Antiretroviral Therapy Adherence Among Youth and Young Adults Living With HIV: Single-Arm Pilot Intervention Study (e26861)
Parya Saberi, Nadra Lisha, Xavier Erguera, Estie Hudes, Mallory Johnson, Theodore Ruel, Torsten Neilands. 269

Informed Decision-making for Health Insurance Enrollment: Survey Study (e27477)
Coralys Colón-Morales, Wayne Giang, Michelle Alvarado. 281

XSL•FO
RenderX

Attitudes of Patients and Health Professionals Regarding Screening Algorithms: Qualitative Study (e17971)
Christina Oxholm, Anne-Marie Christensen, Regina Christiansen, Uffe Wiil, Anette Nielsen. 301

Estimation of Psychological Distress in Japanese Youth Through Narrative Writing: Text-Based Stylometric and Sentiment Analyses (e29500)
Masae Manabe, Kongmeng Liew, Shuntaro Yada, Shoko Wakamiya, Eiji Aramaki. 311

Screening Diabetic Retinopathy Using an Automated Retinal Image Analysis System in Independent and Assistive Use Cases in Mexico: Randomized Controlled Trial (e25290)
Alejandro Noriega, Daniela Meizner, Dalia Camacho, Jennifer Enciso, Hugo Quiroz-Mercado, Virgilio Morales-Canton, Abdullah Almaatouq, Alex Pentland. 319

Preferences for Digital Smartphone Mental Health Apps Among Adolescents: Qualitative Interview Study (e14004)
Robert Ribanszki, Jose Saez Fonseca, Joseph Barnby, Kimberly Jano, Fatima Osmani, Soma Almasi, Elias Tsakanikos. 331

Analysis of Hospital Quality Measures and Web-Based Chargemasters, 2019: Cross-sectional Study (e26887)
Kunal Patel, Olena Mazurenko, Eric Ford. 341

Redesigning a Web-Based Stakeholder Consensus Meeting About Core Outcomes for Clinical Trials: Formative Feedback Study (e28878)
Roulla Katiri, Deborah Hall, Derek Hoare, Kathryn Fackrell, Adele Horobin, Nóra Buggy, Nicholas Hogan, Pádraig Kitterick, Core Rehabilitation Outcome Set For Single-Sided Deafness (CROSSSD) Initiative. 355

Association of COVID-19 Risk Misperceptions With Household Isolation in the United States: Survey Study (e30164)
Joseph Ladapo, Jonathan Rothwell, Christina Ramirez. 363

Online Search Trends Influencing Anticoagulation in Patients With COVID-19: Observational Study (e21817)
Amy Worrall, Claire Kelly, Aine O'Neill, Murray O'Doherty, Eoin Kelleher, Anne Cushen, Cora McNally, Samuel McConkey, Siobhan Glavey, Michelle Lavin, Eoghan de Barra. 375

Early Report

Electronic Paper Displays in Hospital Operations: Proposal for Deployment and Implementation (e30862)
Guruprasad Jambaulikar, Andrew Marshall, Mohammad Hasdianda, Chenzhe Cao, Paul Chen, Steven Miyawaki, Christopher Baugh, Haipeng Zhang, Jonathan McCabe, Jennifer Su, Adam Landman, Peter Chai. 28

XSL•FO
RenderX

Original Paper

Patients' Experiences of Using a Self-help App for Posttraumatic Stress Disorder: Qualitative Study

Lisa H G Riisager[1,2], MCS; Anne Bryde Christensen[3], MSc, MRES, PhD; Frederik Bernt Scharff[1], MSc; Ida-Marie T P Arendt[1], MSc; Israa Ismail[1], BSc; Marianne Engelbrecht Lau[1], MSc, MD, PhD; Stine Bjerrum Moeller[1,2], MSc, PhD

[1]Unit for Psychotherapy Research, Stolpegaard Psychotherapy Centre, Mental Health Services Capital Region of Denmark, Gentofte, Denmark

[2]Department of Psychology, University of Southern Denmark, Odense, Denmark

[3]Research Unit for Psychotherapy & Psychopathology, Mental Health Services West, Region Zealand, Slagelse, Denmark

Corresponding Author:
Lisa H G Riisager, MCS
Department of Psychology
University of Southern Denmark
Campusvej 55
Odense, 5230
Denmark
Phone: 45 25300977
Email: lisagr@health.sdu.dk

Abstract

Background: Posttraumatic stress disorder (PTSD) is a common disorder that requires more treatment options. Mobile health (mHealth) app interventions are promising for patients with PTSD, as they can provide easily accessible support, strategies, and information. However, knowledge about mHealth interventions is sparse and primarily based on quantitative studies.

Objective: The aim of this study is to qualitatively explore the experiences of patients with PTSD with regard to using an mHealth app as a stand-alone intervention before commencing psychotherapeutic treatment.

Methods: We conducted semistructured interviews with 14 participants 6 weeks after they received the app. The participants were all referred to PTSD treatment and were waiting to commence psychotherapeutic treatment. During this waiting time, the participants had no contact with the health staff. Interviews were transcribed and were analyzed using thematic analysis.

Results: A total of 3 themes were identified—the use of app, being a patient, and the overall evaluation of the app. The use of the app was described with the subtheme of habits, and the theme of being a patient included the subthemes of having negative experiences with the app and being a part of a research project. The use of the app encompassed how psychological factors and technical problems could interfere with the use of the app. The theme of being a patient depicted that the waiting time before starting treatment was long, and a subgroup of patients experienced feeling worse during this time, which they partly attributed to using the app. Several suggestions for change have been described in the overall evaluation of the app.

Conclusions: The findings in this study revealed that emotional arousal influenced the use of the app and that it was difficult for participants to establish a habit of using the app, thus reflecting the importance of supporting habit formation when implementing an mHealth app in mental health care services. This study makes an important contribution to the field of mHealth research, as it revealed that some participants had negative experiences resulting from using the app, thus reflecting the potential harm of having an mHealth app without the support of a clinician. It is therefore recommended to use a blended care treatment or an approach in which mental health care professionals prescribe an mHealth app for relevant patients to avoid increased suicidal risk.

(JMIR Form Res 2021;5(8):e26852) doi:10.2196/26852

KEYWORDS

app; PTSD; mHealth; qualitative analysis; patient experience; posttraumatic stress disorder; thematic analysis; smartphone; intervention; mobile phone

XSL•FO
RenderX

Introduction

Background

Posttraumatic stress disorder (PTSD) is a common disorder, with a prevalence of approximately 2% in European countries [1]. There are serious personal and societal consequences associated with PTSD, such as poor quality of life, high comorbidity, and increased use of health services [1,2]. Furthermore, PTSD is recognized as a risk factor for suicidal thoughts and behaviors and completed suicide [3,4]. Although the costs of PTSD are widely recognized, logistical and individual barriers, such as a lack of available mental health care services and negative beliefs about help-seeking, are common [5,6].

In Denmark, patients are referred to the public mental health services (MHS) by their general practitioner if a complex mental disorder is suspected. Patients are entitled to have a diagnostic interview within 30 days, but psychotherapeutic treatment may start substantially later depending on the available resources in the given clinic. As such, patients often experience waiting list periods exceeding 6 weeks before commencing treatment. During this time, the patient has minimal support from the MHS and, therefore, there is considerable interest in developing technological treatment alternatives.

Smartphones are owned by 79% of adults in European countries [7] and offer a platform with the potential to reach populations with otherwise limited access to health care [8]. Mobile health (mHealth) apps provide easy access to cost-effective treatment tools that can potentially increase treatment engagement and well-being [5,9]. There is great potential to effectively integrate apps into mental health care, as they can offer interventions, provide psychoeducation, promote self-awareness, and help overcome the self-stigmatization associated with receiving mental health care [10-12]. A systematic review and meta-analysis of the effects of self-management apps for patients with PTSD revealed that no significant difference was found between self-management app–based treatment groups and the control group on waiting list [8]. However, only 6 studies were included, leaving the quality of the evidence base low.

Although there is a growing body of quantitative research on mHealth apps as a supplement or stand-alone treatment [6,13-16], qualitative studies are surprisingly absent. To our knowledge, only 2 studies have used a qualitative approach to investigate patients' experiences using an mHealth app for PTSD. A study on attitudes toward mHealth in a sample of veterans with either PTSD, depression, and/or an alcohol abuse disorder revealed a marked difference in the openness toward using mHealth apps depending on rurality, where rural veterans expressed more negative views than urban veterans [17]. Another study investigating the PTSD Coach app showed that participants found some coping strategies and self-assessment tools useful [18]. Consequently, it is still unclear how patients with PTSD perceive and experience using mHealth apps [8]. There is a need for future research to explore these experiences, as they can provide unique insights into compliance, engagement, adherence, as well as positive and adverse effects. With this exploratory qualitative study, we seek to fill this research gap.

Objectives

To investigate the potential of using mHealth apps for patients with PTSD in MHS, an mHealth app named *PTSD Help* was developed for use as a stand-alone treatment and as a supplement to psychotherapeutic treatment. This study explores the following research question: how do patients diagnosed with PTSD experience the use of an app, PTSD Help, as a stand-alone treatment before psychotherapeutic treatment?

By exploring this aspect qualitatively, we aim to uncover both the experienced benefits and limitations of using the app, which can be used in the process of development and implementation of mHealth apps for PTSD in MHS.

Methods

Context

The study was conducted within the context of the study "The PTSD help app in a Danish PTSD population: A randomized controlled feasibility trial," which investigated the feasibility of implementing the PTSD Help app in the Danish MHS [13]. The participants in the larger study were randomized to either a waiting list control group or a PTSD Help app treatment group. All participants received treatment for their PTSD diagnosis after 6 weeks [13]. The randomized controlled feasibility trial of the PTSD Help app study will be published in the near future. This study used a qualitative study design to explore the patients' experiences using the app as a stand-alone treatment in the waiting period before the commencement of treatment in the MHS.

Intervention: PTSD Help

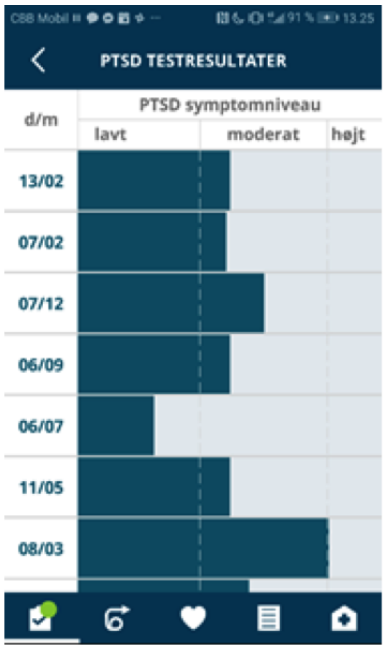
PTSD Help is an mHealth app that includes functionalities such as psychoeducation, emotion regulation tools, a note function, and a crisis plan (Figure 1). Emotion regulation tools offer a range of different interventions, including distraction exercises, grounding exercises, simple body exercises, and calming images accompanied by music. As sleep problems are common in the PTSD population, the app contains two functions to alleviate sleep problems: guided sleep meditation, in which the individual can choose between a male or female voice, and general sleep hygiene advice. Two self-assessment tools are provided in the app: the PTSD checklist for DSM-5 [19] to monitor PTSD symptoms, which is available with an interval of at least 2 weeks to be as close to the clinical use as possible. The second assessment tool aimed to monitor sleep quality (sleep condition indicator) [20], which is available daily (Figure 2). An overall favorite function is provided, giving the patient the opportunity to choose which functions are the most meaningful and bringing these functions to the front page.

To ensure that PTSD Help can be used across treatment modalities, it is designed as a generic intervention and does not include theory-specific elements. This design was chosen as the Danish MHS, which provides a range of different treatment types for patients with PTSD.

XSL•FO
RenderX

Figure 1. The PTSD Help app overall user surface. PTSD: posttraumatic stress disorder.

Figure 2. Overview of the PTSD self-assessment tool (in Danish). PTSD: posttraumatic stress disorder.



Ethical Considerations

Participants provided both verbal and written informed consent before participating in the study and were informed that they could withdraw their consent at any time. Participants were informed that the interviews were being recorded and subsequently anonymized and transcribed verbatim. To ensure anonymity, participants were given a project ID, and audio recordings of the interviews were stored in a secure database administered by the Capital Region of Denmark. The participants' psychotherapeutic treatment commenced shortly after the interview was conducted, thus ensuring minimal treatment gap.

Participants

The study was conducted within the MHS of the Capital Region of Denmark (MHS-CRD), where 14 participants were recruited. Participants included in the study were aged ≥18 years, diagnosed with PTSD in accordance with the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5), were accepted for PTSD treatment, had access to a smartphone with iOS (version 10 or higher) or Android (version 5.0.1. or higher), and provided informed consent. The reasons for exclusion were increased suicidal risk, current episodes of bipolar disorder or psychotic disorder, current substance abuse, inability to understand and/or read Danish, and concurrent psychiatric or psychological treatment of PTSD outside the MHS-CRD [13].

XSL•FO
RenderX

Procedure

Participants gained access to the PTSD Help app after the assessment interview and were given brief verbal general instructions on how to use the different functions in the app. This verbal introduction was based on a manuscript to ensure that all participants received the same information and took approximately 5 minutes. Furthermore, 3 days after gaining access to the app, the participants were contacted by telephone to ensure that there were no technical issues in using the app. There was no further assistance in using the app following the telephone call. Finally, the participants were contacted by telephone 6 weeks after having gained access to the app, and they were invited to participate in semistructured interviews (full description of the recruitment procedure has been provided in the study by Scharff et al [13]). The participants were not prohibited from having contact with health staff while waiting to commence psychotherapeutic treatment; however, this was rarely seen in the MHS-CRD and was therefore not monitored. Interviews were conducted in 2019. The sample was randomly selected because of its strict time schedule.

Materials

The semistructured interview guide aimed at exploring the participants' experiences using the PTSD Help app (see Textbox 1 for concepts covered). The interviews lasted approximately 30 minutes and were conducted over the telephone. The interviews were transcribed verbatim by a graduate psychology student.

Assessments were conducted as part of a larger study. The Mini-International Neuropsychiatric Interview 7.02. [21] was used to establish PTSD diagnosis in accordance with the DSM-5. PTSD symptoms were measured using the PTSD checklist for DSM-5 [19], which is a self-report questionnaire that assesses the presence of the four core DSM-5 PTSD symptom clusters during the past month with a cut-off score between 31 and 33, thus providing knowledge on the severity of symptoms.

Table 1 shows the demographic and clinical characteristics of the total sample of participants. Participants were aged between 20 and 59 years, and the majority were female (8/14, 57%). All participants fulfilled the diagnostic criteria for PTSD according to the International Classification of Diseases, Tenth Revision and had at least one comorbid psychiatric disorder. The most frequently reported comorbid disorder was panic disorder (6/14, 42%), followed by a major depressive episode (3/14, 21%) and agoraphobia (3/14, 21%), thus reflecting a population sample with complex mental health problems.

Textbox 1. Concepts covered in the interviews.

Concepts

- Use of apps in general
- PTSD Help app utility value
- PTSD Help usefulness
- Design
- Suggestions for change
- Disadvantages or negative experiences with using the app

XSL•FO
RenderX

Table 1. Demographic and clinical characteristics of all participants (N=14).

Characteristics	Value
Age (years), mean (SD; range)	39 (13.5; 20-59)
Gender, n (%)	
Female	8 (57)
Male	6 (43)
Diagnosis MINI[a] (ICD-10[b]), n (%)	
PTSD[c] (F43.1 in ICD-10)	14 (100)
PCL-5[d] score, mean (SD)	46.6 (9.3)
Comorbidity using MINI (ICD-10), n (%)	
Major depressive episode (F32.2 in ICD-10)	3 (21)
Major depressive disorder, recurrent (F33.2 in ICD-10)	1 (7)
Agoraphobia (F40.0 in ICD-10)	3 (21)
Social anxiety disorder (F40.1 in ICD-10)	2 (14)
Panic disorder (F41.0 in ICD-10)	6 (42)

[a]MINI: the Mini-International Neuropsychiatric Interview.

[b]ICD-10: International Classification of Diseases, Tenth Revision.

[c]PTSD: posttraumatic stress disorder.

[d]PCL-5: The Posttraumatic Stress Disorder Checklist for the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition.

Data Analysis

Semistructured interviews were conducted by 2 members of the research team (MEL and IMTPA) and a graduate student of psychology. Data were analyzed by first author LHGR and researcher II using thematic analysis under the supervision of an experienced qualitative researcher. Thematic analysis is a qualitative method that allows for the organization, interpretation, and reporting of data [22,23]. The analysis was conducted through 6 phases that adhered to Braun and Clarke's thematic analysis approach [22]. The researchers LHGR and II separately coded the interviews and conducted the thematic analysis. The Lincoln and Guba [24] criteria for trustworthiness were addressed in each phase to ensure the high quality of the analysis. Through researcher triangulation, overlaps were discovered and discussed, and 3 themes were chosen. Themes were chosen according to their relationship to the overall research question.

Results

Overview

Table 2 summarizes the most frequently used functions in the app and the least used functions. The data indicated that all the participants except one (13/14, 92.9%) had used the app and that the mean use was 16.8 (SD 12.3) times during the study period. The total app use ranged from 0 to 39 times, indicating that some participants frequently engaged in the app, whereas one participant never used it during the study period. The most frequently used function in the app was the "Symptoms and strategies" function with a mean of 5.9 (SD 4.4). The least used function was the "Sleep Condition Indicator" with 57.1% (8/14) participants using it with a mean of 1.0 (SD 1.1), followed by the function "Positive activites" with 64.3% (9/14) participants using it with a mean of 1.1 (SD 1.1).

The following themes emerged in the thematic analysis: (1) *use of app*, (2) *being a patient*, and (3) *overall evaluation of the app*. An overview of themes and subthemes is provided in Textbox 2 and presented using verbatim quotes from the participants in the following text.

XSL•FO
RenderX

Table 2. Total app use and the most and least frequently used functions.

App use	Number of times that functions were used, mean (SD; range)	Endorsed, n (%)
Total app use	16.8 (12.3; 0-39)	13 (92.9)
Most used functions		
Symptoms and strategies	5.9 (4.4; 0-15)	12 (85.7)
Crisis plan	5.7 (3.1; 0-13)	13 (92.9)
Help falling asleep	4.1 (5.6; 0-19)	10 (71.4)
Least used functions		
Body exercises	1.2 (0.9; 0-3)	11 (78.6)
Positive activities	1.1 (1.1; 0-3)	9 (64.3)
Sleep condition indicator	1.0 (1.1; 0-3)	8 (57.1)

Textbox 2. Themes and subthemes.

Themes and subthemes

- Use of App
  - Subtheme 1: Habits
- Being a Patient
  - Subtheme 2: Negative experiences with the app
  - Subtheme 3: Being a part of a research project
- Overall Evaluation of the App

Theme 1: Use of the App

Overview

The *use of the app* concerns psychological factors related to app use (such as anxiety, stress, and habits) and technical problems. The psychological factors had a dual function, as they could both motivate as well as disrupt the use of the app. For example, some participants reported using the app during panic attacks, whereas others reported that a high level of anxiety prevented them from using the app:

> *But when you are completely panic-struck, it's not like you think "now I'm going to use it," it's often like it's too late, so it is more when you are about to calm down that it is actually a help, right.* [Participant ID24]

Other participants described similar experiences, where even though they had used the app frequently, a high level of anxiety could prevent them from using the app. For these participants, the thought of using the app when in distress was not an option, as they simply did not remember that they had it in the exact moment:

> *So, do you feel that it has become a habit to use it? Understood in the sense that you remember to use it when you have a panic attack, for example?* [Interviewer]
>
> *No, not quite yet, but it depends a lot on how bad I'm feeling. Because if it's really bad, then it's like you don't really think about anything, then of course I don't think that I need to grab my phone.* [Participant ID21]

Some participants reported using the app to prevent stress, whereas others referred to using the app overall as a distraction tool, which was seen as a positive and useful function. A subgroup of participants experienced technical difficulties, such as missing sound on the sleep meditation or that the mobile phone automatically went to screensaver mode when they used meditation exercises, which caused some participants to stop using the app altogether.

Subtheme 1: Habits

The subtheme of *habits* emerged as participants described their use of the app to explain a low level of app use. One participant described how a high level of anxiety prevented the habitual use of the app, as high arousal interrupted regular use. Other participants reported that the novelty value when using the app in the beginning stimulated frequent use, but that this effect quickly vanished.

> *Well, of course in the beginning I used it to get to know it. And also customize it like in case of emergency, and I have also chosen my favourite functions. And I used it in the beginning, and actually got a little habit of using it a couple of times a day, but then it unfortunately went down the drain.* [Participant ID14]

XSL•FO
RenderX

Several participants reported either forgetting that they had the app or forgetting to use the app in relevant situations. A participant explained how he, in the beginning, tried to use the app as a distraction tool when he had a panic attack but that he quickly forgot he had the app:

> *You forget that you have it on your phone. Because it's not something you think about in everyday life, like "now I have the app" and then I get these feelings where I can feel all this anxiety, it's not that I go to my phone and use it, I think I just forget. That I have this tool available.* [Participant ID16]

Theme 2: Being a Patient

Overview

The theme refers to how the patient experienced having the PTSD Help app during the waiting time until psychotherapeutic treatment started. Within this theme, the subthemes, *negative experiences with the app* and *being a part of a research project*, emerged.

For some participants, the waiting time from when the diagnostic interviews were conducted to when the psychotherapeutic treatment started felt long. One participant stated that the waiting time was too long, and he used the app as a way to not feel completely lost while waiting. Thus, the app was experienced as a supportive tool for some participants during the waiting time.

The participants also discussed their experience of participating in the initial diagnostic interview in the MHS-CRD. For example, one participant explained how there was no time in the initial diagnostic interview to explain the diagnosis, which led the patient to seek this information in the app:

> *...but there has not really been a time when someone has sat down with me and explained what the symptoms are, what I should expect, how the process will be, and here I think the app has been good. Because I can google everything on the internet, but here I know that it comes from a source created by doctors and psychologists. So, it's not just something that's just lying around on the internet.* [Participant ID39]

On the other hand, some participants reported being overwhelmed by the status of being a patient. Having to use the app added to the stress. A participant explained how facing PTSD-related problems would make things worse, and that she did not feel able to engage in the app unsupported:

> *So, from the very first interview, it took almost a month before I had the next two appointments. And during that time, I shut down everything, and didn't want to relate to anything, or think about anything, and I pushed everything away, and when there were four-five days until I had to start, everything got worse up to the maximum again. I only slept two-three hours a night and woke up with anxiety and palpitation and...yes. So, the bottom line is that when I'm forced to deal with it, it gets worse.* [Participant ID29]

The participants who addressed the theme of being a patient all highlighted the waiting time as problematic and that information (either from the app or from a professional) about their diagnosis and the overall treatment process were helpful.

Subtheme 2: Negative Experiences With the App

Although there was an overall positive attitude toward the app in that participants reported becoming more aware of symptoms and sensations that increased their understanding of the disorder, negative experiences from using the app were also reported. Although participants generally saw becoming more aware of their symptoms as clinically beneficial, some experienced it as a worsening of symptoms and exposing vulnerability in a situation of minimal support. One example of a participant reporting using the app regularly reported feeling worse from the increased self-awareness that followed using the intervention tools in the app:

> *So, you may have become more aware of yourself if you can put it that way?* [Interviewer]
>
> *Yes exactly, and what to look for to why you have these symptoms, and what these symptoms mean and that it is not, you know, your own fault, if you can say so.* [Participant]
>
> *Have you noticed if any of your symptoms have changed since you got the app?* [Interviewer]
>
> *I do not think changed, I just think they have become...all in all, worse.* [Participant ID39]

The participants who reported negative experiences all agreed that focusing on their PTSD symptoms without support from a professional caused increased distress.

Subtheme 3: Being a Part of a Research Project

Some participants reported how being a part of a research project was stressful:

> *But what I would also like to say is that I have been to a psychologist before, and now I go to the psychiatric clinic in Hilleroed to find out what treatment I should have, and then at the same time, I should also use the app, and at the same time you get calls, and I get asked a lot of things about the app, so I just think...I'm well aware that you have some of these things, but it might be a little...too much, then I have to use the app, then I go to a psychologist, so it might also take up a little too much time in everyday life, I think...That you are constantly made aware that you have this problem.* [Participant ID16]

Another participant felt almost the opposite, as he explained how contact with the research team helped him relax as he knew he could reach out to the research team if he needed help.

Theme 3: Overall Evaluation of the App

This theme covers how the participants experienced using the app, the design of the app, and its specific functionalities. Within this theme, participants provided several suggestions for changes to the app.

Overall, most participants reported using the app regularly and were generally positive toward it. The design of the app was

XSL•FO
RenderX

reported to be accessible and easy to use, although some participants reported that a few of the functionalities, such as the location of the questionnaires, were not placed intuitively and were therefore confusing. The participants had several suggestions for improvements, such as implementing notifications as reminders, providing psychoeducation for relatives, and personalizing the app.

> *...you're able to personalise the app, where you can put some pictures in it that make you happy, or something like that, right. Because I have learned that on my computer, when it's in sleep mode, it shows some pictures, and for me it has actually been really relaxing to be able to sit there and watch these pictures and think "okay." So, something like that, I don't know...you just relax when seeing something of your own that you have chosen in advance, and maybe also some music that you might like or something like that.* [Participant ID24]

In particular, the suggestion to implement notifications in the app as a reminder for using the exercises, self-assessment, and other features was mentioned several times. There was a broad consensus across participants that the key advantages of the app were that it was a reliable source of information on PTSD and that it worked well as a distraction tool. A small number of participants reported being surprised at how effective some of the emotion regulation tools were (eg, the soothing images).

Discussion

Principal Findings

This qualitative study explored patients' experiences with using the PTSD Help app as a stand-alone treatment before commencing psychotherapeutic treatment. All the interviewed participants except one (13/14, 92.9%) had used at least some functions in the app with the range of use among the participants from 0 to 39 times during the study period. Thus, nearly all participants had experience from engaging with the app to report. The data also indicated a broad range of use among participants.

A total of 3 themes were identified in the analysis— *the use of the app*, *being a patient*, and *the overall evaluation of the app*. Overall, the participants felt comfortable using the app and experienced several functions in the app as helpful. However, the analysis found that a subgroup of participants reported negative experiences with the app. Finally, the wish for a more personalized app emerged.

Our findings showed that some participants reported that increased emotional arousal (eg, anxiety) could lead to both an increase and a decrease in app use depending on the situation. Thus, an optimal level of distress seemed to promote the use of the app and increase individuals' ability to use emotional regulation strategies when needed. One could speculate that to remember to use the app efficiently, at least some distress is necessary. This issue is important to note, as it raises the question of the existence of an optimal load of symptoms when using an app as a stand-alone treatment. Previous studies on the effectiveness of mobile apps have focused on symptom improvement rather than patterns of use [6]. Our study suggests that it is possible that patients who are less impaired by their symptoms will benefit the most and be the most engaged users of the app, whereas patients with more severe symptoms will, to a higher degree, refrain from using the app.

What is new is our finding that a subgroup of participants reported negative experiences with the app. Although this applies only to a limited number of participants, it must be taken seriously, as increased symptom severity in PTSD has been associated with increased suicidal risk [3,4]. This finding emphasizes the potential general challenge of app interventions as a stand-alone treatment. In contrast to psychotherapy, where the clinician's continuous evaluation of the patient, including risks, benefits, and goals, guides which intervention is used, mHealth apps offer the same functions to all patients, potentially leaving vulnerable patients unassisted in choosing and applying interventions [25]. Although the PTSD Help app offers a crisis management plan and strategies for coping with suicide risk, the use of these functions is dependent on the patient, who may have limited insight and, as such, risk not using the functions when needed or not being able to use the functions unassisted. Consequently, it could be argued that to be able to support this vulnerable subgroup of patients, a blended care intervention, including clinician support, is crucial.

Habits were raised as a subtheme in the way the participants used the app. As using the app is a prerequisite to obtain an effect, it is relevant to understand patterns of use and factors involved in the participants' use of the app. One perspective on patterns of use comes from research on habits and habit formation [26]. Developing habits is a dynamic process in which behavioral control is initially goal dependent but shifts to context dependence as the behavior is repeated [27]. In our study, the participants were responsible for defining their goal on their own and received limited contextual instructions in using the app, and the instructions were somewhat vague (eg, 'when you are feeling distressed'). The analysis found that participants reported how forgetfulness and the use of the app not becoming habitual caused unintentional nonadherence. Accordingly, a recent systematic review focusing on medication adherence across chronic medical conditions found that habit strength was strongly correlated with medication adherence [28]. Focusing on psychotherapy, a study on a guided self-help intervention using cognitive bias modification training found a significant treatment effect in patients with depression who reported having created a habitual use of the practiced self-help response [29]. These findings suggest that supporting habit formation through clinician-supported interventions on goal definition and contextual instructions appears to be important when implementing an mHealth app and could have had an impact on the frequency of use of the PTSD Help app.

In relation to the previous suggestion to use a blended care approach to support patients with potential suicidal thoughts in using the app, a blended care approach could also strengthen habit formation and thus potentially be of more help. This would also apply to vulnerable patients with an increased risk of suicide. If using the app successfully becomes a habitual behavior, it will automatically be activated in the relevant context and, therefore, not be dependent on available mental

XSL•FO
RenderX

resources, which may be limited in situations with high emotional arousal.

This study also revealed how the waiting time until treatment starts could feel long and stressful for the patient and result in the patient not engaging in the app, and clinician support could alleviate this. Previous studies have shown that including clinician support in internet-based treatments increases the effect size and patient adherence [30,31]. A study on a clinician-supported version of the PTSD Coach app found promising results in terms of acceptability and feasibility among patients and clinicians [32]. The procedure in the present study did not include clinician support, but it included a technical support call after 3 days, which some participants perceived as a clinical support call, which in turn might have encouraged participants to be more engaged in using the app after this call. Our findings suggest that providing participants with more contextual instructions and at least some clinician support on a regular basis may be important to make the use of the app habitual, which could prove pivotal in optimizing the effect of an app.

The participants' evaluation of the app was closely related to suggestions for change, in which a wish to be able to personalize the app was prominent. This request is consistent with previous qualitative studies on the internet and mobile-based interventions [33,34]. Although the app offers personalization of the content to some extent, this finding echoes the need for more personalized material and interventions. With the PTSD Help app, this could mean changing the focus from the individual choosing his or her own preferred functions in the app to the app recommending relevant interventions based on a data collection of the individual's specific distribution of symptoms. As the app already contains self-monitoring tools, in line with recommendations [9], it seems plausible to use this collected data in an automated tailoring algorithm. A variety of functions should be offered in the app to ensure that it holds content of relevance to every participant with a variety of individual needs and preferences.

Regarding age, all the participants were aged ≤60 years, making it difficult to generalize the results past that age. However, the majority of patients treated in MHS in Denmark are aged ≤60 years, as demonstrated in demographic data from a large sample of 2473 patients without a diagnosis of psychosis in MHS-CRD, where the average age was 33.0 years [35]. Thus, our results cannot be generalized to older adults but can mirror the average age seen in MHS in Denmark.

Limitations

Some limitations influence the strength of the conclusions drawn from the results of this study. The sample was based on a convenience sampling principle, because of a strict time limit for recruitment. The most optimal sampling strategy would have been a selected sample, which would have allowed the information from this study to be applied beyond its settings [36]. Furthermore, no participant in the study was older than 59 years, making it difficult to generalize our findings to older adults.

There is a possibility that there was a positive sampling bias, as those participants who agreed to participate generally had a positive attitude toward the app. Two participants were asked to participate but declined because of a reported negative attitude toward the app. One could speculate that this positive attitude is not necessarily the general attitude of all participants using the app. This could also explain how the data on the total app use showed that one participant did not engage in the app at all during the study period, but this was not revealed in the interviews.

The study covered participants' experiences of using the app and acknowledged that there are multiple factors, such as technological, psychological, and social factors, which affect the experience. These factors may not be fully covered by the semistructured interview in this study, as they are not always apparent to the participant or the participant may not wish to reveal them to a member of the research team.

Future Implications

In light of the current COVID-19 pandemic and the urgent need for developing internet-based treatment services that ensure treatment commitment and adherence, it is important to cover the patient's experience of using mHealth apps, as this will provide the research field with essential information about usability, acceptability, and possible negative effects. This qualitative study makes an important contribution to the field of mHealth research, showing that providing an mHealth app to patients with PTSD without clinician support should be done with caution. Our findings stress that personalization and active interactions could be beneficial to integrate in future mHealth apps to ensure continuous engagement and positive experiences with the treatment. On the basis of these findings, we provide two suggestions when implementing an mHealth app in an MHS with a PTSD population: (1) using a blended care treatment with clinician support to provide assistance and regular psychiatric assessment as well as to stimulate habitual use and (2) using an approach where mental health care professionals *prescribe* an mHealth app for relevant patients. Using one of these suggestions when implementing an mHealth app for patients with PTSD in MHS will enable clinicians to identify patients who are at risk of experiencing worsening symptoms from using the app unassisted. Future studies could investigate the association between use and outcome to determine if a dose-response mechanism exists and which functions patients seem to engage the most with signaling, which might be more important than others.

Conclusions

This qualitative study examined how patients with PTSD experienced the mHealth app PTSD Help as a stand-alone treatment before commencing psychotherapeutic treatment. The overall findings were that the majority of participants expressed that they had a positive experience using the app, but that a subgroup of participants reported negative experiences using the app. The lack of habitual use of the app was also raised to explain a low level of use and not using the app in relevant situations.

XSL•FO
RenderX

There are undoubtedly many advantages in implementing mHealth apps in MHS. Our findings demonstrate the need for research on which patients can benefit and which patients are at risk of experiencing worsening of symptoms from having an mHealth app as a stand-alone treatment. We hope that this study will serve as a springboard for future mHealth studies to form hypotheses about which patients in particular can benefit from using an mHealth app (either as a stand-alone treatment or as a supplement to treatment) and explore this in a mixed methods approach to provide new fine-grained insights into the research field. As a consequence of the participants' wish for a more personalized app, we also encourage further investigation of the possibility of developing such an app using a tailoring algorithm to recommend interventions to the user.

Acknowledgments

This study was supported by Council of the Danish Victims Fund under grant number 18-610-00018.

Authors' Contributions

LHGR, FBS, and MEL designed the study. MEL, IMTPA, and the graduate student conducted the interviews. LHGR and II performed the thematic analysis under the supervision of ABC. LHGR drafted the paper under the supervision of SBM, all authors read the final version of the manuscript and contributed to its revision.

Conflicts of Interest

MEL and FBS were involved in the development of the app.

References

1. Atwoli L, Stein DJ, Koenen KC, McLaughlin KA. Epidemiology of posttraumatic stress disorder: prevalence, correlates and consequences. Curr Opin Psychiatry 2015 Jul;28(4):307-311 [FREE Full text] [doi: 10.1097/YCO.0000000000000167] [Medline: 26001922]
2. Kessler RC, Sonnega A, Bromet E, Hughes M, Nelson CB. Posttraumatic stress disorder in the National Comorbidity Survey. Arch Gen Psychiatry 1995 Dec;52(12):1048-1060. [doi: 10.1001/archpsyc.1995.03950240066012] [Medline: 7492257]
3. Nock MK, Hwang I, Sampson N, Kessler RC, Angermeyer M, Beautrais A, et al. Cross-national analysis of the associations among mental disorders and suicidal behavior: findings from the WHO World Mental Health Surveys. PLoS Med 2009 Aug;6(8):e1000123 [FREE Full text] [doi: 10.1371/journal.pmed.1000123] [Medline: 19668361]
4. Gradus J, Qin P, Lincoln A, Miller M, Lawler E, Sørensen HT, et al. Posttraumatic stress disorder and completed suicide. Am J Epidemiol 2010 Mar 15;171(6):721-727. [doi: 10.1093/aje/kwp456] [Medline: 20160171]
5. Van Ameringen M, Turna J, Khalesi Z, Pullia K, Patterson B. There is an app for that! The current state of mobile applications (apps) for DSM-5 obsessive-compulsive disorder, posttraumatic stress disorder, anxiety and mood disorders. Depress Anxiety 2017 Jun;34(6):526-539. [doi: 10.1002/da.22657] [Medline: 28569409]
6. Kuhn E, Kanuri N, Hoffman J, Garvert D, Ruzek, Taylor CB. A randomized controlled trial of a smartphone app for posttraumatic stress disorder symptoms. J Consult Clin Psychol 2017 Mar;85(3):267-273. [doi: 10.1037/ccp0000163] [Medline: 28221061]
7. Internet use by individuals. Eurostat. URL: https://ec.europa.eu/eurostat/documents/2995521/7771139/9-20122016-BP-EN.pdf [accessed 2021-08-02]
8. Goreis A, Felnhofer A, Kafka J, Probst T, Kothgassner O. Efficacy of Self-Management Smartphone-Based Apps for Post-traumatic Stress Disorder Symptoms: A Systematic Review and Meta-Analysis. Front Neurosci 2020;14:3 [FREE Full text] [doi: 10.3389/fnins.2020.00003] [Medline: 32038153]
9. Bakker D, Kazantzis N, Rickwood D, Rickard N. Mental Health Smartphone Apps: Review and Evidence-Based Recommendations for Future Developments. JMIR Ment Health 2016 Mar;3(1):e7 [FREE Full text] [doi: 10.2196/mental.4984] [Medline: 26932350]
10. Erbes C, Stinson R, Kuhn E, Polusny M, Urban J, Hoffman J, Ruzek, et al. Access, Utilization, and Interest in mHealth Applications Among Veterans Receiving Outpatient Care for PTSD. Military Medicine 2014 Nov;179(11):1218-1222 [FREE Full text] [doi: 10.7205/milmed-d-14-00014]
11. Austin S, Jansen J, Petersen C, Jensen R, Simonsen E. Mobile App Integration Into Dialectical Behavior Therapy for Persons With Borderline Personality Disorder: Qualitative and Quantitative Study. JMIR Ment Health 2020 Jun 11;7(6):e14913 [FREE Full text] [doi: 10.2196/14913] [Medline: 32525488]
12. Price M, Yuen E, Goetter E, Herbert JD, Forman EM, Acierno R, et al. mHealth: a mechanism to deliver more accessible, more effective mental health care. Clin Psychol Psychother 2014;21(5):427-436 [FREE Full text] [doi: 10.1002/cpp.1855] [Medline: 23918764]

XSL•FO
RenderX

13. Scharff F, Lau M, Riisager L, Møller SB, Salimi ML, Gondan M, et al. The PTSD help app in a Danish PTSD population: research protocol of a randomized controlled feasibility trial. Pilot Feasibility Stud 2020;6:92 [FREE Full text] [doi: 10.1186/s40814-020-00633-x] [Medline: 32617173]
14. Miner A, Kuhn E, Hoffman JE, Owen JE, Ruzek JI, Taylor CB. Feasibility, acceptability, and potential efficacy of the PTSD Coach app: A pilot randomized controlled trial with community trauma survivors. Psychol Trauma 2016 May;8(3):384-392. [doi: 10.1037/tra0000092] [Medline: 27046668]
15. Faurholt-Jepsen M, Frost M, Christensen EM, Bardram JE, Vinberg M, Kessing LV. The effect of smartphone-based monitoring on illness activity in bipolar disorder: the MONARCA II randomized controlled single-blinded trial. Psychol Med 2020 Apr;50(5):838-848. [doi: 10.1017/S0033291719000710] [Medline: 30944054]
16. O'Toole M, Arendt M, Pedersen C. Testing an App-Assisted Treatment for Suicide Prevention in a Randomized Controlled Trial: Effects on Suicide Risk and Depression. Behav Ther 2019 Mar;50(2):421-429. [doi: 10.1016/j.beth.2018.07.007] [Medline: 30824256]
17. Connolly SL, Miller CJ, Koenig CJ, Zamora KA, Wright PB, Stanley RL, et al. Veterans' Attitudes Toward Smartphone App Use for Mental Health Care: Qualitative Study of Rurality and Age Differences. JMIR Mhealth Uhealth 2018 Aug 22;6(8):e10748 [FREE Full text] [doi: 10.2196/10748] [Medline: 30135050]
18. Cernvall M, Sveen J, Bergh Johannesson K, Arnberg F. A pilot study of user satisfaction and perceived helpfulness of the Swedish version of the mobile app PTSD Coach. Eur J Psychotraumatol 2018;9(Suppl 1):1472990 [FREE Full text] [doi: 10.1080/20008198.2018.1472990] [Medline: 29805783]
19. Blevins CA, Weathers FW, Davis MT, Witte TK, Domino JL. The Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5): Development and Initial Psychometric Evaluation. J Trauma Stress 2015 Dec;28(6):489-498. [doi: 10.1002/jts.22059] [Medline: 26606250]
20. Espie CA, Kyle SD, Hames P, Gardani M, Fleming L, Cape J. The Sleep Condition Indicator: a clinical screening tool to evaluate insomnia disorder. BMJ Open 2014 Mar 18;4(3):e004183 [FREE Full text] [doi: 10.1136/bmjopen-2013-004183] [Medline: 24643168]
21. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry 1998;59 Suppl 20:22-33;quiz 34. [Medline: 9881538]
22. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
23. Nowell L, Norris J, White D, Moules N. Thematic Analysis. International Journal of Qualitative Methods 2017 Oct 02;16(1):160940691773384 [FREE Full text] [doi: 10.1177/1609406917733847]
24. Lincoln YS, Guba EG, Pilotta JJ. Naturalistic inquiry. International Journal of Intercultural Relations 1985 Jan;9(4):438-439. [doi: 10.1016/0147-1767(85)90062-8]
25. Torous J, Roberts LW. The Ethical Use of Mobile Health Technology in Clinical Psychiatry. J Nerv Ment Dis 2017 Jan;205(1):4-8. [doi: 10.1097/NMD.0000000000000596] [Medline: 28005647]
26. Marien H, Custers R, Aarts H. Understanding the formation of human habits: an analysis of mechanisms of habitual behaviour. In: Verbanken B, editor. The Psychology of Habit: Theory, Mechanisms, Change, and Contexts. Cham: Springer; 2018.
27. Mazar A, Wood W. Defining habit in psychology. In: Verbanken B, editor. The Psychology of Habit: Theory, Mechanisms, Change, and Contexts. Cham: Springer; 2018:13-29.
28. Badawy SM, Shah R, Beg U, Heneghan MB. Habit Strength, Medication Adherence, and Habit-Based Mobile Health Interventions Across Chronic Medical Conditions: Systematic Review. J Med Internet Res 2020 Apr 28;22(4):e17883 [FREE Full text] [doi: 10.2196/17883] [Medline: 32343250]
29. Watkins ER, Taylor RS, Byng R, Baeyens C, Read R, Pearson K, et al. Guided self-help concreteness training as an intervention for major depression in primary care: a Phase II randomized controlled trial. Psychol Med 2012 Jul;42(7):1359-1371 [FREE Full text] [doi: 10.1017/S0033291711002480] [Medline: 22085757]
30. Cuijpers P, Donker T, van Straten A, Li J, Andersson G. Is guided self-help as effective as face-to-face psychotherapy for depression and anxiety disorders? A systematic review and meta-analysis of comparative outcome studies. Psychol Med 2010 Dec;40(12):1943-1957. [doi: 10.1017/S0033291710000772] [Medline: 20406528]
31. Pedersen MK, Mohammadi R, Mathiasen K, Elmose M. Internet-based cognitive behavioral therapy for anxiety in an outpatient specialized care setting: A qualitative study of the patients' experience of the therapy. Scand J Psychol 2020 Dec;61(6):846-854. [doi: 10.1111/sjop.12665] [Medline: 32705703]
32. Possemato K, Kuhn E, Johnson E, Hoffman J, Brooks E. Development and refinement of a clinician intervention to facilitate primary care patient use of the PTSD Coach app. Transl Behav Med 2017 Mar;7(1):116-126 [FREE Full text] [doi: 10.1007/s13142-016-0393-9] [Medline: 27234150]
33. O'Mahen H, Grieve H, Jones J, McGinley J, Woodford J, Wilkinson E. Women's experiences of factors affecting treatment engagement and adherence in internet delivered Behavioural Activation for Postnatal Depression. Internet Interventions 2015 Mar;2(1):84-90 [FREE Full text] [doi: 10.1016/j.invent.2014.11.003]

XSL•FO
RenderX

34. Fleischmann R, Harrer M, Zarski A, Baumeister H, Lehr D, Ebert D. Patients' experiences in a guided Internet- and App-based stress intervention for college students: A qualitative study. Internet Interv 2018 Jun;12:130-140 [FREE Full text] [doi: 10.1016/j.invent.2017.12.001] [Medline: 30135777]
35. Fenger M, Mortensen EL, Poulsen S, Lau M. No-shows, drop-outs and completers in psychotherapeutic treatment: demographic and clinical predictors in a large sample of non-psychotic patients. Nord J Psychiatry 2011 Jun;65(3):183-191. [doi: 10.3109/08039488.2010.515687] [Medline: 20854221]
36. Malterud K. Qualitative research: standards, challenges, and guidelines. Lancet 2001 Aug 11;358(9280):483-488. [doi: 10.1016/S0140-6736(01)05627-6] [Medline: 11513933]

Abbreviations

DSM-5: The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
mHealth: mobile health
MHS: mental health services
MHS-CRD: Mental Health Services of the Capital Region of Denmark
PTSD: posttraumatic stress disorder

Edited by G Eysenbach; submitted 30.12.20; peer-reviewed by S Austin, T Alessa; comments to author 26.01.21; revised version received 23.02.21; accepted 16.05.21; published 04.08.21.

Please cite as:
Riisager LHG, Christensen AB, Scharff FB, Arendt IMTP, Ismail I, Lau ME, Moeller SB
Patients' Experiences of Using a Self-help App for Posttraumatic Stress Disorder: Qualitative Study
JMIR Form Res 2021;5(8):e26852
URL: https://formative.jmir.org/2021/8/e26852
doi:10.2196/26852
PMID:34346896

©Lisa H G Riisager, Anne Bryde Christensen, Frederik Bernt Scharff, Ida-Marie T P Arendt, Israa Ismail, Marianne Engelbrecht Lau, Stine Bjerrum Moeller. Originally published in JMIR Formative Research (https://formative.jmir.org), 04.08.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Physiologic Response to the Pfizer-BioNTech COVID-19 Vaccine Measured Using Wearable Devices: Prospective Observational Study

Alexander G Hajduczok[1*], MD; Kara M DiJoseph[1*], DO; Brinnae Bent[2], PhD; Audrey K Thorp[1], DO; Jon B Mullholand[1], MD; Stuart A MacKay[1], DO; Sabrina Barik[1], MD; Jamie J Coleman[3], MD; Catharine I Paules[4], MD; Andrew Tinsley[5], MD

[1]Division of Internal Medicine, Department of Medicine, Penn State Health Milton S. Hershey Medical Center, Hershey, PA, United States
[2]Department of Biomedical Engineering, Duke University, Durham, NC, United States
[3]Department of Surgery, Denver Health Medical Center, Denver, CO, United States
[4]Division of Infectious Diseases, Department of Medicine, Penn State Health Milton S. Hershey Medical Center, Hershey, PA, United States
[5]Division of Gastroenterology, Department of Medicine, Penn State Health Milton S. Hershey Medical Center, Hershey, PA, United States
[*]these authors contributed equally

Corresponding Author:
Alexander G Hajduczok, MD
Division of Internal Medicine
Department of Medicine
Penn State Health Milton S. Hershey Medical Center
500 University Drive
Hershey, PA, 17033
United States
Phone: 1 7168634928
Email: alexander.hajduczok@jefferson.edu

Abstract

Background: The Pfizer-BioNTech COVID-19 vaccine uses a novel messenger RNA technology to elicit a protective immune response. Short-term physiologic responses to the vaccine have not been studied using wearable devices.

Objective: We aim to characterize physiologic changes in response to COVID-19 vaccination in a small cohort of participants using a wearable device (WHOOP Strap 3.0). This is a proof of concept for using consumer-grade wearable devices to monitor response to COVID-19 vaccines.

Methods: In this prospective observational study, physiologic data from 19 internal medicine residents at a single institution that received both doses of the Pfizer-BioNTech COVID-19 vaccine was collected using the WHOOP Strap 3.0. The primary outcomes were percent change from baseline in heart rate variability (HRV), resting heart rate (RHR), and respiratory rate (RR). Secondary outcomes were percent change from baseline in total, rapid eye movement, and deep sleep. Exploratory outcomes included local and systemic reactogenicity following each dose and prophylactic analgesic use.

Results: In 19 individuals (mean age 28.8, SD 2.2 years; n=10, 53% female), HRV was decreased on day 1 following administration of the first vaccine dose (mean –13.44%, SD 13.62%) and second vaccine dose (mean –9.25%, SD 22.6%). RHR and RR showed no change from baseline after either vaccine dose. Sleep duration was increased up to 4 days post vaccination, after an initial decrease on day 1. Increased sleep duration prior to vaccination was associated with a greater change in HRV. Local and systemic reactogenicity was more severe after dose two.

Conclusions: This is the first observational study of the physiologic response to any of the novel COVID-19 vaccines as measured using wearable devices. Using this relatively small healthy cohort, we provide evidence that HRV decreases in response to both vaccine doses, with no significant changes in RHR or RR. Sleep duration initially decreased following each dose with a subsequent increase thereafter. Future studies with a larger sample size and comparison to other inflammatory and immune biomarkers such as antibody response will be needed to determine the true utility of this type of continuous wearable monitoring in regards to vaccine responses. Our data raises the possibility that increased sleep prior to vaccination may impact physiologic

XSL•FO
RenderX

responses and may be a modifiable way to increase vaccine response. These results may inform future studies using wearables for monitoring vaccine responses.

Trial Registration: ClinicalTrials.gov NCT04304703; https://www.clinicaltrials.gov/ct2/show/NCT04304703

(JMIR Form Res 2021;5(8):e28568) doi:10.2196/28568

KEYWORDS

COVID-19; wearable devices; remote physiologic monitoring; heart rate; heart rate variability; respiratory rate; sleep; REM sleep; deep sleep; wearable; vaccine; monitoring; respiratory; physiological; cohort

Introduction

The COVID-19 pandemic has had a substantial global impact resulting in over 165 million infections and nearly 3.5 million deaths worldwide [1,2]. Vaccines are required to end the pandemic. The first vaccine to receive emergency use authorization for prevention of COVID-19 infection by the United States Food and Drug Administration was the BNT162b2 messenger RNA (mRNA) COVID-19 vaccine (Pfizer-BioNTech COVID vaccine) that encodes the spike protein of the SARS-CoV-2 virus [3,4]. Following preliminary studies with this mRNA vaccine showing neutralizing antibody response, a phase 3 randomized clinical trial demonstrated that the Pfizer-BioNTech vaccine was safe and had an efficacy of 95% in reducing risk of contracting COVID-19 compared to placebo [4-6].

An estimated 21% of US adults report using wearable devices that objectively measure physiologic parameters [7]. Although marketed for personal use, the widespread nature and convenience of these devices allows health care professionals to monitor physiologic changes in real time [8]. The WHOOP Strap 3.0 has been externally validated for tracking of heart rate variability (HRV), resting heart rate (RHR), respiratory rate (RR), and sleep stage duration [9]. HRV is determined by the subtle variation in the time between successive heart beats, thus HRV is a measure of the balance between the sympathetic and parasympathetic nervous system and their composite effects on heart rate [10].

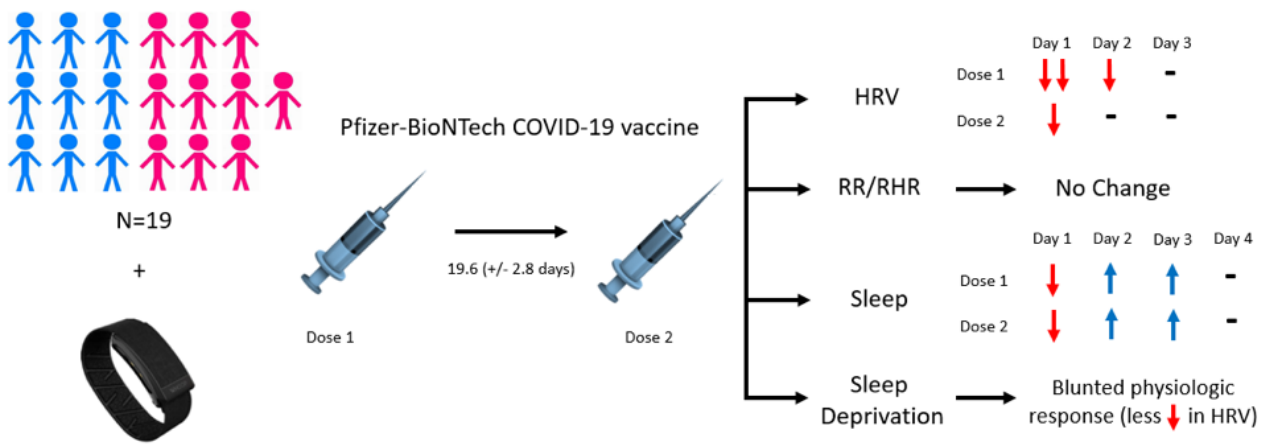
A recent study using the WHOOP device was able to track physiologic changes, specifically an increase in nocturnal RR and decrease in HRV, in individuals who reported COVID-19 infection. These changes were noted 2 days before symptom onset in 20% of individuals and in 80% of the cohort after symptom onset [11]. Other studies have used HRV and RR measured by wearable devices to prospectively and retrospectively predict and identify COVID-19 infection (confirmed by positive testing) [12-14]. Therefore, we postulated that it would be possible to track an array of physiologic responses following COVID-19 vaccination. This wearable remote monitoring strategy could serve as a proof of concept and guide design of future studies of the physiologic and immune responses to vaccines.

Methods

Study Population

Internal medicine residents at Penn State Hershey Medical Center previously enrolled in a clinical trial (NCT04304703) using the WHOOP Strap 3.0 to measure physiologic parameters were used for this analysis (Multimedia Appendix 1, Figure 1) [15].

Figure 1. A total of 19 participants, 53% female, who were vaccinated with two doses of the Pfizer-BioNtech COVID-19 vaccine (mean time between doses 19.6, SD 2.8 days), transmitted continuous physiologic data via the WHOOP device. Changes from baseline were observed in HRV and were most pronounced on day 1 and 2 for dose 1 and only day 1 for dose 2. RR and RHR were unaffected following vaccination. Sleep duration initially decreased on day 1 post vaccine dose 1 and dose 2, with a compensatory increase from days 2 to 4, prior to return to baseline. Sleep deprivation was associated with a blunted HRV response, and premedication was associated with less change in RR and increases in REM and deep sleep percentages. HRV: heart rate variability; REM: rapid eye movement; RHR: resting heart rate; RR: respiratory rate.



XSL•FO
RenderX

Study Design

The primary objective of this prospective observational study was to determine the physiologic changes following the first and second doses of the Pfizer-BioNTech COVID-19 vaccine. Primary outcomes were percent change from baseline in HRV, RR, and RHR for days 1 to 6 following each vaccine dose. Secondary outcomes were percent change from baseline duration of total, rapid eye movement (REM), and deep sleep. Exploratory outcomes included analysis of local and systemic reactogenicity (type and duration) associated with vaccination and prophylactic analgesic use.

Study Procedures

Internal medicine residents were given a WHOOP Strap 3.0 to wear to measure physiologic parameters and sleep [15] (for full details, see Multimedia Appendix 2 [9,16-19]). Eligible participants were surveyed to disclose their vaccination dates for the novel Pfizer-BioNTech COVID-19 vaccine along with local and systemic reactogenicity and analgesic use following each vaccine dose.

Inclusion criteria for this analysis were patients concurrently enrolled in a clinical trial (NCT04304703) using the WHOOP device and who transmitted at least 80% of physiologic data during the study period including at least 24 of 45 days prior to dose one (to establish baseline metrics) and all data for the 6 days following vaccine dose one, dose two, or both. These data cutoffs were chosen based on published data using the WHOOP device for establishing a change from baseline in RR [11]. Local and systemic reactogenicity was graded as mild, moderate, or severe based on guidelines from the Centers for Disease Control and Prevention [1,20]. Patients were excluded if they did not or were unable to disclose the dates of vaccination (Multimedia Appendix 3). Data were blinded to study investigators for analysis. Recorded demographics included age, gender, comorbidities, and year of residency training.

Data collection was approved by the Institutional Board Review at Penn State Hershey Medical Center (STUDY14522).

Statistical Analysis

We defined a significant change from baseline to be greater than 5% a priori. This cutoff was set based on recent findings in two studies: (1) changes in RR and other physiologic parameters in COVID-19–positive individuals, which were used to develop a predictive algorithm for COVID-19 infection risk stratification [11], and (2) precision measurements of heart rate, RR, HRV, and REM sleep stage duration using the WHOOP device were found to have less than 10% error [9].

The percent change of each metric for each participant in the data set was averaged together for the overall percent change of that metric for each day (d; equation 1). In equation 1, b_n is given as the average of the metric from the baseline period for participant n, and x_n is the value of the metric on the day (d) being calculated post vaccine dose.

To determine the effect of sleep for the week leading up to the vaccine on the physiological effects of the vaccine, we computed Pearson correlations between hours of sleep in the 7 days prior to vaccine dose 1 and the percent changes of the physiological measurements post vaccine dose 1 [21].

Symptoms were aggregated and the density of the self-reported duration of symptoms was calculated [22]. Postvaccination reaction severity was compared to changes in physiologic parameters by Pearson correlations.

Results

Baseline Characteristics

A total of 19 participants met inclusion and exclusion criteria for this analysis; 18 individuals for dose 1 and 13 for dose 2 (Multimedia Appendix 3). Participants were 53% (n=10) female, with an age range of 26 to 35 years and a mean and median age of 28.8 (SD 2.2) and 29 years, respectively. No comorbidities were reported in 74% (n=14) of participants (Multimedia Appendix 1).

Baseline metrics were collected for all participants up to 45 days prior to vaccination dose 1 (Table 1). Mean baselines were as follows: RHR 63.09 (SD 6.36) bpm, HRV 52.09 (SD 21.58) ms, and RR 16.27 (SD 1.23) respirations per minute (rpm). Although interindividual variability in metrics had a wider range, intraindividual variability was much lower, most notably in nocturnal RR, with an intraindividual SD of mean 0.37 (SD 0.12) rpm.

Table 1. Baseline physiological and sleep metrics intraindividual mean and SD.

Metric	Intraindividual mean, mean (SD)	Intraindividual SD, mean (SD)
Heart rate variability (ms)	52.09 (21.58)	13.16 (6.94)
Resting heart rate (bpm)	63.09 (6.36)	4.95 (1.50)
Respiratory rate (rpm)	16.27 (1.23)	0.37 (0.12)
Sleep (hours)	6.71 (0.58)	1.48 (0.39)
REM[a] sleep (%)	21.99 (6.71)	6.49 (1.23)
Deep sleep (%)	19.10 (2.15)	3.81 (0.66)

[a]REM: rapid eye movement.

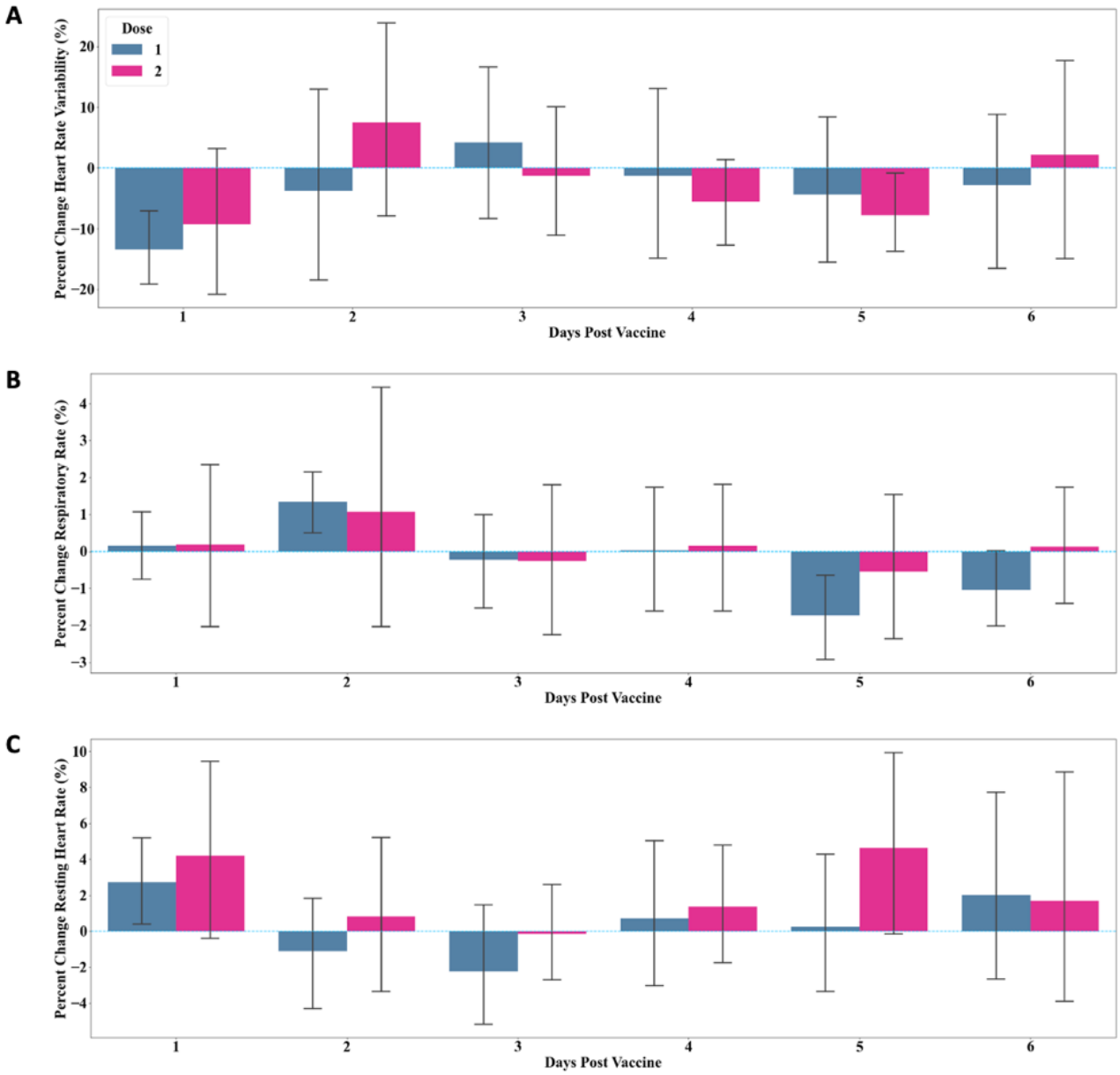
XSL•FO
RenderX

Physiologic Response to COVID-19 Vaccination by Dose

For dose 1 (n=18), there was a reduction in HRV on day 1 (mean percent change –13.44%, SD 13.62%). HRV returned to baseline by day 3 and remained at baseline thereafter (Figure 2A, blue; Table 2). There was no significant change in RHR and RR compared to baseline in the 6 days following vaccination (Figure 2B, C, blue; Table 2).

For dose 2 (n=13), HRV decreased on day 1 (mean percent change –9.25%, SD 22.69%) but quickly normalized to baseline by day 2 (Figure 2A, magenta; Table 3). Similar to dose 1, there was no significant change in RHR and RR in response to dose 2, with both metrics remaining at baseline from day 1 to day 6 (Table 3).

Figure 2. Percent change from baseline in (A) heart rate variability, (B) respiratory rate, and (C) resting heart rate, measured 6 days following COVID-19 vaccine dose 1 (blue) and 2 (magenta). Data is reported as mean (SD).



XSL•FO
RenderX

Table 2. Percent changes from baseline in physiological and sleep metrics for 6 days postvaccine dose 1 (n=18).

Metrics	Day 1, mean percent change (SD)	Day 2, mean percent change (SD)	Day 3, mean percent change (SD)	Day 4, mean percent change (SD)	Day 5, mean percent change (SD)	Day 6, mean percent change (SD)
Heart rate variability	–13.44 (13.62)	–3.74 (34.63)	4.21 (27.23)	–1.32 (30.39)	–4.35 (26.79)	–2.80 (27.46)
Resting heart rate	2.73 (5.50)	–1.10 (6.93)	–2.23 (7.31)	0.72 (8.80)	0.26 (8.68)	2.02 (11.48)
Respiratory rate	0.16 (1.95)	1.34 (1.98)	–0.23 (2.70)	0.02 (3.81)	–1.73 (2.56)	–1.04 (2.29)
Hours of sleep	–8.41 (22.96)	5.00 (18.27)	9.41 (21.60)	7.74 (17.81)	3.21 (24.38)	–3.21 (27.54)
Percent of REM[a] sleep	–4.94 (37.65)	–6.53 (30.06)	–5.13 (33.34)	–6.70 (19.62)	–2.70 (31.10)	–16.98 (29.41)
Percent of deep sleep	9.64 (26.30)	3.08 (23.00)	4.11 (12.66)	4.58 (15.28)	–6.05 (19.58)	–2.38 (17.53)

[a]REM: rapid eye movement.

Table 3. Percent changes from baseline in physiological and sleep metrics for 6 days postvaccine dose 2 (n=13).

Metrics	Day 1, mean percent change (SD)	Day 2, mean percent change (SD)	Day 3, mean percent change (SD)	Day 4, mean percent change (SD)	Day 5, mean percent change (SD)	Day 6, mean percent change (SD)
Heart rate variability	–9.25 (22.69)	7.48 (32.44)	–1.30 (19.44)	–5.56 (13.70)	–7.76 (13.12)	2.19 (30.22)
Resting heart rate	4.20 (9.42)	0.82 (8.27)	–0.15 (5.15)	1.37 (6.17)	4.63 (10.38)	1.70 (12.83)
Respiratory rate	0.19 (4.10)	1.07 (6.44)	–0.26 (4.00)	0.15 (3.22)	–0.54 (3.63)	0.13 (3.06)
Hours of sleep	–2.10 (26.18)	5.33 (17.71)	6.06 (23.84)	9.22 (28.37)	–4.58 (20.45)	0.49 (12.30)
Percent of REM[a] sleep	13.79 (45.88)	–8.73 (39.57)	12.67 (36.41)	0.64 (37.99)	6.12 (32.02)	–8.35 (43.06)
Percent of deep sleep	4.00 (25.02)	–1.70 (19.21)	3.56 (20.31)	–6.01 (22.09)	3.42 (21.31)	–11.37 (24.40)

[a]REM: rapid eye movement.

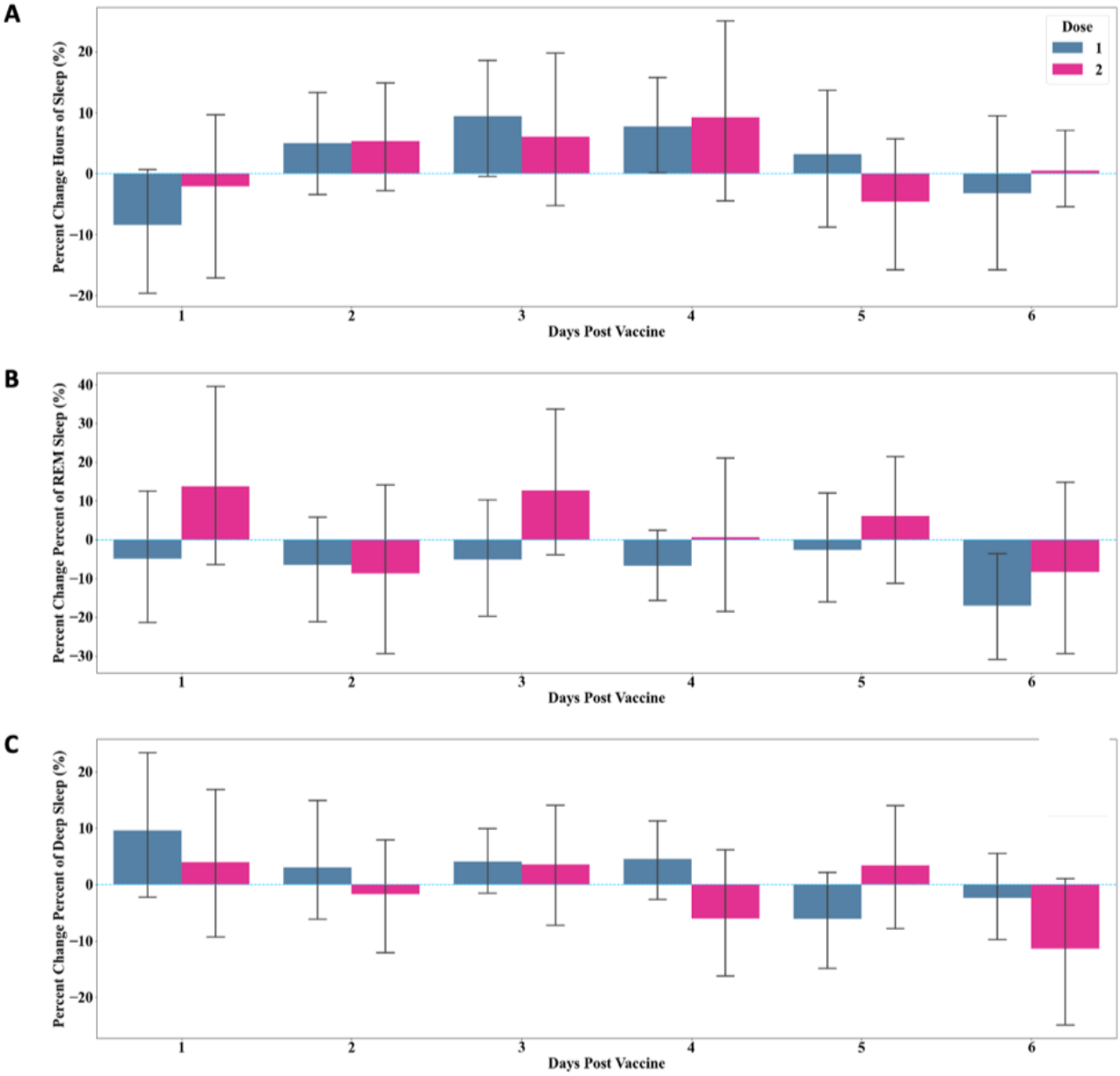
Postvaccination Changes in Sleep

Total sleep duration, REM, and deep sleep duration (in hours) were measured for all participants for 6 days following vaccine administration. Total sleep duration followed a similar overall pattern for both vaccine doses: an initial decrease was observed on day 1 (dose 1: mean –8.41%, SD 22.96%; dose 2: mean –2.1%, SD 26.8%) followed by an increase above baseline on days 2, 3, and 4, with subsequent return to baseline on days 5 to 6 (Figure 3A; Tables 2 and 3). The change in sleep duration peaked on day 3 following dose 1 and day 4 following dose 2. Total sleep duration was proportional to time in bed and thus showed similar trends in response to vaccine dose 1 and dose 2.

Patterns of change in REM and deep sleep did not follow the same pattern as total sleep duration but showed greater variability overall (Figure 3B, C; Tables 2 and 3). Total sleep cycles and sleep disturbances had no correlation with changes in physiologic metrics following either vaccine dose.

XSL•FO
RenderX

Figure 3. Percent change from baseline in (A) total sleep duration, (B) REM sleep duration, and (C) deep sleep duration, measured 6 days following COVID-19 vaccine dose 1 (blue) and 2 (magenta). Data is reported as mean (SD). REM: rapid eye movement.



Sleep Impact on HRV Changes

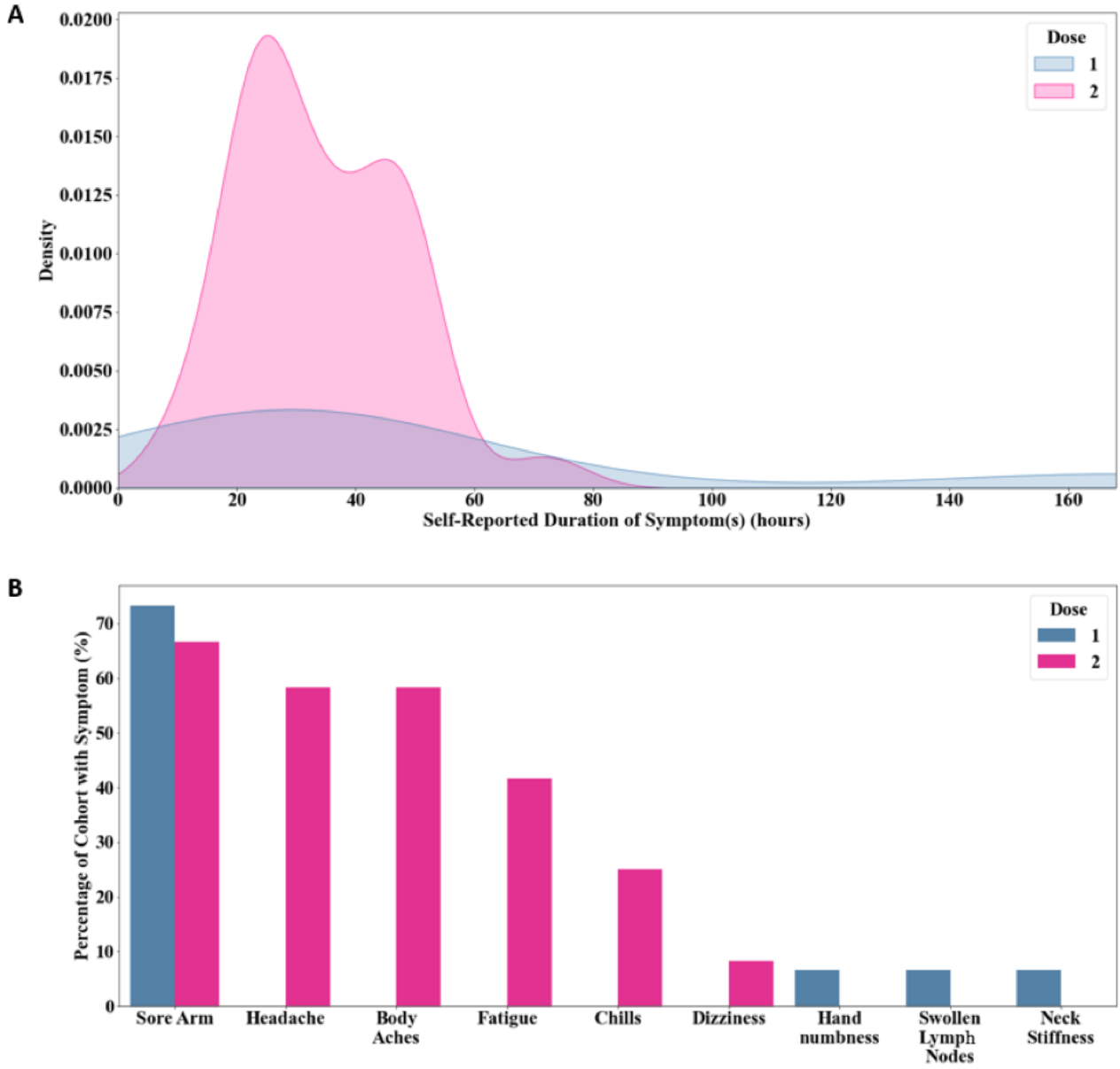
Sleep duration was evaluated 7 days preceding vaccine administration to establish an individualized baseline. During the baseline assessment period, participants slept, on average, 6 hours and 43 (SD 35) minutes per night, of which 21.99% (1 hour and 28 minutes) was REM sleep and 19.1% (1 hour and 17 minutes) was deep sleep. A greater amount of sleep in the 7 days prior to receiving the first dose of the vaccine was moderately correlated with a higher percent change in HRV the 2 days following vaccine dose 1 (Pearson R=0.570 day 1; R=0.494 day 2).

Postvaccination Symptoms

An array of local and systemic reactions to vaccination were reported by participants, ranging from arm soreness to fatigue and body aches. A greater frequency and duration of symptoms were reported following dose 2 (Figure 4A, B). Arm soreness was reported in more than 60% of participants for both doses. The majority of symptoms subsided by hour 60 post vaccination (Figure 4A). The mean symptom duration following dose 1 was 49.7 (SD 49.2) hours, which decreased to 34.1 (SD 13.3) hours for dose 2. The most frequent symptom duration after dose 1 and dose 2 was 24 hours. Overall, postvaccination reactogenicity would be classified as mild to moderate, as no severe adverse reactions such as angioedema or other allergic reactions requiring urgent treatment were reported [1,20]. Presence of postvaccination reactogenicity did not show a strong correlation with changes in sleep or other physiologic parameters.

XSL•FO
RenderX

Figure 4. (A) Self-reported symptom duration following dose 1 and dose 2 of the COVID-19 vaccine. (B) Local and systemic reactogenicity experienced by participants by vaccine dose.



Analgesic Effects on HRV, Sleep, and Postvaccination Symptoms

None of the 19 participants premedicated with analgesic medications (ibuprofen or acetaminophen) prior to dose 1; however, 7 of the 13 (54%) participants premedicated prior to dose 2 (Multimedia Appendices 4 and 5). Overall changes in HRV were the same in both groups (premedication vs no premedication; Multimedia Appendix 5). Those who did not premedicate had a greater response (increase) in RR on day 1 and day 2, but overall RR was unaffected when both groups were analyzed together (Multimedia Appendix 4). RHR had a slightly greater increase on day 1 for those who did premedicate (Multimedia Appendix 4). The group that premedicated had both a greater initial decrease and compensatory increase in total sleep duration (Multimedia Appendix 5). This group also had higher percentage of REM and deep sleep in the days after receiving dose 2, which were most prominent on day 1 (Multimedia Appendix 5).

The duration of all reported symptoms between the groups were similar: participants without medication experienced symptoms for a mean of 30.0 (SD 13.4) hours, and participants who self-medicated prior to dose 2 experienced symptoms for 37.7 (SD 10.0) hours. There was no significant difference in symptom severity among the two groups [1,20].

Discussion

Principal Findings

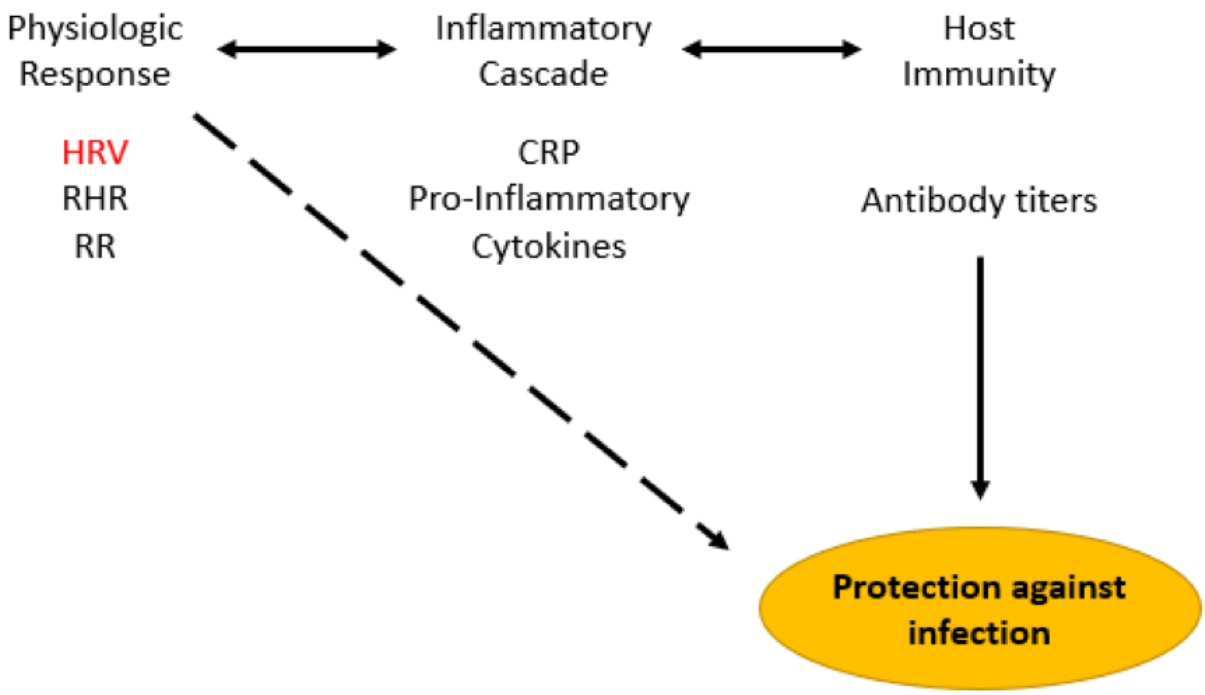
In this small observational study in a relatively young and healthy cohort of participants, we provide evidence that consumer-grade wearable devices can be used to measure physiologic response to COVID-19 vaccination (Figure 1). HRV change from baseline was the most prominent signal in our

XSL•FO
RenderX

study population, while RHR and RR were unaffected (Figure 2).

Decreases in HRV, a surrogate of autonomic tone, have been shown to predict negative clinical outcomes following severe infections [23-25]. HRV decreases have also been correlated with an increased C-reactive protein (CRP) within the first 2 days following administration of the influenza A vaccine [26]. Lower magnitude CRP elevations have been associated with increased risk of infection, suggesting that greater HRV decreases and CRP increases would equate to a protective inflammatory or immunologic response [26,27]. Thus, we propose that a higher percent change in HRV may equate to a more robust immune response to COVID-19 vaccination [14,26,28-30]. The directionality of HRV change (decrease) is significant, as this is suggestive of increased parasympathetic tone due to generation of an immune response to the vaccine [14,28]. Further investigation of HRV response to vaccination could provide a useful surrogate marker for immune system activation (Figure 5) [31]. This could be accomplished in a randomized controlled trial (RCT) of vaccination versus placebo with wearable tracking of HRV in comparison to CRP levels, antibody titers, and protection against infection.

Figure 5. Hypothetical connection between physiologic response measured by wearables (HRV, RHR, RR), inflammatory response (serum CRP and proinflammatory cytokine levels), and host immunity dictated by antibody response to vaccination. Wearable monitoring of physiologic metrics could potentially be a simple and effective way to track the efficacy of vaccine-mediated protection against infection (dashed arrow). HRV highlighted as this parameter showed the greatest changes in this study. CRP: C-reactive protein; HRV: heart rate variability; RHR: resting heart rate; RR: respiratory rate.



There was relatively no change in both RR and RHR in response to vaccination (Figure 2B, C). This is of particular interest given that spikes in RR are clinically relevant in prediction of COVID-19 infection and progression of disease [1,11-13]. Although we did not directly collect data on oxygen saturation or hypoxia, it is likely that changes in RR are specific to COVID-19 infection, which has a predilection for pulmonary pathology. Thus, this change would not be expected with vaccination as, unlike natural disease, it does not impact pulmonary function.

Interestingly, there was a moderate correlation between change in HRV and amount of sleep prior to vaccination (greater sleep was associated with a greater decrease in HRV). Sleep deprivation is known to have a significant impact on viral susceptibility and blunted adaptive immune response in the presence of viral vaccines [32-38]. Decreased antibody titers and overall immune response have been observed in vaccinated participants that are sleep deprived, most notably in response to the influenza and hepatitis A vaccine [34-39]. A recent study of 2884 health care workers showed that 1-hour longer sleep duration was associated with 12% lower odds of COVID-19 infection [40]. Our data demonstrate that sleep duration impacts physiologic response to COVID-19 vaccination and, if correlated with immune response in further studies, could be leveraged to potentiate the effectiveness of vaccination in general.

Limitations and Future Studies

This is a small observational study in a specific cohort of participants with no control arm. A larger powered study is needed for formal statistical analysis of physiologic changes and to control for baseline demographics. The lack of a control group institutes bias; an RCT with a placebo arm (no vaccine) would allow for comparison of physiologic metrics to a control group. However, our use of *percent change from baseline* in our outcomes would help overcome intraindividual variability confounding of results (ie, individuals may have a greater magnitude change in parameters simply because they have a higher baseline, which is accounted for by using percent change

XSL•FO
RenderX

from established baseline). Last, this population is known to have a greater degree of sleep deprivation secondary to duty hours and clinical demands, which may be a confounder and reduce the generalizability of the results [41].

Incorporation of biomarkers such as CRP is needed to corroborate association with physiologic changes (Figure 5) as previously observed in other vaccine studies [26,29,42,43]. Despite only exploring the response to the Pfizer-BioNTech COVID-19 vaccine, this study further confirms the feasibility of using wearable remote physiologic data to monitor responses to vaccines. This simple method to track vaccine responses would be useful for future novel vaccines. The key will be to determine how well remotely monitored physiologic metrics can predict inflammatory response, vaccination antibody titers, and ultimately protection from infection. This may provide a noninvasive method for individualized prediction of vaccine efficacy.

Conclusion

Wearable devices are now widely available to everyday consumers, and as technology has advanced, they are being more widely used to capture medical data [7,8,12,13,44]. This study is a proof of concept for this remote monitoring strategy to capture physiologic response to COVID-19 vaccines. If correlated with immune response and vaccine efficacy in future studies, this approach could be leveraged in the general population to predict response to vaccines. Our data also raises the possibility that increased sleep prior to vaccination may impact physiologic responses. This warrants further study and is a potentially modifiable factor to optimize vaccine response.

Acknowledgments

This study was supported by the 2020-2021 Penn State Hershey Department of Medicine House Staff Award granted to AGH and the 2020 Penn State Hershey Health Systems Science Innovation Grant awarded to AGH.

Authors' Contributions

AGH and KMD contributed equally to this study. AGH, KMD, and BB designed the research study and refined the data collection plan; BB analyzed the data with input from AGH and KMD; AGH, KMD, and BB wrote the manuscript; AKT, JBM, SB, SAM, JJC, CIP, and AT edited the manuscript.

Conflicts of Interest

CIP is a consultant for Axle Informatics for work not related to this manuscript. CIP is also the site PI for the Adaptive COVID-19 Treatment Trial (ACTT) and the ACTIV-5 / Big Effect Trial (BET-B) for the Treatment of COVID-19 which receive funding from the National Institutes of Health. This work is not related to the submitted manuscript.

Multimedia Appendix 1
Baseline characteristics of study participants.
[DOCX File , 13 KB - formative_v5i8e28568_app1.docx]

Multimedia Appendix 2
Description of physiologic metrics.
[DOCX File , 112 KB - formative_v5i8e28568_app2.docx]

Multimedia Appendix 3
Study flowchart. A total of 38 participants were initially enrolled in the study; 33 were screened for inclusion and exclusion criteria, yielding 19 participants that were included for final analysis (18 participants for dose 1 and 13 participants for dose 2).
[PNG File , 102 KB - formative_v5i8e28568_app3.png]

Multimedia Appendix 4
Premedication subgroup analysis. Percent change from baseline in (A) heart rate variability, (B) respiratory rate, and (C) resting heart rate, measured 6 days following COVID-19 vaccine dose 2 is reported for participants who premedicated prior to vaccination (purple) and participants who did not premedicate (green). Data is reported as mean (SD).
[PNG File , 133 KB - formative_v5i8e28568_app4.png]

Multimedia Appendix 5
Premedication subgroup analysis. Percent change from baseline in (A) total sleep duration, (B) rapid eye movement sleep duration, and (C) deep sleep duration, measured 6 days following COVID-19 vaccine dose 2 is reported for participants who premedicated prior to vaccination (purple) and participants who did not premedicate (green). Data is reported as mean and SD.
[PNG File , 133 KB - formative_v5i8e28568_app5.png]

XSL•FO
RenderX

References

1. COVID-19. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/coronavirus/2019-ncov/index.html) [accessed 2003-04-21]
2. COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). Johns Hopkins Coronavirus Resource Center. URL: https://coronavirus.jhu.edu/map.html [accessed 2003-04-21]
3. Pardi N, Hogan MJ, Porter FW, Weissman D. mRNA vaccines - a new era in vaccinology. Nat Rev Drug Discov 2018 Apr;17(4):261-279 [FREE Full text] [doi: 10.1038/nrd.2017.243] [Medline: 29326426]
4. Polack FP, Thomas SJ, Kitchin N, Absalon J, Gurtman A, Lockhart S, C4591001 Clinical Trial Group. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. N Engl J Med 2020 Dec 31;383(27):2603-2615 [FREE Full text] [doi: 10.1056/NEJMoa2034577] [Medline: 33301246]
5. Sahin U, Muik A, Vogler I, Derhovanessian E, Kranz LM, Vormehr M, et al. BNT162b2 induces SARS-CoV-2-neutralising antibodies and T cells in humans. medRxiv. Preprint posted online on December 11, 2020.
6. Wang Z, Schmidt F, Weisblum Y, Muecksch F, Barnes CO, Finkin S, et al. mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating variants. Nature 2021 Apr;592(7855):616-622. [doi: 10.1038/s41586-021-03324-6] [Medline: 33567448]
7. Vogels EA. About one-in-five Americans use a smart watch or fitness tracker. Pew Research Center. 2020. URL: https://www.pewresearch.org/fact-tank/2020/01/09/about-one-in-five-americans-use-a-smart-watch-or-fitness-tracker/ [accessed 2021-03-04]
8. Greiwe J, Nyenhuis SM. Wearable technology and how this can be implemented into clinical practice. Curr Allergy Asthma Rep 2020 Jun 06;20(8):36 [FREE Full text] [doi: 10.1007/s11882-020-00927-3] [Medline: 32506184]
9. Berryhill S, Morton CJ, Dean A, Berryhill A, Provencio-Dean N, Patel SI, et al. Effect of wearables on sleep in healthy individuals: a randomized crossover trial and validation study. J Clin Sleep Med 2020 May 15;16(5):775-783. [doi: 10.5664/jcsm.8356] [Medline: 32043961]
10. Stein PK, Bosner MS, Kleiger RE, Conger BM. Heart rate variability: a measure of cardiac autonomic tone. Am Heart J 1994 May;127(5):1376-1381. [doi: 10.1016/0002-8703(94)90059-0] [Medline: 8172068]
11. Miller DJ, Capodilupo JV, Lastella M, Sargent C, Roach GD, Lee VH, et al. Analyzing changes in respiratory rate to predict the risk of COVID-19 infection. PLoS One 2020;15(12):e0243693 [FREE Full text] [doi: 10.1371/journal.pone.0243693] [Medline: 33301493]
12. Mishra T, Wang M, Metwally AA, Bogu GK, Brooks AW, Bahmani A, et al. Pre-symptomatic detection of COVID-19 from smartwatch data. Nat Biomed Eng 2020 Dec;4(12):1208-1220. [doi: 10.1038/s41551-020-00640-6] [Medline: 33208926]
13. Hirten RP, Danieletto M, Tomalin L, Choi KH, Zweig M, Golden E, et al. Use of physiological data from a wearable device to identify SARS-CoV-2 infection and symptoms and predict COVID-19 diagnosis: observational study. J Med Internet Res 2021 Feb 22;23(2):e26107 [FREE Full text] [doi: 10.2196/26107] [Medline: 33529156]
14. Hasty F, García G, Dávila CH, Wittels SH, Hendricks S, Chong S. Heart rate variability as a possible predictive marker for acute inflammatory response in COVID-19 patients. Mil Med 2020 Nov 18:e34 [FREE Full text] [doi: 10.1093/milmed/usaa405] [Medline: 33206183]
15. Remote physiologic monitoring of resident wellness and burnout. ClinicalTrials.gov. 2020. URL: https://clinicaltrials.gov/ct2/show/NCT04304703 [accessed 2021-03-04]
16. Shaffer F, Ginsberg JP. An overview of heart rate variability metrics and norms. Front Public Health 2017;5:258. [doi: 10.3389/fpubh.2017.00258] [Medline: 29034226]
17. Herzig D, Eser P, Omlin X, Riener R, Wilhelm M, Achermann P. Reproducibility of heart rate variability is parameter and sleep stage dependent. Front Physiol 2017;8:1100. [doi: 10.3389/fphys.2017.01100] [Medline: 29367845]
18. Yasuma F, Hayano J. Respiratory sinus arrhythmia: why does the heartbeat synchronize with respiratory rhythm? Chest 2004 Mar;125(2):683-690. [doi: 10.1378/chest.125.2.683] [Medline: 14769752]
19. Silvani A, Dampney RAL. Central control of cardiovascular function during sleep. Am J Physiol Heart Circ Physiol 2013 Dec;305(12):H1683-H1692 [FREE Full text] [doi: 10.1152/ajpheart.00554.2013] [Medline: 24097430]
20. Local reactions, systemic reactions, adverse events, and serious adverse events: Pfizer-BioNTech COVID-19 vaccine. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/vaccines/covid-19/info-by-product/pfizer/reactogenicity.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fvaccines%2Fcovid-19%2Finfo-by-manufacturer%2Fpfizer%2Freactogenicity.html [accessed 2021-03-04]
21. Pearson correlation coefficient and p-value for testing non-correlation. SciPy. URL: https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.pearsonr.html [accessed 2021-02-27]
22. Plot univariate or bivariate distributions using kernel density estimation. seaborn. URL: https://seaborn.pydata.org/generated/seaborn.kdeplot.html [accessed 2021-02-27]
23. Ahmad S, Ramsay T, Huebsch L, Flanagan S, McDiarmid S, Batkin I, et al. Continuous multi-parameter heart rate variability analysis heralds onset of sepsis in adults. PLoS One 2009 Aug 14;4(8):e6642 [FREE Full text] [doi: 10.1371/journal.pone.0006642] [Medline: 19680545]

XSL•FO
RenderX

24. Kovatchev BP, Farhy LS, Cao H, Griffin MP, Lake DE, Moorman JR. Sample asymmetry analysis of heart rate characteristics with application to neonatal sepsis and systemic inflammatory response syndrome. Pediatr Res 2003 Dec;54(6):892-898. [doi: 10.1203/01.PDR.0000088074.97781.4F] [Medline: 12930915]
25. de Castilho FM, Ribeiro ALP, Nobre V, Barros G, de Sousa MR. Heart rate variability as predictor of mortality in sepsis: a systematic review. PLoS One 2018;13(9):e0203487 [FREE Full text] [doi: 10.1371/journal.pone.0203487] [Medline: 30204803]
26. Lanza GA, Barone L, Scalone G, Pitocco D, Sgueglia GA, Mollo R, et al. Inflammation-related effects of adjuvant influenza A vaccination on platelet activation and cardiac autonomic function. J Intern Med 2011 Jan;269(1):118-125. [doi: 10.1111/j.1365-2796.2010.02285.x] [Medline: 20964738]
27. McDade TW, Borja JB, Kuzawa CW, Perez TLL, Adair LS. C-reactive protein response to influenza vaccination as a model of mild inflammatory stimulation in the Philippines. Vaccine 2015 Apr 21;33(17):2004-2008 [FREE Full text] [doi: 10.1016/j.vaccine.2015.03.019] [Medline: 25795257]
28. Perring S, Jones E. Assessment of changes in cardiac autonomic tone resulting from inflammatory response to the influenza vaccination. Clin Physiol Funct Imaging 2012 Nov;32(6):437-444. [doi: 10.1111/j.1475-097X.2012.01147.x] [Medline: 23031064]
29. Tsai MY, Hanson NQ, Straka RJ, Hoke TR, Ordovas JM, Peacock JM, et al. Effect of influenza vaccine on markers of inflammation and lipid profile. J Lab Clin Med 2005 Jun;145(6):323-327. [doi: 10.1016/j.lab.2005.03.009] [Medline: 15976761]
30. Young HA, Benton D. Heart-rate variability: a biomarker to study the influence of nutrition on physiological and psychological health? Behav Pharmacol 2018 Apr;29(2 and 3-Spec Issue):140-151 [FREE Full text] [doi: 10.1097/FBP.0000000000000383] [Medline: 29543648]
31. Williams DP, Koenig J, Carnevali L, Sgoifo A, Jarczok MN, Sternberg EM, et al. Heart rate variability and inflammation: a meta-analysis of human studies. Brain Behav Immun 2019 Aug;80:219-226. [doi: 10.1016/j.bbi.2019.03.009] [Medline: 30872091]
32. Prather AA, Janicki-Deverts D, Hall MH, Cohen S. Behaviorally assessed sleep and susceptibility to the common cold. Sleep 2015 Sep 01;38(9):1353-1359 [FREE Full text] [doi: 10.5665/sleep.4968] [Medline: 26118561]
33. Spiegel K, Sheridan JF, Van Cauter E. Effect of sleep deprivation on response to immunization. JAMA 2002 Sep 25;288(12):1471-1472. [doi: 10.1001/jama.288.12.1471-a] [Medline: 12243633]
34. Taylor DJ, Kelly K, Kohut ML, Song K. Is insomnia a risk factor for decreased influenza vaccine response? Behav Sleep Med 2017;15(4):270-287 [FREE Full text] [doi: 10.1080/15402002.2015.1126596] [Medline: 27077395]
35. Lange T, Perras B, Fehm HL, Born J. Sleep enhances the human antibody response to hepatitis A vaccination. Psychosom Med 2003;65(5):831-835. [doi: 10.1097/01.psy.0000091382.61178.f1] [Medline: 14508028]
36. Prather AA, Pressman SD, Miller GE, Cohen S. Temporal links between self-reported sleep and antibody responses to the influenza vaccine. Int J Behav Med 2021 Mar;28(1):151-158. [doi: 10.1007/s12529-020-09879-4] [Medline: 32236831]
37. Benedict C, Brytting M, Markström A, Broman J, Schiöth HB. Acute sleep deprivation has no lasting effects on the human antibody titer response following a novel influenza A H1N1 virus vaccination. BMC Immunol 2012 Jan 04;13:1 [FREE Full text] [doi: 10.1186/1471-2172-13-1] [Medline: 22217111]
38. Prather AA, Hall M, Fury JM, Ross DC, Muldoon MF, Cohen S, et al. Sleep and antibody response to hepatitis B vaccination. Sleep 2012 Aug 01;35(8):1063-1069 [FREE Full text] [doi: 10.5665/sleep.1990] [Medline: 22851802]
39. Haspel JA, Anafi R, Brown MK, Cermakian N, Depner C, Desplats P, et al. Perfect timing: circadian rhythms, sleep, and immunity - an NIH workshop summary. JCI Insight 2020 Jan 16;5(1):e131487. [doi: 10.1172/jci.insight.131487] [Medline: 31941836]
40. Kim H, Hegde S, LaFiura C, Raghavan M, Luong E, Cheng S, et al. COVID-19 illness in relation to sleep and burnout. BMJ Nutr Prev Health 2021;4(1):132-139 [FREE Full text] [doi: 10.1136/bmjnph-2021-000228] [Medline: 34308120]
41. Morales J, Yáñez A, Fernández-González L, Montesinos-Magraner L, Marco-Ahulló A, Solana-Tramunt M, et al. Stress and autonomic response to sleep deprivation in medical residents: a comparative cross-sectional study. PLoS One 2019;14(4):e0214858 [FREE Full text] [doi: 10.1371/journal.pone.0214858] [Medline: 30947295]
42. Christian LM, Porter K, Karlsson E, Schultz-Cherry S. Proinflammatory cytokine responses correspond with subjective side effects after influenza virus vaccination. Vaccine 2015 Jun 26;33(29):3360-3366 [FREE Full text] [doi: 10.1016/j.vaccine.2015.05.008] [Medline: 26027906]
43. Whelton SP, Narla V, Blaha MJ, Nasir K, Blumenthal RS, Jenny NS, et al. Association between resting heart rate and inflammatory biomarkers (high-sensitivity C-reactive protein, interleukin-6, and fibrinogen) (from the Multi-Ethnic Study of Atherosclerosis). Am J Cardiol 2014 Mar 15;113(4):644-649 [FREE Full text] [doi: 10.1016/j.amjcard.2013.11.009] [Medline: 24393259]
44. Guk K, Han G, Lim J, Jeong K, Kang T, Lim E, et al. Evolution of wearable devices with real-time disease monitoring for personalized healthcare. Nanomaterials (Basel) 2019 May 29;9(6):813 [FREE Full text] [doi: 10.3390/nano9060813] [Medline: 31146479]

XSL•FO
RenderX

Abbreviations

CRP: C-reactive protein
HRV: heart rate variability
mRNA: messenger RNA
RCT: randomized controlled trial
REM: rapid eye movement
RHR: resting heart rate
rpm: respirations per minute
RR: respiratory rate

Edited by G Eysenbach; submitted 13.03.21; peer-reviewed by M Emery, A Kirkham, Y Albadrany, A Naser, R Boyapati; comments to author 30.03.21; revised version received 12.06.21; accepted 14.06.21; published 04.08.21.

Please cite as:
Hajduczok AG, DiJoseph KM, Bent B, Thorp AK, Mullholand JB, MacKay SA, Barik S, Coleman JJ, Paules CI, Tinsley A
Physiologic Response to the Pfizer-BioNTech COVID-19 Vaccine Measured Using Wearable Devices: Prospective Observational Study
JMIR Form Res 2021;5(8):e28568
URL: https://formative.jmir.org/2021/8/e28568
doi:10.2196/28568
PMID:34236995

©Alexander G Hajduczok, Kara M DiJoseph, Brinnae Bent, Audrey K Thorp, Jon B Mullholand, Stuart A MacKay, Sabrina Barik, Jamie J Coleman, Catharine I Paules, Andrew Tinsley. Originally published in JMIR Formative Research (https://formative.jmir.org), 04.08.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Early Report

Electronic Paper Displays in Hospital Operations: Proposal for Deployment and Implementation

Guruprasad D Jambaulikar[1], MBBS, MPH; Andrew Marshall[1,2], MD; Mohammad Adrian Hasdianda[1,2], MD, MSc, MMSc; Chenzhe Cao[3], MPH; Paul Chen[1,2], MD, MBA; Steven Miyawaki[1], BA; Christopher W Baugh[1,2], MD, MBA; Haipeng Zhang[3,4], DO, MMSc; Jonathan McCabe[1], RN; Jennifer Su[5], MSc; Adam B Landman[1,2,6], MD, MS, MIS, MHS; Peter Ray Chai[1,2,4,7,8], MD, MMS

[1]Department of Emergency Medicine, Brigham and Women's Hospital, Boston, MA, United States
[2]Department of Emergency Medicine, Harvard Medical School, Boston, MA, United States
[3]Brigham Digital Innovation Hub, Brigham and Women's Hospital, Boston, MA, United States
[4]Department of Psychosocial Oncology and Palliative Care, Dana Farber Cancer Institute, Boston, MA, United States
[5]E Ink Corporation, Billerica, MA, United States
[6]Mass General Brigham Information Systems, Somerville, MA, United States
[7]The Koch Institute for Integrated Cancer Research, Massachusetts Institute of Technology, Cambridge, MA, United States
[8]The Fenway Institute, Boston, MA, United States

Corresponding Author:
Peter Ray Chai, MD, MMS
Department of Emergency Medicine
Brigham and Women's Hospital
75 Francis St
Boston, MA, 02115
United States
Phone: 1 617 732 5640
Email: pchai@bwh.harvard.edu

Abstract

Background: Display signage is ubiquitous and essential in hospitals to serve several clerical, operational, and clinical functions, including displaying notices, providing directions, and presenting clinical information. These functions improve efficiency and patient engagement, reduce errors, and enhance the continuity of care. Over time, signage has evolved from analog approaches such as whiteboards and handwritten notices to digital displays such as liquid crystal displays, light emitting diodes, and, now, electronic ink displays. Electronic ink displays are paper-like displays that are not backlit and show content by aligning microencapsulated color beads in response to an applied electric current. Power is only required to generate content and not to retain it. These displays are very readable, with low eye strain; minimize the emission of blue light; require minimal power; and can be driven by several data sources, ranging from virtual servers to electronic health record systems. These attributes make adapting electronic ink displays to hospitals an ideal use case.

Objective: In this paper, we aimed to outline the use of signage and displays in hospitals with a focus on electronic ink displays. We aimed to assess the advantages and limitations of using these displays in hospitals and outline the various public-facing and patient-facing applications of electronic ink displays. Finally, we aimed to discuss the technological considerations and an implementation framework that must be followed when adopting and deploying electronic ink displays.

Methods: The public-facing applications of electronic ink displays include signage and way-finders, timetables for shared workspaces, and noticeboards and bulletin boards. The clinical display applications may be smaller form factors such as door signs or bedside cards. The larger, ≥40-inch form factors may be used within patient rooms or at clinical command centers as a digital whiteboard to display general information, patient and clinician information, and care plans. In all these applications, such displays could replace analog whiteboards, noticeboards, and even other digital screens.

Results: We are conducting pilot research projects to delineate best use cases and practices in adopting electronic ink displays in clinical settings. This will entail liaising with key stakeholders, gathering objective logistical and feasibility data, and, ultimately, quantifying and describing the effect on clinical care and patient satisfaction.

XSL•FO
RenderX

Conclusions: There are several use cases in a clinical setting that may lend themselves perfectly to electronic ink display use. The main considerations to be studied in this adoption are network connectivity, content management, privacy and security robustness, and detailed comparison with existing modalities. Electronic ink displays offer a superior opportunity to future-proof existing practices. There is a need for theoretical considerations and real-world testing to determine if the advantages outweigh the limitations of electronic ink displays.

(JMIR Form Res 2021;5(8):e30862) doi:10.2196/30862

KEYWORDS

electronic ink; patient satisfaction; display systems; whiteboards; hospital; deployment; proposal; implementation; communication; engagement; efficiency; usage

Introduction

Background on Hospital-Based Signage

Display signage is ubiquitous in hospitals and is important for daily hospital operations and clinical care [1]. Signage serves a range of clerical, operational, and clinical functions, including waypoint finding; displaying directions, notices, and bulletins; and presenting and organizing clinical schedules. The use of signs in clinical care became standard practice after a 2001 report by the Institute of Medicine emphasized patient-centered care [2]. Clinicians demonstrated that whiteboards increased patient engagement, thereby increasing satisfaction, mitigating common medical errors, and enhancing continuity of care [3-5].

Over the past decade, plasma and liquid crystal display (LCD) televisions have replaced analog surfaces to convey information. These dynamic digital alternatives enable both public-facing and clinical display systems to refresh and display data longitudinally or continuously. Public-facing displays serve general needs, including providing directions and maps, bulletins, notices, and general information. These may require intermittent or continuous changes, depending on the purpose they serve. Displays serving as bulletin boards may require the ability to display multiple messages, either requiring significant space when presented as analog or requiring continuous refreshes when digitally presented. Clinical display systems present individualized patient data to clinicians or patients. These require more ad hoc changes when analog, creating a risk of inaccurate information due to lapses in manual updates. Alternatively, digital clinical displays lend themselves to a more automated approach through electronic health record (EHR) linkage. As the cost of digitizing display systems has decreased, there is renewed interest in replacing physical signage in hospital systems with a host of digital display options.

Electronic display screens have garnered attention as an effective alternative to traditional paper or printed signage. With increasing network integration, LCD and light emitting diode (LED) screens have demonstrated value despite their cost [6]. As part of a data network (ie, “internet of things”), electronic display screens are more autonomous, allowing hospitals to provide accurate real-time information without much, if any, human manipulation. Electronic display screens can be remotely updated from a central hub, reducing the time and manpower needed for maintenance. They may be more environmentally sustainable by reducing paper and chemical waste [6]. There has already been traction in adopting electronic displays for public-facing information such as hospital waypoint finding, announcements, event postings, public health messaging, and cafeteria menus. These screens also have the potential to replace standard dry-erase boards as an individualized clinical display, standardizing information that patients receive daily or, in some settings such as the clinic or emergency department, multiple times each day. Each hospital will face its own unique challenges regarding the cost of physical devices, energy use concerns, integration difficulties, and privacy concerns for displays of patient information. This can make adoption of patient-facing electronic displays difficult [7].

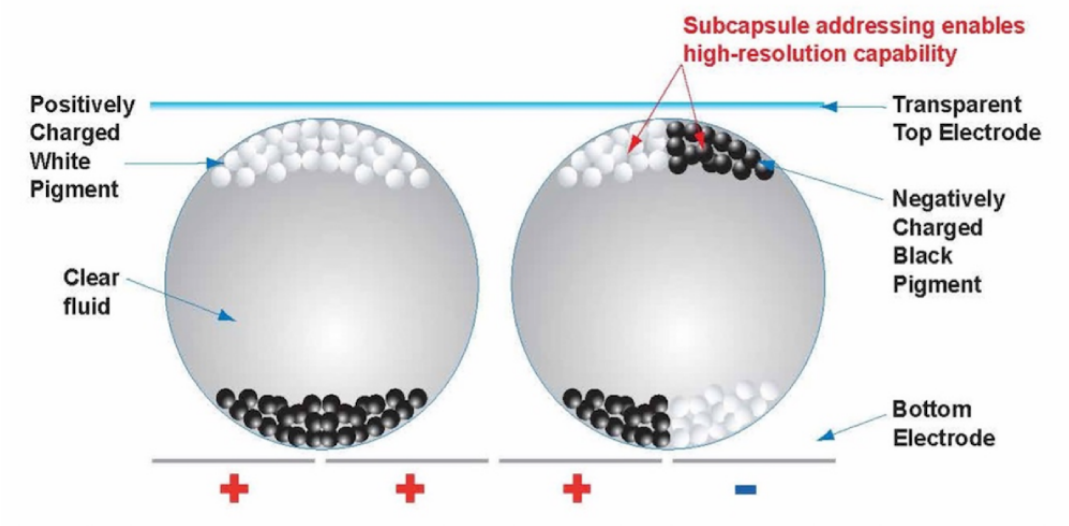
One potential alternative to traditional signage is the use of electronic ink displays. Electronic ink is made of microcapsules of black pigments and white pigments suspended in a clear fluid. When a positive or negative electric field is applied, corresponding particles move to the top of the microcapsule where they become visible to the viewer. These systems are low power, easy to read, and can be manufactured at scale. Commercially used in eReaders and store placards and, thus, already known to many consumers, the adoption of electronic ink displays may serve as a viable platform to deliver information in the hospital setting. In this paper, we described the use of electronic ink, a low-power, high-resolution display screen that can be adapted to multiple hospital uses. We presented several use cases for electronic ink displays and described an implementation schema for hospital systems seeking to explore the use of these displays both for public and clinical functions.

Electronic Ink

A formative development of paper-like electrophoretic surfaces was initiated at Xerox Palo Alto Research Center in 1974, under the pseudonym “Gyricon” [8,9]. A thin layer of transparent plastics with millions of small charged beads, akin to the dry powdery substances found in toner, rotates to present one colored side to the viewer when voltage is applied (Figure 1). The image, once set, does not change shape until new voltage patterns are applied.

XSL•FO
RenderX

Figure 1. Working principle of electronic ink displays. Image provided by E Ink.



In the mid-1990s, a professor at the Massachusetts Institute of Technology Media Lab tasked 2 of his students to take the early work of Gyricon and create a variant of electrophoretic ink that would allow for more precise image rendering and the ability to scale up to mass production. That project was commercialized in 1997 as E Ink [8]. Unlike Gyricon's bichromal electrophoretic inks, E Ink instead uses microcapsules containing black and white pigments suspended in a clear fluid. These microcapsules are 50 microns in size and impart a more precise method of manipulating ink particles to create high-resolution, crisp displays. A core benefit of all electrophoretic inks is that they only need power to change an image, not to maintain it. This results in low-power displays in comparison with standard LCD panels (ie, 25-50 mW vs 100-200 mW for 1.5-2.5–inch displays) [10].

The end result is a readable, "paper-like" display that does not require backlighting and consumes significantly less power than typical LCD or LED screens. For the display of information that is static and does not require constant refreshes, electronic ink displays provide a potential alternative to more heavily powered LCD screens. This enables electronic ink technology to be used in places where a power source may not be available or constant backlighting is not required. Advances in both flexible screens and coloration enabled paper-like electronic ink screens to be applied to a variety of consumer facing signage such as grocery store signs, advertisements, and identification badges [11,12].

Advantages of Electronic Ink Displays

There are several advantages to electronic ink displays compared with conventional LED or LCD screens. Because of the paper-like quality of electronic ink displays, they simulate analog paper surfaces and do not require backlighting, unlike conventional screens. This may make reading electronic ink displays easier by eliminating glare. Given the lack of backlighting, electronic ink displays emit no heat, decreasing the need for cooling systems or vents; this is a distinct advantage for display screens that are integrated into wall alcoves. Electronic ink displays are also lightweight and robust. Because they lack a liquid polymer or crystal layer and light source, the thickness and weight of electronic ink displays are markedly reduced. These features render them convenient to carry, transport, and install, and they are generally less susceptible to damage when dropped. This also allows electronic ink displays to be placed in locations where the weight and structural requirements of LCD panels may have been prohibitive.

Once the electronic ink has been set with the use of charges, it remains static until a refresh is triggered. This enables electronic ink displays to minimize power consumption, compared with other screen types [10]. Displays can, therefore, be powered over long periods of time using minimal power, even with conventional direct current batteries rather than requiring alternating-current wall plugs. Given the minimal power required to refresh a screen or display new information, electronic ink displays can be powered using power over ethernet. This allows for data to flow to the electronic ink display from a central controller while eliminating the need to use additional power infrastructures in existing hospital spaces. The end result for hospital groups is that, depending on the application, electronic ink displays may require less infrastructure support (eg, power over ethernet instead of requiring new electrical outlets), consume less energy, and emit little to no heat. These advantages may result in significant cost savings over the course of the device's lifetime. Finally, each electronic ink display can be loaded with an integrated operating system. This provides flexibility to the platform and can eliminate the need to mirror the display of a tethered computer. By operating as standalone devices, electronic ink screens can be installed in a variety of locations and operational settings.

Disadvantages of Electronic Ink Displays

There are several important limitations of electronic ink screens to consider before implementation in a clinical setting. First, due to a lower screen refresh rate compared with LED or LCD screens, electronic ink displays are not ideal for displaying video, animations, or rapidly changing information. Because of a set number of ink pigments suspended in electronic ink, colors are also limited, compared with the wide palette of colors available in LCD and LED displays. However, recent advances in microencapsulation have enabled combinations of colors

XSL•FO
RenderX

with white pigment to enable color electronic ink displays. With continued development, future iterations of these systems may produce different colors that advance the available color palette for electronic ink displays, such as the four-pigment system, Advanced Color ePaper, that is in development at E Ink [13].

Methods

Potential Applications in Hospital Operations

Numerous potential applications for electronic ink displays exist in a hospital setting. We classify applications into 2 categories: public-facing displays and clinical display systems (Table 1).

Table 1. Potential applications of electronic ink displays to improve hospital operations.

Criteria	Type of display	
	Public-facing displays	Clinical displays
Potential applications	• Notices and bulletins for employees, patients, and caregivers • Way-finders and maps • Location tags and placards • Guidelines and regulations • Public education and awareness messages	• Door signs and bedside cards • Digital communication boards in patient rooms • Tablet devices with access to medical records

Public-Facing Applications

Signage can be used to provide wayfinding and information to visitors in the hospital setting. Hospital settings are often complex and growing and can encompass several buildings built in different eras, creating wayfinding challenges. Easy-to-follow signage and legible directions are key to reducing stress and improving satisfaction for patients when navigating these campuses. Shared clinical spaces, offices, and conference rooms may be used for different purposes at different times and require frequent updates to avoid confusion. Electronic ink displays offer the opportunity to replace some of these screens, providing easily understandable signage that can be updated from a central location on demand.

Portable, lightweight electronic ink displays can also be used instead of noticeboards. Often pinboards and television screens are used to display bulletins and general information to visitors. If linked to a centralized location on the local network, updating electronic ink displays require less manual work than these traditional signage methods. They are potentially lower in total ownership cost and more energy efficient than LED and LCD screens for this purpose, and installation may be easier, especially, for locations that do not require intensive refreshing of screens. Since they are able to be attached to mobile carts, they are more portable, making it less challenging to transport them across different indoor or outdoor locations around the hospital. During the COVID-19 pandemic, signage has been used to provide reminders to hospital staff and visitors of important public health measures such as social distancing, hand hygiene, and mask wearing. The rapid pace at which guidelines continue to change around public health measures, testing requirements, and other COVID-19–related interventions suggests that the use of electronic ink displays could be an effective, low-energy method to provide up-to-date information to hospital visitors.

Patient-Facing Applications

Electronic ink also has utility for clinical displays. In their simplest form, electronic ink displays can be used as a door sign or bedside card, available in several sizes and resolutions (Figure 2). They are easily installed with standard wall anchors, or they can be hung on a hospital bed; they are powered for several months with AA, AAA, or rechargeable batteries. These small electronic ink displays can contain information about patients, including their location or next steps in their plan (eg, travel to radiology department for an x-ray or a cardiac stress test). When integrated with radiofrequency identification or low-energy Bluetooth beacons, electronic ink displays can also function as a wayfinding application. These applications may improve efficiency, help provide on-demand data about a patient's itinerary, and relocate on-demand changes in scheduling or clinical information to the bedside. Digital door signs and bedside cards can be programmed to display as little or as much information as desired and can also be used to display patient safety elements such as precautions, allergies, and alerts.

XSL•FO
RenderX

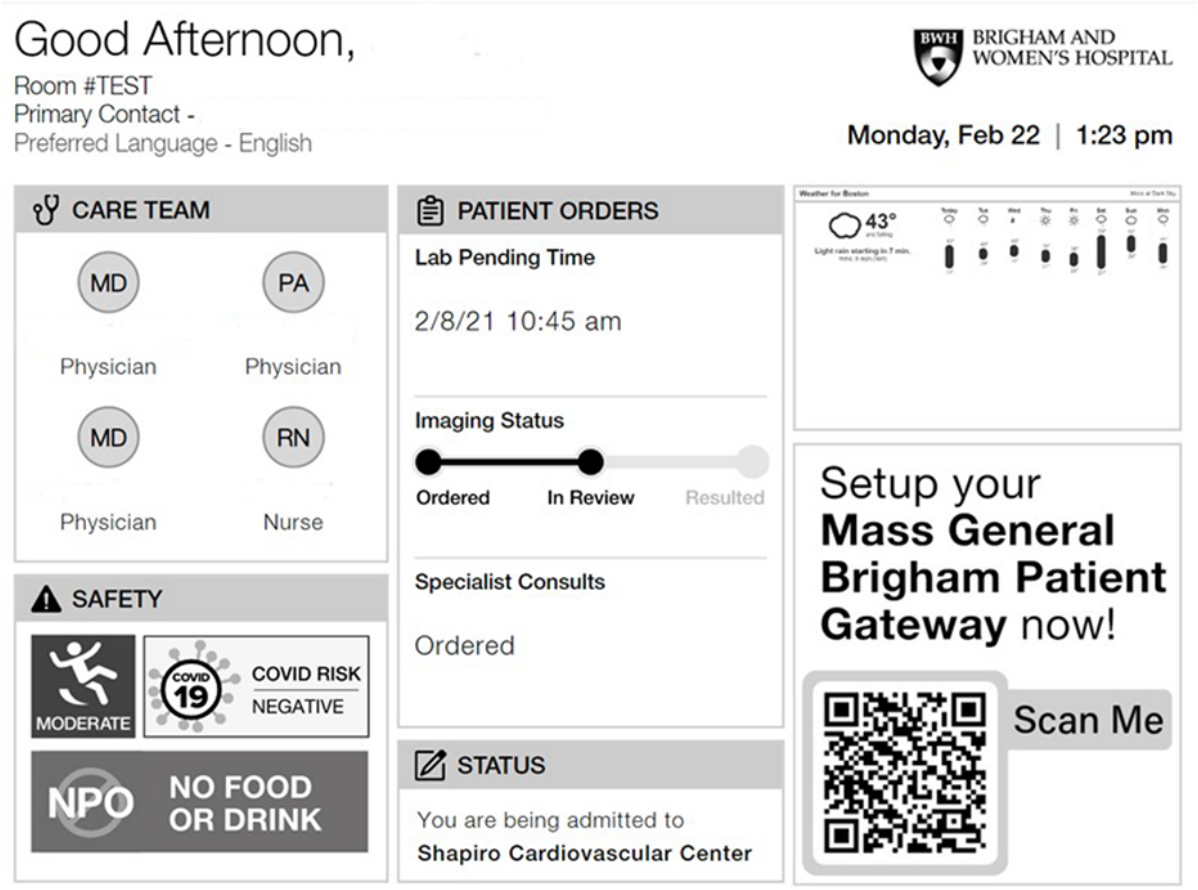
Figure 2. Diverse applications using electronic ink displays in the health care setting. (A) Door sign. (B) Bedside card. (C) Patient room information board. (D) Visitor signage. Images provided by E Ink.

Dry-erase boards have been used in the hospital to orient patients to the date, provide information regarding their care team members, and even communicate pain scores [14-16]. In addition, paper and laminated signs have been used outside of patient rooms to convey information regarding isolation status, fall risk, and need for personal protective equipment [4]. In some instances, hospitals display daily plans to inpatients and their families, and, for acutely ill patients in the emergency department, whiteboards can be used by the teams to coordinate care. A significant drawback to the use of paper or traditional whiteboards is the need to manually and consistently update them at set intervals. In busy hospital settings, this can result in outdated or inaccurate information, creating new safety risks that the tool was originally intended to mitigate. As an alternative to either a whiteboard or a backlit screen such as a television screen, electronic ink screens can be used to display patient information (Figure 3). Since electronic ink screens do not emit light, they are less distracting to patients who are trying to sleep or rest during their stay. This may be instrumental in preventing delirium caused by persistent light stimuli from LCD screens. By providing an indirect stimulus concerning the course of their clinical care, it is possible that electronic ink displays can help provide orientation to patients with prolonged stays in the hospital, thereby addressing disorientation and delirium [17,18]. Patients are becoming increasingly more comfortable with the use of technology to enhance their care. In a 2012 survey study of an urban emergency department, approximately 90% of patients preferred technology-based behavioral interventions [19]. Communication boards can be configured to display data customized for each unique clinical environment. Basic functions include displaying information to orient patients, such as date, time, and names and roles of the current clinical team. Information on diagnostic tests and imaging, as well as final disposition, may help guide patients and their families to understand their clinical course and anticipate potential events that may happen in the hospital. For patient-facing screens, additional functionality may include displaying local weather, transit schedules, cafeteria menus, and important hospital notices, thereby providing on-demand information to patients and their families. Strong communication and efficient information delivery have been associated with improved patient satisfaction, which may, ultimately, influence hospital choice and improve quality of care [20,21].

The success of electronic ink displays in hospitals is contingent on their integration with existing information systems. Legislative and regulatory changes have promoted significant advances in interoperability. These advances have created standard methods for connecting applications to exposed application programing interfaces (APIs) in EHRs, allowing for seamless data exchange. Using this exchange, electronic ink devices may be configured to display real-time information directly from EHRs. Additional discussion on protocols to directly exchange information can be found under the subheading “Technological Considerations” in the Discussion section of this paper.

XSL•FO
RenderX

Figure 3. Electronic ink communication board. NPO: nothing by mouth.



Results

We are planning several pilot studies to evaluate the use of electronic ink screens in various hospital settings. We hope to use these pilot studies to help delineate the best practices and use cases for electronic ink screens in clinical care and hospital operations. These pilot studies will also help develop the information security infrastructure and support needed to manage multiple electronic ink screens at the same time. This programming architecture also permits custom displays of different data on various screens. At this time, electronic ink displays are deployed in our emergency department to display information about a patient's emergency stay, and we are investigating their effect on patient satisfaction. Other use cases will include wayfinding through smaller display screens affixed to hospital beds, two-way asynchronous communication between patients and clinicians through the electronic ink screens, and patient identification in the operating room.

Discussion

Technological Considerations

Connectivity

Ensuring secure and reliable network connectivity is an important consideration in electronic ink display deployments. Electronic ink displays can be deployed in several different manners. The simplest manner is with limited network connectivity, requiring manual entry of information to be displayed. One way to deploy this is to create a virtual environment—a server-based command control for network displays—on the hospital network, which is directly connected to the display. Eventually, the most sustainable and future-proof method will be to connect these displays to data sources that are configured to update specified information such as EHRs automatically and as close to real time as possible. This would add the most value to the efficiency and accuracy of clinical displays inside and outside patient rooms.

Deploying a large number of electronic ink displays in a hospital environment has different considerations compared with the implementation of LCD display screens. Unlike LCD screens that require a persistent internet connection to display information, electronic ink screens only require data connection to change the information displays. Depending on the intended information, electronic ink screens may, therefore, require less internet bandwidth compared with LCD screens. The nature of the intermittent data of electronic ink displays may also pose less of a network security risk, as there is no maintained continuous internet connection; instead, the central architecture pings each electronic ink display only when new data are displayed. Until network infrastructure can handle the traffic, the use of dedicated networks in the form of virtual machines may help ensure efficient display performance and also minimize the chance of unintended disruptions to the primary hospital network.

XSL•FO
RenderX

Privacy and Security

When considering the use of electronic ink displays, hospital systems should understand the potential privacy and security risks. Privacy breaches can occur when unencrypted data that contain protected health information (PHI) are intercepted or when the incorrect data are transmitted to an electronic ink display. For example, manual updates of electronic ink displays in patient rooms could inadvertently expose other individuals in the room to PHI, or, if a patient's information is displayed to the incorrect screen, this could result in a privacy violation. Patient privacy is a prime concern when PHI is transmitted to displays from EHRs. To protect against breaches in patient privacy, data transmission should avoid use of PHI or use the minimum PHI necessary for care delivery. PHI or other sensitive information that needs to be transmitted should use modern encryption protocols. Given that electronic ink displays are visible to anyone in the physical room, masking PHI on the screen could also be considered, such as displaying patient initials instead of full names.

Despite best practices, there remains the possibility that manually entered data in an electronic ink display could inadvertently be erroneous or expose PHI to another patient if an incorrect display is loaded by mistake. These errors are significantly mitigated through the use of automated checks and integrated rules in the electronic medical record that may prevent the display of data in inappropriate locations. Notwithstanding these technical mitigation strategies, we recommend additional physical mitigation, including a physical shut-off button to wipe the screen in case individuals find unauthorized information displayed on the screen.

Network security can be compromised when smart displays are used as an entry point to hospital networks or used as a distributed denial of service attack. Understanding these risks and creating strategies to effectively mitigate them are central to safely deploying these technologies in hospitals (Table 2).

Table 2. Privacy and security considerations for electronic ink displays.

Potential concern	Potential solutions
Privacy and data breaches	• Communicate deidentified data where possible • Suspend continuous network connection and data transfer when not required • Enable electronic ink displays to communicate via independent networks to minimize integration with hospital networks unless necessary • Limit network access by only delegating accounts that need access and following institutional password requirements • Establish a log system to audit and document evidence of undesired activity • Develop a risk-mitigation plan and an incident-response policy that may be implemented in case of emergency
Distributed denial of service attacks	• Establish a log system to audit and document evidence of undesired activity • Displays placed on separate networks with intermittent limited access to hospital servers only as required • Ensure server updates and vulnerabilities are addressed
Failure	• Staggered adoption with careful testing of failure rates • Initial use in conjunction with existing standard practices • Develop a risk-mitigation plan and an incident-response policy that may be implemented in case of emergency

Open firewall ports, used to deliver data to screens, may provide a portal to enter a hospital network and interdict critical health information or conduct malicious attacks against hospital infrastructure [22]. To alleviate these risks, smart display systems can be programmed to only connect to a hospital network for a brief period, during which refreshes are performed and data are transmitted. As is best practice, smart devices should be isolated from main hospital networks.

In order to transmit sensitive notifications containing protected health information, electronic ink displays should securely connect to encrypted wireless networks using Wi-Fi Protected Access 2 connection. Wi-Fi Protected Access-enterprise encryption systems, which require a user to enter a unique username and password to log into the network, provide an even higher layer of security necessary for networks that transmit confidential information. This way, even if a hacker learns the password of one device, they cannot compromise the entire system. Another option is to use two-factor authentication to validate the administrator who is accessing the configuration of the electronic ink screen. This additional layer of security may help mitigate bot-based hacking attempts.

Malicious users with control over smart devices could also conduct distributed denial of service attacks by sending rapid triggers from the devices, in an attempt to overwhelm the hospital network [23,24]. As noted previously, limiting continuity of network transmission should address this issue. Finally, good stewardship of smart devices is paramount to protecting security, such as limiting access only to highly trained staff and continually encouraging good practices.

Systems Integration

The success of electronic ink devices in the hospital relies heavily on the ability to effectively update the information displayed from a central location. For public-facing signage, this likely requires the establishment of a centralized dashboard for content management. Patient-facing displays, on the other hand, likely require integration of EHR for maximum efficiency.

XSL•FO
RenderX

Improvements in interoperability and a nationwide focus on providing patients access to their entire medical records have made it feasible to propose patient-facing applications that integrate EHRs directly. As a part of the Meaningful Use Act [25], hospitals were incentivized to provide access to health care information to patients. The more recent information-blocking provision of the 21st Century Cures Act [26] removed nearly all barriers to patients' ability to access their hospital records. Electronic ink displays may act as a vehicle to provide on-demand access to pertinent patient information that could reduce barriers to accessing health records as well as assisting hospitals in satisfying the requirements of the Meaningful Use Act [27]. For example, emergency department patients can obtain their personal laboratory results or understand a status update from a consultant who has been asked to evaluate them.

Exchange of clinical information between electronic ink devices and EHRs should be built around widely accepted standards. Fast Healthcare Interoperability Resources [28] provides an open-source and widely accepted means for packaging and delivering clinical data. Although some form of health information exchange is required as a part of the Meaningful Use Act, each institution's API may differ. The ideal systems integration uses the APIs health level 7 or Fast Healthcare Interoperability Resources, enabling and automating information exchange on electronic ink screens [29].

A Framework for Electronic Ink Display Deployment

For health care organizations that seek to deploy electronic ink displays, we recommend steps grounded in the plan-do-study-act (PDSA) method (Figure 4) [30]. The PDSA method is often used to accelerate quality improvement initiatives by rapidly testing changes through structured planning, implementation, observation, and iterative improvements based on pilot studies [30]. For successful deployment of a novel technology, it may be helpful to form a centralized committee of network specialists, hospital administration, and clinical experts who understand outcomes surrounding the use of electronic ink displays and information security, to ensure that all stakeholders required to successfully deploy an electronic ink display can assemble and map key tasks prior to implementation. Establishing a central data server that can query and modify displays in a secure manner will be critical to the success of pilot studies, particularly, for electronic ink displays, and should be developed prior to deployment of electronic ink displays in the clinical setting.

Figure 4. A proposed framework to deploy and evaluate the impact of electronic ink displays in a hospital setting.

As part of the "plan" phase, the first step in an electronic ink display implementation is to plan for the infrastructure components. Since the displays will require network connectivity, hospital information security officers should be engaged to perform a risk assessment of the technology and mitigate any high-priority risks identified. It is also important to limit the initial scope of work for any display during the process of deployment. This permits a gradual rollout of electronic ink displays and allows adequate space for implementation issues to be resolved by the study team. Several key stakeholders, including clinical and nonclinical staff, patients, and patient advocates, should discuss and finalize the data elements that make the most sense in a given scenario. Key discussion points should center around important identifying information that may be displayed on the screen.

As part of the "do" phase, the team should select limited, yet important, use cases that would benefit from a brief pilot run. Benchmarks for success of the pilot study and the duration of the study should be established in advance, to provide clear parameters and expectations around deployment. Despite selecting individual pilot studies, a central resource of information technology and device and programming expertise should help govern and manage electronic ink displays. This will ensure that there is continuity around different projects and

XSL•FO
RenderX

that data from projects remain unified in a central location. Additionally, this mechanism will allow for seamless transfer of operating systems and platforms to additional investigations planned by the study team.

As part of the "study" phase, the pilot study should be evaluated to assess feasibility, return on investment, and user experience of the electronic ink displays. Investigators may consider the use of validated measures such as the System Usability Scale or Net Promoter score to understand usability, acceptability, and satisfaction associated with implementation of the system [31-34]. In addition, hardware stability, power consumption and cost, data security and quality, and network load should be monitored during the pilot study.

As part of the "act" phase, lessons learned from the pilot study, including technical, workflow, and other components of the socio-technical model for health information technology, should be carefully reviewed [35]. Based on the pilot studies completed, a use case for widespread implementation should be developed. Messaging with and training of the staff who will operate and interact with electronic ink displays should occur prior to the date when the displays become activated. To prepare for larger rollouts, processes and protocols should be adjusted based on these lessons learned. A dedicated governance process team, including the project team, information technology staff, and institutional leaders, is essential for identifying the appropriateness of employing smart displays for any use case. It is also important to understand the reliability of the technology and the impact it has on technical infrastructures, so it can be safely used for its desired tasks.

Conclusions

Electronic ink displays may be a valuable tool to help optimize hospital operations and communications. They have a variety of use cases for both patients and staff. Key technical considerations for a successful deployment include an appropriate network connectivity, a robust content management process, and a careful configuration that minimizes privacy and security risks. The PDSA framework may guide hospitals, starting with a small pilot study and iteratively refining the process until eventually scaling to the entire organization. Electronic ink displays present a tremendous opportunity to future-proof existing analog processes while overcoming some of the disadvantages of commonly used digital modalities. The promise might outweigh the minor limitations, but there is still a need for testing in real-world health care environments to rigorously evaluate and determine the impacts of electronic ink displays.

Acknowledgments

PRC was funded by NIH K23DA044874, R44DA051106, Hans and Mavis Psychosocial Foundation, and E Ink Corporation. AM was funded by E Ink Corporation and NIH 5T15LM007092-27. Work described in this investigation was funded by an investigator-initiated research grant from E Ink Corporation. This work was supported by the Brigham and Women's Hospital Digital Innovation Hub.

Conflicts of Interest

PRC has has a financial interest in Biobot Analytics, a company engaged in the collection and analysis of wastewater to develop epidemiological data. PRC's interests were reviewed and are managed by Brigham and Women's Hospital and Mass General Brigham in accordance with their conflict of interest policies. ABL was previously a member of the Abbott Medical Device Cybersecurity Council. JS is an employee of E Ink Corporation.

References

1. Bardram J, Hansen T, Soegaard M. Large interactive displays in hospitals: Motivation, examples, and challenges. 2005 Presented at: Proc SIGCHI Conf Hum Factors Comput Syst Citeseer; 2006; Montreal, Canada URL: https://www.semanticscholar.org/paper/Large-Interactive-Displays-in-Hospitals-—-%2C-%2C-and-Bardram-Hansen/d0fff501ef26f51d316e3ad5f75658cafd3ff044
2. Committee on Quality of Health Care in America, Institute of Medicine. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington (DC): National Academies Press; 2001.
3. Tan M, Hooper Evans K, Braddock CH, Shieh L. Patient whiteboards to improve patient-centred care in the hospital. Postgrad Med J 2013 Oct;89(1056):604-609. [doi: 10.1136/postgradmedj-2012-131296] [Medline: 23922397]
4. Goyal A, Glanzman H, Quinn M, Tur K, Singh S, Winter S, et al. Do bedside whiteboards enhance communication in hospitals? An exploratory multimethod study of patient and nurse perspectives. BMJ Qual Saf 2020 Oct;29(10):1-2. [doi: 10.1136/bmjqs-2019-010208] [Medline: 31694874]
5. Singh S, Fletcher KE, Pandl GJ, Schapira MM, Nattinger AB, Biblo LA, et al. It's the writing on the wall: Whiteboards improve inpatient satisfaction with provider communication. Am J Med Qual 2011 Sep 24;26(2):127-131. [doi: 10.1177/1062860610376088] [Medline: 20870743]
6. Grant AE, Meadows JH, editors. Communication Technology Update and Fundamentals. New York: Routledge; 2012.
7. A Complete Guide to Healthcare Digital Signage. BLM Technologies. 2020. URL: https://resources.blmtechnology.com/hubfs/Healthcare%20Digital%20Signage%20White%20Paper%20-%20BLM%20Technologies.pdf [accessed 2021-07-27]
8. Crawford G. A bright new page in portable displays. IEEE Spectr 2000 Oct;37(10):40-46. [doi: 10.1109/6.873916]

XSL•FO
RenderX

9. Kholghi Eshkalak S, Khatibzadeh M, Kowsari E, Chinnappan A, Jayathilaka W, Ramakrishna S. Overview of electronic ink and methods of production for use in electronic displays. Optics & Laser Technology 2019 Sep;117:38-51. [doi: 10.1016/j.optlastec.2019.04.003]
10. Fernández M, Casanova E, Alonso I. Review of Display Technologies Focusing on Power Consumption. Sustainability 2015 Aug 11;7(8):10854-10875. [doi: 10.3390/su70810854]
11. Chen Y, Au J, Kazlas P, Ritenour A, Gates H, McCreary M. Electronic paper: Flexible active-matrix electronic ink display. Nature 2003 May 08;423(6936):136-136. [doi: 10.1038/423136a] [Medline: 12736673]
12. Henzen A, van de Kamer J, Nakamura T, Tsuji T, Yasui M, Pitt M, et al. 13.2: Development of Active Matrix Electronic Ink Displays for Handheld Devices. SID Symposium Digest 2003;34(1):176. [doi: 10.1889/1.1832232]
13. E Ink Color Technologies. E Ink Corporation. URL: https://www.eink.com/color-technology.html [accessed 2021-05-20]
14. Wahlberg K, Lambirth S, Gardner Z. Improving patients' ability to identify their physicians through the use of physician facecards and whiteboards. BMJ Open Qual 2019 Apr 29;8(2):e000606. [doi: 10.1136/bmjoq-2018-000606] [Medline: 31206066]
15. Sehgal NL, Green A, Vidyarthi AR, Blegen MA, Wachter RM. Patient whiteboards as a communication tool in the hospital setting: a survey of practices and recommendations. J Hosp Med 2010 Apr;5(4):234-239. [doi: 10.1002/jhm.638] [Medline: 20394030]
16. Rutherford P, Lee B, Greiner A. Transforming Care at the Bedside Internet. Gordon, editor. URL: http://www.ihi.org:80/resources/Pages/IHIWhitePapers/TransformingCareattheBedsideWhitePaper.aspx [accessed 2021-05-05]
17. Angel C, Brooks K, Fourie J. Standardizing Management of Adults with Delirium Hospitalized on Medical-Surgical Units. Perm J 2016;20(4):16-002 [FREE Full text] [doi: 10.7812/TPP/16-002] [Medline: 27644045]
18. Bauernfreund Y, Butler M, Ragavan S, Sampson EL. TIME to think about delirium: improving detection and management on the acute medical unit. BMJ Open Qual 2018;7(3):e000200. [doi: 10.1136/bmjoq-2017-000200] [Medline: 30167472]
19. Ranney ML, Choo EK, Wang Y, Baum A, Clark MA, Mello MJ. Emergency department patients' preferences for technology-based behavioral interventions. Ann Emerg Med 2012 Aug;60(2):218-27.e48. [doi: 10.1016/j.annemergmed.2012.02.026] [Medline: 22542311]
20. Sun B, Adams J, Orav E, Rucker D, Brennan T, Burstin H. Determinants of patient satisfaction and willingness to return with emergency care. Ann Emerg Med 2000 May;35(5):426-434. [Medline: 10783404]
21. Thompson D, Yarnold P, Williams D, Adams S. Effects of actual waiting time, perceived waiting time, information delivery, and expressive quality on patient satisfaction in the emergency department. Ann Emerg Med 1996 Dec;28(6):657-665. [doi: 10.1016/s0196-0644(96)70090-2] [Medline: 8953956]
22. Gordon W, Wright A, Glynn R, Kadakia J, Mazzone C, Leinbach E, et al. Evaluation of a mandatory phishing training program for high-risk employees at a US healthcare system. J Am Med Inform Assoc 2019 Jun 01;26(6):547-552 [FREE Full text] [doi: 10.1093/jamia/ocz005] [Medline: 30861069]
23. Chai PR, Zhang H, Baugh CW, Jambaulikar GD, McCabe JC, Gorman JM, et al. Internet of Things Buttons for Real-Time Notifications in Hospital Operations: Proposal for Hospital Implementation. J Med Internet Res 2018 Aug 10;20(8):e251 [FREE Full text] [doi: 10.2196/jmir.9454] [Medline: 30097420]
24. Sonar K, Upadhyay H. A survey: DDOS attack on Internet of Things. International Journal of Engineering Research and Development 2014;10(11):58-63.
25. Promoting Interoperability Programs. Centers for Medicare & Medicaid Services. 2018. URL: https://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms [accessed 2021-07-26]
26. Officer of the Commissioner. 21st Century Cures Act. US Food & Drug Administration. 2020. URL: https://www.fda.gov/regulatory-information/selected-amendments-fdc-act/21st-century-cures-act [accessed 2021-07-26]
27. Information BONCHITCPOCDIRTTCOVID1PHEI. Fed Regist. 2020. URL: https://www.federalregister.gov/documents/2020/11/04/2020-24376/information-blocking-and-the-onc-health-it-certification-program-extension-of-compliance-dates-and [accessed 2021-05-10]
28. Braunstein ML. Healthcare in the Age of Interoperability: The Promise of Fast Healthcare Interoperability Resources. IEEE Pulse 2018;9(6):24-27. [doi: 10.1109/MPUL.2018.2869317] [Medline: 30452344]
29. Health Intersect Pty Ltd. Resources For Health: A Fresh Look Proposal Internet. 2011. URL: http://www.healthintersections.com.au/?p=502 [accessed 2021-05-10]
30. Berwick D, Bauchner H, Fontanarosa PB. Innovations in Health Care Delivery. JAMA 2015 Aug 18;314(7):675. [doi: 10.1001/jama.2015.9257]
31. Bangor A, Kortum PT, Miller JT. An Empirical Evaluation of the System Usability Scale. International Journal of Human-Computer Interaction 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]
32. Reichheld F. The One Number You Need to Grow. Harv Bus Rev Internet. 2003 Dec 1. URL: https://hbr.org/2003/12/the-one-number-you-need-to-grow [accessed 2021-06-01]
33. Net Promot Netw. What Is Net Promoter? Internet. URL: https://www.netpromoter.com/know/ [accessed 2021-04-20]
34. Brooke J. SUS: a retrospective. J Usability Stud 2013;8(2):40.

XSL•FO
RenderX

35. Sittig DF, Singh H. A new sociotechnical model for studying health information technology in complex adaptive healthcare systems. Qual Saf Health Care 2010 Oct;19 Suppl 3:i68-i74 [FREE Full text] [doi: 10.1136/qshc.2010.042085] [Medline: 20959322]

Abbreviations

API: application programing interface
EHR: electronic health record
LCD: liquid crystal display
LED: light emitting diode
PDSA: plan-do-study-act
PHI: protected health information

Edited by G Eysenbach; submitted 01.06.21; peer-reviewed by Y Nishimura; comments to author 27.06.21; revised version received 30.06.21; accepted 06.07.21; published 04.08.21.

Please cite as:
Jambaulikar GD, Marshall A, Hasdianda MA, Cao C, Chen P, Miyawaki S, Baugh CW, Zhang H, McCabe J, Su J, Landman AB, Chai PR
Electronic Paper Displays in Hospital Operations: Proposal for Deployment and Implementation
JMIR Form Res 2021;5(8):e30862
URL: https://formative.jmir.org/2021/8/e30862
doi:10.2196/30862
PMID:34346904

©Guruprasad D Jambaulikar, Andrew Marshall, Mohammad Adrian Hasdianda, Chenzhe Cao, Paul Chen, Steven Miyawaki, Christopher W Baugh, Haipeng Zhang, Jonathan McCabe, Jennifer Su, Adam B Landman, Peter Ray Chai. Originally published in JMIR Formative Research (https://formative.jmir.org), 04.08.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Mobile Health for Smoking Cessation Among Disadvantaged Young Women During and After Pregnancy: User-Centered Design and Usability Study

Marloes E Derksen[1], MSc; Monique WM Jaspers[2], Prof Dr; Sander van Strijp[1], MSc; Mirjam P Fransen[1], PhD

[1]Department of Public and Occupational Health, Amsterdam Public Health Research Institute, Amsterdam UMC, University of Amsterdam, Amsterdam, Netherlands

[2]Department of Medical Informatics, Amsterdam Public Health Research Institute, Amsterdam UMC, University of Amsterdam, Amsterdam, Netherlands

Corresponding Author:
Marloes E Derksen, MSc
Department of Public and Occupational Health
Amsterdam Public Health Research Institute
Amsterdam UMC, University of Amsterdam
Meibergdreef 9
Amsterdam, 1105 AZ
Netherlands
Phone: 31 20 5664494
Email: m.e.derksen@amsterdamumc.nl

Abstract

Background: Smoking prevalence during and after pregnancy remains high among socioeconomically disadvantaged women. Mobile health (mHealth) apps with game and social support elements seem promising to support smoking cessation.

Objective: This study aims to describe the user-centered design and usability evaluation of Kindle, an mHealth app with game and social support elements, to support disadvantaged young women during and after pregnancy through the first stages of smoking cessation.

Methods: Disadvantaged women (n=9), members of their social networks (n=4), and nurses supporting these women (n=51) were informants throughout the iterative prototype development of Kindle according to the International Organization for Standardization 9241-11:2018. Specific phases included understanding the context of use through secondary analysis of qualitative interview data (phase 1), establishing the user and organizational requirements (phase 2), production of design solutions (phase 3), and usability inspection of the prototype through a heuristic evaluation (3 experts) along with user testing by a think aloud method (5 disadvantaged women and 5 nurses; phase 4). Usability problems were categorized according to the principles of the Healthcare Information and Management Systems Society.

Results: Phase 1 resulted in an understanding of the VoorZorg program and the needs of VoorZorg nurses and clients (eg, focus on early stages of change and building new supportive networks to aid clients in smoking cessation). In phase 2, we established requirements (n=22; eg, mHealth app, secure communication between nurses and clients, easy-to-use interfaces, inclusion of game elements, and tailoring at early stages of change in smoking cessation). Phase 3 resulted in a prototype of Kindle, combining the interface for nurses and clients, including the following functionalities: personal goal setting with earning points; secured chat function between nurses and other clients; and tips, diary, and profile creation. The heuristic evaluation and thinking aloud method in phase 4 revealed 78 usability problems in the interfaces. Most usability problems concerned *simplicity* (eg, unclear clickable button) and *naturalness* (eg, unclear icon).

Conclusions: The user-centered design and usability testing of the mHealth app Kindle yielded useful insights. The involvement of end users, specifically socioeconomically disadvantaged women during and after their pregnancy, resulted in a prototype that met their needs and requirements (eg, mHealth app, secure communication between nurses and clients, easy-to-use interfaces, inclusion of game elements, and tailoring to the early stages of change in smoking cessation) to achieve readiness for smoking cessation. Moreover, the usability evaluation by end users and experts revealed unique usability problems for this population. These insights allow for further optimization of Kindle and encourage future studies to engage disadvantaged populations in all phases of mHealth intervention design and usability testing.

XSL•FO
RenderX

(JMIR Form Res 2021;5(8):e24112) doi:10.2196/24112

KEYWORDS

think aloud; heuristic evaluation; usability; mHealth; game elements; smoking prevention; user-centered design; mobile phone

Introduction

Background

Tobacco smoking among pregnant women accounts for a substantial proportion of preventable morbidity and mortality [1]. Smoking cessation among pregnant women not only benefits their own health but also reduces the risks of miscarriage, preterm birth, and low birth weight [2]. Moreover, cessation of smoking after pregnancy prevents their offspring from secondhand smoke exposure and, consequently, from diseases linked to secondhand smoke exposure, such as sudden infant death syndrome and respiratory diseases [2]. Strong predictors of smoking prevalence among women in Europe and the United States are low levels of educational attainment, health literacy, and socioeconomic status [3-5].

Mobile health (mHealth) apps appear to have positive effects on smoking cessation [6,7], and the inclusion of multiple game elements seems particularly promising [8]. In general, pregnant women have been found to frequently use eHealth and mHealth [9] and consider mHealth as a useful and playful tool [10]. In particular, the functionalities to interact with other mothers are useful, yet the quality of these web-based communities has been criticized by women. Moreover, women have indicated that interactive functionalities could be enriched by a direct chat with their health care professionals in addition to face-to-face care [10] and that they prefer easy-to-use interfaces [10]. Women have also signaled that they are more likely to be influenced by pregnancy-related information retrieved from an mHealth pregnancy app than widespread internet use [9]. However, women lack mHealth apps that allow personalization and have concerns about data security [10].

Although smoking during pregnancy is a problem among disadvantaged women (ie, those with a lower educational level, unplanned pregnancies, and additional risk factors [3]), the positive effects of mHealth interventions on smoking cessation are not evident in disadvantaged populations, as these interventions show only few improvements in health outcomes in disadvantaged populations [11]. Another review of disadvantaged patients with diabetes showed that mHealth interventions should be improved in terms of access, design, and usability [12]. As mHealth interventions are typically designed with minimal involvement of end users [13], the effectiveness of mHealth interventions among disadvantaged populations might be improved by a user-centered design approach. In a user-centered design approach, users influence how a design takes shape by providing input at subsequent design phases, typically during requirement gathering and usability testing. The added value of a user-centered approach has been demonstrated by a meta-analysis of randomized controlled trials on serious games for health lifestyle promotion. The effectiveness of participatory design depends on roles (eg, informants and co-design) and game design elements (eg, game levels or challenges) [14]. However, to the best of our knowledge, few studies have followed a user-centered design approach in the design of mHealth apps for disadvantaged populations [15,16].

In the Netherlands, the most disadvantaged pregnant women are offered a preventive care program called VoorZorg, which is supplementary to standard maternal care in the Netherlands [17]. The program resembles the nurse-family partnership developed in the United States [18] and the family-nurse partnership implemented in the United Kingdom [19]. These women are supported on multiple domains (eg, personal development and health promotion) [20] by certified, specialized nurses during home visits lasting 2.5 years.

On enrollment in VoorZorg (at 16-28 weeks of gestation), 43% of the women smoked. This reduced to 33% at 32 weeks of gestation and to 48% at 8 weeks after delivery [20]. VoorZorg nurses find it hard to support these women in smoking cessation due in large part to the use of support methods that do not fit the needs of women. Moreover, women are poorly motivated to stop smoking because of multiple stressors and other challenges they face [21]. Disadvantaged women appear to be in the early stages of change for smoking cessation, and their social networks mainly play a negative role in their smoking cessation efforts [22]. These insights reveal a misalignment between these women's contexts and traditional action-oriented interventions for smoking cessation. Without planned interventions, these women will remain stuck in the early stages [23]. VoorZorg nurses and their clients are thus in need of an innovative intervention to move women through early stages of change with supportive social networks to stimulate them to quit smoking. As such an intervention is still nonexistent and recognizing the use of mHealth apps seems promising, we developed and evaluated a smartphone prototype with game elements to support disadvantaged young women during and after pregnancy with smoking cessation named *Kindle*.

Objectives

This paper aims to report the user-centered design process and usability evaluation of Kindle by disadvantaged women and health care professionals and provides insights and recommendations regarding the design of mHealth apps for disadvantaged user populations.

Methods

Design

In this user-centered design study, we included women from the VoorZorg program, members of their social networks, and VoorZorg nurses as informants in the design of Kindle. The study was conducted from June 2017 to May 2018 in the Netherlands. The design team members were designers affiliated with Waag (an organization developing innovative, inclusive technology for society), researchers (MD, PhD student Public Health; MF, PhD Public Health; SS, Master of Science student

XSL•FO
RenderX

Medical Informatics; and MJ, professor Medical Informatics and human factors engineering expert), and a VoorZorg program representative of the Netherlands Centre for Preventive Youth Health. Kindle was inspected for usability by 3 human factors engineering experts under the supervision of MJ and tested by representative end users in June 2018.

Informed consent was obtained from all participants in this study. The Medical Ethics Review Committee of Amsterdam University Medical Centers confirmed that the Medical Research Involving Human Subjects Act does not apply to this study, and therefore, no official approval of the committee was required. We used the Statement on Reporting of Evaluation Studies in the Health Informatics framework to report our study [24].

In the iterative user-centered design process of Kindle, we followed the standards of the International Organization for Standardization 9241-11:2018, which supports the identification and planning of effective human-centered design activities. These design activities entail four phases: (1) understanding and specifying the context of use, (2) specifying the user and organizational requirements, (3) production of design solutions, and (4) evaluating design solutions [25].

The usability of the Kindle was assessed as part of the fourth phase. Usability is generally defined as "the extent to which a system, product, or service can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use" [25]. As no usability method is effective in all circumstances, a combination of usability methods that complement each other is generally preferred [26]. One expert-based usability inspection method is heuristic evaluation. During heuristic evaluation, a small number of human factors engineering experts evaluate a user interface according to a set of heuristics, which likely results in high-quality results over short periods [26,27]. The Healthcare Information and Management Systems Society (HIMSS) principles are commonly used to categorize usability problems. These principles include simplicity, naturalness, consistency, minimizing cognitive load, efficient interactions, forgiveness and feedback, effective use of language, effective information presentation, and preservation of context [28]. The think aloud method is a low-threshold, user-based usability testing method [29]. Think aloud entails end users performing tasks with the user interface while verbalizing what they are doing and provides insight into the causes of usability problems encountered by end users, thereby providing suggestions for redesign [26].

Phase 1: Understand and Specify the User Context

We performed a secondary analysis on unpublished, empirical data from our earlier qualitative studies among VoorZorg nurses [21], clients, and the social networks of clients [22]. MD inductively coded text fragments describing the user context, which were then discussed with MF and MJ.

Phase 2: Specify User and Organizational Requirements

Overview

For the specification of user and organizational requirements, we used the same secondary analysis of our qualitative interviews in phase 1. Moreover, we held two intervention design inquiry sessions, one among VoorZorg nurses and one among clients and members of their social networks.

Participant Recruitment

Women who participated in VoorZorg, members of their social networks, and VoorZorg nurses were involved as informants in the user-centered design process of Kindle (ie, phases 2, 3, and 4). Participant recruitment started by informing managers of Youth Health Care Organizations executing the VoorZorg program and asking their permission to contact their nurses. The nurses were then informed about the study during a conference. Subsequently, nurses were asked via email whether they were willing to participate and willing to invite their clients for this study. In general, the target population of VoorZorg consists of women who at enrollment are (1) up to 28 weeks of gestation of their first (live born) child, (2) aged <26 years, (3) lower educated, (4) proficient in Dutch, and (5) have minimally one additional risk factor (eg, alcohol or drug use, financial difficulties, domestic violence, and psychosocial symptoms). Less than 1% of the births per Dutch municipality qualify to enroll in VoorZorg [17]. Most of these women (98%) had four or more risk factors [30]. Next, all participating clients were asked, by the researcher (MD), to invite members of their social networks to design sessions.

Inclusion was based on consecutive sampling. Women (ie, clients) were included when they were registered in the VoorZorg program and were in any of the stages of change in smoking cessation [23]. Members of the social networks of these clients were included when they were related to clients meeting the inclusion criteria. For nurses to be included, they must have worked as a VoorZorg nurse in the Netherlands for a minimum of 6 months.

Response and Characteristics

The mean age of the clients (n=9) was 24 years (SD 4.29). Most clients had a child aged <1 year (n=6) or had a child aged >1 year (n=3) and were in the early stages of change in smoking cessation [23], smoking 2.5-20 cigarettes per day. Most pregnancies were unplanned, with one client being pregnant during the study. Clients either had a low educational level (ie, primary education, prevocational secondary education, years 1-3 of higher secondary education, and vocational secondary education level 1) or intermediate level of education (ie, years 4-6 of higher secondary education and no higher vocational or university education) [31]. Participating members of the social networks were partners (n=1), family members or household members (n=1), and friends (n=2). They also had low or intermediate educational levels and were all current smokers, smoking 5-27.5 cigarettes per day. Nurses were all women and, on average, 53 years of age (SD 10.55) with 9 years (SD 2.61) of experience as a VoorZorg nurse, and 37 out of 97 (38%) of their clients were current smokers (Table 1).

XSL•FO
RenderX

Table 1. Sample characteristics of end users.

End users	Age (years), mean (SD)	Stage of change quitting smoking	Number of cigarettes smoked per day, mean (SD)	Planned pregnancy	Age of child (years)	Educational level
Clients	24 (4.29)		7.33 (6.29)			
Client 1		__[a]	—	—	<1	—
Client 2		Contemplation	20	No	<1	Low
Client 3		Action	—	Yes	>1	Low
Client 4		Contemplation	5	No	<1	Intermediate
Client 5		Precontemplation	—	No	<1	Low
Client 6		Contemplation	2.5	No	<1	Intermediate
Client 7		Precontemplation	3.5	Yes	>1	Intermediate
Client 8		Preparation	6.5	No	<1	Intermediate
Client 9		Preparation	10	Yes	>1; pregnant	Intermediate
Social network	31 (12.42)		19.17 (12.33)			
Member 1		Contemplation	25	N/A[b]	N/A	Low
Member 2		—	—	N/A	N/A	—
Member 3		Precontemplation	27,5	N/A	N/A	Low
Member 4		Preparation	5	N/A	N/A	Intermediate
Nurses	53 (10.55)	N/A	N/A	N/A	N/A	N/A

[a]Missing data.

[b]N/A: not applicable.

We intended to recruit 8 clients and members of their social networks in design rounds 2, 4, and 5. Approximately half of the recruited participants (n=9) did not show up at these design rounds (phases 2 and 3). On the basis of the standards in the field of usability end user testing [32,33], we intended and recruited 5 end users to evaluate the usability of each interface of Kindle (participation rate 100%; phase 4). Owing to the geographical disparity of end users, some design rounds entailed multiple sessions at different locations. Approximately half of the participants took part in multiple phases, of which 2 nurses participated in two rounds of phase 3 (Table 2).

XSL•FO
RenderX

Table 2. Sample participation of end users per design phase and round.

End users	Phase 2		Phase 3			Phase 4
	Round 1	Round 2	Round 3	Round 4	Round 5	Usability
Clients						
Client 1		✓				
Client 2		✓		✓		✓
Client 3		✓		✓		✓
Client 4		✓			✓	✓
Client 5		✓			✓	✓
Client 6				✓		
Client 7				✓		✓
Client 8				✓		
Client 9					✓	
Social network						
Member 1		✓		✓		
Member 2		✓				
Member 3		✓				
Member 4				✓		
Nurses						
Nurse 1	✓		✓			✓
Nurse 2	✓		✓		✓	
Nurse 3	✓		✓		✓	
Nurse 4	✓		✓			✓
Nurse 5	✓		✓			✓
Nurse 6					✓	
Nurse 7						✓
Nurse 8						✓

Procedure Design Sessions

All design sessions of Kindle (ie, phases 2, 3, and 4) with end users took place at easily accessible locations (ie, at Waag, at Youth Health Care Organizations, and at the clients' home). Each session was prepared and guided by 1 or 2 designers, whereas 1 or 2 researchers observed sessions but did not actively participate in it. After each round, designers, researchers, and the VoorZorg program representative discussed the output of the sessions in a project group. Researchers advised on how to proceed to the next sessions, assimilating the output of the sessions with evidence-based behavior change techniques. The VoorZorg program representative advised on, in conformance with output of the sessions, VoorZorg organizational requirements and opportunities. Thereafter, the designers prepared the next session with end users.

Clients and members of their social environment received a €10 (US $11.8) gift card for their participation in each session; clients received an extra €5 (US $5.9) for every member of their social environment they brought to the session. Organizations were compensated for their nurses' time participating in the design sessions (€25 [US $29.6] per participating nurse per session). Except for the first design session, this session was part of the nurses' biennial training.

Specification of User Requirements

User requirements were specified in the first two rounds of design inquiry sessions (next to the secondary analysis of our qualitative interviews; phase 1). The first round consisted of nurses (n=51) divided into five user groups. In these sessions, nurse's perspectives on the requirements of the intervention to be developed were gathered. The session started with brainstorming why their clients quit smoking and relapse. Next, nurses were introduced to game elements and were encouraged to think of game elements that might aid their clients' smoking cessation or prevent relapse. Finally, nurses were asked to give their perspectives on user requirements other than the inclusion of specific game elements of the intervention to be developed.

The second round was held with clients (n=5) and members of their social networks (n=3). Participants shared their reasons for smoking and alternative activities they could undertake and were encouraged to translate these activities into an app.

XSL•FO
RenderX

Phase 3: Produce Design Solutions to Meet User Requirements

Preliminary design solutions were created and assessed by end users, starting with an exploration by nurses (n=5; round 3). Next, clients (n=5) and members of their social networks (n=2) assessed the paper mock-up of Kindle (round 4). Finally, an improved paper mock-up was assessed by both nurses (n=3) and clients (n=3; round 5).

Phase 4: Evaluate Against Requirements

Heuristic Evaluation

A total of 3 human factors engineering experts received a guideline containing background information about VoorZorg, why Kindle was developed, its end users, and aims. Moreover, the guideline explained how to install and open the app on smartphones and entailed instructions on performing the usability inspection. No training was given to the experts on how to use the Kindle before the heuristic evaluation. We instructed the experts to systematically evaluate both the nurse and client interfaces of Kindle by freely exploring the functionalities. Experts were asked to describe usability flaws in detail and classify them according to the HIMSS principles [28]. Experts were instructed to rate the severity of the problems they encountered according to the Nielsen five-point Likert severity rating (0 indicating no usability problem to 4 indicating a usability catastrophe) [34]. Experts were encouraged to write comments to further explain the rationale for their ratings.

The results of the usability inspections by all 3 experts were merged according to functionality and HIMSS principles [28]. The average severity score was calculated when multiple experts identified the same usability problem.

Think Aloud Method

All end users (ie, nurses and clients) had prior experience with smartphone apps and were informants during the design phase of Kindle. However, none of the participants had prior experience with the prototype. Data collection lasted 20-50 minutes and took place at clients' homes and at the nurses' workplace at Youth Health Care Organizations. Participants used the smartphones of the evaluators to perform the usability test of Kindle, which was video recorded via a third-party smartphone app (ie, AZ Screen Recorder [by Hecorat Global Technology], downloaded from the Google Play Store that recorded the smartphone screen, audio, and user inputs). We explained to participants that they would use the app by performing specific tasks provided by an evaluator (MD or SS). Participants were instructed to verbalize their thoughts while performing the tasks. Before the actual usability test, participants were given a warm-up task to practice thinking aloud. The warm-up task was to add a specific contact and contact details to the contacts list while thinking aloud. None of the participants had difficulty verbalizing their thoughts during the warm-up task.

The tasks of the usability test were based on real-life scenarios in which testing of all the main functions of the app was covered. The series of tasks were always conducted in a fixed order across participants (Textbox 1). The evaluator reminded participants to continue thinking aloud when they stopped doing so. If a participant was not able to complete a task after three attempts, the evaluator provided a clue.

Textbox 1. Think aloud usability tasks.

Nurse Interface

- Create profile
- Create group
- Manage group
- Manage personal goals of clients
- Use chat functionalities
- Read and add tips (ie, advice)

Client Interface

- Create profile
- Create personal goals
- Use chat functionalities
- Read and add tips (ie, advice)
- Use personal diary

To reveal and describe usability issues in detail, one of the think aloud evaluators (SS) analyzed the videos by coding participants' utterances and user input per task (ie, functionality). A usability issue was reported when a participant was not able to complete the instructed task in her first attempt. The reported usability problems were then categorized according to the HIMSS usability principles by a think aloud evaluator (SS) [28]. Next, the think aloud protocols were merged by the end user group to provide an overview of usability problems per functionality. These were then merged with the findings from the heuristic evaluation, after which all usability problems from the heuristic evaluation and think aloud method were discussed,

XSL•FO
RenderX

and recommendations were made to resolve each usability problem (SS, MD, and MJ).

Results

Outcomes Phase 1: Understanding and Specification of the User Context

A secondary analysis of qualitative interview data revealed that clients generally indicated good relationships with their VoorZorg nurse. During home visits, clients and nurses kept in touch via WhatsApp, with varied intensity (eg, some clients and nurses only communicated concerning appointments and others would regularly ask personal and medical questions via WhatsApp). Most clients interviewed did not have a job nor were they currently enrolled in education. A number of clients did not live on their own but, for example, lived with their parents or in assisted living facilities. Clients had limited social networks and were normally not in contact with other VoorZorg clients. All interviewed clients had a smartphone and access to the internet. Only a few clients had a tablet, laptop, computer, or game console.

Outcomes Phase 2: Specification of User and Organizational Requirements

The user and organizational requirements that we identified during the first two rounds of Kindle's design sessions with nurses, clients, and clients' social networks were divided into design requirements and functionality requirements (Textbox 2).

Textbox 2. User and organizational requirements.

Design

- Mobile health app
- Easy to use, simple use of language, little use of texts, and visualizing content due to lower health literacy of clients
- Social media–like design
- Not necessarily be presented as smoking cessation intervention
- Harmonizing VoorZorg values (ie, no advising, judging, patronizing, pedantic tone, and yet following and endorsing clients)
- App with multiple functionalities to digitalize aspects of the VoorZorg program
- Usable for both nurses and clients
- Nonaddictive or time consuming
- No costs for clients
- Not childish
- Positive focus

Functionalities

- Enabling secured communication between nurses and clients (ie, social support)
- Enabling anonymous communication between clients (ie, peer contact)
- Tailored at early stages of change in smoking cessation (ie, precontemplation, contemplation, and preparation [23])
- Focus on gaining control over life
- Providing a way of dealing with stressors and boredom
- Arousing intrinsic motivation for smoking cessation
- Providing clients with self-understanding and building self-efficacy in clients (ie, social support)
- Rewarding and acknowledging clients' efforts (ie, social support and game element)
- Challenging (game element)
- Earning points or compliments efforts (ie, social support and game element)
- Providing information

Outcomes Phase 3: Production of Design Solutions to Meet User and Organizational Requirements

The design sessions resulted in a preliminary prototype, named *Kindle* (Textbox 3; Figures 1-6), meeting all user and organizational requirements.

XSL•FO
RenderX

Textbox 3. Intervention characteristics of Kindle.

Name

- Kindle

Type of Intervention

- Mobile health app

Aim

- To support women through the first stages of smoking cessation

Targeted Determinants

- Increasing clients' readiness for smoking cessation
- Creating a supportive social network for clients
- Increasing clients' self-efficacy in obtaining personal goals
- Increasing clients' knowledge and self-efficacy (eg, tips)
- Improving communication with nurse (eg, secured chatting)
- Processing difficulties in life (eg, diary)

Setting

- Developed for use in a care setting and at clients' home

Nurse Interface Functionalities

- Set up a profile by entering their name and choosing an avatar representing themselves and choosing from taking a picture with their smartphone camera or an image from their smartphone gallery (Figure 1).
- Manage clients from the admin panel. Nurses can add and delete clients to and from Kindle. Moreover, nurses can block clients from participating in the group chat (Figure 2).
- Endorse and reward clients for their progress in obtaining their goals by assigning hearts (ie, heart-shaped points).
- Communicate with clients via secured private chat and group chat (Figure 3; ie, secured server). All messages in the chat functionality could be loved by tapping a heart-shaped button (ie, similar to the "like" functionality on social media), by which clients were empowered in their contributions to the chat.
- Create tips or moderate tips shared by clients (Figure 4).

Client Interface Functionalities

- Set up a profile by entering their name and choosing an avatar representing themselves and choosing from taking a picture with their smartphone camera or an image from their smartphone gallery (Figure 1).
- Formulate personal goals (ie, "heart desires"), by which they could work on resolving barriers for smoking cessation and build self-efficacy in obtaining personal goals. Women can select a category (ie, being a mother, healthy lifestyle, my child, work and leisure, safety, finances, talking and listening, family and friends, and help) and then enter their personal goal (Figure 5). Clients could enter three active personal goals to work on at the same time.
- View their personal goal attainment progress (ie, 50 hearts represent an obtained goal).
- Communicate with nurse via secured private chat and group chat. All messages in the chat functionality could be "loved."
- Read and create tips by and for other clients.
- Write private posts in their digital diary; clients could also add images to their posts (Figure 6).

Game Elements [35]

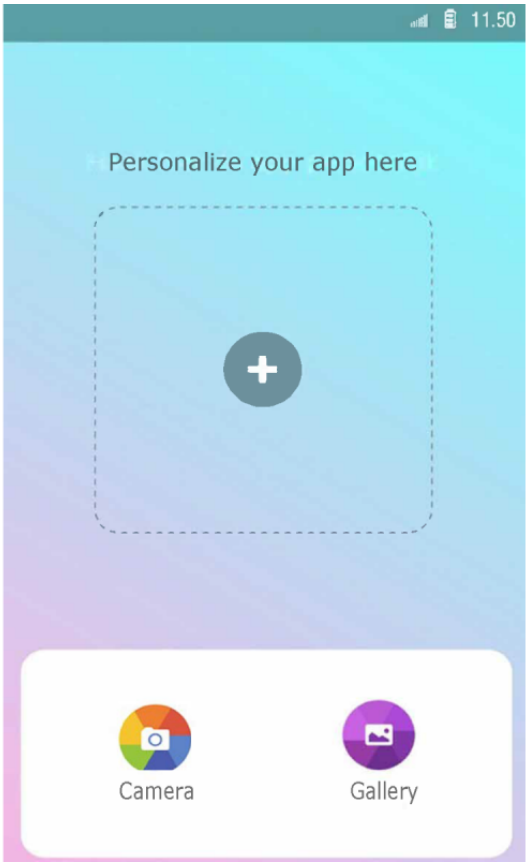
- Avatar creation (ie, setting up profile)
- Player management features (ie, personal goals and progress)
- Intermittent rewards (ie, earning hearts with progress in personal goals)
- Social utility (ie, tips)
- Support network (ie, chat)
- User input (ie, tips and diary)

XSL•FO
RenderX

Development Stage

- Early—the prototype had limited functionality, being a series of screenshots that were linked together via clickable buttons

Figure 1. Example screenshot of the profile creation section where the user can choose an avatar (identical in both interfaces).



XSL•FO
RenderX

Figure 2. Example screenshot of the admin panel in the nurse interface of Kindle.

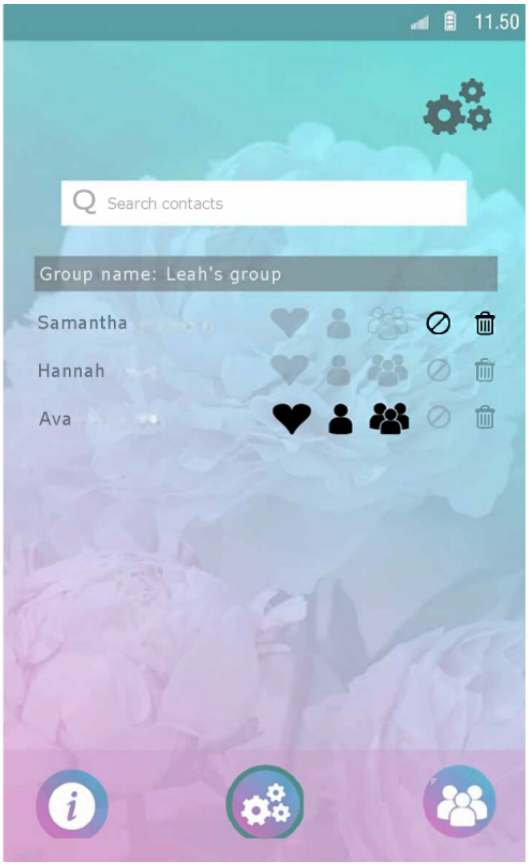


Figure 3. Example screenshot of the group chat functionality in the nurse interface of Kindle.

XSL•FO
RenderX

Figure 4. Example screenshot of the tip functionality of the nurse interface of Kindle.

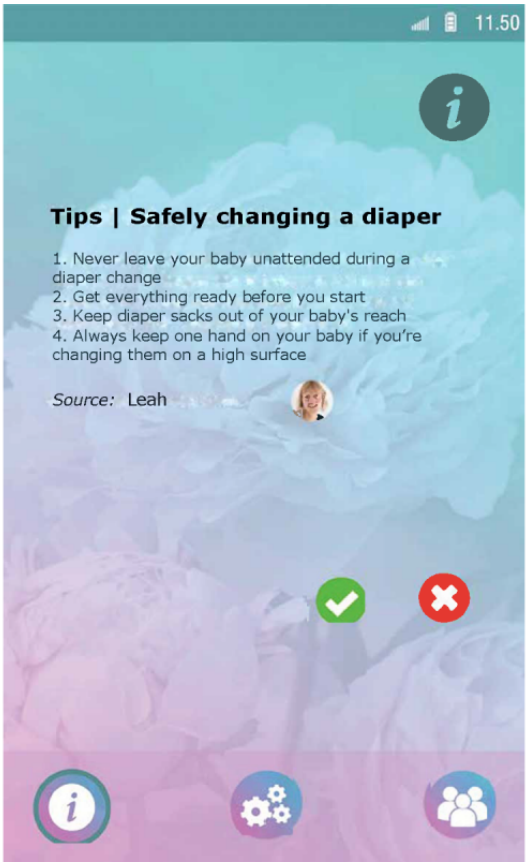
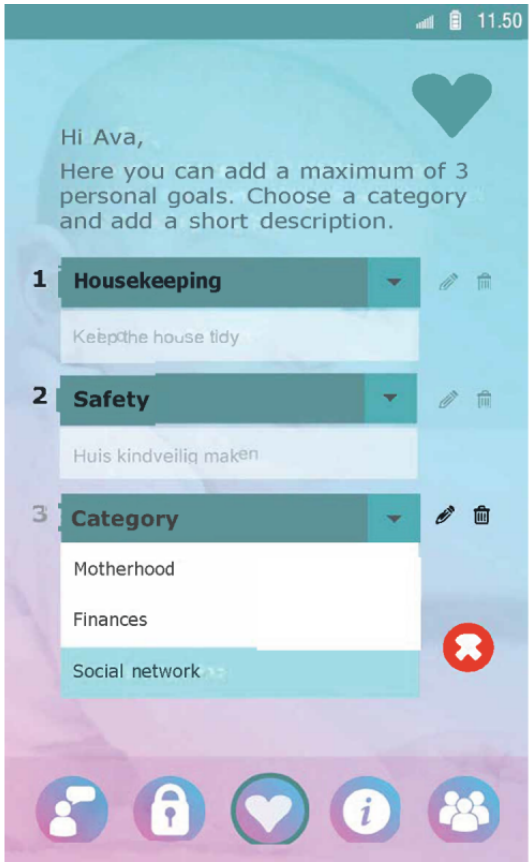
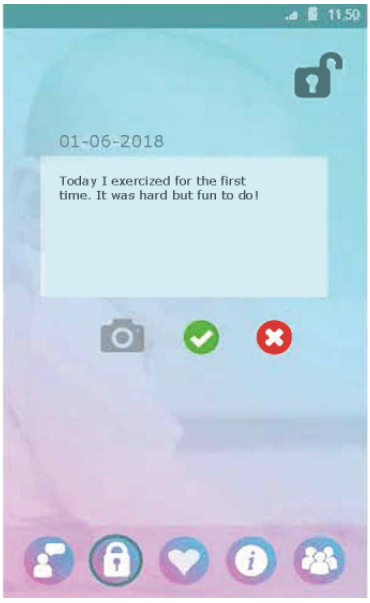


Figure 5. Example screenshot of the goal setting functionality in the client interface of Kindle.



XSL•FO
RenderX

Figure 6. Example screenshot of the diary functionality of the client interface of Kindle.

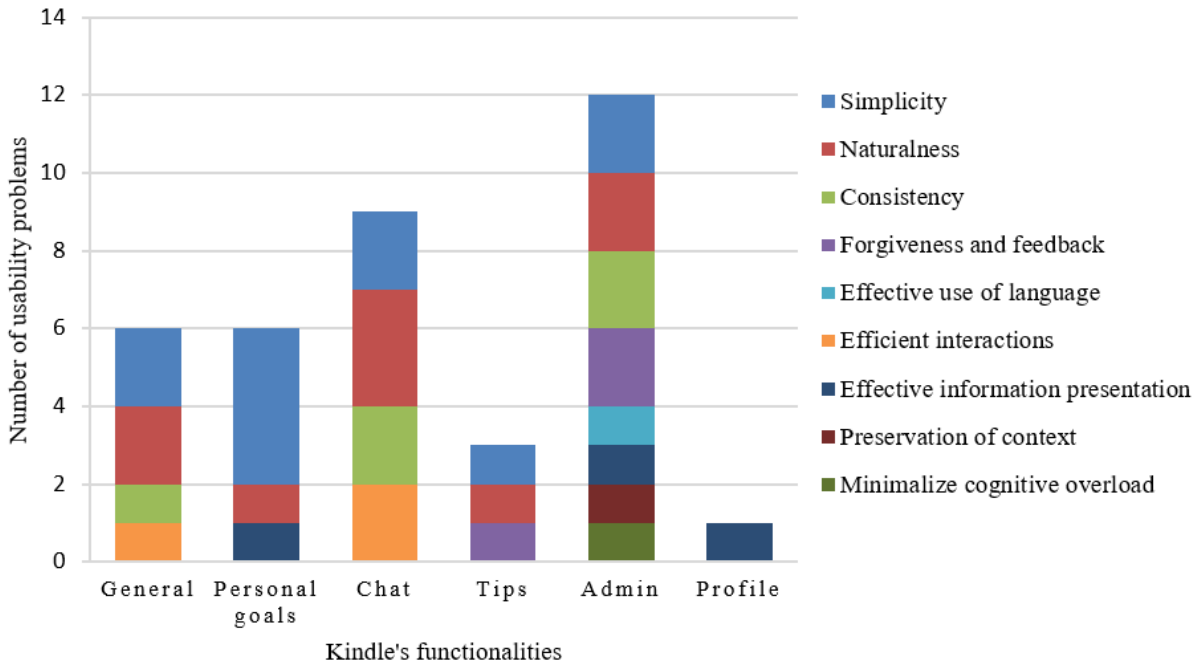


Outcomes Phase 4: Evaluation Against Requirements

Usability Problems Within the Nurse Interface of Kindle

We found 37 usability problems within the nurse interface of Kindle (Figure 7). We identified the general problems and problems related to functionalities. Most usability problems (n=12) were found in the admin functionality (eg, issues of consistency—using the same icon or button with different meanings), followed by the chat function (n=9; eg, issues for efficient interactions—it is unclear when messages are sent). Most usability issues revealed by both evaluation methods concerned violation of the simplicity of the HIMSS principles (eg, the private chat function is hidden in the admin menu) and naturalness (eg, unclear icons). In total, 24 of 37 (65%) potential usability problems were detected in the heuristic evaluation, 7 of 37 (19%) usability problems were detected in the think aloud method, and 6 of 37 (16%) usability problems were detected by both heuristic evaluation and think aloud. The mean severity of usability problems detected through heuristic evaluation was rated rated as 1.8 (SD 1.00), reflecting that the usability problems found by experts were, on average, minor. We provide a complete overview of the usability problems of the nurse interface of Kindle, per the HIMSS principle found through heuristic evaluation and think aloud and provide a recommendation to solve the issue (Multimedia Appendix 1, Table S1).

Figure 7. The number of usability problems within the nurse interface per functionality and principle.

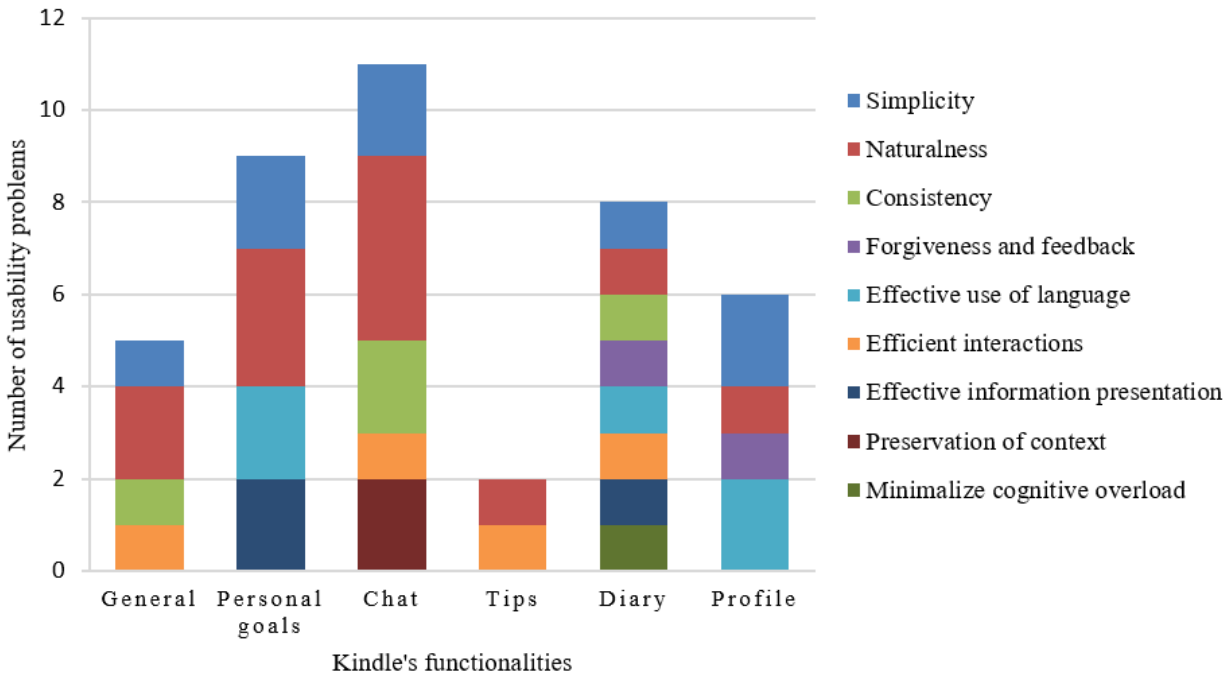


XSL•FO
RenderX

Usability Problems Within the Client Interface of Kindle

In total, 41 usability problems within the client interface of Kindle were discovered (Figure 8). We identified general problems and problems based on their functionality. Most usability problems (n=11) were found in the chat functionality (eg, issue of consistency—the group chat does not have the heart icon next to the input field, unlike the private chat), followed by the personal goals function (n=9; eg, issue of simplicity—the numbers above the golden heart icons are unclear). Most usability issues by both evaluation methods concerned violation of the HIMSS principles simplicity (eg, it is not clear that the heart is a clickable button to give a *like*) and naturalness (eg, it is not clear that the lock icon in the navigation bar represents a diary). In total, 31 of 41 (76%) potential usability problems were detected in the heuristic evaluation, 4 of 41 (10%) usability problems were detected in the think aloud method, and 6 of 41 (15%) usability problems were detected by both heuristic evaluation and think aloud. The mean severity of usability problems detected through heuristic evaluation was rated as 1.8 (SD 0.81), reflecting that the usability problems found by experts were minor. A complete overview of the usability problems of the client interface of Kindle, per the HIMSS principle found through heuristic evaluation and thinking aloud, and a recommendation to solve the issue can be found in Multimedia Appendix 1, Table S2.

Figure 8. The number of usability problems within the client interface per functionality and principle.



Recommendations to Improve the Usability of Kindle

For each usability problem, researchers SS and MD and human factors engineering expert MJ made recommendations to improve the usability of the nurses and client interface of the Kindle prototype (Multimedia Appendix 1, Tables S1 and S2). A final iteration round following the recommendations resulted in a final version of Kindle (Textbox 4; Figures 9-11). This final version was evaluated in a pilot study.

XSL•FO
RenderX

Textbox 4. Final version of intervention characteristics.

Name

- Kindle

Intervention Type

- Mobile health app

Aim

- To support women through the first stages of smoking cessation

Targeted Determinants

- Increasing clients' readiness for smoking cessation
- Creating a supportive social network for clients
- Increasing clients' self-efficacy in obtaining personal goals
- Improving communication with nurse (eg, secured chatting)

Setting

Developed for use in a care setting and at clients' home

Nurse Interface Functionalities

- Manage clients from the admin panel. Nurses can add and delete clients to and from Kindle. Moreover, nurses can block clients from participating in the group.
- Endorse and reward clients for their progress in obtaining their goals by assigning hearts (ie, heart-shaped points).
- Communicate with their clients via a secured private chat and group chat (ie, secured server). All messages in the chat functionality could be "loved" by tapping a heart-shaped button (ie, similar to the "like" functionality on social media), by which clients were empowered in their contributions to the chat.

Client Interface Functionalities

- Formulate personal goals (ie, "heart desires"), by which they could work on resolving barriers for smoking cessation and build self-efficacy in obtaining personal goals. Women can select a category (ie, being a mother, healthy lifestyle, my child, work and leisure, safety, finances, talking and listening, family and friends, and help) and then enter their personal goal.
- View their goal attainment progress (ie, 50 hearts represent an obtained goal).
- Communicate with their nurses via a secured private chat and in a group chat with peers (ie, other clients). All messages in the chat functionality could be "loved."

Game Elements [35]

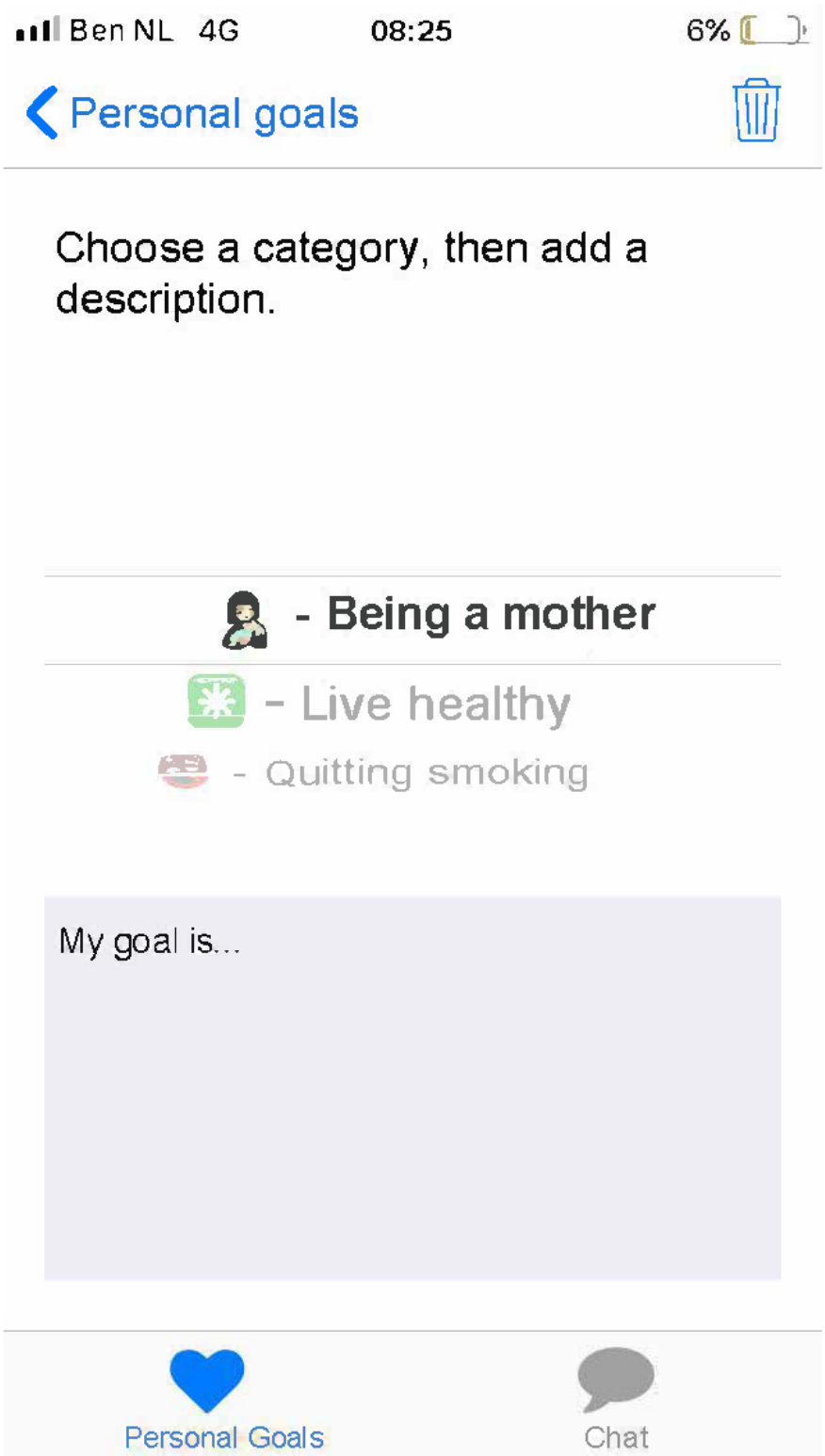
- Player management features (ie, personal goals)
- Intermittent rewards (ie, earning hearts with progress in personal goals)
- Support network (ie, chat)

Development Stage

- Advanced: fully functional for pilot implementation and evaluation.

XSL•FO
RenderX

Figure 9. Example screenshot of the goal setting functionality in the client interface of Kindle.



XSL•FO
RenderX

Figure 10. Example screenshot of the goal setting functionality in the nurse interface of Kindle.

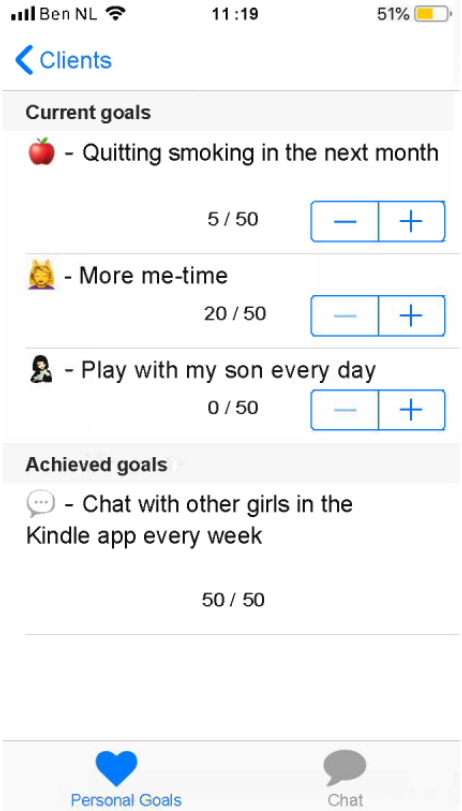
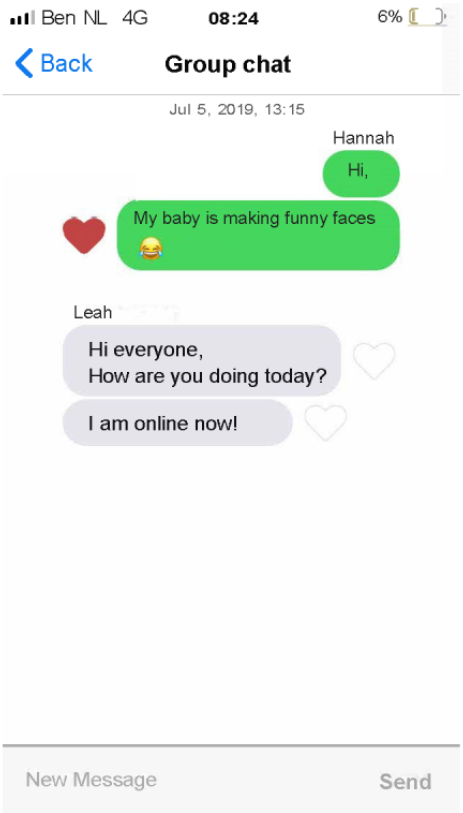


Figure 11. Example screenshot of the group chat functionality in the client interface of Kindle.



XSL•FO
RenderX

Discussion

Principal Findings

In this paper, we describe the user-centered design and usability evaluation of an mHealth app (*Kindle*) that supports disadvantaged young women during and after pregnancy by moving through the first stages of smoking cessation. Disadvantaged women, members of their social networks, and nurses were informants throughout the phases of the iterative prototype design. In the first phase of the intervention design, secondary analysis of qualitative interview data revealed that nurses and clients keep in touch through WhatsApp during home visits and that all interviewed clients had smartphones and internet access but usually possessed no other devices. The clients were not in contact with other clients. In phase 2, we established user and organizational requirements from the secondary interview data and design sessions with end users. The main requirement was that the intervention should be an mHealth app, offering secure communication between nurses and clients. Moreover, the intervention should be tailored to the early stages of change in smoking cessation, include game and social support elements, and have easy-to-use interfaces. In phase 3, the Kindle prototype with game elements was developed technically. Kindle combines a nurse and client interface and includes the following functionalities: personal goal setting with earning points, chat function with a nurse and other clients, tips, and admin function or diary and profile creation. Prototype usability (phase 4) was evaluated by a combination of heuristic evaluation among experts and think aloud sessions among end users (ie, nurses and disadvantaged women). We found 78 usability problems for both interfaces. Most usability problems concerned violation of the principles of *simplicity* and *naturalness* and were found in the chat (both interfaces), admin (nurse interface), and goal setting (client interface) functionalities. Following the recommendations from the usability evaluation, a final iteration round resulted in a final version of Kindle.

Comparison With Prior Work

The first phase of our user-centered design was devoted to understanding and specifying the user context, resulting in a specific focus on the early stages of smoking behavior change [23]. This is in contrast to action stage–oriented smoking cessation apps, which are widely available or being developed [7,36,37]. The use of the transtheoretical model in interventions is associated with positive effects on health behavior [38]. According to an inventory by Paige et al [37], processes of change that aid people in moving through stages of behavioral change are widely applied in mHealth apps for smoking cessation, including Kindle. Moreover, in accordance with other research [8,39-42], the first phase highlighted the importance of supportive social networks for smoking cessation.

By involving both end user groups (clients and nurses), we were able to identify key user and organizational requirements (ie, phase 2) and to incorporate them into Kindle's design solution (ie, phase 3). This coherence of the mHealth app with user objectives or requirements has been identified as one of the critical factors for smoking cessation mHealth apps [43]. Moreover, we involved end users as informants in our design process, which has been found to be more effective in changing behavioral determinants [14]. The involvement of socioeconomically disadvantaged populations thus appears to be a feasible and effective strategy in mHealth design.

One of the user requirements of Kindle that we identified was the use of game elements and social interactions. Previous research has also suggested the potential of social features in serious games for smoking cessation [8]. Moreover, pregnant women in other studies also highlighted the usefulness and playfulness of social interaction functionalities within mHealth [10]. Yet, among smoking cessation mHealth apps, few have a game or social nature similar to that of Kindle [37,44,45]. Kindle further comprises a unique, secured chat functionality, whereas only 16% of mHealth smoking cessation apps integrated web-based communication exclusive to the app [37]. This is striking because peer-to-peer communication and communication with an advisor (eg, health care professional) is associated with more effective eHealth interventions [38,46] and is generally highly preferred among pregnant women in web-based apps [10].

Goal setting and rewards were included as game elements in Kindle. These functionalities are also regularly found in smoking cessation mHealth apps [44,47] and appear to have a significant positive impact on health behavior [38]. However, only 15% of mHealth apps for smoking cessation have a progress tracking feature similar to that of Kindle [44]. As we did not find evidence on the effectiveness of functionalities or game elements concerning sharing or creating tips and keeping a digital diary, these elements were not incorporated in the final version of Kindle.

Aligning interventions to the low health literacy levels of clients was another user and organization requirement. Similar to Kindle, most mHealth apps for smoking cessation incorporated plain usage of language as part of health literacy considerations [37].

Finally, it was also a user and organizational requirement that the app be free of charge for end users. Kindle will be freely available, similar to many other smoking cessation mHealth apps [37,44,45,47].

In our usability evaluation, human factors engineering experts inspected and end users tested for each interface (ie, clients and nurses) of Kindle through heuristic evaluation and think aloud, respectively, which has now become a general practice in usability evaluations [48]. The types of problems detected in our study differed according to each evaluation method. The think aloud method with end users disclosed more critical usability problems, concerning being able to actually use the app as intended, whereas the heuristic evaluation among experts mainly resulted in the disclosure of less severe, noncritical problems, concerning the ease of use of the app or a less optimal user experience. These findings are in accordance with earlier research [49,50] and demonstrate that the combination of expert and user usability methods was truly complementary, and result in surplus value in the design of a usable app [26,49].

XSL•FO
RenderX

Most usability problems were found in the chat functionality of both interfaces, the admin function in the nurse interface, and goal setting in the client interface and concerned *simplicity* and *naturalness* issues. We formulated recommendations to resolve these issues, so that Kindle, similar to other smoking cessation mHealth apps, obtains good scores on functionality and esthetics [45]. According to participants in co-designing a self-management mHealth intervention, an app's usability and intuitiveness might be negatively affected by the inclusion of numerous functionalities [51]. Consequently, in the final iteration round, Kindle is expected to benefit from fewer functionalities, whereas the remaining ones (ie, goal setting and chat) should follow usability standards.

Usability evaluations among disadvantaged populations are scarce [12]. However, disadvantaged populations may reveal unique usability problems in terms of the content and functionalities of interventions [52]. This was also reflected in our study, where disadvantaged women revealed approximately 10% (4/41) of the usability problems with the client interface. Nurses detected more usability problems (approximately 7/37, 19%), yet these were mainly found in the admin function, which was not a functionality of the client interface. This relatively low number of problems might be a positive side effect of our user-centered approach to the design of Kindle, which is expected to resolve potential usability issues in early stage versions of the intervention.

Strengths and Limitations

A strength of our study was the involvement of end users throughout all phases of the user-centered design and usability evaluation of Kindle. The involvement and input of disadvantaged women in the design sessions were highly valuable. In this way, we were able to meet their (and organizational) requirements. This will likely result in higher acceptance of the implementation of Kindle as an intervention and, consequently, is expected to support its effectiveness. Another strength of our study was the triangulation of methods in the fourth phase of Kindle's design. We used two types of usability evaluation methods to detect usability problems in our prototype, which provided us with a more complete overview of usability problems, as only approximately 15% (6/37 and 6/41) of the usability problems found by think aloud and heuristic evaluation overlapped.

Our study also had limitations. First, the design sessions were attended by fewer clients than intended and recruited. The clients often did not show up to a session they had confirmed to attend. Involving disadvantaged populations in research is challenging [5,53]. Moreover, with less than 1% of the births per Dutch municipality qualifying to enroll in VoorZorg, our target population is very small. Nonetheless, we succeeded in fulfilling multiple rounds of design with mixed compositions of participants. In qualitative health research and usability end user tests, smaller sample sizes are acceptable, as they provide higher information power [32,33,54].

Another limitation was the involvement of a limited number of members of women's social networks in most rounds of intervention design. Identical to clients' no shows, we were not able to recruit as many members of social networks as intended. However, during Kindle's design process, we found that existing social networks mainly had a negative role in clients' smoking cessation efforts [22], and clients wanted support from peers.

Moreover, we evaluated the usability of a prototype with limited functionality. The think aloud method was based on certain real-life tasks that covered all the functionalities. This may have highlighted other usability problems that would have occurred during free use of the app. The limited functionality of the prototype may also have resulted in an incomplete insight into usability problems. However, conducting a usability evaluation using heuristic evaluation and think aloud is common in early system design phases, as insights can be used to redesign the system [26].

Practical and Research Implications

Our study adds to the limited existing research following and reporting on all phases of user-centered design of mHealth interventions aimed at disadvantaged populations and a small fraction of studies that report the results of their usability evaluation [48]. Our study indicates that disadvantaged women are capable of participating in all phases of the intervention design. Their input has been valuable in detecting their needs and important usability problems while performing tasks to evaluate Kindle's usability. However, the attendance of disadvantaged women in the design sessions was less than intended and recruited. This implies that more research is needed to gain insight into how disadvantaged populations can be involved in all user-centered design processes and usability evaluations of mHealth interventions aimed at these populations. This may help achieve improved intervention reach, adoption, and implementation among disadvantaged populations.

Another research implication stems from the attendance of multiple clients in the design sessions. This showed the added value of connecting clients with other clients (ie, peers), rather than involving existing social networks of women in the intervention design. The social interactions were positive, supported clients, and inspired both end users and designers to incorporate aspects of these interactions in the design solutions. More research is needed on the effectiveness of such social components of digital interventions on health behavior change.

Practically, we will use the results of this study to pilot test Kindle in the VoorZorg context. Finally, we aim to implement Kindle in the nationwide VoorZorg program and other Dutch care settings that encompass intensive support from health care professionals with disadvantaged clients.

Conclusions

The user-centered design and usability evaluation of Kindle provided valuable insights for improving its first design. By involving health care professionals and socioeconomically disadvantaged, young women during and after their pregnancy (ie, end users), we were able to gain insight into their context, needs, and requirements. Consequently, together with the end users, we were able to meet their requirements to achieve readiness for smoking cessation in our first design solution. We evaluated the usability of the prototype through experts and end users, which revealed unique usability problems for this population. These insights allow for further optimization of

XSL•FO
RenderX

Kindle, and we encourage future studies to engage disadvantaged populations in mHealth intervention design and usability testing.

Acknowledgments

This work was supported by the Fonds NutsOhra-Healthy Future Nearby (grant 102253). The authors thank Professor Anton Kunst for his contributions to the research project and his critical review of the manuscript. The authors thank the Netherlands Centre for Preventive Youth Health, Youth Healthcare Organizations and Waag for their cooperation and contribution to this study.

Authors' Contributions

MD contributed to conceptualization, data curation, methodology, formal analysis, investigation, software, validation, visualization, and writing of original draft preparation. MJ contributed to conceptualization; methodology; resources; supervision; validation; and writing, reviewing, and editing the manuscript. SS contributed to formal analysis; investigation; data curation; and writing, reviewing, and editing the manuscript. MF contributed to conceptualization, methodology, resources, funding acquisition, project administration, supervision, validation, and reviewing and editing the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Overview of usability problems per user interface.
[DOCX File , 44 KB - formative_v5i8e24112_app1.docx]

References

1. Core Indicators of the Health and Care of Pregnant Women and Babies in Europe in 2015. European Perinatal Health Report. 2018. URL: https://www.europeristat.com/images/EPHR2015_Euro-Peristat.pdf [accessed 2021-07-13]
2. The Health Consequences of Smoking? 50 Years of Progress: a Report of the Surgeon General. Atlanta, GA: US Department of Health and Human Services. Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 2014.
3. Smedberg J, Lupattelli A, Mårdby AC, Nordeng H. Characteristics of women who continue smoking during pregnancy: a cross-sectional study of pregnant women and new mothers in 15 European countries. BMC Pregnancy Childbirth 2014 Jun 25;14:213 [FREE Full text] [doi: 10.1186/1471-2393-14-213] [Medline: 24964728]
4. Kondracki A. Prevalence and patterns of cigarette smoking before and during early and late pregnancy according to maternal characteristics: the first national data based on the 2003 birth certificate revision, United States, 2016. Reprod Health 2019 Sep 13;16(1):142 [FREE Full text] [doi: 10.1186/s12978-019-0807-5] [Medline: 31519184]
5. Tobacco and Inequities: Guidance for Addressing Inequities in Tobacco-related Harm. World Health Organisation. 2014. URL: http://www.euro.who.int/__data/assets/pdf_file/0005/247640/tobacco-090514.pdf [accessed 2021-07-13]
6. Marcolino MS, Oliveira JAQ, D'Agostino M, Ribeiro AL, Alkmim MB, Novillo-Ortiz D. The impact of mhealth interventions: systematic review of systematic reviews. JMIR Mhealth Uhealth 2018 Jan 17;6(1):e23 [FREE Full text] [doi: 10.2196/mhealth.8873] [Medline: 29343463]
7. Do HP, Tran BX, Le Pham Q, Nguyen LH, Tran TT, Latkin CA, et al. Which eHealth interventions are most effective for smoking cessation? A systematic review. Patient Prefer Adherence 2018;12:2065-2084 [FREE Full text] [doi: 10.2147/PPA.S169397] [Medline: 30349201]
8. Derksen ME, van Strijp S, Kunst AE, Daams J, Jaspers MW, Fransen MP. Serious games for smoking prevention and cessation: a systematic review of game elements and game effects. J Am Med Inform Assoc 2020 May 1;27(5):818-833 [FREE Full text] [doi: 10.1093/jamia/ocaa013] [Medline: 32330255]
9. Wallwiener S, Müller M, Doster A, Laserer W, Reck C, Pauluschke-Fröhlich J, et al. Pregnancy eHealth and mHealth: user proportions and characteristics of pregnant women using Web-based information sources-a cross-sectional study. Arch Gynecol Obstet 2016 Nov;294(5):937-944. [doi: 10.1007/s00404-016-4093-y] [Medline: 27084763]
10. Goetz M, Müller M, Matthies LM, Hansen J, Doster A, Szabo A, et al. Perceptions of patient engagement applications during pregnancy: a qualitative assessment of the patient's perspective. JMIR Mhealth Uhealth 2017 May 26;5(5):e73 [FREE Full text] [doi: 10.2196/mhealth.7040] [Medline: 28550005]
11. Stowell E, Lyson M, Saksono H, Wurth R, Jimison H, Pavel M, et al. Designing and Evaluating Mhealth Interventions for Vulnerable Populations: a Systematic Review. In: CHI Conference on Human Factors in Computing Systems. 2018 Presented at: CHI'18; April 21–26, 2018; Montréal, Canada. [doi: 10.1145/3173574.3173589]

XSL•FO
RenderX

12. Mayberry LS, Lyles CR, Oldenburg B, Osborn CY, Parks M, Peek ME. Mhealth interventions for disadvantaged and vulnerable people with type 2 diabetes. Curr Diab Rep 2019 Nov 25;19(12):148 [FREE Full text] [doi: 10.1007/s11892-019-1280-9] [Medline: 31768662]
13. Eyles H, Jull A, Dobson R, Firestone R, Whittaker R, Te Morenga L, et al. Co-design of mhealth delivered interventions: a systematic review to assess key methods and processes. Curr Nutr Rep 2016 Jul 4;5(3):160-167. [doi: 10.1007/s13668-016-0165-7]
14. DeSmet A, Thompson D, Baranowski T, Palmeira A, Verloigne M, de Bourdeaudhuij I. Is participatory design associated with the effectiveness of serious digital games for healthy lifestyle promotion? A meta-analysis. J Med Internet Res 2016 Apr 29;18(4):e94 [FREE Full text] [doi: 10.2196/jmir.4444] [Medline: 27129447]
15. Weinmann T, Radon K, Sukalla F, Gerlich J, Barth S, Nowak D, et al. Developing an awareness campaign to reduce second hand smoke among disadvantaged families-a participatory m-health approach. Int J Environ Res Public Health 2018 Sep 6;15(9):1945 [FREE Full text] [doi: 10.3390/ijerph15091945] [Medline: 30200621]
16. Gordon M, Henderson R, Holmes JH, Wolters MK, Bennett IM, SPIRIT (Stress in Pregnancy: Improving Results with Interactive Technology) Group. Participatory design of ehealth solutions for women from vulnerable populations with perinatal depression. J Am Med Inform Assoc 2016 Jan 5;23(1):105-109 [FREE Full text] [doi: 10.1093/jamia/ocv109] [Medline: 26342219]
17. Mejdoubi J, van den Heijkant S, Struijf E, van Leerdam F, HiraSing R, Crijnen A. Risicofactoren voor kindermishandeling bij jonge hoogrisicozwangeren: design van het VoorZorgonderzoek. Jeugdgezondszorg-tijdschr 2013 May 1;45(2):26-31. [doi: 10.1007/s12452-013-0007-6]
18. Olds D. The nurse-family partnership: an evidence-based preventive intervention. Infant Ment Health J 2006 Jan;27(1):5-25. [doi: 10.1002/imhj.20077] [Medline: 28640426]
19. Robling M, Bekkers M, Bell K, Butler CC, Cannings-John R, Channon S, et al. Effectiveness of a nurse-led intensive home-visitation programme for first-time teenage mothers (Building Blocks): a pragmatic randomised controlled trial. Lancet 2016 Jan;387(10014):146-155. [doi: 10.1016/s0140-6736(15)00392-x]
20. Mejdoubi J, van den Heijkant SC, van Leerdam FJ, Crone M, Crijnen A, HiraSing RA. Effects of nurse home visitation on cigarette smoking, pregnancy outcomes and breastfeeding: a randomized controlled trial. Midwifery 2014 Jun;30(6):688-695. [doi: 10.1016/j.midw.2013.08.006] [Medline: 24041564]
21. Derksen ME, Kunst AE, Jaspers MW, Fransen MP. Barriers experienced by nurses providing smoking cessation support to disadvantaged, young women during and after pregnancy. Health Soc Care Community 2019 Nov 25;27(6):1564-1573 [FREE Full text] [doi: 10.1111/hsc.12828] [Medline: 31446634]
22. Derksen ME, Kunst AE, Murugesu L, Jaspers MW, Fransen MP. Smoking cessation among disadvantaged young women during and after pregnancy: exploring the role of social networks. Midwifery 2021 Jul;98:102985 [FREE Full text] [doi: 10.1016/j.midw.2021.102985] [Medline: 33761432]
23. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. Am J Health Promot 1997 Sep;12(1):38-48. [doi: 10.4278/0890-1171-12.1.38] [Medline: 10170434]
24. Brender J, Talmon J, de Keizer N, Nykänen P, Rigby M, Ammenwerth E. Stare-hi - statement on reporting of evaluation studies in health informatics: explanation and elaboration. Appl Clin Inform 2013;4(3):331-358 [FREE Full text] [doi: 10.4338/ACI-2013-04-RA-0024] [Medline: 24155788]
25. Ergonomics of Human-system Interaction-part 11: Usability: Definitions and Concepts, 2018. ISO. URL: https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en [accessed 2021-07-13]
26. Jaspers MW. A comparison of usability methods for testing interactive health technologies: methodological aspects and empirical evidence. Int J Med Inform 2009 May;78(5):340-353. [doi: 10.1016/j.ijmedinf.2008.10.002] [Medline: 19046928]
27. Nielsen J, Molich R. Heuristic Evaluation of User Interfaces. In: Proceedings of the SIG CHI conference on Human factors in computing systems. 1990 Presented at: CHI'90; April 1-5, 1990; Seattle Washington USA. [doi: 10.1145/97243.97281]
28. Belden J, Grayson R, Barnes J. Defining and Testing EMR Usability: Principles and Proposed Methods of Emr Usability Evaluation and Rating. Healthcare Information and Management Systems Society (HIMSS). 2009. URL: http://hdl.handle.net/10355/3719 [accessed 2021-07-13]
29. Boren T, Ramey J. Thinking aloud: reconciling theory and practice. IEEE Trans Profess Commun 2000;43(3):261-278. [doi: 10.1109/47.867942]
30. Mejdoubi J, Heijkant SV, Struijf E, Leerdam FV, Crijnen A, HiraSing R. The identification of pregnant women at risk for child abuse: methodology. Gynecol Obstet Res Open J 2015 May 12;2(1):18-25. [doi: 10.17140/GOROJ-2-105]
31. Opleidingsniveau. Statistics Netherlands (CBS). 2019. URL: https://www.cbs.nl/nl-nl/nieuws/2019/33/verschil-levensverwachting-hoog-en-laagopgeleid-groeit/opleidingsniveau [accessed 2021-07-13]
32. Nielsen J, Landauer T. 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: INTERACT'93; April 24-29, 1993; Amsterdam, The Netherlands URL: https://doi.org/10.1145/169059.169166 [doi: 10.1145/169059.169166]
33. Nielsen J. How Many Test Users in a Usability Study? NN Group. 2012. URL: https://www.nngroup.com/articles/how-many-test-users/ [accessed 2021-07-13]

XSL•FO
RenderX

34. Nielsen J. Severity Ratings for Usability Problems. Papers and Essays. 1995. URL: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwi4ym3ePxAhWKaCsKHU6RCKYQFnoECAMQAA&url=https%3A%2F%2Fwww.nngroup.com%2Farticles%2Fhow-to-rate-the-severity-of-usability-problems%2F&usg=AOvVaw0oh_cFDUuiOjEWHMS6nccp [accessed 2021-07-15]
35. King D, Delfabbro P, Griffiths M. Video game structural characteristics: a new psychological taxonomy. Int J Ment Health Addiction 2009 Apr 7;8(1):90-106. [doi: 10.1007/s11469-009-9206-4]
36. Oosterveen E, Tzelepis F, Ashton L, Hutchesson MJ. A systematic review of eHealth behavioral interventions targeting smoking, nutrition, alcohol, physical activity and/or obesity for young adults. Prev Med 2017 Jun;99:197-206. [doi: 10.1016/j.ypmed.2017.01.009] [Medline: 28130046]
37. Paige S, Alber J, Stellefson M, Krieger J. Missing the mark for patient engagement: mHealth literacy strategies and behavior change processes in smoking cessation apps. Patient Educ Couns 2018 May;101(5):951-955 [FREE Full text] [doi: 10.1016/j.pec.2017.11.006] [Medline: 29153592]
38. Webb T, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010 Feb 17;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]
39. Blok DJ, de Vlas SJ, van Empelen P, van Lenthe FJ. The role of smoking in social networks on smoking cessation and relapse among adults: a longitudinal study. Prev Med 2017 Jun;99:105-110 [FREE Full text] [doi: 10.1016/j.ypmed.2017.02.012] [Medline: 28216381]
40. Simons-Morton BG, Farhat T. Recent findings on peer group influences on adolescent smoking. J Prim Prev 2010 Aug;31(4):191-208 [FREE Full text] [doi: 10.1007/s10935-010-0220-x] [Medline: 20614184]
41. Phua J. The reference group perspective for smoking cessation: an examination of the influence of social norms and social identification with reference groups on smoking cessation self-efficacy. Psychol Addict Behav 2013 Mar;27(1):102-112. [doi: 10.1037/a0029130] [Medline: 22732053]
42. Naslund JA, Kim SJ, Aschbrenner KA, McCulloch LJ, Brunette MF, Dallery J, et al. Systematic review of social media interventions for smoking cessation. Addict Behav 2017 Oct;73:81-93 [FREE Full text] [doi: 10.1016/j.addbeh.2017.05.002] [Medline: 28499259]
43. El-Hilly AA, Iqbal SS, Ahmed M, Sherwani Y, Muntasir M, Siddiqui S, et al. Game on? Smoking cessation through the gamification of mhealth: a longitudinal qualitative study. JMIR Serious Games 2016 Oct 24;4(2):e18 [FREE Full text] [doi: 10.2196/games.5678] [Medline: 27777216]
44. Rajani NB, Weth D, Mastellos N, Filippidis FT. Use of gamification strategies and tactics in mobile applications for smoking cessation: a review of the UK mobile app market. BMJ Open 2019 Jun 17;9(6):e027883 [FREE Full text] [doi: 10.1136/bmjopen-2018-027883] [Medline: 31213452]
45. Lee J, Dallery J, Laracuente A, Ibe I, Joseph S, Huo J, et al. A content analysis of free smoking cessation mobile applications in the USA. J Smok Cessat 2019 Apr 23;14(4):195-202. [doi: 10.1017/jsc.2019.6]
46. Morrison LG, Yardley L, Powell J, Michie S. What design features are used in effective e-health interventions? A review using techniques from Critical Interpretive Synthesis. Telemed J E Health 2012 Mar;18(2):137-144. [doi: 10.1089/tmj.2011.0062] [Medline: 22381060]
47. Thornton L, Quinn C, Birrell L, Guillaumier A, Shaw B, Forbes E, et al. Free smoking cessation mobile apps available in Australia: a quality review and content analysis. Aust N Z J Public Health 2017 Dec;41(6):625-630. [doi: 10.1111/1753-6405.12688] [Medline: 28749591]
48. 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]
49. Peute LW, Jaspers MW. Usability evaluation of a laboratory order entry system: cognitive walkthrough and think aloud combined. Stud Health Technol Inform 2005;116:599-604. [Medline: 16160323]
50. Yen P, Bakken S. A comparison of usability evaluation methods: heuristic evaluation versus end-user think-aloud protocol - an example from a web-based communication tool for nurse scheduling. AMIA Annu Symp Proc 2009 Nov 14;2009:714-718 [FREE Full text] [Medline: 20351946]
51. Jessen S, Mirkovic J, Ruland CM. Creating gameful design in mhealth: a participatory co-design approach. JMIR Mhealth Uhealth 2018 Dec 14;6(12):e11579 [FREE Full text] [doi: 10.2196/11579] [Medline: 30552080]
52. 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]
53. Moore LW, Miller M. Initiating research with doubly vulnerable populations. J Adv Nurs 1999 Nov;30(5):1034-1040. [doi: 10.1046/j.1365-2648.1999.01205.x] [Medline: 10564401]
54. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. Qual Health Res 2016 Nov;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]

XSL•FO
RenderX

Abbreviations

HIMSS: Healthcare Information and Management Systems Society
mHealth: mobile health

Edited by G Eysenbach; submitted 04.09.20; peer-reviewed by M Wallwiener, I Schiering; comments to author 06.10.20; revised version received 26.11.20; accepted 31.05.21; published 04.08.21.

Please cite as:
Derksen ME, Jaspers MWM, van Strijp S, Fransen MP
Mobile Health for Smoking Cessation Among Disadvantaged Young Women During and After Pregnancy: User-Centered Design and Usability Study
JMIR Form Res 2021;5(8):e24112
URL: https://formative.jmir.org/2021/8/e24112
doi:10.2196/24112
PMID:34346895

©Marloes E Derksen, Monique WM Jaspers, Sander van Strijp, Mirjam P Fransen. Originally published in JMIR Formative Research (https://formative.jmir.org), 04.08.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Mobile App (WHEELS) to Promote a Healthy Lifestyle in Wheelchair Users With Spinal Cord Injury or Lower Limb Amputation: Usability and Feasibility Study

Dirk Hoevenaars[1,2*], MSc; Jasmijn F M Holla[2,3,4*], PhD; Leonie te Loo[3], MSc; Johan M Koedijker[3], PhD; Sarah Dankers[5], MSc; Han Houdijk[6,7], PhD; Bart Visser[4,5], PhD; Thomas W J Janssen[1,2,4], PhD; Sonja de Groot[1,2,4], PhD; Marije Deutekom[3,4,8], PhD; WHEELS Study Group[9]

[1]Faculty of Behavioural and Movement Sciences, Vrije Universiteit, Amsterdam, Netherlands
[2]Amsterdam Rehabilitation Research Center, Reade, Amsterdam, Netherlands
[3]Faculty of Health, Sports and Social Work, Inholland University of Applied Sciences, Haarlem, Netherlands
[4]Center for Adapted Sports Amsterdam, Amsterdam Institute of Sport Science, Amsterdam, Netherlands
[5]Faculty of Health, Amsterdam University of Applied Sciences, Amsterdam, Netherlands
[6]Center for Human Movement Sciences, University Medical Center Groningen, University of Groningen, Groningen, Netherlands
[7]Department of Research and Development, Heliomare Rehabilitation Center, Wijk aan Zee, Netherlands
[8]Faculty of Sports and Nutrition, Amsterdam University of Applied Sciences, Amsterdam, Netherlands
[9]see Acknowledgments
[*]these authors contributed equally

Corresponding Author:
Jasmijn F M Holla, PhD
Faculty of Health, Sports and Social Work
Inholland University of Applied Sciences
Blijdorplaan 15
Haarlem, 2015 CE
Netherlands
Phone: 31 621115227
Email: jasmijn.holla@inholland.nl

Abstract

Background: Maintaining a healthy lifestyle is important for wheelchair users' well-being, as it can have a major impact on their daily functioning. Mobile health (mHealth) apps can support a healthy lifestyle; however, these apps are not necessarily suitable for wheelchair users with spinal cord injury or lower limb amputation. Therefore, a new mHealth app (WHEELS) was developed to promote a healthy lifestyle for this population.

Objective: The objectives of this study were to develop the WHEELS mHealth app, and explore its usability, feasibility, and effectiveness.

Methods: The WHEELS app was developed using the intervention mapping framework. Intervention goals were determined based on a needs assessment, after which behavior change strategies were selected to achieve these goals. These were applied in an app that was pretested on ease of use and satisfaction, followed by minor adjustments. Subsequently, a 12-week pre-post pilot study was performed to explore usability, feasibility, and effectiveness of the app. Participants received either a remote-guided or stand-alone intervention. Responses to semistructured interviews were analyzed using content analysis, and questionnaires (System Usability Score [SUS], and Usefulness, Satisfaction, and Ease) were administered to investigate usability and feasibility. Effectiveness was determined by measuring outcomes on physical activity, nutrition, sleep quality (Pittsburgh Sleep Quality Index), body composition, and other secondary outcomes pre and post intervention, and by calculating effect sizes (Hedges *g*).

Results: Sixteen behavior change strategies were built into an app to change the physical activity, dietary, sleep, and relaxation behaviors of wheelchair users. Of the 21 participants included in the pilot study, 14 participants completed the study. The interviews and questionnaires showed a varied user experience. Participants scored a mean of 58.6 (SD 25.2) on the SUS questionnaire, 5.4 (SD 3.1) on ease of use, 5.2 (SD 3.1) on satisfaction, and 5.9 (3.7) on ease of learning. Positive developments in body composition were found on waist circumference (P=.02, g=0.76), fat mass percentage (P=.004, g=0.97), and fat-free mass percentage (P=.004,

XSL•FO
RenderX

g=0.97). Positive trends were found in body mass (P=.09, g=0.49), BMI (P=.07, g=0.53), daily grams of fat consumed (P=.07, g=0.56), and sleep quality score (P=.06, g=0.57).

Conclusions: The WHEELS mHealth app was successfully developed. The interview outcomes and usability scores are reasonable. Although there is room for improvement, the current app showed promising results and seems feasible to deploy on a larger scale.

(JMIR Form Res 2021;5(8):e24909) doi:10.2196/24909

KEYWORDS

mHealth; mobile app; lifestyle; usability; feasibility; wheelchair users; spinal cord injury; lower limb amputation

Introduction

A healthy lifestyle is known to be beneficial for a person's well-being and happiness in many ways [1]. Healthy lifestyle behavior is especially important for wheelchair users owing to its major impact on their daily level of functioning [2]. Nevertheless, physical inactivity, obesity, and low vitality are common among wheelchair users with spinal cord injury (SCI) or lower limb amputation (LLA), which increase the risk of secondary health problems such as cardiovascular disease and can cause a reduced quality of life [3-9]. Therefore, it is important for wheelchair users to be encouraged to achieve and maintain a healthy lifestyle during and after inpatient rehabilitation [10-13].

Despite encouragement during inpatient rehabilitation, it appears to be difficult for wheelchair users with SCI or LLA to adopt or maintain the recommended physical activity levels and healthy diet after discharge [12,14-16]. Environmental factors such as fitness centers not having accessible toilets and personal factors such as lack of knowledge and motivation can play a role in this lack of physical activity. One of the barriers for maintaining a healthy lifestyle is the lack of professional guidance after discharge. Wheelchair users with SCI or LLA clearly express their need for such support [17,18], and guided interventions have shown positive effects for behavioral change and maintenance [19-21]. To save costs and time, mobile health (mHealth) tools could be used to support the professional in providing this additional guidance, which has been shown to be an effective method for changing behavior [22]. mHealth tools focus mainly on personal determinants of behavior. Although they cannot remove existing barriers in the physical or societal environment, they can increase the knowledge and skills to cope with these barriers.

mHealth provides the opportunity to stimulate, support, and monitor a healthy lifestyle at the individual and group levels [11]. Given that smartphones have become an integral part of our lives, mHealth seems to be a promising tool supporting healthy changes in physical activity, sedentary behavior, diet, relaxation, and sleep. A healthy lifestyle refers to the combination of healthy physical activity with appropriate dietary, relaxation, and sleep behaviors. Successful self-management apps have been developed for people with chronic conditions [23,24]. People tend to value mHealth apps that support goal-setting, and provide information and advice, feedback, self-monitoring tools, social support, and reinforcement [25]. The use of an appropriate combination of techniques to change lifestyle-related determinants mediates the potential effectiveness of an mHealth app [26,27]. However, determinants of physical activity, nutrition, and sleep can vary among populations and are different in individuals with a disability [28]. To date, there is no mHealth app designed specifically for wheelchair users with SCI or LLA to target all of these behaviors simultaneously. Therefore, in the Wheelchair ExercisE and Lifestyle Study (WHEELS) project, an existing lifestyle app for healthy able-bodied people was adapted for wheelchair users with SCI and LLA based on the intervention mapping protocol [29].

Targeting behavior change specifically for wheelchair users includes overcoming additional social barriers, which was taken into account during development of the WHEELS app [30]. An intervention targeting the combination of physical activity and dietary behavior seems to be superior to an intervention targeting physical activity or diet alone in weight management and improving health [31,32]. Because poor sleep quality and lack of sleep have a negative association with weight regulation, it seems to be of added value to also target resting and sleep behavior [33-35]. Given the positive relations between physical activity, diet, and sleep behavior, as well as combining multiple healthy lifestyle features, a lifestyle app was designed in which physical activity, dietary, sleep, and resting behaviors were targeted simultaneously [35-37].

To evaluate this combined lifestyle app, a usability and effectiveness study was performed in wheelchair users with SCI or LLA. A multicomponent intervention, which is a combination of different intervention components such as an app combined with counseling, seems to be more effective than a stand-alone app intervention [25]. This raised the question as to whether this would also be the case among wheelchair users. Therefore, both a stand-alone and a remote-guided version of the mHealth intervention were applied during the intervention period. In the remote-guided version, personal guidance from a lifestyle coach was offered throughout the intervention period. The aims of this study were to: (1) describe the development of the WHEELS mHealth app; and (2) explore its usability, feasibility, and effectiveness.

Methods

Development of the mHealth Intervention Using Intervention Mapping

The WHEELS lifestyle app was developed using the intervention mapping framework for planning theory- and evidence-based health promotion programs [38]. This framework consists of six steps: (1) perform a needs assessment and state

XSL•FO
RenderX

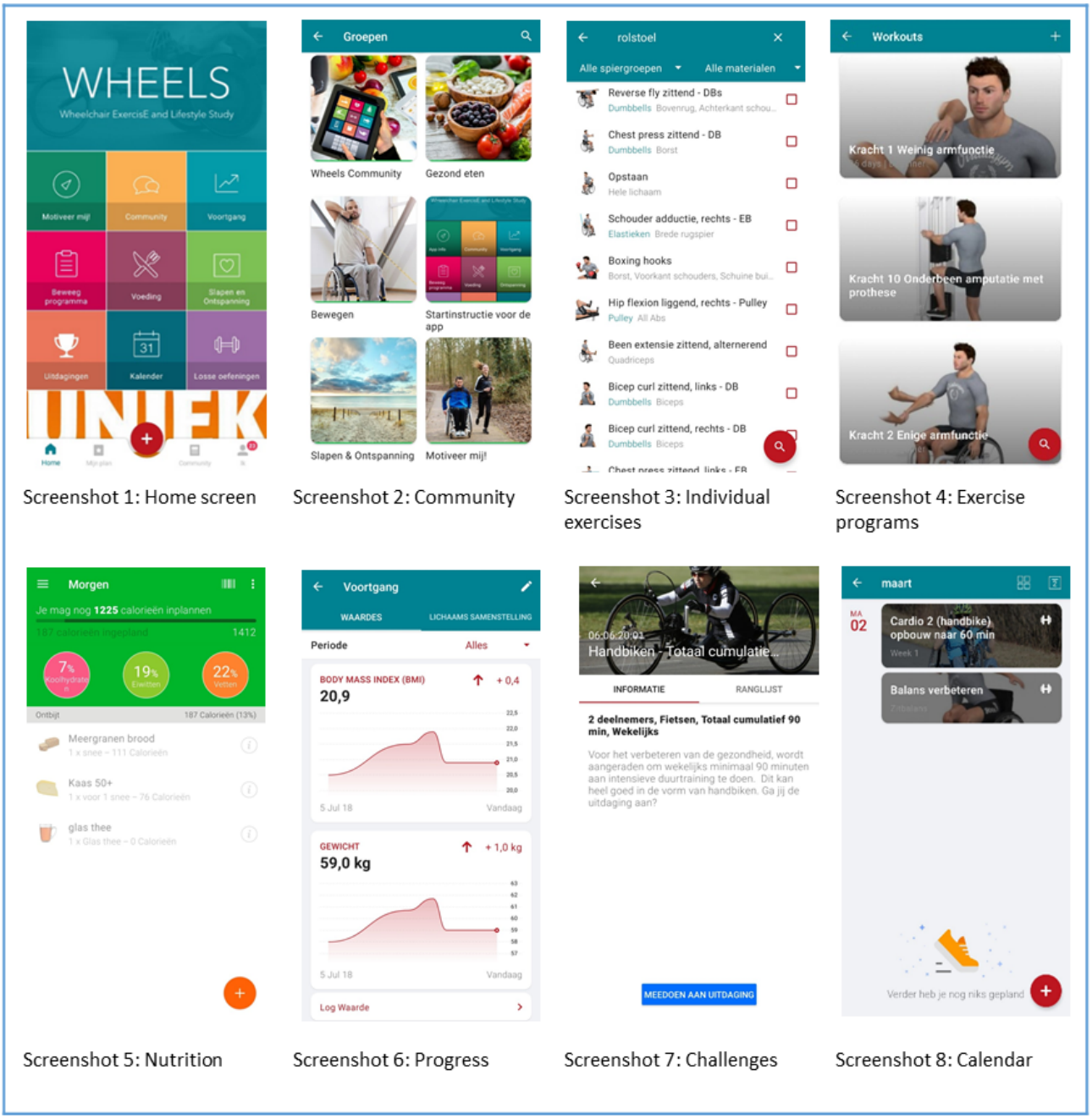
intervention goals; (2) construct matrices of change objectives; (3) choose theory- and evidence-based behavior change methods and practical applications to deliver them; (4) pretest, refine, and produce the program; (5) develop an implementation plan; and (6) create an evaluation plan. These six intervention mapping steps to plan a mobile lifestyle intervention for wheelchair users with SCI or LLA, focusing particularly on steps 1 to 4, are presented in detail in Multimedia Appendix 1.

App Description and Content

The WHEELS app is targeted toward wheelchair users with SCI or LLA to help them comply with the scientific exercise guidelines for adults with SCI [7], achieve a healthy energy balance, and achieve a healthy balance between exercise and sleep/relaxation. In the app, wheelchair users are guided to these intervention goals by providing them knowledge and a format for setting personally meaningful subgoals and the functionalities described below. To work with the app, the user creates an account that they can personalize. Personal characteristics are used to provide feedback within the app (eg, height, body mass, and activity level are used to estimate resting energy expenditure). When logging in, users can navigate through the app from the home screen as shown in screenshot 1 in Figure 1. From the home screen, users can navigate to the "Individual exercises" and "Exercise program" tiles where exercises can be performed and scheduled in their personal agenda (Figure 1, screenshots 3 and 4). In the "Food" part, users receive an overview of their energy balance based on their nutrition intake and energy expenditure (Figure 1, screenshot 5). Nutrition plans and goals can be created (eg, losing weight) in which the app would guide the user through some steps toward a reasonable nutrition plan, resulting in a suggested daily energy intake. In the "Sleep & Relaxation" environment, exercises and knowledge are offered on balancing physical and mental load and relaxation. Behind the "Progress" tile, users are able to obtain insight and track changes in predefined health and fitness parameters such as weight and BMI. In addition, the "Community" section includes start instructions and four groups (exercise, nutrition, sleep and relaxation, lifestyle change tool) in which information and tips are posted over time (Figure 1, screenshot 2). A fifth group allows users to ask questions, share experiences and tips, and interact with other users. Finally, users are able join various challenges in which they can compare performances with each other (eg, a weekly 90-minute handcycle challenge) (Figure 1, screenshot 7).

XSL•FO
RenderX

Figure 1. Structure overview of the WHEELS app.



Usability and Feasibility Study

Participants

To evaluate the usability and feasibility of the app, potential end users were invited to participate in the study. Recruitment took place by advertisement at patient associations gatherings, on social media, and within the rehabilitation centers Reade (Amsterdam, the Netherlands) and Heliomare (Wijk aan Zee, the Netherlands). Potential participants were included when all of the following criteria were met: chronic SCI (including spina bifida) or LLA (>1 year), wheelchair-dependent for longer (>500 m) distances, 18 years or older, sufficient knowledge of the Dutch language, and access to a smartphone or tablet connected to the internet. Potential participants were excluded when one of the following criteria was met: insufficient understanding of technology to benefit from the app; limited functioning in the arm/hand to operate a smartphone or tablet; presence of progressive disorders that can influence the outcomes; presence of psychiatric disorders; and negative outcome to unsupervised exercise based on a medical screening, including a graded exercise test. The target was to include 15 individuals with SCI and 15 individuals with LLA resulting in a sufficient sample size with a possible 10% dropout rate, based on the literature [39]. All participants provided written informed consent and ethical approval was granted by the local Medical Ethical Committee of Slotervaart Ziekenhuis-Reade (METC nr. P1761).

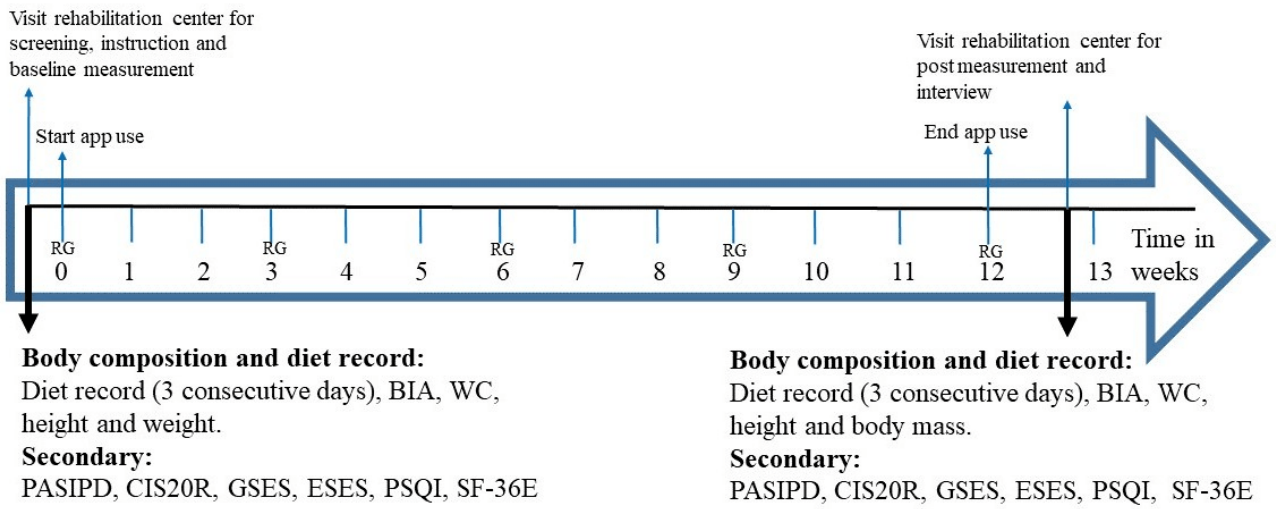
Study Design and Protocol

Participants were asked to participate in a 12-week pre-post pilot study focusing on the usability and feasibility of the app. Block randomization stratified by disorder with a block size of one was used to equally allocate the participants to a stand-alone

XSL•FO
RenderX

or remote-guided intervention group. The remote-guided group received guidance from a lifestyle coach during the 12-week intervention (Figure 2). At the start of the study, the stand-alone group received an individual explanation and demonstration of the app from the researcher. During the study, this group was allowed to consult the researcher with any questions or difficulties regarding use of the app. The remote-guided group also received support at the start of the study, but additionally received remote guidance, consisting of an additional face-to-face consultation (30 minutes) at the start of the intervention and 10-15 minute contact moments after 3, 6, 9, and 12 weeks by phone, app, or email. The purpose of these contact moments was to discuss progress and to adjust the goals or the program if necessary. The two lifestyle coaches providing the supervision were 4th year students in Functional Exercise Therapy who were trained in motivational interviewing, assisted by an experienced rehabilitation professional from Reade or Heliomare who they could consult with any questions.

Figure 2. Schematic overview of measurements. BIA: bioimpedance analysis; CIS20R: Checklist Individual Strength; ESES: Exercise Self-Efficacy Scale for spinal cord injury; GSES: General Self-Efficacy Scale; PASIPD: Physical Activity Scale for Individuals with Physical Disabilities; PSQI: Pittsburgh Sleep Quality Index; RG: remote guidance; SF-36E: Short-Form 36 health survey; WC: waist circumference.



Owing to the design and nature of the study, it was not possible to blind participants or researchers. The participants kept a 3-day diet record during the same weeks as the scheduled pre and post measurements, which took place at a rehabilitation center. In addition, body composition was measured before and after the intervention, and participants were invited through email to complete an online questionnaire in Qualtrics to measure lifestyle and health-related quality of life. During the last visit, a semistructured interview was conducted and recorded.

Qualitative Evaluation: Usability and Feasibility

Topics evaluated to assess usability were: (1) usefulness, (2) ease of use, (3) satisfaction, (4) ease of learning, (5) motivation, (6) adherence, and (7) goals. In addition, semistructured interviews and questionnaires (System Usability Scale [SUS] [40], and Usefulness, Satisfaction, and Ease of Use [USE] [41]) were used to gain insight into usability and feasibility.

The interviews were conducted by the same two researchers and lecturers in Sports Studies (LtL and JK) who instructed the participants on how to use the app at the start of the study. The interviews took place following the postintervention measurements in the rehabilitation center in a private room, lasting on average 30 minutes and were audio-recorded with consent of the participants. An interview guide, developed by the research team and partially based on the SUS and USE questionnaires and broader literature on behavior change [41-43], was used to structure the interview. After starting with the question “How experienced are you with using smartphone apps?” the following topics were discussed: goals, adherence, motivation, ease of use, satisfaction, and usability. The complete interview guide is provided in Multimedia Appendix 2.

The SUS questionnaire is a 10-item Likert scale providing an overall subjective assessment on usability of a product. An 11-point Likert scale was used with a score ranging from 0 (“strongly disagree”) to 10 (“strongly agree”). A total score was calculated by rescaling the score to a total of 100, with a higher score indicating higher perceived usability [40]. A SUS score of 70 is considered to be average based on a wide range of interfaces [44].

Additionally, three out of four dimensions of the USE questionnaire were used to gain insight in the dimensions “ease of use,” “ease of learning,” and “satisfaction.” Each dimension is composed of 11, 4, and 7 items, respectively. The dimension “usefulness” was left out as it overlapped strongly with the SUS questionnaire. An 11-point Likert scale was used with a score ranging from 0 (“strongly disagree”) to 10 (“strongly agree”) and averaged for each dimension [41].

Quantitative Evaluation

Nutritional Habits

The diet record took place on 3 consecutive days with one weekend day, which is considered one of the most reliable methods of dietary assessment [45]. Uncertainties about

XSL•FO
RenderX

registered diet records (ie, unclear handwriting, unclear food proportions) were solved by contacting the participant. Diet records were analyzed based on the nutrition values of the recorded products as shown in the Dutch nutrient database Nederlands Voedingsstoffenbestand version 2019/6.0 [46] and averaged over at least 2 available days.

Body Composition

Body mass was determined to the nearest 0.1 kilogram by deducting the mass of the wheelchair from the total mass of the participant and wheelchair combined, measured with a wheelchair weighing scale (RS1010, Allscales Europe used at Reade; Detecto 6550 used at Heliomare). Height and waist circumference (WC) were measured with a tape measure to the nearest 0.5 centimeters with the participants in supine position on a treatment table. WC was determined by taking the average of three measurements. BMI (body mass/height^2) was calculated with the proposed equation of Himes [47] in the case of LLA, based on the relative body segmental mass determined by Osterkamp [48] to adjust for lost body mass. Fat mass and fat-free mass were measured with a Bodystat 1500MDD (used at Reade) bioelectrical impedance analysis (BIA) device or with a Bodystat 500 Touch (used at Heliomare) BIA device in supine position. The BIA electrodes were placed at the right side of the body after the participant was in supine position. In case of unilateral LLA on the right side, the left side was measured. Electrodes were attached on the hands and feet according to the user manual. The BIA formula of Kyle et al [49] was used to calculate the fat-free mass and fat mass percentages based on the measured reactance and resistance.

Physical Activity

Physical activity was measured by the Dutch Physical Activity Scale for Individuals with Physical Disabilities (PASIPD), a 12-item 7-day recall self-reported questionnaire evaluating physical activity level in individuals with a physical disability. The PASIPD outcome is the metabolic equivalents of task (MET) hours spent per day and was calculated according to the method of Washburn et al [50]. The score in MET (hours/day) can range from 0 to 182.3 and can differentiate significantly among various physical activity levels.

Self-Efficacy

The General Self-Efficacy Scale and Exercise Self-efficacy scale, which are valid and reliable 10-item questionnaires (4-point Likert scale, total scores range from 10 to 40, where a higher score represents a higher self-efficacy), were used to assess the general self-efficacy and coping ability in daily life and exercise self-efficacy [51-56].

Fatigue

The Checklist Individual Strength (CIS20R), a reliable self-reported questionnaire (20 questions answered on a 7-point Likert scale, total scores range from 20 to 140 where a higher score represents more severe fatigue), was used to measure multiple aspects of fatigue [56,57].

Sleep Quality

The Pittsburgh Sleep Quality Index (PSQI) is a valid self-reported questionnaire to evaluate overall sleep quality [58]. The questionnaire consists of 19 questions resulting in a global score between 0 and 21, which consists of seven component scores (ie, sleep quality, sleep onset latency, sleep duration, sleep efficiency, sleep disturbances, use of sleeping medication, and daytime dysfunction). A higher global score represents worse sleep quality. The PSQI is considered as a reliable and valid method to evaluate sleep quality [59].

Quality of Life

The short-form health survey enabled (SF-36E) questionnaire was constructed to measure health-related quality of life on eight different dimensions of health for individuals with a mobility impairment [60]. The eight dimensions are physical functioning, social functioning, physical role emotional role, mental health, vitality, bodily pain, general health, and health transition. Dimension scores are rescaled to a 0-100 score where a higher score represents a higher quality of life. The Dutch version of the SF-36E is considered as a reliable and valid tool for individuals with chronic disabilities [61].

Data Analysis

Qualitative Evaluation

The audio recordings were transcribed verbatim and analyzed using a content analysis approach [62]. After familiarizing with the data by listening to the interviews and reading the transcripts, initial codes were identified by labeling text segments (open coding). The open codes were a mix of inductive codes that arose from reading the transcripts and deductive codes that arose from the study aims and interview guide. The open codes were then organized in a thematic map by comparing them and categorizing them into codes and subcodes. Finally, the codes and subcodes were integrated into core categories or main themes, derived from the topics of the interview guide based on the SUS and USE questionnaires, and the broader literature on behavior change [40,41,43]. The coding was carried out by one researcher (JH), who discussed and agreed on the themes and codes with a second researcher (LtL). They discussed their findings with the research team to ensure reliability of coding and data interpretation. The transcripts were analyzed with MAXQDA version 11 (VERBI GmbH).

Quantitative Evaluation

All quantitative data were analyzed with IBM SPSS software (Version 26). Pre and postintervention changes were compared using a paired-sample *t* test. Normality assumptions were checked with the Kolmogorov-Smirnov test. If normality was violated, a Wilcoxon signed-rank test was performed. Significance was accepted at $P<.05$. Effect sizes were determined by Hedges *g*, except when assumptions were violated and the effect size was determined by $z/(\sqrt{n})$ [63]. Effect sizes can be interpreted based on the following: $g<0.2$ indicates a very small effect size, $g=0.2$-0.5 indicates a small effect size, $g=0.5$-0.8 indicates a medium effect size, and $g>0.8$ indicates a large effect size.

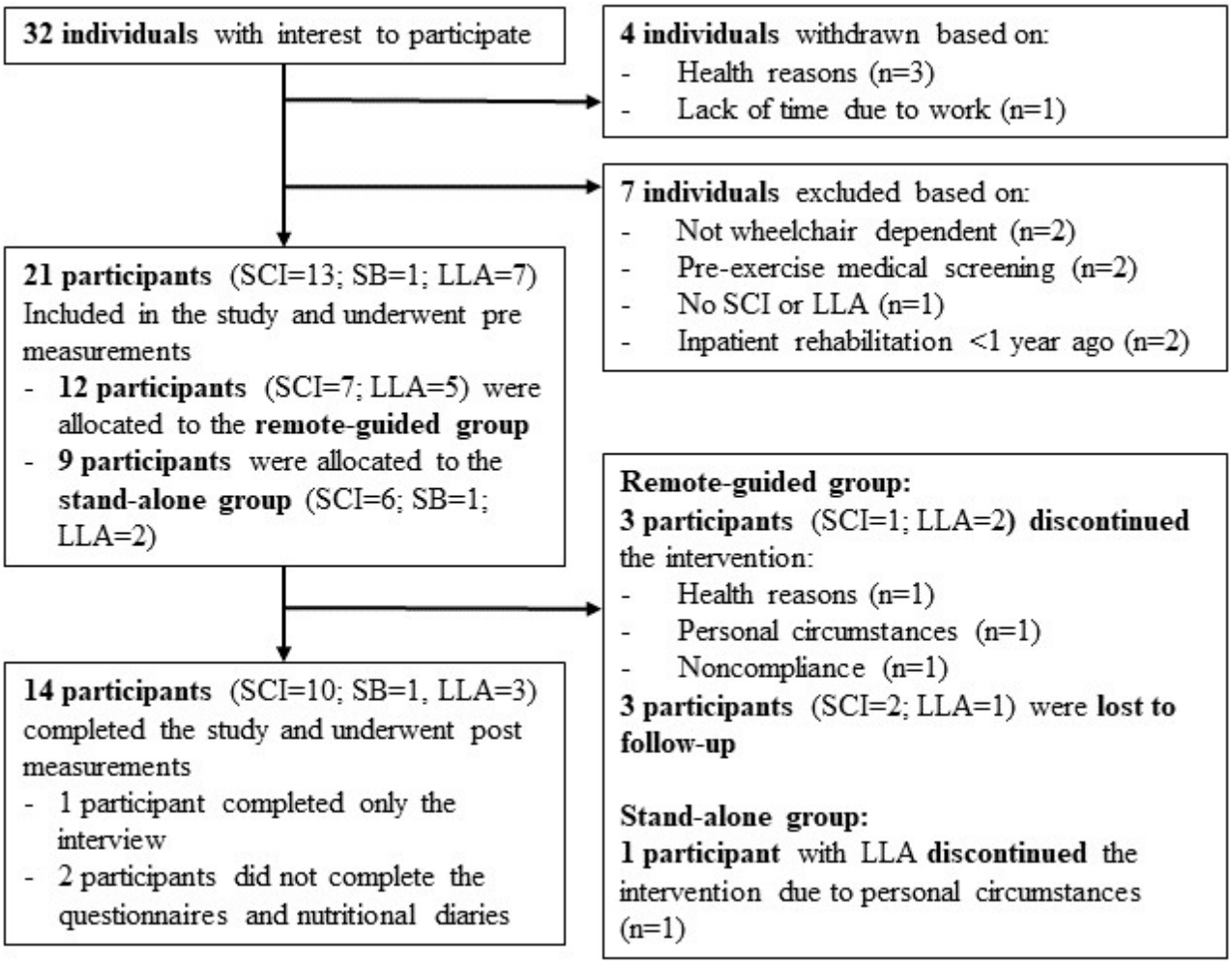
XSL•FO
RenderX

Results

Participants

Twenty-one participants were included in the process and effect evaluation study, 11 of whom completed all pre- and postmeasurements during the intervention period, as shown in Figure 3. One participant completed only the interview at the postmeasurement, and two participants did not complete the nutritional diaries and questionnaires at the postmeasurement. Demographic information of the participants who completed the intervention period are summarized in Table 1. No significant differences in demographic characteristics were found between the remote-guided and stand-alone groups. Because of the small group sizes, the results of the total population are presented. The results of statistical comparisons within the remote-guided and stand-alone groups are presented in Multimedia Appendix 3.

Figure 3. Flow chart of participant inclusion. LLA: lower limb amputation; SB: spina bifida; SCI: spinal cord injury.



XSL•FO
RenderX

Table 1. Demographic characteristics of the participants.

Characteristics	Baseline (N=21)	Participants that completed the postintervention interview		
		Remote-guided (n=6)	Stand-alone (n=8)	Total group (n=14)
Gender (female/male), n	14/7	4/2	5/3	9/5
Spinal cord injury, n				
Total	14	4	7	11
Tetraplegic	1	0	1	1
Paraplegic	13	4	6	10
Lower limb amputation, n				
Total	7	2	1	3
Unilateral	4	2	0	2
Bilateral	3	0	1	1
Age (years), mean (SD)	51.6 (11.9)	55.5 (11.7)	54.1 (11.8)	54.7 (11.3)
Height (m), mean (SD)	1.68 (0.20)	1.74 (0.11)	1.64 (0.24)	1.69 (0.20)
Body mass (kg), mean (SD)	87.6 (21.5)	88.7 (19.8)	80.1 (19.4)	83.8 (19.3)
BMI (kg/m^2), mean (SD)	33.7 (13.1)	30.0 (5.9)	31.6 (11.5)	30.9 (9.2)
Time since injury (years), mean (SD)	15.5 (15.8)	10.7 (6.3)	22.1 (22.5)	17.2 (17.8)

Qualitative Evaluation

Prior Experience

Most participants (n=11) indicated they were reasonably to very experienced smartphone app users. Three participants did not consider themselves experienced with smartphones, including one participant who indicated that they mainly use a tablet, which was also used to run the WHEELS app. Six participants had no experience, two had minimal experience, and six had more extensive experience with lifestyle apps prior to participation in this study.

Themes

Overview

The codes and subcodes that emerged from the content analysis could be clearly classified under the dimensions of usability and feasibility of interest for this study. Therefore, the topics of the interview guide largely corresponded to the main themes used to categorize the results: (1) motivation and lifestyle goals, (2) app use and adherence, (3) satisfaction, (4) usefulness, (5) ease of learning, (6) ease of use, and (7) needs and suggestions for improvement. The themes were gathered and presented together for the app in general, and for the specific physical activity and exercise, food, sleep and relaxation, and community sections of the app.

Motivation and Lifestyle Goals

Motivations mentioned to participate in the pilot study were incentive to work on a healthy lifestyle, gaining insight into/becoming aware of physical activity and nutritional behavior, discovering new exercises, critically testing the app, and making suggestions for improvement. Participants mentioned at least one goal they hoped to achieve with support of the app. Ten of the 14 participants indicated weight loss as a lifestyle goal, 8 participants had goals related to increasing physical activity/exercise, 4 indicated healthy energy and/or food intake as lifestyle, 3 mentioned more overarching goals such as staying or becoming healthy and fit, and one participant indicated that he had a goal related to rest and relaxation, which was to fall asleep better.

> *Well, what I said: on the one hand to maintain fitness and also to maintain weight, because with age, weight goes on rather than it goes off. Especially when you sit all day, it is more difficult than when you walk, I think. So I would rather keep that stable and yes, in the positive case, I could also lose some weight.* [Participant 5, female, 59 years old, LLA]

Ten participants reported having partially achieved one or more lifestyle goals, five of whom indicated that they had lost weight. “And yes the food, I was very busy with losing weight. I succeeded, so it works well in that respect.” [Participant 9, female, 42 years old, SCI]

Different reasons were mentioned for why lifestyle goals were not (or only partially) achieved, such as bad weather (cold and wet); hay fever; personal circumstances, including the death of a loved one; laxity; stubbornness; shoulder injury; or having set too ambitious goals. Additionally, the goal of improving falling asleep was difficult to target specifically. “I wanted to lose 5 kilos and in the period I was working on this I already realized that this was a bridge too far.” [Participant 4, male, 67 years old, SCI]

Partially achieving the lifestyle goals was not always entirely attributed to use of the app. Some participants indicated that participating in the pilot study, and therefore consciously working on a healthy lifestyle, already provided sufficient incentive to pursue existing lifestyle goals. Nevertheless, the majority indicated that the app had influenced their physical activity and dietary behavior, particularly contributing to raising

XSL•FO
RenderX

awareness of, and providing insight into, daily energy and nutrient intake.

> *Because normally, when there are some things on the table you really have no idea how many calories are in it. And if you look a little further in the app, you can also see how many carbohydrates and other macronutrients are in it. So it was firstly to build up a bit of discipline: what will I shove in during the day. And secondary, a piece of awareness was built up. Then you have eaten a nice pizza, and suddenly get 1200 calories on your plate, and then you are not allowed to eat anything for the rest of the day.* [Participant 1, male, 64 years old, SCI]

In addition, the participants explained that the app provided incentive discipline to exercise regularly and to eat healthy, with pop-up reminders to exercise contributing to this.

> *I thought I should honestly fill in what I eat and drink and then you see your own overview… Then [when the overview shows a surplus in calories] you think: another day tomorrow, I have to fix this right away.* [Participant 17, female, 57 years old, SCI]

> *I definitely do less now without the app. Because the app sort of said: now it's time for your weekly gym exercise.* [Participant 19, male, 39 years old, SCI]

Finally, it was mentioned that the app provided direction and tools to adopt a healthy lifestyle, such as by offering exercises.

> *Look, I sit in a wheelchair and I don't do anything else. But now you see oh, I can do that and I can do that and I can do that… And now I have come this far, also together with my physiotherapist, that I get out of that wheelchair. That I look at what kind of standing exercises and suchlike I could do, and that motivates me enormously.* [Participant 16, female, 70 years old, LLA]

App Use and Adherence

Six participants reported having used the app daily, two had used the app extensively during the 12-week intervention period, but not daily, whereas five had used the app only in the beginning (2-5 weeks), and one had used the app irregularly. Three reasons were given by several participants for having used (certain parts of) the app less or not at all. The first reason was dissatisfaction with the functioning or ease of use of the app: “If at some point you feel that it is not working in the way you would like, yes, then you think forget it.” [Participant 1]

The second reason was personal circumstances such as an injury, illness, and psychological stress: “I have also seen the cardio program. There was something about building up biking, but just because I was not in good shape, I didn't start doing that.” [Participant 2, female, 64 years old, SCI]

The third reason given was laziness:

> *I did not use that [food part]*
>
> Interviewer: *Okay and what was the reason for that?*
>
> *Yes, laxity, ease. Maybe also the stubbornness that you think: yes, I know how to lose weight.* [Participant 11, male, 61 years old, SCI]

The food diary was used most intensively, which was completed almost daily by six participants. Four others also used the food part but less intensively or quit after a few weeks. Another participant only tested the food part for usability.

Four participants used the exercise database and preprogrammed exercise routines daily or frequently during the 12 weeks. One participant also used these parts of the app regularly, although less frequently. Four participants used the physical activity and exercise part in the beginning but stopped after a few weeks. In addition, one participant had only explored the exercise database and training programs and another participant only tested the physical activity and exercise part for usability. Competitive activities were mentioned as one of the reasons for no or little use of the exercise part: “In the beginning I also used the app, but because I went to the gym twice a week, I stopped using it.” [Participant 2]

A second reason given to stop using the exercise part was that the app was no longer needed because the exercises were known and could be performed without the app.

> *Yes, those were just example exercises and then I thought: oh yes, that is a fun one, oh I am going to do that, and: oh, that is also a fun one that I am going to do and then you have four or five [exercises]. Yes, then I really don't need that app anymore. Because then I already know what to do. Then I no longer have to look at that app every day.* [Participant 5]

Three participants had used the sleep and relaxation part. The first participant performed the relaxation exercises several times, the second had read the information about sleep and relaxation, and the third explained that he still used the tips to relax because he slept rather poorly and the tips helped him to relax. A fourth participant only tested the sleep and relaxation part for usability. A frequently mentioned reason for why the sleep and relaxation part had not been used was that it was not needed because participants did not experience stress or sleeping problems: “Not looked at [sleep and relaxation part] because I am sufficiently balanced and relaxed.” [Participant 11]

Three participants used the informative community groups. They were alerted by email to new messages in which lifestyle information and tips were shared and read them. The interactive part of the community in which participants could ask each other questions, and share experiences and tips was hardly used. One participant expressed that it was unfortunate that hardly any interaction had started. Reasons given for not being active in the community were unwillingness to brag and incomparability.

> *It is more because it is so incomparable to each other. Look, someone who just got out of rehabilitation may be very proud to have handcycled 10 kilometers. While I think, yes, when I say I have done 20 kilometers... I think that...I don't feel the need to proclaim it or anything. I think, yes, you are not comparable.* [Participant 5]

XSL•FO
RenderX

Finally, the calendar was used by several participants to plan their exercises. Two participants reported having used the progress registration part by regularly registering their weight and WC. A few participants started the handcycling challenge; however, no participant completed this challenge because it was unclear how the time that was handcycled for this challenge could be recorded.

Satisfaction

Overall, participants were satisfied with the WHEELS app. Several participants indicated that they wanted to continue using the app. It was further noted they were happy that there is finally a lifestyle app suitable for wheelchair users, that the combination of attention to healthy nutrition and physical activity/exercise is nice, and that the app has something in it for everyone.

> *I think that because you have a lot of different things in it, you have a very large target group. Some things may not be helpful to me, but that's not to say it's not useful for the app.* [Participant 9]

When asked if they would recommend the app to others, 12 participants answered yes, one participant would only recommend the food part, and one person would not recommend the app at all: “Maybe, I haven't thought about it. But there is, yeah, it's a good way to start up. Until you get into a routine.” [Participant 19]

With regard to the exercise database and preprogrammed exercise routines, the participants indicated that there was great variation in exercises and that the animation provided a clear example of the desired implementation: “I have to tell you, I think it looks super cool. Also the exercises that are offered, I think the variation in exercises is very good.” [Participant 11]

Some participants were less satisfied with the exercise database because they had difficulty finding the right exercises: “And then you see so many exercises in that list that you actually do not know which one to choose. And that was a problem for me.” [Participant 7, male, 54 years old, SCI]

Regarding the food part, the participants indicated that they liked that that the app provided insight into their daily energy and nutrient intake. Furthermore, they explained that the food product list was very extensive, as almost every food could be found in the list and then easily added to the food diary.

> *I found the food diary very useful, you can see exactly what you eat in calories and protein.* [Participant 8, female, 34 years old, SCI]
>
> *The nice thing about the food app is that no matter what you eat or drink, you can always find it somewhere and it has a calorie number.* [Participant 1]

The relaxation part was rarely used, making it difficult to indicate whether the participants were satisfied with the content. Two participants indicated that the exercises were perhaps a bit too spiritual: “Well it often comes across as very spiritual, so to speak, while, yes, while that might raise an aversion.” [Participant 9]

With regard to the community aspect, a few participants stated that it was motivating that lifestyle tips were regularly posted. Opinions were divided about the community group in which experiences and questions could be shared: some considered this to be of added value, whereas others did not feel the need for it.

Usefulness

The participants indicated that the app is particularly useful for (recent) wheelchair users who have little physical activity and exercise experience, providing tools and inspiration to start exercising and develop a healthy lifestyle.

> *Because I see a lot of wheelchair users around me who are simply aimless. Who can have a huge hold on that [the app]. Especially if you are not physically active or are starting to be physically active.* [Participant 11]

More experienced wheelchair users with a more physically active lifestyle had less need for the complete app, but found some parts useful, in most cases concerning the food part to gain insight into dietary intake, as described in the previous section on satisfaction.

Ease of Learning

Opinions differed on how easy it was to learn how to use the app. Some participants quickly became skillful, others needed a few weeks, and still others gave up using the exercise part because they could not learn to work with it quickly. Most interviews showed that learning to work with the app takes some time, and not all participants had the patience and will to spend time on this.

> *The first period is quite a lot of investing and maybe I could have got a bit more out of it. But then you expect all users to use it, invest and maybe benefit from it afterward, and I think that is too much to ask.* [Participant 17]

Some participants indicated they could not repeat actions they had previously performed with the app, such as scheduling an exercise in the calendar. This indicates it was not easy for all participants to remember how to use the app. Some participants also indicated they had asked family members to help them learn to work with the app. This also reveals that learning to work with the app was not experienced as easy by all participants.

> *Actually during the first 2 or 3 weeks of use, you don't know all the tips at once, because that is too much, but then the advantage is that my daughter is very handy and serious with that [app use]. So then I got another tip and I could do a bit more. Yes, perfect that app.*” [Participant 12, female, 61 years old, LLA]

The part that participants most often reported as unsuccessful or difficult to master was putting together a training schedule themselves.

> *Well, what I said earlier about making such a [exercise] program of your own and then...Yes that. I could not completely figure that out, and I must honestly say that at some point I also think: well, never mind.* [Participant 5]

XSL•FO
RenderX

Finally, the participants hardly reported any problems with learning to work with the food part. This part seemed to work quite intuitively.

Ease of Use

The participants generally found the food part easier to use than the exercise part. The vast majority labeled the food part as easy to use, whereas the exercise part was described as easy to use by about half of the participants.

Three factors emerged that negatively affected the ease of use of the exercise part. Not all participants were well aware of the distinction and coherence between the exercise database with which personal training routines could be compiled, the preprogrammed exercise routines, and the calendar.

> *Well with those exercise programs I found it a bit vague at first because you also have two boxes [tiles at the start screen that direct to the different parts of the app] with exercise or something.* [Participant 8]

Second, the participants indicated they had difficulty finding suitable exercises in the exercise database. Several participants were unaware of the possibility or did not know how to use the search box to find specific exercises.

> *For example, if I had done something new at the physio, it was sometimes a bit of a search in that whole list of: What suits this best? And then indeed there were sometimes whole laundry lists with exercises in it...I think that at some point it just missed its target...[Participant 4]*

Third, it was not clear to all participants how activities could be registered to end up in the activity diary or the activity stream.

> *And for example I had put an exercise in the calendar and then I thought oh then it will automatically register that I did that, but that is not the case. Then you had to click that again.* [Participant 8]

The participants generally found the food part clear to use. Several participants indicated that they liked the fact that the food product list was so extensive that it was easy to log consumed products in the food diary: “Yes, I actually did not come across a product that I could not find [in the food product list].” [Participant 14]

The participants found the time it took to keep a food diary less user-friendly. Several participants were not aware of the possibility to save frequently eaten meals such as a fixed breakfast. This means that the individual products that make up the meal only need to be entered once. The next time, the entire meal can be added to the food diary with one click. Partly because participants were not aware of this possibility, they indicated that they were tired of having to import the same products every time.

> *At some point you are a bit done with reentering those foods every time. I always eat a bowl of curd with muesli in the morning. Yes, the next day I had to scan it again or had to go back and find out where that curd is. And I thought that was a bit of a disadvantage.* [Participant 7]

Needs and Suggestions for Improvement

The participants expressed several needs and suggestions for improvement that were largely related to the difficulties they experienced during app use. The most frequently mentioned needs and suggestions related to personalization, user instruction or remote guidance, and improving insight into energy expenditure by connecting a wrist-worn activity monitor to the app. No difference in suggestions was found between the remote-guided and stand-alone groups.

In the exercise part, better personalization could make it easier to find and select suitable exercises. In addition, it was suggested that the preprogrammed exercise routines could be better tailored to the individual needs and functional capabilities: “However, I would appreciate the programs to be a little bit more distinct to your abilities. That would have been perfect.” [Participant 19]

With regard to the food part, the participants expressed their need for personalized daily energy intake advice: “If you want to use it for people with a spinal cord injury, you really have to adjust the calorie advice, because we are actually allowed to consume fewer calories.” [Participant 7]

The user instruction was not found and used by all participants. Some participants indicated they needed a help desk for questions and problems with using the app.

> *That I couldn't find back what I had done before was the most frustrating...That I didn't know how I had done it before and why I couldn't find it again…Someone who offered me guidance, by phone or every 2 weeks face-to-face, could have definitely helped me.* [Participant 17]

In addition, several participants would have liked to have had a coach to help them on their way, have questions answered, and discuss suitable activities and progress on their goals. One of the participants who had received remote guidance explained that it had been of little added value to her because the coach barely reflected on what she had done. She would have liked the coach to monitor her and to discuss progress on her goals with her: “Yes someone who says: you went over your [calorie] goal all week. Do you have an idea how you will solve this in the coming week?” [Participant 14, female, 42 years old, SCI]

Finally, it was revealed that, in combination with lifestyle guidance by a coach, the app would probably have been used more intensively: “I think I would have used it under supervision more, the WHEELS app.” [Participant 17]

As a third and last suggestion, several participants explained that they would like to have better insight into their rate of exertion and energy expenditure during activities. Connecting an activity monitor for wheelchair users to the app has been suggested as an opportunity to meet this need.

> *Yes, a pedometer too, for a wheelchair user. It is nice to know what your muscles have done, what your heart rate has been and how many calories or energy you have consumed…I would rather put on a wristband for that and will also enter some extra information.* [Participant 17]

XSL•FO
RenderX

Quantitative Evaluation

Usability and Feasibility

The average SUS score for all participants was 58.6, which is below the general average interface SUS score of 70 [44]. Participants showed little difference in scores between the remote-guided and stand-alone groups in SUS score and USE dimension scores. All SUS scores and USE dimensions scores can be found in Table 2 for each group and all participants.

Table 2. Descriptive statistics of usability and feasibility questionnaires.

Questionnaire	Remote-guided (n=4)		Stand-alone (n=7)		All (N=11)	
	Mean (SD)	Range[a]	Mean (SD)	Range	Mean (SD)	Range
System Usability Scale	56.0 (5.8)	14.0	60.1 (32.2)	84.0	58.6 (25.2)	84.0
Ease of use	5.0 (1.7)	4.3	5.6 (3.8)	9.7	5.4 (3.1)	9.7
Satisfaction	5.1 (2.2)	4.8	5.2 (3.7)	9.7	5.2 (3.1)	9.7
Ease of learning	5.5 (3.2)	7.0	6.1 (4.2)	9.5	5.9 (3.7)	9.5

[a]Represents the spread; the difference between the maximum and minimum values.

Nutritional Intake

Nutritional intake changes of 10 participants over the 12-week intervention period are shown in Table 3. The quality of the diet records of two participants was not adequate (ie, portion size or ingredients used were not described in sufficient detail to obtain macronutrients correctly) to analyze at least 2 recorded days at pre and postmeasurement, and these records were therefore excluded. No significant differences were found over time within the whole group, or within the remote-guided or stand-alone groups separately, in total calorie count or macronutrients. However, there was a positive trend (with medium effect sizes) toward a reduction in fat consumed and relative alcohol intake. Other outcomes showed either small or very small effect sizes. Full results of the separate groups are presented in Table S1 in Multimedia Appendix 3.

Table 3. Nutritional intake pre and post the 12-week intervention based on diet records (n=10).

Daily average consumed	Pre, mean (SD)	Post, mean (SD)	*P* value	Effect size
Kilocalories	1920 (531)	1637 (377)	.14[a]	0.46
Protein (g)	79.8 (21.8)	74.1 (13.5)	.37	0.28
Fat (g)	82.9 (31.4)	69.6 (22.9)	.07[a]	0.56
Carbohydrates (g)	179.8 (61.3)	154.7 (63.3)	.24[a]	0.37
Alcohol (g)	12.7 (13.3)	7.9 (10.7)	.16[a]	0.44
Protein (%)	16.9 (3.7)	18.6 (3.6)	.14[a]	0.47
Fat (%)	38.4 (8.5)	38.2 (8.6)	.88	0.05
Carbohydrates (%)	37.7 (8.1)	37.3 (10.3)	.88[a]	0.05
Alcohol (%)	4.5 (4.7)	3.4 (5.0)	.09[a]	0.53

[a]Nonparametric test used due to violation of normality with corresponding effect size.

Body Composition

For the whole group, all body composition outcomes showed favorable changes over the intervention period, which was significant for WC, fat mass percent, fat-free mass, and fat-free mass percent, and a trend toward significance for body mass and BMI (Table 4). Large effect sizes were found for fat mass percent and fat-free mass percent, whereas all other outcomes showed medium effect sizes except for body mass. Results for each group separately are shown in Table S2 in Multimedia Appendix 3, demonstrating slightly better results in favor of the remote-guided group compared to the stand-alone group.

XSL•FO
RenderX

Table 4. Body composition changes pre and post the 12-week intervention (n=13).

Body composition	Pre, mean (SD)	Post, mean (SD)	*P* value[a]	Effect size
Body mass (kg)	83.0 (20.0)	81.6 (20.6)	.09	0.49
BMI (kg/m^2)	29.7 (7.8)	29.1 (7.8)	.07	0.53
WC[b] (cm)	103.9 (13.5)	101.4 (13.8)	.02	0.76
FM[c] (kg)	32.1 (10.1)	26.7 (9.2)	.004	0.96
FM (%)	39.2 (9.5)	33.4 (10.5)	.004	0.97
FFM[d] (kg)	50.9 (15.4)	54.9 (17.7)	.02	0.70
FFM (%)	60.8 (9.5)	66.6 (10.5)	.004	0.97

[a]Paired-sample *t* test.
[b]WC: waist circumference.
[c]FM: fat mass.
[d]FFM: fat-free mass.

Questionnaire Outcomes

No significant changes in questionnaire outcomes were found over time (Table 5), although the PSQI results showed a favorable trend toward better sleep quality. A medium effect size was found in sleep quality. Small effect sizes were found on the CIS20R and four subcategories of the SF-36E. No clear differences were found between groups, as shown in Table S3 in Multimedia Appendix 3.

Table 5. Results from questionnaires pre and post the 12-week intervention (n=12).

Questionnaire	Pre, mean (SD)	Post, mean (SD)	*P* value	Effect size
PASIPD[a] (MET h/day)	23.8 (15.1)	22.4 (12.2)	.64	0.14
GSES[b]	34.1 (4.3)	33.6 (4.2)	.69	0.11
ESES[c]	33.0 (3.8)	32.4 (4.4)	.50[d]	0.07
CIS20R[e]	72.3 (11.6)	69.6 (7.7)	.18[d]	0.39
PSQI[f]	8.0 (2.7)	6.7 (2.2)	.06	0.57
SF-36E[g]				
Physical functioning	52.1 (18.6)	49.2 (18.2)	.58	0.16
Social functioning	72.9 (17.5)	79.2 (20.9)	.36[d]	0.27
Role limitation physical	57.8 (26.1)	59.4 (18.0)	.84	0.06
Role limitation emotional	79.2 (23.7)	74.3 (24.7)	.53[d]	0.18
Mental health	81.7 (12.6)	74.6 (15.1)	.13	0.46
Energy/vitality	62.5 (14.6)	62.5 (12.2)	>.99	1.00
Pain	59.0 (17.6)	49.1 (21.5)	.11	0.48
General health perceptions	63.8 (18.1)	60.0 (21.3)	.37[d]	0.26

[a]PASIPD: Physical Activity Scale for individuals with Physical Disability.
[b]GSES: General Self-Efficacy Scale.
[c]ESES: Exercise Self-Efficacy Score.
[d]Nonparametric test used due to violation of normality.
[e]CIS20R: Checklist Individual Strength.
[f]PSQI: Pittsburgh Sleep Quality Index.
[g]SF-36E: Short-Form Health Survey 36 Enabled.

XSL•FO
RenderX

Discussion

Principal Findings

This paper describes the development of the WHEELS mHealth app for wheelchair users with SCI or LLA. Additionally, the first insight on usability, feasibility, and effectiveness of the app is provided. The perceived usability and feasibility varied among participants and showed room for improvement. Participants did show a positive development in body composition such as a significant decrease in fat mass, which was often mentioned as a personal lifestyle goal. Combined with a positive trend for sleep quality, and reduced fat and alcohol intake, the app seems promising to improve lifestyle behavior in wheelchair users, with the caveat that no change in physical activity levels were detected. Environmental barriers might have contributed to this, which cannot be influenced by mHealth.

The SUS score and usability scores for ease of use, ease of learning, and satisfaction ranged between the minimal (0) and maximal (10) scores, indicating a varied user experience. The questionnaire outcomes were in line with the interviews, in which some participants were merely positive and experienced no struggle using the app, whereas others mentioned difficulties using the exercise and planning part, for example. These differences in perceived usability could be related to differences in motivation and time spent within the app, and likely influenced the extent to which the app has led to the desired lifestyle behaviors. This is best explained by the Fogg behavior model, which describes that behavior change is related to three elements, motivation, ability, and trigger, where motivation and ability show an inverse relationship [64]. If a certain level of effort and time (motivation) was not put into understanding the app (ability), an individual would not meet the minimal requirements to benefit from the app. Participants who were willing to put more time and effort in becoming familiar with the app, and thus showed more motivation, were more positive about the product and expressed fewer difficulties working with the exercise and planning part. This is in line with earlier research, which shows that a higher level of app engagement is associated with increased intervention effectiveness [65-67]. However, this could also be a flaw of the app, as the required ability might be too high to fully benefit from the app. Therefore, by reducing the ability needed to understand the app, less motivation is needed to continue using the app. Clearer and easier instructions that require little time and effort could possibly reduce the required motivation to meet the ability needed to benefit working with the app.

Another solution for overcoming difficulties with using the app could be found in a remote-guided intervention approach, in which the app is combined with guidance by a lifestyle coach. A remote-guided approach seems to be more promising in achieving improvement in behavioral and health outcomes [25]. Unfortunately, from this study, no conclusions can be drawn regarding differences between those who used the app as a stand-alone intervention and those who used the app with remote guidance. The remote-guided group did show more significant changes than the stand-alone group; however, owing to a larger dropout than expected, the group sizes were small and thus results should be interpreted with caution. Half of the participants were allocated to a remote-guidance group where they received regular phone consultations by students in Functional Exercise Therapy. However, based on interview reports, the effect of the provided phone consultations was limited, possibly due to the lack of experience the students had in motivational interviewing. Previous research suggests more advanced and prolonged training in motivational interviewing is needed to allow embedding of these skills [68]. Another explanation for the limited effect is that nonverbal communication was hardly possible because most consultations took place by phone, which could have reduced the consultation effects due to loss of possible valuable cues and information [69]. Moreover, multiple participants from the stand-alone group indicated that a consultant would have benefitted them in either solving difficulties in working with the app or as an additional motivator and guide in behavior change. The addition of peer health coaches could be of added value, as research shows that individuals with SCI can benefit from this type of support [70]. Therefore, it would be interesting to investigate the effect of using the app in combination with face-to-face guidance of trained peer health coaches (blended).

When taking a closer look at the effect evaluation, significant and favorable changes were seen in measures of body composition. This seems to be in line with other findings such as reduced body mass, reduced fat intake, and reduced relative alcohol intake, although these reductions were not statistically significant. Registered body composition changes were most likely partly caused by nutritional changes. The feature to track nutrition intake raised the participants' awareness of their nutritional intake and triggered them to change dietary habits, which was also mentioned during the interviews. The diet records showed a trend toward a positive change in nutrition behavior. These changes were not statistically significant, possibly due to the small sample size. No changes in physical activity levels were found, which could be caused by factors at the interpersonal, institutional, community, and policy influence levels that were not targeted in the app but are associated with physical activity among wheelchair users [71]. However, the significant increase in fat-free mass would suggest an increase in muscle mass caused by physical activity. Moreover, physical activity was measured with a self-reported questionnaire, which correlates poorly with objective physical activity outcomes in participants who were already relatively active at the start of the study [72]. Therefore, there may have been an increase in physical training that was not reflected in the total PASIPD score.

Limitations and Strengths

The targeted groups, wheelchair users with SCI and LLA, may experience different barriers and facilitators for developing a healthy lifestyle, and when developing the app we expected that the app had to be tailored accordingly. However, the needs assessment showed many similarities, resulting in the use of similar behavior change techniques for both groups. The intention was to use the 16 change strategies in the intervention to influence the main behavioral determinants identified in the development phase. However, it is uncertain whether all 16

XSL•FO
RenderX

strategies were applied as intended during the study. For example, tailoring options were limited due to software limitations, and several participants had not used all parts of the app with the result that they were not exposed to all behavior change strategies. This could have possibly affected the usability and effectiveness outcomes.

The interview yielded suggestions for improvement that could in turn improve the usability, feasibility, and effectiveness of the app. However, owing to high dropout, relevant feedback may have been missed from users who were less satisfied with the intervention or had more difficulty changing their behavior. Unfortunately, in most cases, we were unable to determine the reason for dropout, as this could have provided valuable information. The relative high dropout rate led to a lower sample size than intended and a possible biased user experience. Additionally, this resulted in only one individual with tetraplegia completing the study, making these results less generalizable to the whole SCI community. Nevertheless, despite the low inclusion rate, significant positive changes in body composition were found, which is very promising. However, these results should be interpreted with caution, as a higher possibility of type II errors is present due to the small sample size and multiple performed tests.

Several body composition outcomes were measured with a BIA device. The transformation formula used to calculate measured resistance to body composition outcomes was based on empirical data from the general population. Thus, the validity of the BIA measurements on this specific population could be argued. However, systematic deviation does not have to affect test-retest reliability, which would therefore make the BIA still able to detect changes over time. Additionally, the BIA outcomes seem to be in line with interview outcomes and nutrition diaries. Physical activity levels did not show any changes, which was subjectively measured with the PASIPD questionnaire and, as mentioned above, correlates poorly with objective physical activity outcomes that represent physical activity more accurately [72].

Future Studies

These first results on effectiveness of the WHEELS app seem to be promising for body composition changes, nutritional habits, and sleep quality. Improvements in manual instructions and support regarding use of the app are suggested. A study with a larger sample size and stronger research design, for example a repeated-measures mixed model design, is warranted, which would allow further investigation on the effectiveness of the different parts of the app for improving body composition, dietary behavior, physical activity, and health, and the interaction between stand-alone and remote-guided use of the intervention. In this larger study, it is recommended to measure physical activity objectively (eg, with accelerometry) to be able to conclude whether the app does or does not influence physical activity behavior. Preferably, such a study should be performed in participants who are less active at inclusion compared to the participants of this study. Accelerometry, including heart rate, would be recommended, as it could differentiate among intensity levels and thus provide a more valid physical activity outcome.

Conclusion

This paper describes the development, usability, and feasibility of the WHEELS mHealth app for wheelchair users with SCI or LLA, and provides the first insight into its effectiveness. Although usability could be improved, the app scored reasonably well and seems to be feasible to implement on a larger scale. First results on lifestyle changes seem promising, and effectiveness could possibly increase if the mentioned suggestions for improvement by participants are processed into the app.

Acknowledgments

This research is cofinanced by the National Taskforce for Applied Research SIA, part of the Netherlands Organisation for Scientific Research (NWO), under grant RAAK.PUB03.029. The members of the WHEELS Study Group are:

LE van den Akker, L Alpay, H Bijwaard, M Deutekom, R Doms, JFM Holla (program coordination): Inholland University of Applied Sciences, Haarlem, The Netherlands.

T Dadema, JM Dallinga, S Dankers, RH Engelbert, M Tieland, B Visser, PJM Weijs: Amsterdam University of Applied Sciences, Amsterdam, The Netherlands.

S de Groot, TWJ Janssen: Amsterdam Rehabilitation Research Centre, Reade, Amsterdam, The Netherlands.

H Houdijk, LJ Valent: Heliomare Rehabilitation Centre, Wijk aan Zee, The Netherlands.

LHV van der Woude: University Medical Centre Groningen, Groningen, The Netherlands.

JBJ Bussmann: Erasmus University Medical Centre, Rotterdam, The Netherlands.

N van Schijndel: Virtuagym B.V., Amsterdam, The Netherlands.

Conflicts of Interest

None declared; the involved researchers were distinct from the app developer Virtuagym B.V.

Multimedia Appendix 1

XSL•FO
RenderX

Development process of the WHEELS app.
[DOCX File , 85 KB - formative_v5i8e24909_app1.docx]

Multimedia Appendix 2
Interview guide for the usability and feasibility study of the WHEELS project.
[DOCX File , 22 KB - formative_v5i8e24909_app2.docx]

Multimedia Appendix 3
Tables S1-S3.
[DOCX File , 25 KB - formative_v5i8e24909_app3.docx]

References

1. Kumar K. Importance of healthy life style in healthy living. Juniper Online J Public Health 2017;2(5):11 [FREE Full text] [doi: 10.19080/jojph.2017.02.555596]
2. Pisinger C, Toft U, Aadahl M, Glümer C, Jørgensen T. The relationship between lifestyle and self-reported health in a general population: the Inter99 study. Prev Med 2009 Nov;49(5):418-423. [doi: 10.1016/j.ypmed.2009.08.011] [Medline: 19716843]
3. Pell JP, Donnan PT, Fowkes FG, Ruckley CV. Quality of life following lower limb amputation for peripheral arterial disease. Eur J Vasc Surg 1993 Jul;7(4):448-451. [doi: 10.1016/s0950-821x(05)80265-8] [Medline: 8359304]
4. van den Berg-Emons RJ, Bussmann JB, Stam HJ. Accelerometry-based activity spectrum in persons with chronic physical conditions. Arch Phys Med Rehabil 2010 Dec;91(12):1856-1861. [doi: 10.1016/j.apmr.2010.08.018] [Medline: 21112426]
5. van Langeveld SA, Post MW, van Asbeck FW, Gregory M, Halvorsen A, Rijken H, et al. Comparing content of therapy for people with a spinal cord injury in postacute inpatient rehabilitation in Australia, Norway, and The Netherlands. Phys Ther 2011 Feb;91(2):210-224. [doi: 10.2522/ptj.20090417] [Medline: 21212372]
6. Bragaru M, Dekker R, Geertzen JHB, Dijkstra PU. Amputees and sports: a systematic review. Sports Med 2011 Sep 01;41(9):721-740. [doi: 10.2165/11590420-000000000-00000] [Medline: 21846162]
7. Martin Ginis KA, van der Scheer JW, Latimer-Cheung AE, Barrow A, Bourne C, Carruthers P, et al. Evidence-based scientific exercise guidelines for adults with spinal cord injury: an update and a new guideline. Spinal Cord 2018 Apr;56(4):308-321. [doi: 10.1038/s41393-017-0017-3] [Medline: 29070812]
8. Adriaansen J, Ruijs L, van Koppenhagen CF, van Asbeck FWA, Snoek G, van Kuppevelt D, et al. Secondary health conditions and quality of life in persons living with spinal cord injury for at least ten years. J Rehabil Med 2016 Nov 11;48(10):853-860 [FREE Full text] [doi: 10.2340/16501977-2166] [Medline: 27834436]
9. Myers J, Lee M, Kiratli J. Cardiovascular disease in spinal cord injury: an overview of prevalence, risk, evaluation, and management. Am J Phys Med Rehabil 2007 Feb;86(2):142-152. [doi: 10.1097/PHM.0b013e31802f0247] [Medline: 17251696]
10. Adriaansen JJE, van Asbeck FWA, Lindeman E, van der Woude LHV, de Groot S, Post MWM. Secondary health conditions in persons with a spinal cord injury for at least 10 years: design of a comprehensive long-term cross-sectional study. Disabil Rehabil 2013 Jun;35(13):1104-1110. [doi: 10.3109/09638288.2012.712196] [Medline: 22991949]
11. Boulos MNK, Wheeler S, Tavares C, Jones R. How smartphones are changing the face of mobile and participatory healthcare: an overview, with example from eCAALYX. Biomed Eng Online 2011 Apr 05;10:24 [FREE Full text] [doi: 10.1186/1475-925X-10-24] [Medline: 21466669]
12. van den Berg-Emons RJ, Bussmann JB, Haisma JA, Sluis TA, van der Woude LH, Bergen MP, et al. A prospective study on physical activity levels after spinal cord injury during inpatient rehabilitation and the year after discharge. Arch Phys Med Rehabil 2008 Nov;89(11):2094-2101. [doi: 10.1016/j.apmr.2008.04.024] [Medline: 18996237]
13. Haisma JA, Bussmann JB, Stam HJ, Sluis TA, Bergen MP, Dallmeijer AJ, et al. Changes in physical capacity during and after inpatient rehabilitation in subjects with a spinal cord injury. Arch Phys Med Rehabil 2006 Jun;87(6):741-748. [doi: 10.1016/j.apmr.2006.02.032] [Medline: 16731207]
14. Desveaux L, Goldstein RS, Mathur S, Hassan A, Devlin M, Pauley T, et al. Physical Activity in Adults with Diabetes Following Prosthetic Rehabilitation. Can J Diabetes 2016 Aug;40(4):336-341. [doi: 10.1016/j.jcjd.2016.02.003] [Medline: 27052673]
15. Farkas GJ, Pitot MA, Berg AS, Gater DR. Nutritional status in chronic spinal cord injury: a systematic review and meta-analysis. Spinal Cord 2019 Jan 12;57(1):3-17. [doi: 10.1038/s41393-018-0218-4] [Medline: 30420688]
16. Westerkamp EA, Strike SC, Patterson M. Dietary intakes and prevalence of overweight/obesity in male non-dysvascular lower limb amputees. Prosthet Orthot Int 2019 Jun 20;43(3):284-292. [doi: 10.1177/0309364618823118] [Medline: 30663528]
17. van den Akker LE, Holla JFM, Dadema T, Visser B, Valent LJ, de Groot S, WHEELS-study group. Determinants of physical activity in wheelchair users with spinal cord injury or lower limb amputation: perspectives of rehabilitation professionals

XSL•FO
RenderX

and wheelchair users. Disabil Rehabil 2020 Jul;42(14):1934-1941. [doi: 10.1080/09638288.2019.1577503] [Medline: 30924706]
18. Holla JFM, van den Akker LE, Dadema T, de Groot S, Tieland M, Weijs PJM, WHEELS-study group. Determinants of dietary behaviour in wheelchair users with spinal cord injury or lower limb amputation: Perspectives of rehabilitation professionals and wheelchair users. PLoS One 2020 Jan 31;15(1):e0228465 [FREE Full text] [doi: 10.1371/journal.pone.0228465] [Medline: 32004359]
19. Nooijen CFJ, Stam HJ, Schoenmakers I, Sluis TAR, Post MWM, Twisk JWR, Act-Active Research Group, et al. Working mechanisms of a behavioural intervention promoting physical activity in persons with subacute spinal cord injury. J Rehabil Med 2016 Jul 18;48(7):583-588 [FREE Full text] [doi: 10.2340/16501977-2110] [Medline: 27346837]
20. Littman AJ, Haselkorn JK, Arterburn DE, Boyko EJ. Pilot randomized trial of a telephone-delivered physical activity and weight management intervention for individuals with lower extremity amputation. Disabil Health J 2019 Jan;12(1):43-50. [doi: 10.1016/j.dhjo.2018.08.002] [Medline: 30115584]
21. Ma JK, West CR, Martin Ginis KA. The effects of a patient and provider co-developed, behavioral physical activity intervention on physical activity, psychosocial predictors, and fitness in individuals with spinal cord injury: a randomized controlled trial. Sports Med 2019 Jul;49(7):1117-1131. [doi: 10.1007/s40279-019-01118-5] [Medline: 31119717]
22. Payne HE, Lister C, West JH, Bernhardt JM. Behavioral functionality of mobile apps in health interventions: a systematic review of the literature. JMIR Mhealth Uhealth 2015 Feb 26;3(1):e20 [FREE Full text] [doi: 10.2196/mhealth.3335] [Medline: 25803705]
23. Singh G, MacGillivray M, Mills P, Adams J, Sawatzky B, Mortenson WB. Patients' perspectives on the usability of a mobile app for self-management following spinal cord injury. J Med Syst 2019 Dec 11;44(1):26. [doi: 10.1007/s10916-019-1487-y] [Medline: 31828440]
24. Al Ayubi SU, Parmanto B, Branch R, Ding D. A persuasive and social mHealth application for physical activity: a usability and feasibility study. JMIR Mhealth Uhealth 2014 May 22;2(2):e25 [FREE Full text] [doi: 10.2196/mhealth.2902] [Medline: 25099928]
25. Schoeppe S, Alley S, Van Lippevelde W, Bray NA, Williams SL, Duncan MJ, et al. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. Int J Behav Nutr Phys Act 2016 Dec 07;13(1):127 [FREE Full text] [doi: 10.1186/s12966-016-0454-y] [Medline: 27927218]
26. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010 Feb 17;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]
27. van Genugten L, Dusseldorp E, Webb TL, van Empelen P. Which combinations of techniques and modes of delivery in internet-based interventions effectively change health behavior? A meta-analysis. J Med Internet Res 2016 Jun 07;18(6):e155 [FREE Full text] [doi: 10.2196/jmir.4218] [Medline: 27268104]
28. van der Ploeg HP, van der Beek AJ, van der Woude LHV, van Mechelen W. Physical activity for people with a disability: a conceptual model. Sports Med 2004;34(10):639-649. [doi: 10.2165/00007256-200434100-00002] [Medline: 15335241]
29. Bartholomew L, Markham C, Ruiter R, Fernandez M, Kok G, Parcel G. Planning health promotion programs: an intervention mapping approach. 4th ed. San Fransisco: Jossey-Bass; 2016.
30. Smith EM, Sakakibara BM, Miller WC. A review of factors influencing participation in social and community activities for wheelchair users. Disabil Rehabil Assist Technol 2016;11(5):361-374 [FREE Full text] [doi: 10.3109/17483107.2014.989420] [Medline: 25472004]
31. Bardus M, van Beurden SB, Smith JR, Abraham C. A review and content analysis of engagement, functionality, aesthetics, information quality, and change techniques in the most popular commercial apps for weight management. Int J Behav Nutr Phys Act 2016 Mar 10;13:35 [FREE Full text] [doi: 10.1186/s12966-016-0359-9] [Medline: 26964880]
32. Swift DL, Johannsen NM, Lavie CJ, Earnest CP, Church TS. The role of exercise and physical activity in weight loss and maintenance. Prog Cardiovasc Dis 2014;56(4):441-447 [FREE Full text] [doi: 10.1016/j.pcad.2013.09.012] [Medline: 24438736]
33. Marshall NS, Glozier N, Grunstein RR. Is sleep duration related to obesity? A critical review of the epidemiological evidence. Sleep Med Rev 2008 Aug;12(4):289-298. [doi: 10.1016/j.smrv.2008.03.001] [Medline: 18485764]
34. Wu Y, Zhai L, Zhang D. Sleep duration and obesity among adults: a meta-analysis of prospective studies. Sleep Med 2014 Dec;15(12):1456-1462. [doi: 10.1016/j.sleep.2014.07.018] [Medline: 25450058]
35. Chaput J. Sleeping more to improve appetite and body weight control: dream or reality? Am J Clin Nutr 2015 Jan;101(1):5-6. [doi: 10.3945/ajcn.114.101543] [Medline: 25527744]
36. Rayward AT, Burton NW, Brown WJ, Holliday EG, Plotnikoff RC, Duncan MJ. Associations between changes in activity and sleep quality and duration over two years. Med Sci Sport Exerc 2018;50(12):2425-2432. [doi: 10.1249/mss.0000000000001715]
37. Prendergast KB, Mackay LM, Schofield GM. The clustering of lifestyle behaviours in New Zealand and their relationship with optimal wellbeing. Int J Behav Med 2016 Oct 4;23(5):571-579. [doi: 10.1007/s12529-016-9552-0] [Medline: 26944753]

XSL•FO
RenderX

38. Fernandez ME, Ten Hoor GA, van Lieshout S, Rodriguez SA, Beidas RS, Parcel G, et al. Implementation mapping: using intervention mapping to develop implementation strategies. Front Public Health 2019;7:158. [doi: 10.3389/fpubh.2019.00158] [Medline: 31275915]
39. Viechtbauer W, Smits L, Kotz D, Budé L, Spigt M, Serroyen J, et al. A simple formula for the calculation of sample size in pilot studies. J Clin Epidemiol 2015 Nov;68(11):1375-1379. [doi: 10.1016/j.jclinepi.2015.04.014] [Medline: 26146089]
40. Brooke J. SUS-A quick and dirty usability scale. In: Usability evaluation in industry. London: CRC Press; 1996:189-194.
41. Lund A. Measuring usability with the USE Questionnaire. Usability Interf 2001;8(2):3-6.
42. Brooke J. Usability engineering in office product development. In: Proceedings of the Second Conference of the British Computer Society, human computer interaction specialist group on People and computers: designing for usability. Cambridge: Cambridge University Press; 1986:249-259.
43. Prochaska JO, DiClemente CC. Stages and processes of self-change of smoking: toward an integrative model of change. J Consult Clin Psychol 1983 Jun;51(3):390-395. [doi: 10.1037//0022-006x.51.3.390] [Medline: 6863699]
44. Bangor A, Staff T, Kortum P, Miller J, Staff T. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009;4(3):114-123.
45. Ortega RM, Pérez-Rodrigo C, López-Sobaler AM. Dietary assessment methods: dietary records. Nutr Hosp 2015 Feb 26;31(Suppl 3):38-45 [FREE Full text] [doi: 10.3305/nh.2015.31.sup3.8749] [Medline: 25719769]
46. Nederlands voedingsstoffenbestand (NEVO). Rijksinstituut voor Volksgezondheid en Milieu. 2019. URL: https://nevo-online.rivm.nl/ [accessed 2019-06-30]
47. Himes JH. New equation to estimate body mass index in amputees. J Am Diet Assoc 1995 Jun;95(6):646. [doi: 10.1016/S0002-8223(95)00175-1] [Medline: 7759736]
48. Osterkamp LK. Current perspective on assessment of human body proportions of relevance to amputees. J Am Diet Assoc 1995 Feb;95(2):215-218. [doi: 10.1016/S0002-8223(95)00050-X] [Medline: 7852688]
49. Kyle UG, Genton L, Karsegard L, Slosman DO, Pichard C. Single prediction equation for bioelectrical impedance analysis in adults aged 20--94 years. Nutrition 2001 Mar;17(3):248-253. [doi: 10.1016/s0899-9007(00)00553-0] [Medline: 11312069]
50. Washburn RA, Zhu W, McAuley E, Frogley M, Figoni SF. The physical activity scale for individuals with physical disabilities: development and evaluation. Arch Phys Med Rehabil 2002 Feb;83(2):193-200. [doi: 10.1053/apmr.2002.27467] [Medline: 11833022]
51. Teeuw B, Schwarzer R, Jerusalem M. Dutch Adaptation of the General Self-Efficacy Scale. Berlin; 1994. URL: http://userpage.fu-berlin.de/~health/dutch.htm [accessed 2019-01-07]
52. Barlow JH, Williams B, Wright C. The Generalized Self-Efficacy Scale in people with arthritis. Arthritis Care Res 1996 Jun;9(3):189-196. [doi: 10.1002/1529-0131(199606)9:3<189::aid-anr1790090307>3.0.co;2-#] [Medline: 8971228]
53. Peter C, Cieza A, Geyh S. Rasch analysis of the General Self-Efficacy Scale in spinal cord injury. J Health Psychol 2014 Apr;19(4):544-555. [doi: 10.1177/1359105313475897] [Medline: 23463793]
54. Nooijen CFJ, Post MWM, Spijkerman DCM, Bergen MP, Stam HJ, van den Berg-Emons RJG. Exercise self-efficacy in persons with spinal cord injury: psychometric properties of the Dutch translation of the Exercise Self-Efficacy Scale. J Rehabil Med 2013 Apr;45(4):347-350 [FREE Full text] [doi: 10.2340/16501977-1112] [Medline: 23474694]
55. McAuley E. Self-efficacy and the maintenance of exercise participation in older adults. J Behav Med 1993 Feb;16(1):103-113. [doi: 10.1007/BF00844757] [Medline: 8433355]
56. Vercoulen JH, Swanink CM, Fennis JF, Galama JM, van der Meer JW, Bleijenberg G. Dimensional assessment of chronic fatigue syndrome. J Psychosom Res 1994 Jul;38(5):383-392. [doi: 10.1016/0022-3999(94)90099-x] [Medline: 7965927]
57. Makowiec-Dabrowska T, Koszada-Włodarczyk W. The CIS20R Questionnaire and its suitability for prolonged fatigue studies. Med Pr 2006;57(4):335-345. [Medline: 17133914]
58. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989 May;28(2):193-213. [doi: 10.1016/0165-1781(89)90047-4] [Medline: 2748771]
59. Mollayeva T, Thurairajah P, Burton K, Mollayeva S, Shapiro CM, Colantonio A. The Pittsburgh sleep quality index as a screening tool for sleep dysfunction in clinical and non-clinical samples: A systematic review and meta-analysis. Sleep Med Rev 2016 Feb;25:52-73. [doi: 10.1016/j.smrv.2015.01.009] [Medline: 26163057]
60. Froehlich-Grobe K, Andresen EM, Caburnay C, White GW. Measuring health-related quality of life for persons with mobility impairments: an enabled version of the short-form 36 (SF-36E). Qual Life Res 2008 Jun 22;17(5):751-770. [doi: 10.1007/s11136-008-9342-5] [Medline: 18427950]
61. Aaronson NK, Muller M, Cohen PD, Essink-Bot ML, Fekkes M, Sanderman R, et al. Translation, validation, and norming of the Dutch language version of the SF-36 Health Survey in community and chronic disease populations. J Clin Epidemiol 1998 Nov;51(11):1055-1068. [doi: 10.1016/s0895-4356(98)00097-3] [Medline: 9817123]
62. Green J, Thorogood N. Qualitative methods for health research. 4th ed. London: SAGE Publications; 2018.
63. Tomczak M, Tomczak E. The need to report effect size estimates revisited. An overview of some recommended measures of effect size. Trends Sport Sci 2014;21(1):19-26. [doi: 10.4135/9781473997448]
64. Fogg B. A behavior model for persuasive design. In: Proceedings of the 4th International Conference on Persuasive Technology. Claremont, CA: ACM Press; 2009:1-7.

XSL•FO
RenderX

65. Vandelanotte C, Spathonis KM, Eakin EG, Owen N. Website-delivered physical activity interventions a review of the literature. Am J Prev Med 2007 Jul;33(1):54-64. [doi: 10.1016/j.amepre.2007.02.041] [Medline: 17572313]
66. Gilliland J, Sadler R, Clark A, O'Connor C, Milczarek M, Doherty S. Using a smartphone application to promote healthy dietary behaviours and local food consumption. Biomed Res Int 2015;2015:841368. [doi: 10.1155/2015/841368] [Medline: 26380298]
67. Wang JB, Cadmus-Bertram LA, Natarajan L, White MM, Madanat H, Nichols JF, et al. Wearable sensor/device (Fitbit One) and SMS text-messaging prompts to increase physical activity in overweight and obese adults: a randomized controlled trial. Telemed J E Health 2015 Oct;21(10):782-792 [FREE Full text] [doi: 10.1089/tmj.2014.0176] [Medline: 26431257]
68. Evidence scan: Training professionals in motivational interviewing. The Health Foundation. 2011 Nov. URL: https://www.health.org.uk/sites/default/files/TrainingProfessionalsInMotivationalInterviewing.pdf [accessed 2021-07-05]
69. Hall JA, Horgan G, Murphy NA. Nonverbal communication. Annu Rev Psychol 2019 Jan 04;70:271-294. [doi: 10.1146/annurev-psych-010418-103145] [Medline: 30256720]
70. Houlihan BV, Brody M, Everhart-Skeels S, Pernigotti D, Burnett S, Zazula J, et al. Randomized trial of a peer-led, telephone-based empowerment intervention for persons with chronic spinal cord injury improves health self-management. Arch Phys Med Rehabil 2017 Jun;98(6):1067-1076. [doi: 10.1016/j.apmr.2017.02.005] [Medline: 28284835]
71. Martin Ginis KA, Ma JK, Latimer-Cheung AE, Rimmer JH. A systematic review of review articles addressing factors related to physical activity participation among children and adults with physical disabilities. Health Psychol Rev 2016 Dec;10(4):478-494. [doi: 10.1080/17437199.2016.1198240] [Medline: 27265062]
72. van den Berg-Emons RJ, L'Ortye AA, Buffart LM, Nieuwenhuijsen C, Nooijen CF, Bergen MP, et al. Validation of the Physical Activity Scale for individuals with physical disabilities. Arch Phys Med Rehabil 2011 Jun;92(6):923-928. [doi: 10.1016/j.apmr.2010.12.006] [Medline: 21507382]

Abbreviations

BIA: bioelectrical impedance analysis
CIS20R: Checklist Individual Strength
LLA: lower limb amputation
MET: metabolic equivalents of task
mHealth: mobile health
PASIPD: Physical Activity Scale for Individuals with Physical Disabilities
PSQI: Pittsburgh Sleep Quality Index
SCI: spinal cord injury
SF-36E: Short-Form Health Survey Enabled
SUS: System Usability Scale
USE: Usefulness, Satisfaction, and Ease of use
WC: waist circumference
WHEELS: Wheelchair ExercisE and Lifestyle Study

Edited by G Eysenbach; submitted 12.10.20; peer-reviewed by B Mortenson, S Skeels; comments to author 05.01.21; revised version received 01.03.21; accepted 08.05.21; published 09.08.21.

Please cite as:
Hoevenaars D, Holla JFM, te Loo L, Koedijker JM, Dankers S, Houdijk H, Visser B, Janssen TWJ, de Groot S, Deutekom M, WHEELS Study Group
Mobile App (WHEELS) to Promote a Healthy Lifestyle in Wheelchair Users With Spinal Cord Injury or Lower Limb Amputation: Usability and Feasibility Study
JMIR Form Res 2021;5(8):e24909
URL: https://formative.jmir.org/2021/8/e24909
doi:10.2196/24909
PMID:34379056

©Dirk Hoevenaars, Jasmijn F M Holla, Leonie te Loo, Johan M Koedijker, Sarah Dankers, Han Houdijk, Bart Visser, Thomas W J Janssen, Sonja de Groot, Marije Deutekom, WHEELS Study Group. Originally published in JMIR Formative Research (https://formative.jmir.org), 09.08.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete

XSL•FO
RenderX

bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Design of the Maternal Website EMAeHealth That Supports Decision-Making During Pregnancy and in the Postpartum Period: Collaborative Action Research Study

Isabel Artieta-Pinedo[1*], RN, CNM, PhD; Carmen Paz-Pascual[1*], RN, CNM; Paola Bully[2*], BSc, MSc; Maite Espinosa[1*], BSc, MSc; EmaQ Group[3]

[1]Osakidetza-Basque Health Service, Biocruces-Bizkaia Health Research Institute, Osi Barakaldo-Sestao, Barakaldo, Spain

[2]Paola Bully Methodological and Statistical Consulting, Sopuerta, Spain

[3]Biocruces-Bizkaia Health Research Institute, Barakaldo, Spain

[*]these authors contributed equally

Corresponding Author:
Isabel Artieta-Pinedo, RN, CNM, PhD
Osakidetza-Basque Health Service
Biocruces-Bizkaia Health Research Institute
Osi Barakaldo-Sestao
Lurkizaga S/N
Barakaldo, 48901
Spain
Phone: 34 946006667
Email: misabel.artieta@ehu.eus

Abstract

Background: Despite the benefit maternal education has for women, it needs new tools to increase its effectiveness and scope, in tune with the needs of current users.

Objective: We attempted to develop a multifunctional personalized eHealth platform aimed at the self-management of health in relation to maternity, which can be considered a flexible and adaptable maternal education tool.

Methods: The International Patient Decision Aid Standards (IPDAS) were applied. A website prototype was developed for implementation in the public health system using a collaborative action research process, in which experts and patients participate, with qualitative research techniques, as well as focus groups, prioritization, and consensus techniques.

Results: We have proposed a website that includes (1) systematically updated information related to clinical practice guidelines, (2) interaction between peers and users/professionals, (3) instruments for self-assessment of health needs as a basis for working on counseling, agreement on actions, help in the search for resources, support in decision-making, and monitoring and evaluation of results, and (4) access for women to their clinical data and the option of sharing the data with other health agents. These components, with different access requirements, would be reviewed through iterative cycles depending on the frequency and effectiveness resulting from their use and would be accessible from any digital device.

Conclusions: A website that supports maternal education should contain not only information, but also resources for individual attention and social support. Its usefulness for the health and satisfaction of women should be evaluated in various different environments.

(JMIR Form Res 2021;5(8):e28855) doi:10.2196/28855

KEYWORDS

prenatal education; women; patient decision aid; decision-making; clinical decision support systems; action research and pregnancy; implementation science; health service needs and demands

XSL•FO
RenderX

Introduction

Background

Permanent and sometimes abrupt changes in our societies are undeniable facts, and changes in health, society, and technology often come quickly and simultaneously. Health care tries to respond to increased demand and the current needs of users with the new resources available. In recent years, the examples of services that use new technologies in the care of patients with chronic pathology or multimorbidity have multiplied, and telemedicine has a widely understood potential for training and support in shared decision-making [1,2], using tools including video consultation, mobile messaging, and expert systems that issue responses from clinical practice guides or available resources. Sometimes they are accessible from platforms that are outside the health system but very widespread among the population, such as Facebook [3].

Pregnant and postpartum women represent a demographic that is especially in need of support. The transition to motherhood involves a change of identity, social role, and activity from the known to a new reality [4,5]. In this process, they feel great pressure to do well, which can ultimately reduce their well-being [6]. Stress during pregnancy or in the postpartum period generates physical and psychological responses, which can ultimately lead to disorders such as depression, anxiety, and emotional distress [7]. The effort of professionals to generate preparation tools for childbirth and the promotion of physical and psychological health are common [8,9]. However, there is also agreement on the need to update these tools [7,10]. The use of e-technologies offers advantages, such as accessibility at any time and place, and for a very large population, information can be gathered in a personalized way and the user can review the resource as much as desired.

Women of reproductive age, and more specifically, those in the stage of maternity, represent a population group that is especially likely to benefit from the use of these new technologies. They systematically use the internet as a source of information, including information about their health [11], and frequently base their decisions on the information obtained [12,13].

Previous experiences with the use of telemedicine during pregnancy have given positive results in the prevention of pre-eclampsia [14] and in the management of diabetes [15]. In the postpartum period, online communication between peers has been associated with lower levels of anxiety and feelings of loneliness [16] and attention through a website [17] or a mobile app [18], with a higher perception of self-efficacy, psychological well-being, and satisfaction with the care received. The availability of computer-based educational programs has also been shown to be associated with greater self-efficacy and a more positive attitude toward breastfeeding [19]. However, the determinants of its effectiveness remain to be explored, since the results vary depending on the behavior studied [20], the time of measurement [21], or the population studied [22].

It seems necessary, and indeed urgent, to have a web tool to support decision-making for the maternity stage, which is reliable [23], personalized [24], and based on the women themselves and the initiatives of professionals [25]. The Medical Research Council provides a theoretical framework for the development of this complex intervention [26] and recommends actions that facilitate implementation in the real world. Therefore, it was decided that the website should include people who teach and receive maternal education from the beginning, and a Participatory Action Research study, based on theories of behavioral change, was used. The integrated model proposed by Fishbein et al [27] was considered to be the most appropriate since it takes into account the influence of attitudes, self-confidence, sociodemographic variables, and social norms on the intention to act and ultimately on the achievement of healthy habits. The influence of these variables, in addition, would be modified by barriers and facilitators in the environment. This theoretical model has been successfully applied in the study of sexual behavior [28] and has been used in similar studies to prevent excessive weight gain during pregnancy [29].

The development of this tool must follow a strategy that takes into account the resources, needs, and characteristics of the target population, and for this development, a collaborative environment should be created, involving the population, putting it into practice, implementing it, and evaluating the process through iterative cycles of continuous improvement [30].

Objective

We attempted to develop a multifunctional and personalized eHealth platform, bringing together patients and professionals, which allows them to attend and monitor their health needs and take informed decisions during pregnancy, childbirth, and parenting. Subsequently, its clinical effectiveness and its implementation in routine conditions will be evaluated (usability and acceptability by users and professionals, and impact on the health system). This article focuses on describing the design and development of the tool.

Methods

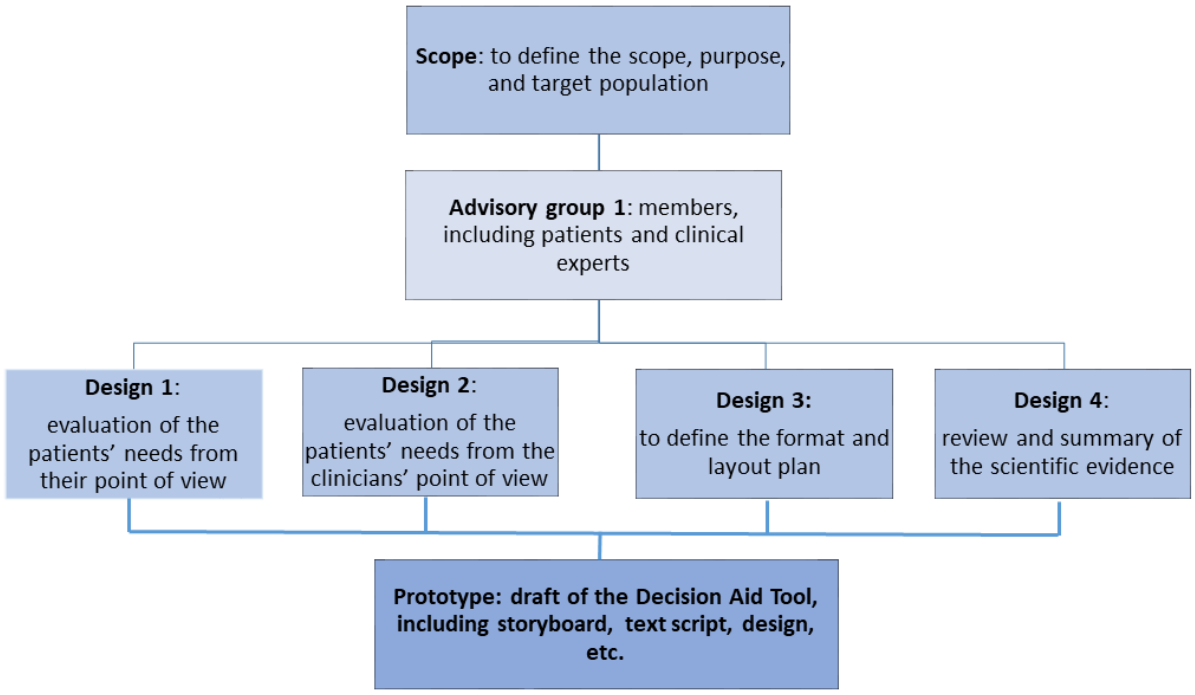
Procedure for the Development of the EMAeHealth Tool

For the development of the “EMAeHealth” tool, the International Patient Decision Aid Standards (IPDAS) were followed (Figure 1) [31]. The shared decision-making (SDM) model is defined as an approach in which health care professionals and patients make decisions together using the best available evidence. The SDM model emphasizes respect for the patient’s autonomy and promotes training so that the patient is involved in the decision-making process.

The study was conducted in Basque Country, a region in the north of Spain with a population of 1.2 million. This region has a public health service, which is free at the point of use, and is universally and readily accessible, with local centers that provide care for people living in the surrounding geographical area. During pregnancy, women are monitored via alternating appointments with their primary care midwife and gynecologist, and in the last trimester, they are invited to attend antenatal education sessions run by one of the midwives at the health center.

XSL•FO
RenderX

Figure 1. Development process of a decision aid tool. Adapted from the report by Perestelo-Pérez L et al [31].



Scope

We propose a flexible maternal education tool that will be adapted and adaptable to each population group, offered universally by the public health system, and that supports women continuously throughout the maternity process.

Advisory Group

A multidisciplinary research team was formed of hospital and primary care midwives, pediatricians, psychologists, and experts in methodology. The process of pregnancy and parenting is so common that the team could hardly distance itself from their own experiences as mothers, which provided a "patient" vision.

Design 1

A qualitative study (focus groups) was carried out following the guidelines of Krueger and Casey [32]. Women who went to see their midwife were recruited consecutively. Two groups were formed (pregnant women and postpartum women), and they were given a questionnaire that allowed each of the two groups to be divided according to income level and education. Finally, 31 women participated in four groups that were homogeneous in terms of parity and socioeconomic status. Each study participant was given the study fact sheet along with an informed consent form, and all of them agreed to audio recordings of the meetings. For the collection of information, three experts in qualitative methods designed a script related to the topics to be discussed (perceived needs regarding pregnancy care and maternal education), using the thematic content analysis method. Three researchers independently read the transcripts of the sessions and arranged the information into the topics to be discussed. They assigned codes to the text segments and regrouped them into categories using ATLAS.ti software (Scientific Software Development GmbH). With these categories and subcategories, and the relationships between them, a conceptual structure was constructed by each of the analysts. It was later triangulated and compared again with the text to give the final results [33].

Design 2

The advisory group was joined by other professionals, including a gynecologist, two pediatric nurses, experts in breastfeeding and health education, epidemiologists, psychologists, sociologists, and experts in qualitative research. The selection of these people was based on the opinion of the research team and on their work experience and previous interest and participation in research work. Finally, a group of nine women and two men was formed (age between 30 and 57 years), with health care experience ranging from 5 to 30 years.

Nominal group techniques were used in the sessions. After a study and individual analysis work, each team member had 5 minutes to explain the main health needs of women during pregnancy and the first year postpartum. After the round of presentations, time was devoted to identifying common ideas. In those in which there was a discrepancy, explanation and defense was allowed, and they were finally ranked by voting [34].

Design 3

In order to encourage participation and implementation of the tool, a meeting was scheduled, to which all active midwives in the health centers of Bizkaia and in the hospitals of Cruces and Basurto were invited. This meeting was attended by 25% of the total 20 midwives. They were shown the results obtained in the previous steps and asked to choose the actions they considered most feasible and relevant from all the proposals for a new framework of continuous personalized maternal education [34].

XSL•FO
RenderX

Design 4

A descriptive study was carried out. Using the most widely used internet search engines in our region (Google, Yahoo, and Bing), the first 25 web pages that appeared when entering words, such as pregnancy, childbirth, postpartum, and breastfeeding, in Spanish and English were selected.

Each of these websites was evaluated by the midwives participating in the study using the LIDA questionnaire, version 1.2 [35] as a measure of reliability, accessibility, and usability, since it is a validated questionnaire that considers the characteristics of the author's reliability, conflicts of interest, references, and relevance common in other studies [36]. Each of the websites received a score on this questionnaire.

Throughout the process, care was taken to ensure the trustworthiness of the data. Several facts give credibility and reliability to the data obtained, including its origin in a team made up of health care professionals with high involvement in their fields and observations over the years, the high degree of agreement between professionals from different fields (pediatrics, gynecology, nursing, and hospital and primary care midwifery), and the transferability that has been sought through representativeness in both women in the focus groups and the midwives who were asked to select the most feasible and priority interventions. We think that the duration, the coherence between the different groups and stages, the congruence with the bibliography, and the inclusion of different types of professionals and women allowed us to reasonably assume the credibility, transferability, dependability, and confirmability of our data.

Both the design of the website and its subsequent evaluation have received a favorable report from the Ethics Committee for Clinical Research of the Basque Country (PI2012072 and PI20200044).

Results

Design 1

It was seen that the focus of women's worries were different based on the stages of pregnancy/postpartum. In early pregnancy, women's main concern was for "everything to go well." As the pregnancy progressed, they needed emotional support and wanted to feel confident and be self-reliant to face their fears of the birth and care of their children. The needs expressed by women were as follows: accurate information that was accessible and suited to their specific life moments; flexible maternal education programs in terms of schedule and content; and greater participation of partners. All of them had a positive opinion of our health system and the role of midwives, although they would like more support after giving birth. Puerperal women reported "excessive pressure" in favor of breastfeeding despite its difficulty [33].

A website allows permanent accessibility and continuity from the beginning of pregnancy to 1 year after delivery, so it would respond to several of the needs expressed. This website would require rigorous stable information based on clinical practice guidelines. It should also allow interaction, in order to address the specific doubts of each moment of the motherhood process (for the woman or her partner).

Design 2

After seven sessions, the group of experts proposed three moments in which important changes in women's health needs can be seen as follows: at the beginning of pregnancy, at the onset of childbirth, and in the first months after childbirth. Three interactive exploratory tools were proposed for identifying potential difficulties, establishing peer-to-peer and health care communication, and negotiating possible solutions to specific needs. Personalized intervention is key, because a woman's interest and resources vary depending on personal characteristics such as age, race, parity, and pregnancy evolution.

The care should support the identification of specific needs, and guide the advice and support for the achievement of personal goals in areas such as choice of the type of delivery, breastfeeding, or parenting [36].

The proposed website should, therefore, include self-assessment tools, preferably appropriate to the three key moments, and would provide possible answers for the need felt, allowing the woman's agreement to be checked and the case to be followed up. For example, answers could be given to doubts about a healthy diet and proper weight gain, or providing resources for the proper management of childbirth anxiety.

Design 3

Casting a vote, all the participant midwives considered that a web format that would allow communication with the women would be the most useful, would offer the most possibilities of meeting the needs expressed by the women regarding continued and accessible care, and would facilitate the solution of occasional problems. They also considered it very necessary to have a channel of communication between the women who live in the same area and share a health system or similar pregnancy protocols, to facilitate the formation of social networks and the solution of frequently asked questions. They believe that this forum could be moderated by a professional.

Design 4

The quality of information available on the internet regarding pregnancy, childbirth, and the postpartum period in general was moderate, poor, or very poor, with rare references to the source of information. Reliability was higher on websites belonging to public bodies, universities, or health companies than on websites belonging to commercial companies, and was higher among all of them than on websites of individuals or small private groups.

It was found that the higher quality websites were up-to-date, with information based on scientific evidence, frequently including videos and personal experiences of other women, and the possibility of interacting with peers and health professionals [36].

"EMAeHealth" Prototype

As a result of the previous analysis and reflections, a prototype named the "EMAeHealth" web portal (Figure 2) was developed. The name stands for "maternal education" in Spanish (educación

XSL•FO
RenderX

maternal). It includes the following *functionalities*: (1) help women to self-evaluate and prioritize their health needs related to pregnancy, childbirth, puerperium, lactation, and child care, identifying people and populations at risk of physical, mental, or social complications; (2) promote training and self-management, establish decision-making support systems, and serve as a guide in identifying available resources; and (3) allow access to content from multiple devices and operating systems, and store the data safely.

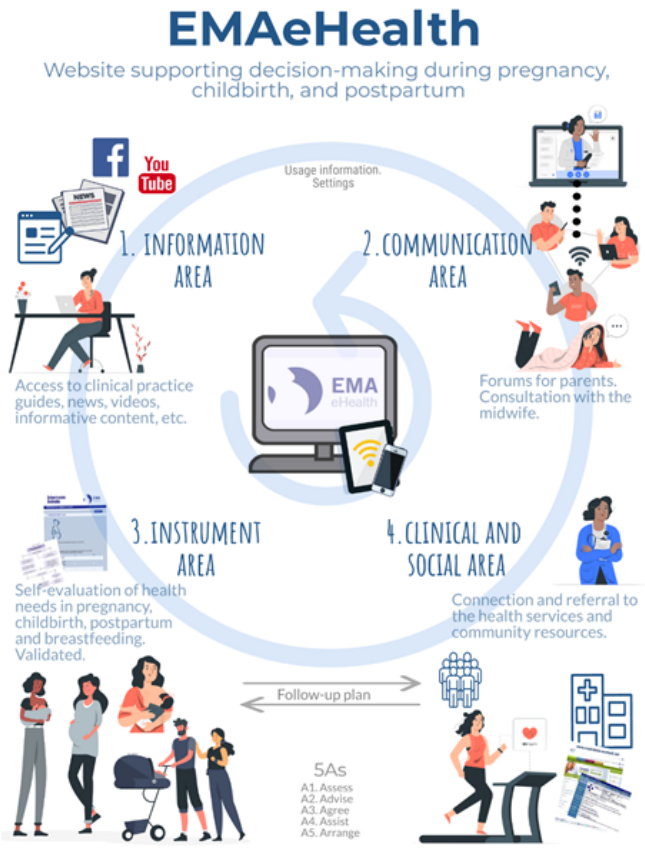
Experts in software engineering propose a *structure* (Figure 3) in which, starting from a home page with the general menu, the different contents are distributed in a modular structure related to the different needs and offers, and with different degrees of accessibility based on the identity of the user (patient, professional, and health system). An initial module would give access to publicly available information and to various self-assessment tools, which would not offer any further resources (EMAportal). A second module would allow access to both peer communication forums and resources available in response to self-assessment, for which the woman would have to identify herself as a member of the group (EMAcommunity). A third module, which is more complex, would allow access to health data and the monitoring of care, for which a login would be needed to guarantee confidentiality (EMAhealthcare). The structure would allow professionals and administrators to get statistics on its use and health outcomes, enabling the modification, elimination, or enhancement of different areas.

Functionalities and applications would be added to this initial structure based on the feedback received (Figure 4).

The platform contains a file or data repository for monitoring information, with a sufficient guarantee of storage security (EMAcore), which allows the transmission of this data between servers of different information systems and their connection between several systems (the Basque Health System *Osakidetza* and other health and social systems) (EMAApi). The complete development of this website would involve the participation of the organization where the woman is registered for the pregnancy, and therefore, the most personal data and interactions should be stored along with the rest of the medical history, thus ensuring confidentiality.

Regarding the *content*, the EMAeHealth portal contemplates interaction with the woman and her partner at various levels as presented below.

Figure 2. Diagram of the EMAeHealth tool prototype for supporting women during pregnancy, chidbirth, postpartum, and breastfeeding/parenting. There are four main areas.



XSL•FO
RenderX

Figure 3. Architecture and information flow of the EMAeHealth website.

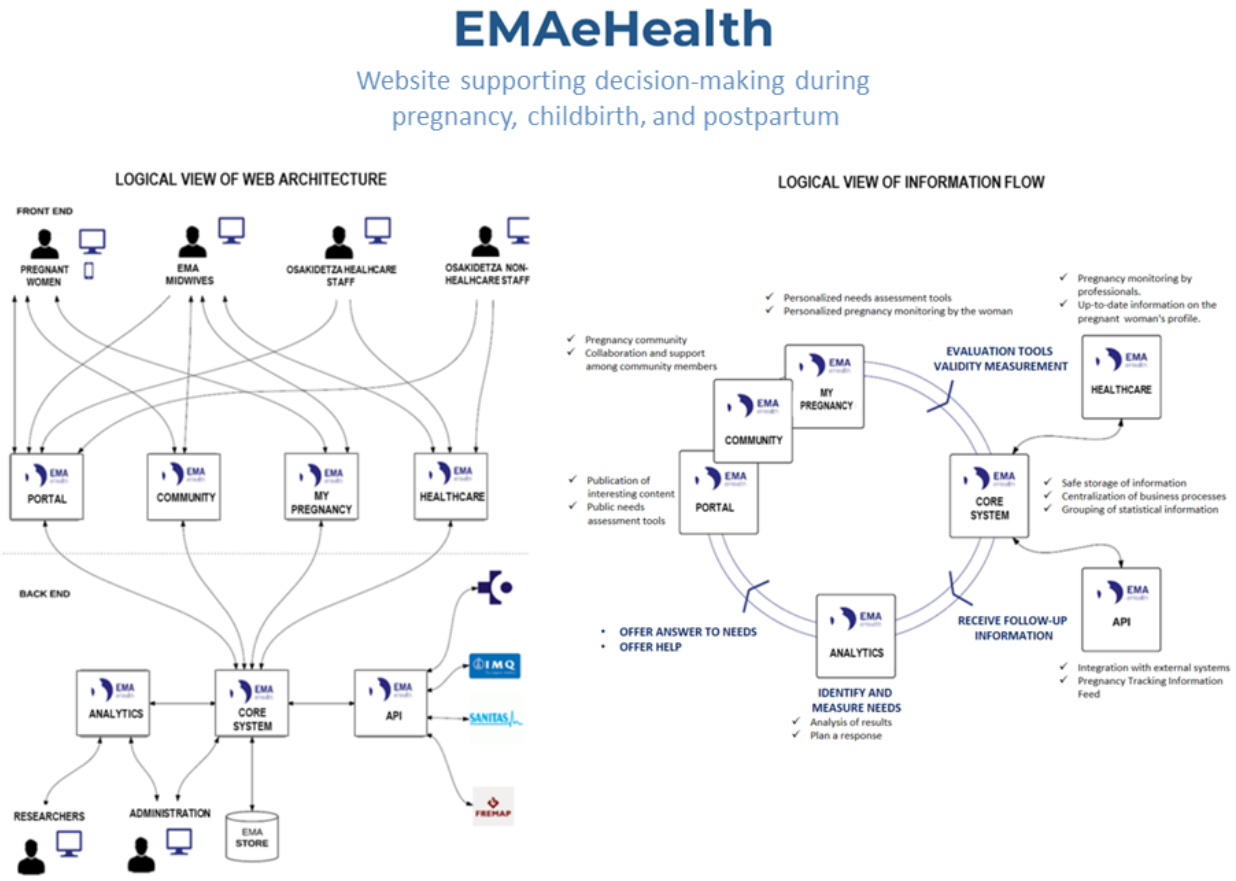
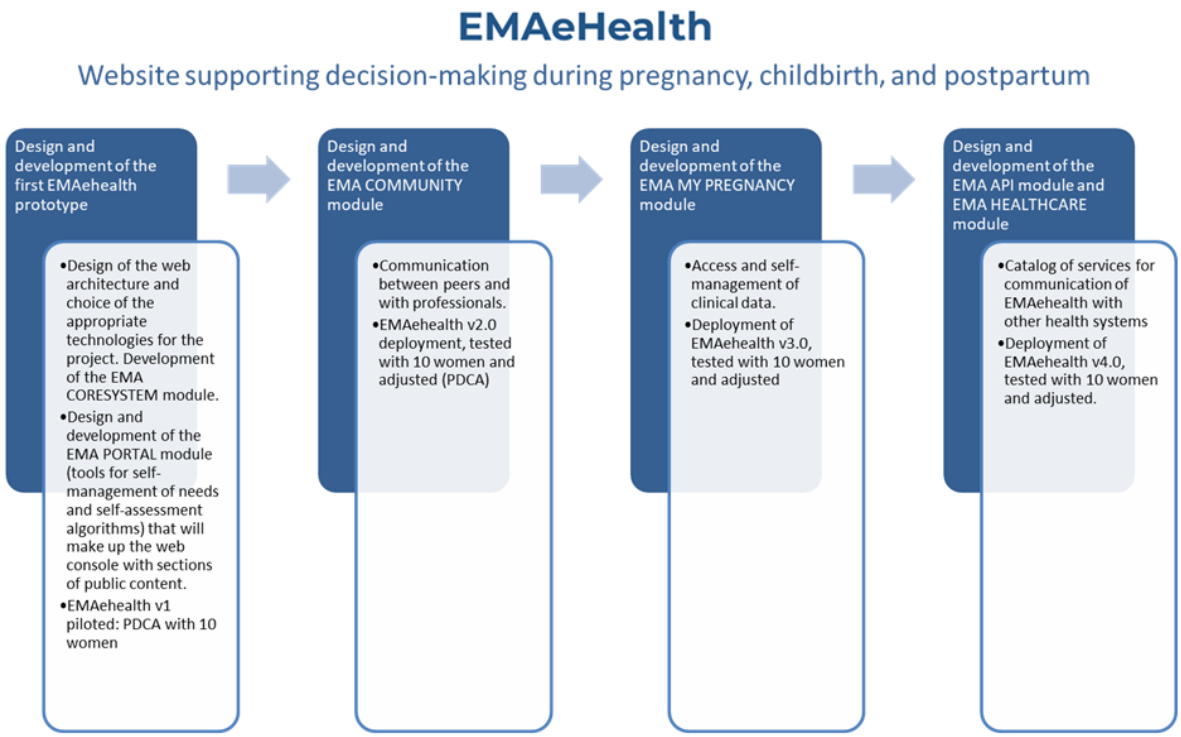


Figure 4. Flowchart of the development of the EMAeHealth tool up to its clinical application.



XSL•FO
RenderX

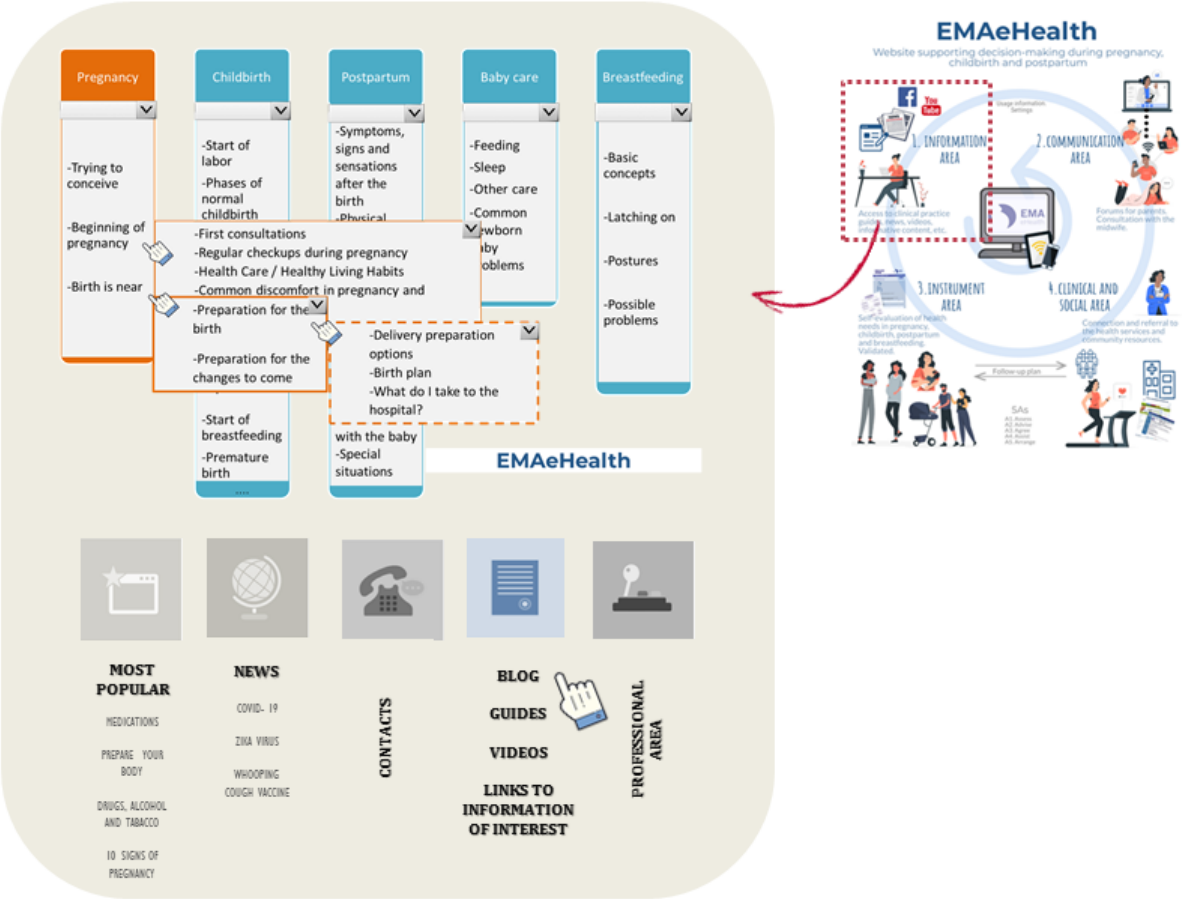
Information Area

This is an area open to the public that will offer information based on evidence and be constantly updated by health service professionals (Figure 5). The website will inform users about the maternity process, including the desire for pregnancy; first, second, and third trimesters; labor and delivery; immediate and late postpartum; and breastfeeding. It would also include care of the newborn, paying special attention to the vaccination calendar during the baby's first year of life by sending reminders about the dates and following up on the correct vaccination. In each section, the physiology of the process, healthy lifestyle, possible health problems, common remedies, and available resources will be described.

Current news will also be posted such as changes in legislation related to the process and the effects of the Zika virus and COVID-19. Applications, such as calculators, estimates of appropriate weight, an exercise calendar, and healthy diets will be offered.

It will be presented in multiple formats and support blogs, videos, interviews, articles, and data suitable for different types of population. The portal will have a data monitoring and management system that will allow the research team to collect browsing statistics on the different sections and adjust the content if necessary.

Figure 5. Contents in the EMAeHealth information area. Information based on available clinical evidence, selected and reviewed by professionals.



Communication Area

This is an area that will require user login and will allow the female user to contact other women from a close environment and in a situation similar to her situation, through forums or conversations, or consult with her health care professional. Peer interaction is one of the most highly valued functions among users of digital tools [37,38], and it has been shown to increase postpartum mental well-being [18]. The regular use of social networks to not only receive but also contribute knowledge can favor a greater connection to the tool and therefore greater access to information based on evidence and greater possibility of interaction with the health care professional [38]. A professional will moderate the tool on a rotating basis and answer questions. There will be an administrator who can block anyone who distorts the content or alters relationships. In a project like this, part of the work of midwives would be focused on non-face-to-face interventions, so work plans should also be adapted to this. Primary care midwives currently provide much of the service through prearranged appointments, and this time could be split between face-to-face consultations and online interaction.

On the other hand, the areas that need identification will only be available to the woman during pregnancy and the first year postpartum, or until the end of the pregnancy in the case of abortion. In this way, the women included share close experiences in time and space, which would predictably make the contact more useful. Access would be reactivated for each pregnancy for a similar period.

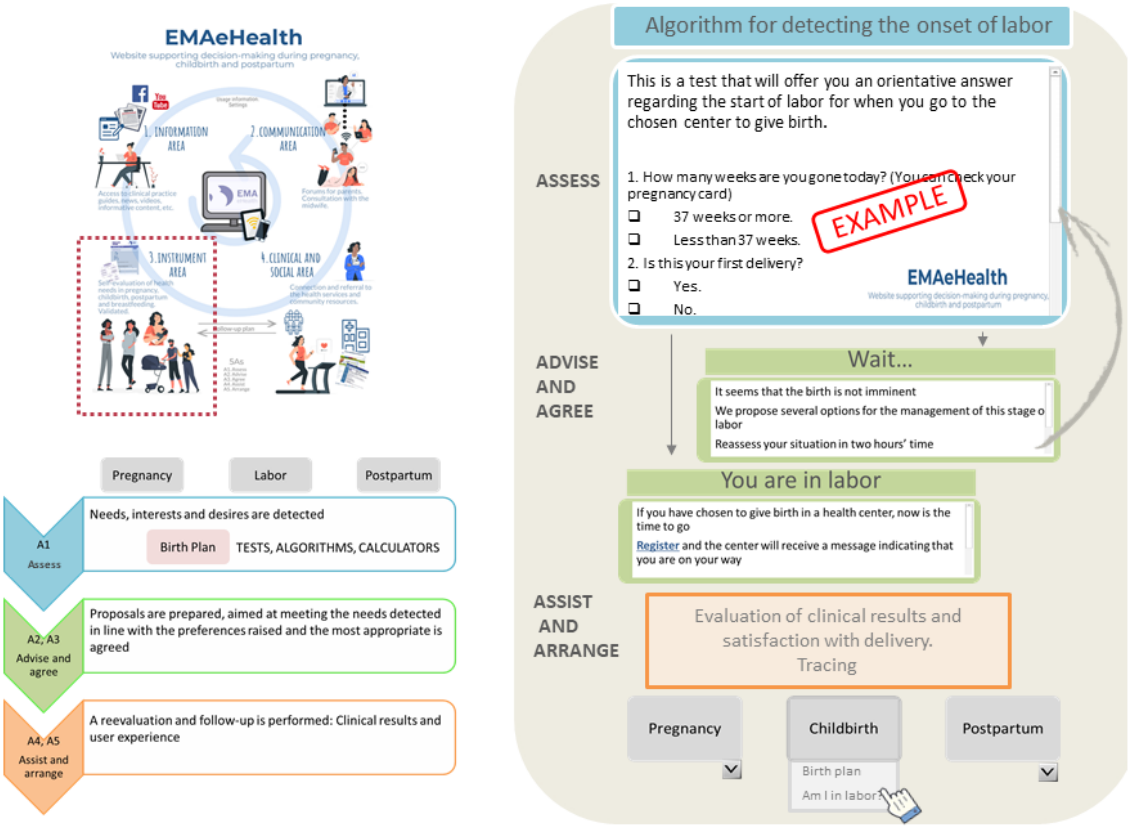
XSL•FO
RenderX

Health Self-Management Area

This area has valid and reliable *self-assessment instruments* to check or reflect on one's own health needs (Figure 6). It consists of scales and questionnaires that cover preferences, personality traits, culture, or beliefs that can determine a decision and the satisfaction obtained with it. As suggested by the group of experts, these self-assessment tools could be presented at least at three key moments as follows: start of pregnancy, onset of birth, and postpartum period. In this space, we would work directly with individual health needs, applying the following five As: assess (self-assessment and prioritization), advise (training/support), agree (planning actions), assist (guides to available resources), and arrange (carrying out monitoring and evaluating the results). Following this procedure, the woman responds to a validated questionnaire on her needs and preferences related to childbirth, and depending on the results, the most suitable alternatives and their availability are offered. Once the woman selects her favorite, the information on how to access it is provided, and follow-up and feedback on the result achieved are agreed. This process will often be mixed, combining the responses from the digital tool and the midwife who attends to the woman.

The same strategy will be followed with all those topics that, based on the information collected on the web, demonstrate their relevance and interest for the user population (eg, food, newborn care, and pelvic floor rehabilitation). The validation process of the self-assessment tools initially included is currently being carried out with the support of the Basque Government (Basque Government Health Department; grant number: 2018111087).

Figure 6. Contents in the EMAeHealth self-assessment instruments area. The algorithm is currently under development and validation, and thus, it may be subject to change.



Clinical Data Area

This area will require user login, and the woman will be able to see her clinical data, add the most recent items, and share them with other professionals if she wishes.

The safe storage of the data and feedback will allow the information presented to be analyzed and adjusted. The website will provide the team with the necessary interface to be able to include, validate, and evaluate new elements in each area (EMAAnalysis). Finally, the website should be accessible from any digital device (eg, phone, tablet, computer, and smart TV) and for any company or organization (eg, health and social) that the woman wishes to use and allow access, naturally with a strict guarantee of data protection.

Discussion

Principal Findings

Despite the benefit that current maternal education can bring to women, new tools are needed to increase its effectiveness and reach. Proposals for renewal arise and are permanently evaluated [39-41]. Digital tools have proven to be useful, and at times like the present, they have been revealed as essential, which is why it is necessary for health care workers to explore new areas of care and support for women [38].

The creation of a collaborative environment, the involvement of the affected population, and the application of international standards to aid patient decision-making have resulted in a

XSL•FO
RenderX

proposal for a digital tool that not only reports back, as is usual with these tools, but also interacts with the woman, allowing her to self-assess her health needs in order to take decisions, self-manage, and respond to these needs.

The resources included in our prototype are, in addition to exhaustive information, communication systems between peers and with professionals, validated self-assessment instruments, follow-up plans based on clinical practice guidelines, and connection with health and social resources in the community. The website will allow women to access their clinical data, and its use will be enhanced by making it accessible from any type of device, whether tablet, mobile, or computer, safely.

This is a prototype and a proposed tool. The next step is to run a pilot test in order to find out its clinical effectiveness, usability, and acceptability by users and professionals, and to see its impact on the health system.

Comparison With Prior Work

Interventions are usually designed, and then, their degree of implementation in clinical practice, their acceptability, and their usefulness are evaluated. A study related to the acceptance of telemedicine [42] concluded that it is essential to take into account the individual characteristics of the end user in the design of the tools and their usefulness as perceived by professionals when handling them in clinical practice. These conclusions are in line with our tool design methodology and with models such as the CFIR (Consolidated Framework for Implementation Research) [43] or the TDF (Theoretical Domains Framework) [44], and the IPDAS methodology [31], which have guided our work. In addition, the internal and external contexts where it will be implemented and the implementation process itself has also been taken into account. We have not found any studies on websites aimed at pregnant or postpartum women that describe, step by step, how the web design was carried out to facilitate subsequent implementation.

Regarding content, in general, websites dedicated to pregnancy offer information on the process and resources available for care, focusing on specific aspects, such as suitable weight gain, giving up toxins, and increasing vaccination rates [45]. In the postpartum period, they also address specific problems, such as the possibility of screening for anxiety and depression [46], which is then followed up with psychoeducational interventions, peer support, and psychological therapy. Very few tools offer comprehensive care for women during pregnancy and in the postpartum period, with the exception of the NHS website.

Limitations

The fact that we looked at the specific needs of pregnant and postpartum women in Basque Country plays in favor of the adaptability of the tool to our context; however, it might not be generalizable to other populations with different characteristics, thus requiring adaptation when extrapolating it to other contexts.

The maintenance of this type of tool requires the involvement of professionals and the organization in which they work, that is, the involvement of health service managers. The job stability of professionals in the public health system favors the continuity of the tool, as well as their participation in its design. However, pressure on health care resulting from COVID-19 may constitute a barrier to its implementation and maintenance. The involvement of the organization might also be diminished by other priorities at this difficult moment.

Conclusions

Although it is true that some health websites are not widely used or are declining in use [41], the “EMAeHealth” platform we propose has shown, since its inception, some characteristics that would facilitate both its adoption and maintenance, as well as its effectiveness in increasing the quality of the care provided [31].

First, its design and development are based on needs expressed by users and health professionals, which will probably improve its implementation [47]. Its flexibility and adaptability will also facilitate its use.

Second, it is proposed as a resource that would be part of the public system. Websites from official organizations have proven to be more rigorous and reliable [36]. Additionally, they can eliminate any hint of conflict of interest that could occur on privately financed websites and favor more universal use, eliminating possible differences based on the socioeconomic level [48,49]. This tool makes it possible to link support for shared decision-making with the use of clinical practice guidelines as a source of evidence and as a basis for recommendations on care or self-care, which would boost acceptance and adoption of the latter by professionals.

Third, it is a customized tool. Among all patients generally, in addition to their physical and mental well-being, health care has to attend to patients’ values, beliefs, or capacities [50], but this is a particularly substantial part of attention in maternity. Pregnancy is a physiological stage in principle, but also a time when self-perception, parenting styles, and other elements are the most relevant variables, and in which other actors, such as partners and family members, with their own customs, convictions, or even ideologies (cultural and social baggage), are directly involved.

Naturally, the website described here should not be seen as a substitute for personal attention [2,31]. Research shows that, in addition to regularly consulting the internet, women go to midwives and consider them their best source of information [12,13,51]. Although useful, these web tools are not sufficient to meet all the needs of women at this stage [52]. Their full potential would be achieved as support options, serving women who wish to take a more active role in managing their health.

In conclusion, health professionals must participate in the adaptation of new technologies in daily practice [53], which attend to the needs of the population. Involving the public in the design phase will facilitate the implementation of EMAeHealth and similar resources as generators of a higher quality of life for women, without increasing costs in health systems.

XSL•FO
RenderX

Acknowledgments

In order to facilitate the usability and subsequent implementation of the EMAeHealth tool in routine practice, the participation, from the beginning, of health and nonhealth professional experts in the area of maternal health, including patients, has been considered. Their contribution has been continuous in the design of the prototype and in the evaluation of the content. This work was co-financed in part by the Basque Government Department of Health (grant 2018111087) and by the Carlos III Health Institute, through the project "PI13/02632" (co-financed by the European Regional Development Fund/European Social Fund). EMA-Q Group is made up of: Amaia Maquibar, Ana Cristina Fernandez, Angela Rodriguez, Covadonga Perez, David Moreno, Gemma Villanueva, Gloria Gutierrez de Teran, Gorane Lozano, Ines Cabeza, Itziar Estalella, Jesus Sanchez, Kata Legarra, M Jesus Mulas, M Jose Trincado, Marie Pierre Gagnon, Mercedes Saenz de Santamaria, Monica Blas, Pilar Amorrortu and Sonia Alvarez.

Authors' Contributions

IAP and CPP are responsible for the project; they designed the study, supervised all the stages of its development, and contributed to the interpretation of the results and the writing of the manuscript. ME coordinated part of the fieldwork, the systematic search, screening, review, and selection of the articles, and together with PB, contributed ideas for the interpretation of the data and contributed to the writing of the first draft of the manuscript. All the authors contributed ideas, participated in the collection of information, and reviewed and approved the final draft of the manuscript.

Conflicts of Interest

None declared.

References

1. Pascual M, Muñoz A. PITES-ISA: New services based on Telemedicine and e-Health aimed at interoperability, patient safety and decision support. Madrid, Spain: Instituto de Salud Carlos III: Unidad de Investigación en Telemedicina y e-Salud; Oct 2017.
2. de Molina-Férnandez MI, Raigal-Aran L, de la Flor-Lopez M, Prata P, Font-Jimenez I, Valls-Fonayet F, et al. The effectiveness of a digital shared decision-making tool in hormonal contraception during clinical assessment: study protocol of a randomized controlled trial in Spain. BMC Public Health 2019 Sep 04;19(1):1224 [FREE Full text] [doi: 10.1186/s12889-019-7572-9] [Medline: 31484579]
3. Transforming health and care through technology. NHS Digital. URL: https://digital.nhs.uk/about-nhs-digital/corporate-information-and-documents/transforming-health-and-care-through-technology [accessed 2021-07-18]
4. Laney EK, Hall MEL, Anderson TL, Willingham MM. Becoming a Mother: The Influence of Motherhood on Women's Identity Development. Identity 2015 May 15;15(2):126-145. [doi: 10.1080/15283488.2015.1023440]
5. Mercer RT. Becoming a mother versus maternal role attainment. J Nurs Scholarsh 2004;36(3):226-232. [doi: 10.1111/j.1547-5069.2004.04042.x] [Medline: 15495491]
6. das Neves Carvalho JM, Ribeiro Fonseca Gaspar MF, Ramos Cardoso AM. Challenges of motherhood in the voice of primiparous mothers: initial difficulties. Invest Educ Enferm 2017 Oct;35(3):285-294 [FREE Full text] [doi: 10.17533/udea.iee.v35n3a05] [Medline: 29767909]
7. Pan W, Gau M, Lee T, Jou H, Liu C, Wen T. Mindfulness-based programme on the psychological health of pregnant women. Women Birth 2019 Feb;32(1):e102-e109. [doi: 10.1016/j.wombi.2018.04.018] [Medline: 29752225]
8. McLeish J, Redshaw M. Mothers' accounts of the impact on emotional wellbeing of organised peer support in pregnancy and early parenthood: a qualitative study. BMC Pregnancy Childbirth 2017 Jan 13;17(1):28 [FREE Full text] [doi: 10.1186/s12884-017-1220-0] [Medline: 28086827]
9. Buultjens M, Murphy G, Robinson P, Milgrom J, Monfries M. Women's experiences of, and attitudes to, maternity education across the perinatal period in Victoria, Australia: A mixed-methods approach. Women Birth 2017 Oct;30(5):406-414. [doi: 10.1016/j.wombi.2017.03.005] [Medline: 28389170]
10. McMillan A, Barlow J, Redshaw M. Birth and beyond: a review of the evidence about antenatal education. ResearchGate. 2009. URL: https://www.researchgate.net/publication/258369289_Birth_and_beyond_a_review_of_the_evidence_about_antenatal_education [accessed 2020-03-01]
11. O'Higgins A, Murphy OC, Egan A, Mullaney L, Sheehan S, Turner MJ. The use of digital media by women using the maternity services in a developed country. Ir Med J 2014;107(10):313-315. [Medline: 25556255]
12. Dalhaug EM, Haakstad LAH. What the Health? Information Sources and Maternal Lifestyle Behaviors. Interact J Med Res 2019 Jul 05;8(3):e10355 [FREE Full text] [doi: 10.2196/10355] [Medline: 31278731]
13. Ghiasi A. Health information needs, sources of information, and barriers to accessing health information among pregnant women: a systematic review of research. J Matern Fetal Neonatal Med 2021 Apr;34(8):1320-1330. [doi: 10.1080/14767058.2019.1634685] [Medline: 31216921]

XSL•FO
RenderX

14. van den Heuvel JFM, Kariman SS, van Solinge WW, Franx A, Lely AT, Bekker MN. SAFE@HOME - Feasibility study of a telemonitoring platform combining blood pressure and preeclampsia symptoms in pregnancy care. Eur J Obstet Gynecol Reprod Biol 2019 Sep;240:226-231. [doi: 10.1016/j.ejogrb.2019.07.012] [Medline: 31330428]
15. Alqudah A, McMullan P, Todd A, O'Doherty C, McVey A, McConnell M, et al. Service evaluation of diabetes management during pregnancy in a regional maternity hospital: potential scope for increased self-management and remote patient monitoring through mHealth solutions. BMC Health Serv Res 2019 Sep 13;19(1):662 [FREE Full text] [doi: 10.1186/s12913-019-4471-9] [Medline: 31514743]
16. Shorey S, Chee CYI, Ng ED, Lau Y, Dennis C, Chan YH. Evaluation of a Technology-Based Peer-Support Intervention Program for Preventing Postnatal Depression (Part 1): Randomized Controlled Trial. J Med Internet Res 2019 Aug 29;21(8):e12410 [FREE Full text] [doi: 10.2196/12410] [Medline: 31469084]
17. Jiao N, Zhu L, Chong YS, Chan WS, Luo N, Wang W, et al. Web-based versus home-based postnatal psychoeducational interventions for first-time mothers: A randomised controlled trial. Int J Nurs Stud 2019 Nov;99:103385. [doi: 10.1016/j.ijnurstu.2019.07.002] [Medline: 31442783]
18. Chan KL, Leung WC, Tiwari A, Or KL, Ip P. Using Smartphone-Based Psychoeducation to Reduce Postnatal Depression Among First-Time Mothers: Randomized Controlled Trial. JMIR Mhealth Uhealth 2019 May 14;7(5):e12794 [FREE Full text] [doi: 10.2196/12794] [Medline: 31094354]
19. Abuidhail J, Mrayan L, Jaradat D. Evaluating effects of prenatal web-based breastfeeding education for pregnant mothers in their third trimester of pregnancy: Prospective randomized control trial. Midwifery 2019 Feb;69:143-149. [doi: 10.1016/j.midw.2018.11.015] [Medline: 30513445]
20. O'Leary ST, Narwaney KJ, Wagner NM, Kraus CR, Omer SB, Glanz JM. Efficacy of a Web-Based Intervention to Increase Uptake of Maternal Vaccines: An RCT. Am J Prev Med 2019 Oct;57(4):e125-e133. [doi: 10.1016/j.amepre.2019.05.018] [Medline: 31471001]
21. Hildingsson I, Rubertsson C. Childbirth experiences among women with fear of birth randomized to internet-based cognitive therapy or midwife counseling. J Psychosom Obstet Gynaecol 2020 Sep;41(3):205-214. [doi: 10.1080/0167482X.2019.1634047] [Medline: 31244352]
22. Joshi A, Amadi C, Meza J, Aguire T, Wilhelm S. Evaluation of a computer-based bilingual breastfeeding educational program on breastfeeding knowledge, self-efficacy and intent to breastfeed among rural Hispanic women. Int J Med Inform 2016 Jul;91:10-19. [doi: 10.1016/j.ijmedinf.2016.04.001] [Medline: 27185505]
23. Sacks S, Abenhaim HA. How evidence-based is the information on the internet about nausea and vomiting of pregnancy? J Obstet Gynaecol Can 2013 Aug;35(8):697-703. [doi: 10.1016/S1701-2163(15)30859-8] [Medline: 24007704]
24. Kraschnewski JL, Chuang CH, Poole ES, Peyton T, Blubaugh I, Pauli J, et al. Paging "Dr. Google": does technology fill the gap created by the prenatal care visit structure? Qualitative focus group study with pregnant women. J Med Internet Res 2014 Jun 03;16(6):e147 [FREE Full text] [doi: 10.2196/jmir.3385] [Medline: 24892583]
25. Carter J, Sandall J, Shennan A, Tribe R. Mobile phone apps for clinical decision support in pregnancy: a scoping review. BMC Med Inform Decis Mak 2019 Nov 12;19(1):219 [FREE Full text] [doi: 10.1186/s12911-019-0954-1] [Medline: 31718627]
26. O'Cathain A, Croot L, Duncan E, Rousseau N, Sworn K, Turner KM, et al. Guidance on how to develop complex interventions to improve health and healthcare. BMJ Open 2019 Aug 15;9(8):e029954 [FREE Full text] [doi: 10.1136/bmjopen-2019-029954] [Medline: 31420394]
27. Fishbein M, Yzer MC. Using Theory to Design Effective Health Behavior Interventions. Commun Theory 2003 May;13(2):164-183. [doi: 10.1111/j.1468-2885.2003.tb00287.x]
28. Stevens R, Gilliard-Matthews S, Dunaev J, Todhunter-Reid A, Brawner B, Stewart J. Social Media Use and Sexual Risk Reduction Behavior Among Minority Youth: Seeking Safe Sex Information. Nurs Res 2017;66(5):368-377 [FREE Full text] [doi: 10.1097/NNR.0000000000000237] [Medline: 28858145]
29. Olson CM, Groth SW, Graham ML, Reschke JE, Strawderman MS, Fernandez ID. The effectiveness of an online intervention in preventing excessive gestational weight gain: the e-moms roc randomized controlled trial. BMC Pregnancy Childbirth 2018 May 09;18(1):148 [FREE Full text] [doi: 10.1186/s12884-018-1767-4] [Medline: 29743026]
30. WHO guideline: recommendations on digital interventions for health system strengthening. World Health Organization. URL: https://apps.who.int/iris/bitstream/handle/10665/311941/9789241550505-eng.pdf?ua=1 [accessed 2020-04-16]
31. Perestelo-Pérez L, Pérez-Ramos J, Rivero-Santana A, Carballo-González D, Serrano AP. Manual con criterios de evaluación y validación de las Herramientas de Ayuda para la Toma de Decisiones. Ministerio de Sanidad, Servicios Sociales e Igualdad. URL: http://ipdas.ohri.ca/SESCS_2013_HATD.pdf [accessed 2021-07-18]
32. Krueger RA, Casey MA. Focus Groups: A Practical Guide for Applied Research. Thousands Oak, CA: SAGE Publications Inc; 2009.
33. Paz Pascual C, Artieta Pinedo I, Grandes G, Espinosa Cifuentes M, Gaminde Inda I, Payo Gordon J. [Perceived needs of women regarding maternity. Qualitative study to redesign maternal education]. Aten Primaria 2016 Dec;48(10):657-664 [FREE Full text] [doi: 10.1016/j.aprim.2015.12.004] [Medline: 27039971]

XSL•FO
RenderX

34. Artieta-Pinedo I, Paz-Pascual C, Grandes G, Espinosa M. Framework for the establishment of a feasible, tailored and effective perinatal education programme. BMC Pregnancy Childbirth 2017 Feb 08;17(1):58 [FREE Full text] [doi: 10.1186/s12884-017-1234-7] [Medline: 28178926]
35. The LIDA Instrument: Minervation validation instrument for health care web sites. Minervation. URL: http://www.minervation.com/wp-content/uploads/2011/04/Minervation-LIDA-instrument-v1-2.pdf [accessed 2021-07-18]
36. Artieta-Pinedo I, Paz-Pascual C, Grandes G, Villanueva G, Q Group E. An evaluation of Spanish and English on-line information sources regarding pregnancy, birth and the postnatal period. Midwifery 2018 Mar;58:19-26. [doi: 10.1016/j.midw.2017.12.002] [Medline: 29277038]
37. Lupton D, Maslen S. How Women Use Digital Technologies for Health: Qualitative Interview and Focus Group Study. J Med Internet Res 2019 Jan 25;21(1):e11481 [FREE Full text] [doi: 10.2196/11481] [Medline: 30681963]
38. Vamos CA, Merrell L, Detman L, Louis J, Daley E. Exploring Women's Experiences in Accessing, Understanding, Appraising, and Applying Health Information During Pregnancy. J Midwifery Womens Health 2019 Jul 03;64(4):472-480. [doi: 10.1111/jmwh.12965] [Medline: 31050386]
39. Fuentes-Peláez N, Amorós P, Molina M, Jané M, Martínez-Bueno C. The Design of a Maternal Education Program Based on Analysis of Needs and Collaborative Work. Revista de cercetare si interventie sociala 2013;42:50-67 [FREE Full text]
40. Martínez-Bueno C, Amorós-Martí P, del Val-García JL, Burballa-Rived M, Icart-Isern MT. Evaluación de la implementación del programa de Preparación para el Nacimiento en Cataluña: características de las matronas y fidelidad al programa. Matronas Prof 2018;19(4):126-134 [FREE Full text]
41. Pålsson P, Kvist LJ, Persson EK, Kristensson Hallström I, Ekelin M. A survey of contemporary antenatal parental education in Sweden: What is offered to expectant parents and midwives' experiences. Sex Reprod Healthc 2019 Jun;20:13-19. [doi: 10.1016/j.srhc.2019.01.003] [Medline: 31084812]
42. Harst L, Lantzsch H, Scheibe M. Theories Predicting End-User Acceptance of Telemedicine Use: Systematic Review. J Med Internet Res 2019 May 21;21(5):e13117 [FREE Full text] [doi: 10.2196/13117] [Medline: 31115340]
43. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci 2009 Aug 07;4:50 [FREE Full text] [doi: 10.1186/1748-5908-4-50] [Medline: 19664226]
44. Michie S, Johnston M, Abraham C, Lawton R, Parker D, Walker A. Making psychological theory useful for implementing evidence based practice: a consensus approach. Qual Saf Health Care 2005 Feb 01;14(1):26-33 [FREE Full text] [doi: 10.1136/qshc.2004.011155] [Medline: 15692000]
45. Hussain T, Smith P, Yee LM. Mobile Phone-Based Behavioral Interventions in Pregnancy to Promote Maternal and Fetal Health in High-Income Countries: Systematic Review. JMIR Mhealth Uhealth 2020 May 28;8(5):e15111 [FREE Full text] [doi: 10.2196/15111] [Medline: 32463373]
46. Hussain-Shamsy N, Shah A, Vigod SN, Zaheer J, Seto E. Mobile Health for Perinatal Depression and Anxiety: Scoping Review. J Med Internet Res 2020 Apr 13;22(4):e17011 [FREE Full text] [doi: 10.2196/17011] [Medline: 32281939]
47. Gagnon M, Ndiaye MA, Larouche A, Chabot G, Chabot C, Buyl R, et al. Optimising patient active role with a user-centred eHealth platform (CONCERTO+) in chronic diseases management: a study protocol for a pilot cluster randomised controlled trial. BMJ Open 2019 Apr 02;9(4):e028554 [FREE Full text] [doi: 10.1136/bmjopen-2018-028554] [Medline: 30944143]
48. Dalton JA, Rodger D, Wilmore M, Humphreys S, Skuse A, Roberts CT, et al. The Health-e Babies App for antenatal education: Feasibility for socially disadvantaged women. PLoS One 2018 May 16;13(5):e0194337 [FREE Full text] [doi: 10.1371/journal.pone.0194337] [Medline: 29768407]
49. Bennett IM, Chen J, Soroui JS, White S. The contribution of health literacy to disparities in self-rated health status and preventive health behaviors in older adults. Ann Fam Med 2009 May;7(3):204-211 [FREE Full text] [doi: 10.1370/afm.940] [Medline: 19433837]
50. Nijagal M, Wissig S, Stowell C, Olson E, Amer-Wahlin I, Bonsel G, et al. Standardized outcome measures for pregnancy and childbirth, an ICHOM proposal. BMC Health Serv Res 2018 Dec 11;18(1):953 [FREE Full text] [doi: 10.1186/s12913-018-3732-3] [Medline: 30537958]
51. Camacho-Morell F, Esparcia J. Influence and use of information sources about childbearing among Spanish pregnant women. Women Birth 2020 Jul;33(4):367-376. [doi: 10.1016/j.wombi.2019.08.003] [Medline: 31447406]
52. Arcia A, Stonbraker S, Warner ERA. Continuing Education Module-Information Needs and Information-Seeking Processes of Low-Income Pregnant Women in Relation to Digital Maternity Education Resources. J Perinat Educ 2019 Jul 01;28(3):151-162 [FREE Full text] [doi: 10.1891/1058-1243.28.3.151] [Medline: 31341374]
53. WHITE PAPER: On Artificial Intelligence - A European approach to excellence and trust. European Commission. 2020. URL: https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf [accessed 2020-03-01]

Abbreviations

IPDAS: International Patient Decision Aid Standards
SDM: shared decision-making

XSL•FO
RenderX

Edited by G Eysenbach; submitted 16.03.21; peer-reviewed by M Mehrabi, K Narwaney, R Koppel; comments to author 05.05.21; revised version received 17.05.21; accepted 23.05.21; published 09.08.21.

Please cite as:
Artieta-Pinedo I, Paz-Pascual C, Bully P, Espinosa M, EmaQ Group
Design of the Maternal Website EMAeHealth That Supports Decision-Making During Pregnancy and in the Postpartum Period: Collaborative Action Research Study
JMIR Form Res 2021;5(8):e28855
URL: https://formative.jmir.org/2021/8/e28855
doi:10.2196/28855
PMID:34383670

©Isabel Artieta-Pinedo, Carmen Paz-Pascual, Paola Bully, Maite Espinosa, EmaQ Group. Originally published in JMIR Formative Research (https://formative.jmir.org), 09.08.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Artificial Intelligence–Based Chatbot for Anxiety and Depression in University Students: Pilot Randomized Controlled Trial

Maria Carolina Klos[1*], PhD candidate, MA; Milagros Escoredo[2*], MA; Angela Joerin[2*], MA; Viviana Noemí Lemos[1*], PhD; Michiel Rauws[2*], MBA; Eduardo L Bunge[3*], PhD

[1]Interdisciplinary Center for Research in Health and Behavioral Sciences (CIICSAC), Universidad Adventista del Plata (UAP)., National Scientific and Technical Research Council (CONICET)., Libertador San Martín, Entre Ríos, Argentina

[2]X2AI inc., San Francisco, CA, United States

[3]Department of Psychology, Palo Alto University, Palo Alto, CA, United States

[*]all authors contributed equally

Corresponding Author:
Maria Carolina Klos, PhD candidate, MA
Interdisciplinary Center for Research in Health and Behavioral Sciences (CIICSAC), Universidad Adventista del Plata (UAP).
National Scientific and Technical Research Council (CONICET).
25 de Mayo 99
Libertador San Martín, Entre Ríos, 3103
Argentina
Phone: 54 3435064263
Email: mcarolinaklos@gmail.com

Abstract

Background: Artificial intelligence–based chatbots are emerging as instruments of psychological intervention; however, no relevant studies have been reported in Latin America.

Objective: The objective of the present study was to evaluate the viability, acceptability, and potential impact of using Tess, a chatbot, for examining symptoms of depression and anxiety in university students.

Methods: This was a pilot randomized controlled trial. The experimental condition used Tess for 8 weeks, and the control condition was assigned to a psychoeducation book on depression. Comparisons were conducted using Mann-Whitney *U* and Wilcoxon tests for depressive symptoms, and independent and paired sample *t* tests to analyze anxiety symptoms.

Results: The initial sample consisted of 181 Argentinian college students (158, 87.2% female) aged 18 to 33. Data at week 8 were provided by 39 out of the 99 (39%) participants in the experimental condition and 34 out of the 82 (41%) in the control group. On an average, 472 (SD 249.52) messages were exchanged, with 116 (SD 73.87) of the messages sent from the users in response to Tess. A higher number of messages exchanged with Tess was associated with positive feedback ($F_{2,36}$=4.37; P=.02). No significant differences between the experimental and control groups were found from the baseline to week 8 for depressive and anxiety symptoms. However, significant intragroup differences demonstrated that the experimental group showed a significant decrease in anxiety symptoms; no such differences were observed for the control group. Further, no significant intragroup differences were found for depressive symptoms.

Conclusions: The students spent a considerable amount of time exchanging messages with Tess and positive feedback was associated with a higher number of messages exchanged. The initial results show promising evidence for the usability and acceptability of Tess in the Argentinian population. Research on chatbots is still in its initial stages and further research is needed.

(JMIR Form Res 2021;5(8):e20678) doi:10.2196/20678

KEYWORDS

artificial intelligence; chatbots; conversational agents; mental health; anxiety; depression; college students

Introduction

The most prevalent disorders in Argentina are anxiety (16.4%) and mood (12.3%) disorders. The average age for the onset of these conditions is 20 years [1]. The Pan American Health Organization (PAHO) and the Argentinian Ministry of Health have highlighted the importance of optimizing health care services for individuals who are not receiving any form of

XSL•FO
RenderX

psychological care [2]. Furthermore, the epidemiological data collected in Argentina emphasizes the need for strategies that prevent delays to treatment access [1]. Behavioral intervention technologies (BITs) are a novel and effective delivery format that can expand the mental health services offered and facilitate early access to those in need [3]. Chatbots are examples of BITs that represent an opportunity for addressing delays associated with access to treatment for depression and anxiety [4]. However, no studies on the use of chatbots for analyzing depression and anxiety have been conducted in Argentina.

Chatbots developed using artificial intelligence (AI) are emerging in the field of psychology [5]. Currently, there are two chatbots that have addressed anxiety and depressive symptoms, Woebot [6] and Tess [7]. Woebot is a chatbot based on the cognitive behavioral approach with evidence for the reduction of anxiety and depressive symptoms in students during a follow-up after 2 weeks. Fulmer et al [7] reported a reduction in depressive and anxiety symptoms in college students using Tess, a chatbot that provides support and psychoeducation through an integrative approach. Although the research completed by Fulmer et al [7] and Fitzpatrick et al [6] reported decreased depressive and anxiety symptoms in college students, these studies were performed in the United States. To the best of our knowledge, there are no studies on chatbots used for addressing mental health disorders in Spanish-speaking populations. Other examples of chatbots with empirical support are Manage Your Life Online (MYLO) that focuses on problem solving [8]; Shim, for well-being based on the cognitive behavioral approach and elements of positive psychology [9]; Tess for pediatric obesity and prediabetes treatment [10]; and Wysa, a chatbot that uses cognitive behavioral therapy, behavioral reinforcement, and mindfulness techniques to support patients with depression [11]. Research studies on chatbots for mental health have several limitations such as small sample sizes and short-term follow-ups [6,7]. Additionally, current chatbots for mental health promotion present several problems, such as the lack of recognition of the emotional tone of users, crisis identification and management, as well as the need for strategies to reduce the frustration arising from feelings of incomprehension by users when the chatbot does not respond accurately.

The present study aims to assess the viability and acceptability of psychological interventions delivered through Tess to college students in Argentina. The objectives of this study were as follows: (1) identify participant flow from recruitment to follow-up; (2) understand aspects related to the usage patterns of Tess, such as the number of messages sent and exchanged; (3) examine the relationship between the feedback provided by the participants and the number of messages exchanged with Tess; and (4) compare the outcomes on depression and anxiety between and within groups among the college students who completed the study. Although the focus of this research was not the effectiveness of the chatbot, comparisons were made between the experimental and control groups to obtain preliminary data for future randomized controlled trials given the importance of obtaining preliminary information about the viability and acceptability of Tess as a means of psychological intervention for college students in Argentina.

Methods

Trial Design

This was a pilot randomized controlled parallel-group trial. The experimental group had access to Tess for 8 weeks and the control group to a psychoeducation electronic book.

Participants

The participants were college students in Entre Ríos, Argentina. The inclusion criteria were as follows: being a resident of Argentina, 18 years or older, and a college student, as well as providing informed consent. Recruitment was conducted through presentations in different university courses. Participants who provided consent were assigned to the experimental or control group by simple randomization conducted through a Python algorithm.

Intervention

Experimental Group

The experimental group utilized Tess, an AI-based chatbot that delivers brief text conversations as comprehensive support for mental health. Tess sends reminders, psychoeducational content, and emotional support responses based on what the users express. Tess combines words and emojis in the messages for providing a more user-friendly experience. Tess responds with prescribed statements to replicate empathetic answers that are appropriate for the emotion or concern expressed by the participants. For example, a participant expressing anxious feelings would be offered a relaxation strategy. The conversations offered by Tess were based on the cognitive behavioral model [12], emotion-focused therapy [13], solution-focused brief therapy [14], and motivational interviewing [15]. Such conversations were developed by mental health experts. After each conversation, Tess asked, “Was our conversation helpful?” If a user responded positively (eg, “yes, thank you”) to an intervention based on cognitive behavioral therapy (CBT) and negatively (eg, “no, not really”) to emotion-focused therapy, Tess would then offer more CBT-based interventions. For users who answered in a negative or neutral manner, Tess would offer alternative interventions.

In the present study, customized conversations for university students in Argentina were elaborated, revised, and tested within the framework of a previous study developed in the United States [7]. During the 8-week intervention for this test, Tess initiated contact asking about the emotions and moods of the participants once a day during the initial weeks and every other day in the following weeks. All the conversations with Tess occurred through Facebook messenger.

Control Group

An electronic psychoeducation book focusing on affective symptoms was provided to the participants in the control group [16]. The provided evidence-based information and resources helped students identify and seek treatment for depressive symptoms.

XSL•FO
RenderX

Engagement and Feedback

Engagement was measured using the number of messages exchanged with Tess. In addition, the dropout rates in the experimental and control groups were analyzed. The perceived feedback of the participants was collected after each conversation with Tess through the following question: "Was our conversation helpful?" The answers from the users were coded as positive, negative, or ambivalent and assigned values of 1, 2, and 3, respectively. For instance, if a user responded saying "yes, thank you," then that response was coded as positive.

Measures

The Patient Health Questionnaire-9 (PHQ-9) [17] is a self-reporting questionnaire comprising 9 items that evaluate the frequency and severity of depressive symptoms during the last 2 weeks. Each of the 9 items is based on the *Diagnostic and Statistical Manual of Mental Disorders-IV* (*DSM-IV*) criteria, which are scored from 0 (not at all) to 3 (nearly every day). The PHQ-9 is one of the most used measures to assess depressive symptoms and has been validated in Argentina with adequate psychometric properties (Cronbach α=.87) [18]. The first 2 items are considered screening criteria (PHQ-2); if these are scored with 0 or 1, then an absence of symptoms is assumed. Scores ranging from 5 to 9 are interpreted as mild, from 10 to 14 as moderate, from 15 to 20 as moderately severe, and over 20 as severe.

The Generalized Anxiety Disorder Scale (GAD-7) [19] is a 7-item self-reporting scale that evaluates the frequency and severity of anxious thoughts and behaviors during the last 2 weeks. Items are based on the diagnostic criteria of the *DSM-IV* and scored from 0 (not at all) to 3 (nearly every day). Rodríguez de Behrends and Brenlla [20] reported an adequate reliability level (Cronbach α=.74) for the Argentinian population.

Ethical Aspects

The study was approved by the Research Ethics Committee of the Faculty of Health Sciences (FCS) of the Universidad Adventista del Plata (UAP), National Registry of Health Research (RENIS, reference number: CE000237), and Ministerial Resolution of the Ministry of Health of the Province of Entre Ríos (reference number: 3999). This resolution is recorded in ACT 1-2019 of the registration of this committee. Participants expressed their consent in a form according to the personal data protection law (Argentine National Law 25.326) through checkbox selection (electronic signature) on a closed form.

Data were collected through Tess. All personally identifiable information was eliminated in the transcriptions downloaded from Tess. The downloaded data were processed and stored using secure servers and were compliant with the Health Insurance Portability and Accountability Act. Upon completion of the study, the control group obtained access to Tess for 8 weeks and both groups were granted free access to Tess for a year. If a participant expressed suicidal ideations, Tess was programmed to provide the National Line of Suicide Prevention numbers, the crisis text line, and 911, and encourage seeking professional help.

Data Analysis

The data collected was entered and analyzed using the Statistical Package for the Social Sciences (Version 20.0; IBM Corporation) [21]. The number of messages exchanged was considered to assess the feasibility and acceptability of Tess. Additionally, the participants' qualitative feedback was analyzed by two researchers (CK and ME) and coded into three categories: positive, negative, and ambivalent. A data analysis protocol was carried out. The treatment of missing data through multiple imputation or plausibility analysis techniques was not possible owing to the high percentage of participants who dropped out of the intervention [22].

A one-factor analysis of variance (ANOVA) was applied to determine if feedback (positive, negative, or ambivalent) impacted the number of interactions that users had with Tess. To examine the baseline characteristics between samples, a *t* test was performed for independent samples to compare anxiety levels. The Mann-Whitney *U* statistic (respecting the ordinal nature of the variables; namely, if the first 2 items were scored 0 or 1, the system did not ask the subsequent items) was used to compare depressive symptoms.

To evaluate the effects between conditions, a *t* test was performed for the independent samples to assess the anxiety symptoms and the Mann-Whitney *U* statistic was used to compare the mean ranges of depression. To assess the longitudinal effects from the baseline to week 8 within conditions, a *t* test was performed for related samples assess the anxiety symptoms and the Wilcoxon test was performed to compare the mean ranges of depression. To complement the significance test, the effect sizes in the intragroup and intergroup tests were calculated. For the *t* tests, the effect size was calculated using Cohen *d*; measures between 0.2 and 0.3 were labeled "small effect," around 0.5 as "moderate effect," and above 0.8 as "large effect" [23]. For the Mann-Whitney *U* and Wilcoxon tests, the *r* formula was calculated based on the *z* scores. The measures between 0.1 and 0.3 were labeled "small effect," between 0.3 and 0.5 as "moderate effect," and above 0.5 as large effect" [23,24].

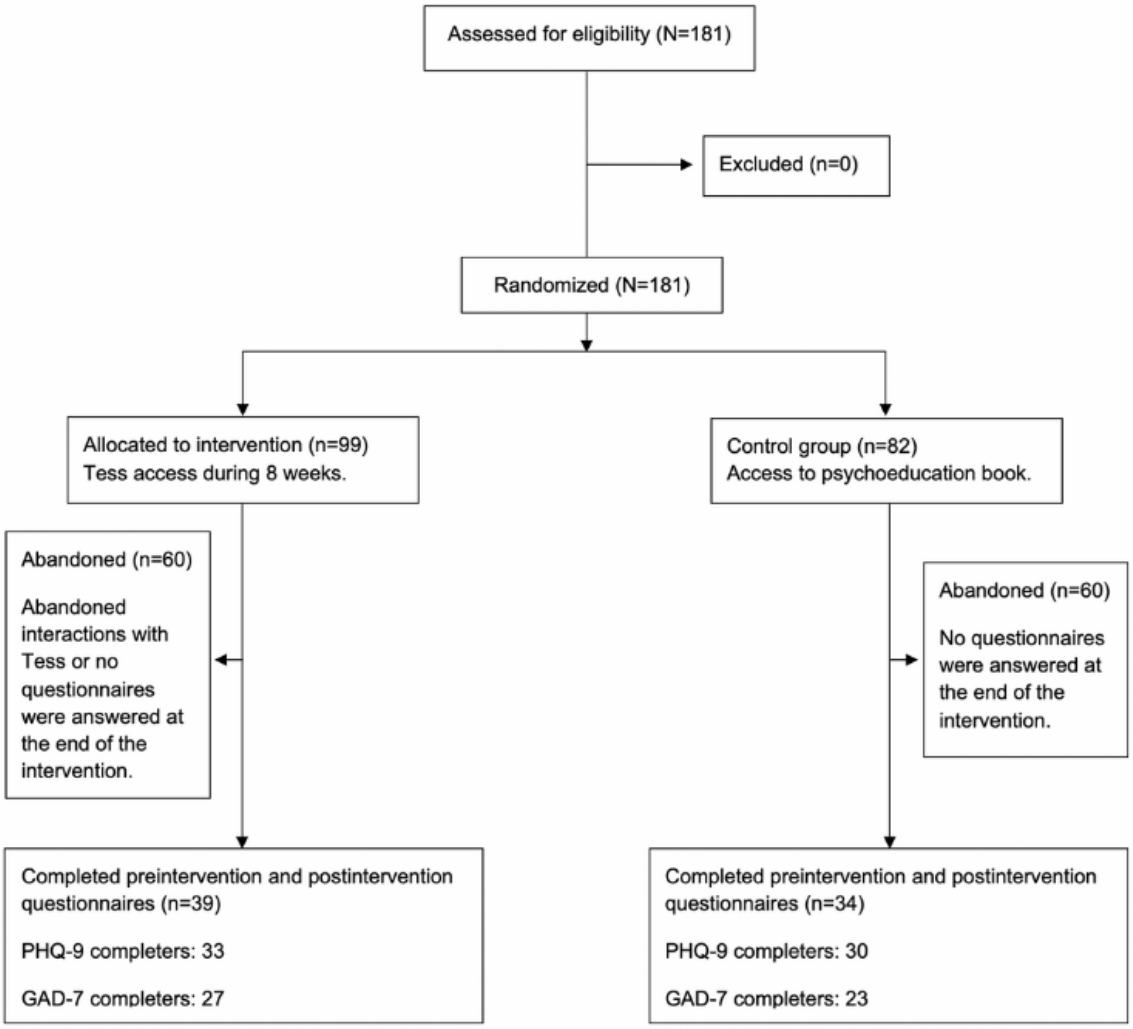
Results

Initial Observations

The initial sample consisted of 181 college students in Argentina, aged 18 to 33, with 158 (87.2%) identifying as female. Among the 181 students, 99 (55%) were randomized to the experimental condition and 82 (45%) to the control condition. Data at week 8 were provided by 39 out of the 99 (39%) participants in the experimental condition and 34 out of the 82 (41%) in the control group. Regarding data on the depressive symptoms, 33 (33%) participants in the experimental condition and 30 (37%) in the control condition provided data at week 8. Regarding data on anxiety symptoms, 27 (27%) participants in the experimental condition and 23 (28%) in the control condition provided data at week 8 (see Figure 1).

XSL•FO
RenderX

Figure 1. CONSORT flow diagram. GAD-7: Generalized Anxiety Disorder Scale-7, PHQ-9: Patient Health Questionnaire-9.



Messages Exchanged

Regarding the participants' engagement with Tess (39/99), after 8 weeks, there was an average of 472 exchanged messages (SD 249.52), where the minimum interaction level involved 162 messages and the maximum involved 1290. More specifically, an average of 116 (SD 73.87) of the exchanged messages were sent from the user to Tess.

Feedback

Feedback from most participants (25/39) at week 8 was coded as positive (ie, *Yes, you really understand me, Tess. Thanks for talking to me. My anxiety has decreased and I can confidently go outside again*.). A minor number of participants (7/39) provided ambivalent (ie, *Not much, but it's ok, I am capable*) or negative (ie, *Sometimes I ask you something and you don't specifically respond to what I asked*.). A one-factor ANOVA was applied to determine if feedback (positive, negative, or ambivalent) impacted the number of interactions that users had with Tess. Results showed that feedback from users was associated with the number of messages exchanged with Tess ($F_{2,36}$=4.37; P=.02). Post hoc contrasts resulting from the Scheffé test showed statistically significant differences between those participants providing positive feedback and those providing negative feedback (P=.04); nevertheless, a higher number of messages exchanged with Tess was associated with positive feedback. No differences were observed between the participants providing ambivalent and positive feedback or ambivalent and negative feedback (See Table 1).

Table 1. Number of interactions per user feedback.

Variable	Positive, mean (SD)	Ambivalent, mean (SD)	Negative, mean (SD)	Statistical values	
				F	*P* value
Interactions	551.24 (52.54)	374.43 (73.11)	287.43 (23.35)	4.37	.02

XSL•FO
RenderX

Potential Impact of Tess on Indicators of Depression and Anxiety

Baseline Characteristics

There were no statistically significant differences at the baseline in the anxiety (t_{48}=.16; *P*=.9) and depression scores (*U*=451.50; *P*=.5) between the experimental and the control groups (See Table 2).

Table 2. Comparison of the average values and ranges for the anxiety and depression variables at the baseline and week 8 between the experimental and control groups.

Variable	Experimental group (n=39)	Control group (n=34)	Statistics	*P* value
Anxiety at baseline, mean (SD)	15.59 (5.30)[a]	15.35 (5.75)[b]	*t*=0.16	.90
Anxiety at week 8, mean (SD)	13.04 (7.12)[a]	16.26 (5.79)[b]	*t*=1.74	.09
Depression at baseline, middle range	33.32[c]	30.55[d]	*U*=451.50	.50
Depression at week 8, middle range	30.58[c]	33.57[d]	*U*=448.00	.48

[a]n=27.
[b]n=23.
[c]n=33.
[d]n=30.

Between-Group Differences

No statistically significant differences were observed between the experimental and the control groups in the average scores for anxiety (t_{48}= 1.74; *P*=.09) or in the average ranges for depression (*U*=448.00; *P*=.48) at week 8 (See Table 2). Regarding the effect sizes, the mean scores for anxiety in the experimental group were lower than for the control group after 8 weeks and the effect size of the intervention was moderate (*d*=.5; 95% CI [-6.96 to.51]). For depressive symptoms, the experimental group reported a lower mean score than the control group and the effect size of the intervention was nonexistent (*r*=.09).

Within-Group Differences

Within the experimental condition, a statistically significant decrease in the symptoms was observed from the baseline to week 8 for the anxiety scores (t_{26}=2.15; *P*=.04); the control condition did not demonstrate any significant changes (t_{22}=1.00; *P*=.33). Regarding depressive symptoms, no significant differences were found either in the experimental condition (*Z*=1.76; *P*=.08) or in the control condition (*Z*=.00; *P*>.99) (See Table 3).

Table 3. Comparison of the average values and ranges within groups for anxiety and depression variables from the baseline to week 8.

Variable and condition	Baseline	Week 8	Statistics	*P* value
Anxiety, mean (SD)				
Experimental	15.59 (5.30)	13.04 (7.12)	*t*=2.15	.04
Control	15.35 (5.75)	16.26 (5.79)	*t*=1.00	.33
Depression, middle range				
Experimental	8.83	7.14	*Z*=1.76	.08
Control	6.50	6.50	*Z*=0.00	.99

Discussion

Important Findings

The use of chatbots (ie, conversational agents) to address mental health conditions may contribute to the treatment of large populations and attend to the needs of those who do not have access to treatment. To the best of our knowledge, there are no studies on the use of chatbots for mental health in Latin America. This trial was intended to evaluate an AI-based chatbot (Tess) in a sample comprising Argentinian college students. The specific objectives were as follows: (1) understand the participant flow from recruitment to follow-up; (2) report aspects related to the usage patterns of Tess, such as the number of messages sent and exchanged; (3) examine participant feedback; and (4) compare the preliminary measures of depression and anxiety.

Regarding the usage patterns of Tess, there are three findings that support a satisfactory level of engagement. First, a considerable number of participants in the experimental (39/99, 39%) and control (34/82, 41%) conditions remained in the study throughout the 8-week study period. The completion rates found in the current study are better than that observed in most unpaid and unsupported Internet-based interventions for depression and anxiety, where 90% of the users withdraw after the first two sessions [25]. Furthermore, in studies using mobile apps, the follow-up completion rates were comparable (53%); the mean percentage of complete “adherers” was 36% for depression and 41% for anxiety [26]. When compared to a chatbot study

XSL•FO
RenderX

for US college students, a lower attrition rate was reported (31% and 9% in the control and experimental conditions, respectively); however, this study compensated participants and the follow-up was at 2 weeks, making it difficult to compare the outcomes [6].

Second, participants in the experimental condition had exchanged a considerable number of messages with Tess (M 472; SD 249.52), and the mean number of messages sent from the user to Tess was 116 (SD 73.87). A previous study on the usage patterns of the depression modules of Tess showed a much lower average number (17.57) of messages sent to Tess by adult users [27]. It is possible that college, younger, and Latinx students are more willing to engage in conversations with chatbots than older populations in the United States. Two previous studies involving college students in the United States did not report the number of messages sent by the user to Tess [6,7]. Regarding the messages exchanged, Fulmer et al [7] reported a comparable number of total messages exchanged during a period of 4 weeks (M 286; SD 104.6), whereas the total number of messages exchanged in the current study was during a period of 8 weeks.

Third, feedback provided by those in the experimental condition was mostly positive (eg, *Yes, you really understand me, Tess. Thanks for talking to me.)*. Among the participants offering negative feedback, there was a predominant dissatisfaction regarding the accuracy of some interventions (eg, *Sometimes I ask you something and you don't specifically respond to what I asked.*). Feedback is a key component for AI-based chatbots as it allows systems to tailor the dialogues to the user. Interestingly, the positive and negative feedbacks were associated with the number of messages exchanged. Users who reported higher satisfaction had the highest number of exchanged messages; it is possible that providing positive feedback could lead to better customization the intervention messages. This finding is relevant as it supports the need to collect user feedback for achieving optimal levels of customization and increasing engagement that could lead to higher intervention doses.

Regarding the impact of Tess on anxiety and depressive symptoms, no statistically significant differences were found between groups. Interestingly, when comparing within-group scores, the experimental group showed a significant decrease in anxiety symptoms after 8 weeks of intervention and a near-significant trend (P=0.07) for depressive symptoms. Analyzing the effect sizes showed that Tess had a moderate effect on anxiety and no effect on depression in the experimental group. These outcomes were unexpected given that previous studies using Tess [7] and another conversational agent called Woebot [6] reported significant reductions in anxiety and depressive symptoms; both studies used a similar control group (a psychoeducation book). Moreover, in the current study, depression was measured using PHQ-9 as a categorical and ordinal variable, whereas Fulmmer et al [7] and Fitzpatrick et al [6] used it as a continuous measure.

The lack of between-group differences could be explained by several factors. First, the current study was underpowered. Second, although the findings of the current study were not statistically significant, the direction of the change observed for anxiety and depression was as expected; therefore, it is possible that low-intensity interventions delivered via chatbots may require a higher dose to yield a between-group effect when delivered to Argentinian students. Third, Tess provides many conversations based on different theoretical approaches, and this may have resulted in less therapeutic power. However, Fulmer et al [7] observed significant effect using similar conversations. Fourth, it is possible that during adaptation of the dialogues from English to Spanish, the quality of the intervention may have been reduced.

Limitations and Future Directions

This pilot study has several limitations. First, the current analysis was conducted with intervention completers; therefore, future studies with larger samples (including completer and intent-to-treat analyses) are needed. Second, only college students from a specific region in Argentina were included in this study, and the socioeconomic aspects of the sample were not assessed; thus, the inclusion of a more diversified sample is suggested. Third, there was a high dropout rate throughout the 8-week period. This is congruent with the findings reported by most studies that use technology-based intervention (see "The Law of Attrition") [28]. A high dropout rate may be due to the limited capacity of most digital interventions to capture the attention and motivation of users. Additionally, high dropout rates in studies with digital interventions were linked to the fact that as access is easy, a lower level of commitment is required from the user to enroll in the study compared to traditional face-to-face interventions. Fourth, as most participants who remained until completion of the study were female, male participation was scarce. Fifth, the control group had access to a psychoeducation book, and there was no information on whether they read it. As chatbot research is in its initial stages, further studies could benefit from offering waitlists rather than self-help books. Although offering a waitlist could present an ethical dilemma, this would mitigate the potential effects of not having an intervention if short-term studies are conducted.

Future chatbot studies may benefit from designing chatbots with more conversations based on a specific therapeutic approach rather than using a few conversations from several approaches. Additionally, analyzing the impact of chatbots as adjuncts to face-to-face psychotherapy and comparing these interventions with face-to-face psychotherapy alone would yield important insights regarding the advancement of research on chatbots for mental health. Finally, simple randomization was used in this study; future studies may consider using unequal randomization (2:1) so that more participants enter the experimental group or a stratified randomization procedure so that participants with similar characteristics can be assigned equally to the experimental and control groups**.**

Conclusions

Students spent a considerable amount of time exchanging messages with Tess and positive feedback was associated with higher numbers of messages exchanges. The initial results showed preliminary evidence regarding the effectiveness of Tess in addressing anxiety symptoms, but there was no significant effect on depressive symptoms in Argentinian college

XSL•FO
RenderX

students. Given the high prevalence of anxiety and depression in Argentinian college students [1] and the need to expand mental health care access, developing affordable strategies such as chatbots may become effective tools to address these needs. AI-based chatbots have the ability to reach higher levels of customization and may thus be of service to educational and mental health care centers aiming to deliver interventions to targeted users that are accessible at any time without geographical restrictions. Additionally, chatbots may be used as standalone resources for those who have no access to treatment or as a complement to traditional treatments. Although the initial evidence on the efficacy of chatbots is promising, research on chatbots is still in its initial stages and presents several limitations. Thus, more robust evidence is needed.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT-eHEALTH checklist (v 1.6.1).
[PDF File (Adobe PDF File), 1170 KB - formative_v5i8e20678_app1.pdf]

References

1. Cía AH, Stagnaro JC, Aguilar Gaxiola S, Vommaro H, Loera G, Medina-Mora ME, et al. Lifetime prevalence and age-of-onset of mental disorders in adults from the Argentinean study of mental health epidemiology. Soc Psychiatry Psychiatr Epidemiol 2018 Feb;53(4):341-350. [doi: 10.1007/s00127-018-1492-3]
2. Framework for the implementation of the regional strategy on mental health 2011. Pan American Health Organization. Washington, D.C: PAHO; 2011. URL: https://iris.paho.org/bitstream/handle/10665.2/3535/PAHO%20framework%20implem%20reg%20strat%20on%20mh%202011.pdf?sequence=1&isAllowed=y [accessed 2019-01-20]
3. Mohr DC, Burns MN, Schueller SM, Clarke G, Klinkman M. Behavioral intervention technologies: evidence review and recommendations for future research in mental health. Gen Hosp Psychiatry 2013 Jul;35(4):332-338 [FREE Full text] [doi: 10.1016/j.genhosppsych.2013.03.008] [Medline: 23664503]
4. Bendig E, Erb B, Schulze-Thuesing L, Baumeister H. The next generation: chatbots in clinical psychology and psychotherapy to foster mental health – a scoping review. Verhaltenstherapie 2019 Aug:1-13. [doi: 10.1159/000501812]
5. Gaffney H, Mansell W, Tai S. Conversational agents in the treatment of mental health problems: mixed-method systematic review. JMIR Ment Health 2019 Oct;6(10):e14166 [FREE Full text] [doi: 10.2196/14166] [Medline: 31628789]
6. Fitzpatrick KK, Darcy A, Vierhile M. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. JMIR Ment Health 2017 Jun;4(2):e19 [FREE Full text] [doi: 10.2196/mental.7785] [Medline: 28588005]
7. Fulmer R, Joerin A, Gentile B, Lakerink L, Rauws M. Using psychological artificial intelligence (Tess) to relieve symptoms of depression and anxiety: randomized controlled trial. JMIR Ment Health 2018 Dec;5(4):e64 [FREE Full text] [doi: 10.2196/mental.9782] [Medline: 30545815]
8. Gaffney H, Mansell W, Edwards R, Wright J. Manage Your Life Online (MYLO): a pilot trial of a conversational computer-based intervention for problem solving in a student sample. Behav Cogn Psychother 2014 Nov;42(6):731-746. [doi: 10.1017/S135246581300060X] [Medline: 23899405]
9. Ly KH, Ly A, Andersson G. A fully automated conversational agent for promoting mental well-being: A pilot RCT using mixed methods. Internet Interv 2017 Dec;10:39-46 [FREE Full text] [doi: 10.1016/j.invent.2017.10.002] [Medline: 30135751]
10. Stephens TN, Joerin A, Rauws M, Werk LN. Feasibility of pediatric obesity and prediabetes treatment support through Tess, the AI behavioral coaching chatbot. Transl Behav Med 2019 May;9(3):440-447. [doi: 10.1093/tbm/ibz043] [Medline: 31094445]
11. Inkster B, Sarda S, Subramanian V. An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: real-world data evaluation mixed-methods study. JMIR Mhealth Uhealth 2018 Nov;6(11):e12106 [FREE Full text] [doi: 10.2196/12106] [Medline: 30470676]
12. Beck J. Cognitive Behavior Therapy: Basics and Beyond, Second Edition. New York City: The Guilford Press; 2011.
13. Greenberg L. Emotion-Focused Therapy: Coaching Clients to Work Through Their Feelings. Washington, DC: American Psychological Association; 2002.
14. Pichot T, Dolan Y. Solution-Focused Brief Therapy: Its Effective Use in Agency Settings. Haworth Marriage and the family. New York: Routledge; 2003.
15. Rollnick S, Miller WR. What is motivational interviewing? Behav Cognit Psychother 2009 Jun;23(04):325. [doi: 10.1017/S135246580001643X]
16. Depression in College Students. National Institute of Mental Health. URL: https://www.nimh.nih.gov/health/publications/espanol/la-depresion-en-la-adolescencia/19-mh-8089s_157885.pdf [accessed 2018-11-10]

XSL•FO
RenderX

17. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9. J Gen Intern Med 2001 Sep;16(9):606-613. [doi: 10.1046/j.1525-1497.2001.016009606.x]
18. Urtasun M, Daray FM, Teti GL, Coppolillo F, Herlax G, Saba G, et al. Validation and calibration of the patient health questionnaire (PHQ-9) in Argentina. BMC Psychiatry 2019 Sep 18;19(1):291-310 [FREE Full text] [doi: 10.1186/s12888-019-2262-9] [Medline: 31533674]
19. Newman MG, Zuellig AR, Kachin KE, Constantino MJ, Przeworski A, Erickson T, et al. Preliminary reliability and validity of the generalized anxiety disorder questionnaire-IV: A revised self-report diagnostic measure of generalized anxiety disorder. Behavior Therapy 2002 Mar;33(2):215-233. [doi: 10.1016/S0005-7894(02)80026-0]
20. Rodríguez de Behrends M. Estudio preliminar de las propiedades psicométricas del Cuestionario de Trastorno de Ansiedad Generalizada-IV (GAD-Q-IV-Generalized Anxiety Disorder Questionnaire-IV) en población de Buenos Aires. VII Congreso Internacional de Investigación y Práctica Profesional en Psicología XXII Jornadas de Investigación XI Encuentro de Investigadores en Psicología del MERCOSUR. Facultad de Psicología-Universidad de Buenos Aires. 2015 Nov. URL: http://www.psi.uba.ar/investigaciones/eventos_cientificos/vii_congreso_memorias/6_tecn_proces_eval.pdf [accessed 2019-01-20]
21. IBM SPSS Statistics for Windows, Version 22. Armonk, NY: IBM Corp; 2013.
22. Fernández-García MP, Vallejo-Seco G, Livácic-Rojas P, Tuero-Herrero E. The (Ir)responsibility of (under)estimating missing data. Front Psychol 2018 Apr;9:556. [doi: 10.3389/fpsyg.2018.00556]
23. Cohen J. Statistical Power Analysis for the Behavioral Sciences (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates; 1988.
24. Tomczak M, Tomczak E. The need to report effect size estimates revisited. An overview of some recommended measures of effect size. Trends in Sport Sciences 2014;1(21):19-25 [FREE Full text]
25. Titov N, Dear BF, Johnston L, Lorian C, Zou J, Wootton B, et al. Improving adherence and clinical outcomes in self-guided internet treatment for anxiety and depression: randomised controlled trial. PLoS One 2013 Jul;8(7):e62873 [FREE Full text] [doi: 10.1371/journal.pone.0062873] [Medline: 23843932]
26. Linardon J, Fuller-Tyszkiewicz M. Attrition and adherence in smartphone-delivered interventions for mental health problems: a systematic and meta-analytic review. J Consult Clin Psychol 2020 Jan;88(1):1-13. [doi: 10.1037/ccp0000459] [Medline: 31697093]
27. Dosovitsky G, Pineda B, Jacobson NC, Chang C, Escoredo M, Bunge EL. Artificial intelligence chatbot for depression: descriptive study of usage. JMIR Form Res 2020 Nov;4(11):e17065 [FREE Full text] [doi: 10.2196/17065] [Medline: 33185563]
28. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]

Abbreviations

AI: artificial intelligence
BITs: behavioral intervention technologies
CBT: cognitive behavioral therapy
***DSM-IV*:** *Diagnostic and Statistical Manual of Mental Disorders-IV*
GAD-7: Generalized Anxiety Disorder Scale-7
MYLO: Manage Your Life Online
PAHO: Pan American Health Organization
PHQ-9: Patient Health Questionnaire-9

Edited by G Eysenbach; submitted 07.07.20; peer-reviewed by E Bendig, M Rahmadiana; comments to author 17.08.20; revised version received 11.11.20; accepted 29.05.21; published 12.08.21.

Please cite as:
Klos MC, Escoredo M, Joerin A, Lemos VN, Rauws M, Bunge EL
Artificial Intelligence–Based Chatbot for Anxiety and Depression in University Students: Pilot Randomized Controlled Trial
JMIR Form Res 2021;5(8):e20678
URL: https://formative.jmir.org/2021/8/e20678
doi:10.2196/20678
PMID:34092548

©Maria Carolina Klos, Milagros Escoredo, Angela Joerin, Viviana Noemí Lemos, Michiel Rauws, Eduardo L Bunge. Originally published in JMIR Formative Research (https://formative.jmir.org), 12.08.2021. This is an open-access article distributed under

XSL•FO
RenderX

the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Experience of Key Stakeholders During the Implementation and Use of Trauma Therapy via Digital Health for Military, Veteran, and Public Safety Personnel: Qualitative Thematic Analysis

Lorraine Smith-MacDonald[1,2], PhD; Chelsea Jones[1,3,4], PhD; Phillip Sevigny[1,5], PhD; Allison White[1,2], MSW; Alexa Laidlaw[1,2], BSc; Melissa Voth[1,2], BEd; Cynthia Mikolas[1,6], BSW; Alexandra Heber[7,8], MD, FRCPC, CCPE; Andrew J Greenshaw[1,9], PhD; Suzette Brémault-Phillips[1,2], PhD

[1]Heroes in Mind, Advocacy and Research Consortium (HiMARC), Faculty of Rehabilitation Medicine, University of Alberta, Edmonton, AB, Canada
[2]Department of Occupational Therapy, Faculty of Rehabilitation Medicine, University of Alberta, Edmonton, AB, Canada
[3]1 Field Ambulance Physical Rehabilitation Department, Canadian Forces Health Services, Department of National Defense, Edmonton, AB, Canada
[4]Leiden University Medical Centre, Leiden University, Leiden, Netherlands
[5]Department of Educational Psychology, Faculty of Education, University of Alberta, Edmonton, AB, Canada
[6]Military Family Resource Centre, Edmonton, AB, Canada
[7]Health Professionals Division, Veterans Affairs Canada, Ottawa, ON, Canada
[8]Department of Psychiatry, University of Ottawa, Ottawa, ON, Canada
[9]Department of Psychiatry, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, AB, Canada

Corresponding Author:
Chelsea Jones, PhD
Heroes in Mind, Advocacy and Research Consortium (HiMARC)
Faculty of Rehabilitation Medicine
University of Alberta
1-94 Corbett Hall
8205 - 114 Street
Edmonton, AB, T6G 2G4
Canada
Phone: 1 7804920404
Email: cweiman@ualberta.ca

Abstract

Background: Exposure to occupational stressors and potentially psychologically traumatic events experienced by public safety personnel (eg, paramedics, police, fire, and correctional officers), military members, and veterans can lead to the development of posttraumatic stress injuries and other mental health disorders. Providing emergency services during COVID-19 has intensified the challenges. Owing to COVID-19 restrictions, mental health service providers offering support to these populations have had to rapidly pivot to use digital versus in-person methods of service delivery.

Objective: This paper aims to explore the experience of mental health service providers regarding digital health service delivery, including the current state of digital mental health service delivery, barriers to and facilitators of the use of digital health for mental health service delivery experienced during the pandemic, and recommendations for implementing and integrating digital health into regular mental health service delivery.

Methods: This embedded mixed-methods study included questionnaires and focus groups with key stakeholders (N=31) with knowledge and experience in providing mental health services. Data analysis included descriptive, quantitative, and qualitative thematic analyses.

Results: The following three themes emerged: being forced into change, daring to deliver mental health services using digital health, and future possibilities offered by digital health. In each theme, participants' responses reflected their perceptions of service providers, organizations, and clients. The findings offer considerations regarding for whom and at what point in treatment digital health delivery is appropriate; recommendations for training, support, resources, and guidelines for digitally delivering trauma therapy; and a better understanding of factors influencing mental health service providers' perceptions and acceptance of digital health for mental health service delivery.

XSL•FO
RenderX

Conclusions: The results indicate the implementation of digital health for mental health service delivery to military members, public safety personnel, and veterans. As the COVID-19 pandemic continues, remote service delivery methods for trauma therapy are urgently needed to support the well-being of those who have served and continue to serve.

(JMIR Form Res 2021;5(8):e26369) doi:10.2196/26369

KEYWORDS

trauma; mental health; telemedicine; therapy; rehabilitation; digital health; psychotherapy; military; veteran; first responder; public safety personnel; teletherapy; psychotherapy; telepsychiatry; mobile phone

Introduction

Background

The COVID-19 pandemic necessitated a sudden shift in the provision of mental health services from in-person to digital health delivery to comply with physical distancing restrictions [1]. This shift required adaptations at the organizational, service provider, and client levels. It was also experienced differently by the general population than by military members (MMs), veterans, and public safety personnel (PSP; eg, paramedics, police, fire, and correctional officers) dealing with symptoms and other effects of operational stress injuries (OSIs). As a result, continuing access to mental health supports for trauma-affected populations, including MMs, veterans, and PSP, was pivoted to digitally delivered trauma therapy.

In addition, MMs and PSP are critical to the COVID-19 pandemic response and management as they respond to medical emergencies, enforce public health orders, and support infection control and prevention initiatives [2]. Although accustomed to performing in complex, ambiguous, high stake, and rapidly changing environments, the context of the COVID-19 pandemic has amplified these factors and led to operational environments that are highly stressful and uncertain [3]. Therefore, supporting MMs and PSP in a timely and responsive manner is critical, particularly as increased occupational and operational stressors are likely to continue during and beyond the COVID-19 pandemic and can lead to OSIs.

Digital Health for MMs, Veterans, and PSP

Creative approaches to mental health service delivery that maximize access to services while minimizing physical contact and proximity are required as a result of the pandemic. Digital health is an umbrella term that encompasses a variety of health technologies such as teletherapy, telemedicine, eHealth, mobile health, and health information technology delivered through various modes such as videoconferencing or telephone and has offered a means by which to address immediate mental health needs. For MMs, veterans, and PSP living in rural or remote areas, digital health can facilitate access to mental health treatment by removing barriers to access associated with transportation or service availability [4]. The ability to access mental health treatment from the comfort and safety of a home environment has been identified as a facilitator for individuals experiencing social anxiety or with concerns regarding privacy and stigma [5-7].

Before the COVID-19 pandemic, the adoption and implementation of telehealth as a means of mental health service provision was slow and difficult to sustain [8], and the response of mental health clinicians to the shift to or adoption of digital health to provide care for MMs, veterans, and PSP was mixed [9]. Barriers to widespread and sustained digital health use were associated with concerns regarding the equivalency of telehealth-delivered services to face-to-face treatment, establishment of a therapeutic alliance, client and service provider acceptance of digital health, technical connectivity challenges, patient privacy and confidentiality, software and equipment availability, usability and reliability, associated costs, and regulatory concerns [1,9-12]. Recommendations associated with digital delivery, in general, emphasize clinician training and supervision around the use of technologies and mental health interventions [13]; development of guidelines and policies for software, web-based platforms; and patient consent [14].

Regulatory colleges for health care professionals have provided some practice standards and guidelines for the use of digital health many of which have been developed specifically in response to COVID-19. Regulatory colleges have offered initial standards and guidance regarding privacy and confidentiality, consent, risk management and safety, technical issues and security of data, ongoing training regarding both professional and technical competency, and possible insurance or jurisdiction issues [15-18]. However, given the rapidly changing landscape, little research has been conducted exploring how these standards have been taken up by frontline service providers and factors that facilitate their implementation. In particular, issues relating to privacy are especially salient given not only the widespread stigma associated with mental health concerns in the military or PSP organizations but also the sensitivity of disclosures that may occur when working with MMs or PSP who have high security clearance.

Despite these barriers, encouraging evidence supporting the use of digital health was identified through a systematic scoping review conducted by members of this research team in advance of this study regarding digital health delivery of trauma therapies for MMs, veterans, and PSP [19]. This review illustrated that the delivery of cognitive processing therapy (CPT), prolonged exposure therapy, and behavioral activation therapy using videoconferencing demonstrated comparable effectiveness to in-person treatment in reducing posttraumatic stress disorder symptoms in MMs and veterans [19]. This scoping review identified significant knowledge gaps. Specifically, the majority of available evidence comes from studies conducted in the United States with active-duty MMs and veterans where service providers predominately used CPT. Far less research has been conducted on other treatment modalities or on the treatment of PSP; practically no research on digital health for MMs, veterans, and PSP has been undertaken in the Canadian context. It is

XSL•FO
RenderX

imperative that these knowledge gaps be addressed to promote continued, effective, and ethical use of digital health technologies during and beyond the pandemic.

Purpose

The purpose of this study is to explore the perspectives of mental health service providers regarding the shift to using digital health to facilitate mental health service provision for MMs, veterans, and PSP as a result of the COVID-19 pandemic. Specific study objectives include the determination of the following:

- The current state of digital health use in mental health service delivery for the target populations.
- Barriers and facilitators experienced by mental health service providers in the transition to the use of digital health.
- Recommendations for the implementation and integration of digital health into regular mental health service delivery by mental health service providers.

Methods

Overview

This study used an embedded mixed-methods design [20] in a community-engaged research setting [21]. The primary method was qualitative and involved focus groups with key stakeholders who had direct knowledge and experience, as well as the most useful and timely information sensitive to the context of the research population and question [22]. Engagement with key stakeholders aimed to explore the history of events, the shift to digital health use, what occurred with the shift to the use of digital health (eg, changes in service delivery, access, practice, policy, and technology use), what needs, problems, and solutions have arisen (eg, digital health approaches), and important considerations going forward [23-25]. Quantitative descriptive data were also collected and nested within the thematically analyzed qualitative data. Ethical approval was obtained from the research ethics board of the University of Alberta before study initiation.

Participants

Study participants included representatives from the following groups: (1) MMs, PSP, and veterans working in peer-support, health or wellness, or mental health service positions; (2) multidisciplinary mental health service providers; (3) organizational leaders and policy and decision-makers from local, provincial, and federal PSP, military, and veteran organizations; and (4) subject matter experts and researchers (eg, in digital health, mental health, technology, privacy, security, and implementation). Although attempts were made to recruit MMs, PSP, and veterans who could primarily be categorized as having lived through the experience of a posttraumatic stress injury or OSIs and who were directly receiving digital health trauma therapy, no participants agreed to take part in the study. For MMs and PSP, this was likely because of deployments and frontline service provision during COVID-19.

Inclusion and Exclusion Criteria

Participants were included and excluded based on the criteria described in Textbox 1.

Textbox 1. Inclusion and exclusion criteria of the participants.

Inclusion Criteria

- If they had population-, service provision–, subject-, or organization-specific expertise
- If they were able to offer insights into real-world service delivery, practice, policy, or technology; privacy and security issues in the Canadian context; practical considerations and experiences associated with the use of digital health for mental health service delivery; or realistic solutions associated with digital health service delivery to military members, public safety personnel, and veterans in Canada
- If they were English speaking
- If they were able to provide informed written consent

Exclusion Criteria

- If they were unfamiliar with the population or use of digital health to deliver mental health services
- If they were unable to contribute meaningfully to an exploration of the use of digital health
- If they were non–English speaking
- If they were unable to provide informed consent

Recruitment

Study participants were recruited through word of mouth and convenience, snowball, and purposeful sampling. Key individuals in military, veteran, and PSP organizations were contacted by phone, and emails were disseminated through relevant networks and organizations with local, provincial, and national reach. Key stakeholders (or their designates) were asked to contact the research team directly and were screened for inclusion. Anonymity and confidentiality were reviewed, and consent was obtained before the commencement of study participation.

Data Collection

Data were collected between August and October 2020. Quantitative data were collected using questionnaires that captured demographic information (eg, age, gender, profession, organizational affiliation, years of service, and province of residence), participants' comfort level with particular digital health platforms and technology, and their perceptions of facilitators of and barriers to digital health service delivery.

XSL•FO
RenderX

Questionnaires were administered on the web via REDCap (Research Electronic Data Capture)—a secure, web-based software platform hosted at the University of Alberta and designed to support data capture for research studies [25]. Qualitative data were then collected through 90-minute focus groups (n=5; 3-6 participants per focus group) and interviews (n=1) conducted and recorded over the phone or encrypted Zoom (Zoom Video Communication Inc) software [26]. Examples of focus groups and interview questions are given in Multimedia Appendix 1. Each focus group was purposely heterogeneous, with respect to professional representation and experience with digital health, to allow for broad cross-talk and pollination of complementary and alternative ideas, experiences, and conversation, resulting in rich comprehensive data. The focus groups and interviews were guided and facilitated by senior members of the research team. Key topics of discussion included the previous and current state of using digital health for mental health service delivery in the midst of COVID-19 pandemic; barriers to, facilitators of, and recommendations for the use of digital health technologies to deliver mental health services to MMs, veterans, and PSP; digital health technological issues, acceptance, and methods of delivery; clinical effectiveness; and needs, including infrastructure and implementation.

Data Analysis

Qualitative and quantitative data were analyzed using standard analytical procedures. Quantitative data were analyzed descriptively and statistically using SPSS (IBM Corp). Audio- or video-recorded focus groups and interviews were transcribed and thematically analyzed [27], both deductively and inductively, following an iterative process. Deductively, initial codes were developed based on focus group topics and study objectives. Inductive coding involved identifying themes that emerged from the data [28]. In total, 3 junior researchers independently conducted open coding for each focus group, after which a senior researcher reviewed and refined the codes. These were then combined and tabulated into preliminary themes. The analysis of the preliminary themes by the collective research team followed, with differences resolved through discussion. A proposed thematic theory underwent collective analysis; the preliminary themes were modified; key quotes were isolated to illustrate the selected themes; and the final thematic narrative was prepared.

Results

Questionnaire Results

The overall focus of this analysis was to explore the experience of mental health service providers working with MMs, veterans, and PSP during the transition from in-person to digital health service delivery of mental health interventions. Demographic details of the study participants who completed the digital health questionnaires (N=31) are shown in Tables 1 and 2. Study participants had diverse professional backgrounds, resided in various regions of Canada, and had varying levels of experience within their respective professions and with digital health.

XSL•FO
RenderX

Table 1. Occupational and geographical characteristics of the study participants (N=31).

Characteristics	Values, n (%)
Organizational affiliation (n=31)	
Alberta Health Services	4 (16)
Canadian Armed Forces or Department of National Defense	5 (20)
Postsecondary institution	4 (16)
Municipal Police Service	2 (8)
Operational Stress Injury Clinic	2 (8)
Canadian Coast Guard	3 (12)
Private Mental Health Clinic	3 (12)
The Royal Canadian Legion	4 (16)
Other	4 (16)
Profession (n=27)	
Paramedic	4 (16)
Psychiatrist	3 (12)
Police officer	2 (8)
Administrator	4 (16)
Psychologist	4 (16)
Physician	2 (8)
Chaplain	2 (8)
Researcher	3 (12)
Safety advisor	1 (4)
Other allied health professional	2 (8)
Province (n=24)	
British Columbia	3 (12)
Alberta	15 (62)
Ontario	3 (12)
Quebec	1 (4)
Nova Scotia	1 (4)
Saskatchewan	1 (4)
Years of service (n=28)	
1-4	1 (4)
5-9	4 (16)
10-14	3 (12)
15-19	8 (24)
20-24	8 (24)
25-29	2 (8)
30-34	1 (4)
≥35	1 (4)

XSL•FO
RenderX

Table 2. Age and gender of the study participants.

Characteristics	Values, n (%)
Age (years; n=23)	
18-34	0 (0)
35-49	14 (60.9)
50-64	9 (39.1)
≥65	0 (0)
Gender (n=24)	
Male	13 (54.2)
Female	11 (45.8)

A variety of means of communication used to facilitate mental health service delivery were identified by participants, including videoconferencing platforms, email, telephone, and text. Table 3 shows the digital health methods that participants felt comfortable using.

Intention to use remote delivery, perceptions of the effectiveness of digital health for mental health service delivery, and concerns about privacy and security were rated using a 7-point Likert scale (Figure 1). The vast majority of participants somewhat agreed, agreed, or strongly agreed that they were likely to use digital health delivery of mental health services if in-person therapy was not available. Most respondents also perceived that mental health services delivered remotely were effective. Over half of the respondents expressed at least some concern regarding the maintenance of security and privacy when delivering mental health services through digital health means. Multiple facilitators of and barriers to digital health delivery of mental health services were also identified by participants in the pre–focus group questionnaires (Tables 4 and 5). Common barriers included lack of stable, interruption-free internet access, lack of personal presence, and challenges in developing a therapeutic relationship, whereas identified facilitators included decreased perception of stigma as well as greater convenience and access to mental health services. Consideration of the abovementioned factors associated with the delivery of trauma-focused therapies to MMs, veterans, and PSP before engagement in focus groups and interviews set the context for interprofessional discussions around using digital health for mental health service delivery.

Table 3. Technology and platforms participants have used and feel comfortable using for digital health (N=31).

Technology and platforms	Values, n (%)
Email—personal	12 (11)
Email—work	11 (10)
Phone—personal	15 (13)
Phone—work	11 (10)
Text—personal	9 (8)
Text—work	4 (4)
Teleconference without video	10 (9)
Google Meets	5 (5)
GoTo Meeting	0 (0)
Cisco Webex	5 (5)
Lifesize	0 (0)
Zoom Business or Enterprise	4 (4)
Zoom Health Care	10 (9)
Zoom	12 (11)
Other[a]	4 (4)

[a]*Other* platforms reported included Skype for Business (n=1), Microsoft Teams (n=2), and FaceTime (n=1).

XSL•FO
RenderX

Figure 1. Participants' perceptions of intention to use digital health, effectiveness, and privacy and security risk (N=24).

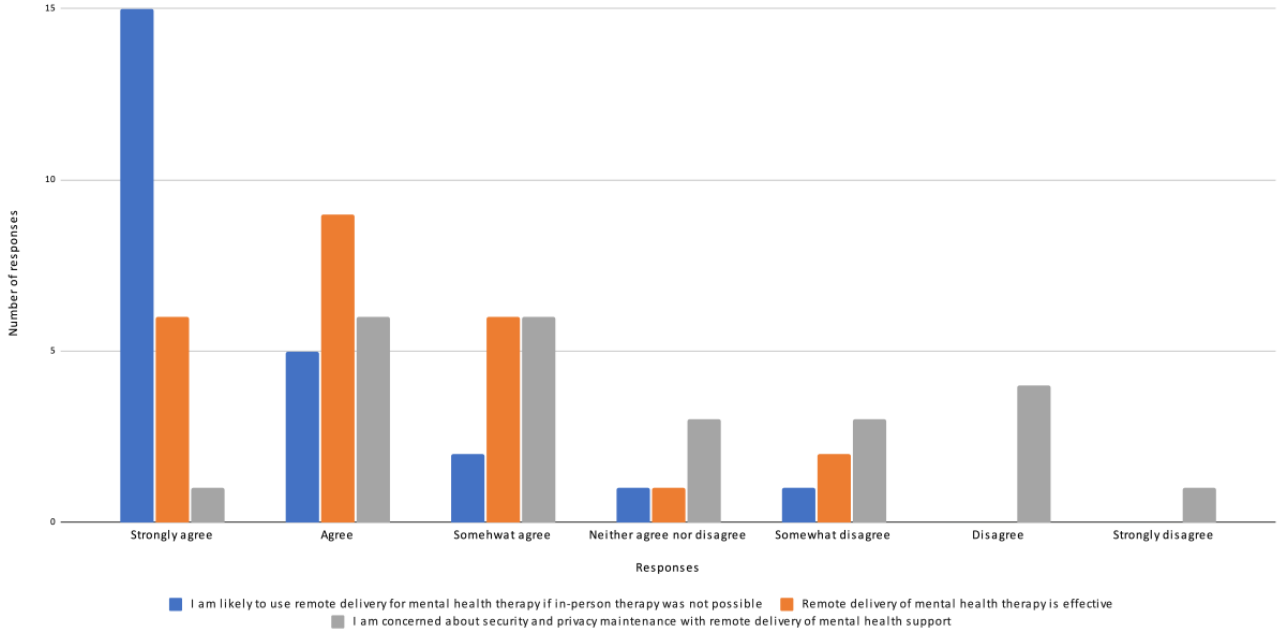


Table 4. Facilitators of digital health delivery (N=24).

What do you think some of the benefits of remote delivery of mental health support are?	Values, n (%)
Not having to travel	22 (23)
Greater convenience	21 (22)
Greater availability of services (ie, more therapists available)	19 (20)
Increased privacy	14 (15)
Reduced mental health stigma from others	17 (18)
Other	3 (3)

Table 5. Barriers to digital health delivery (N=24).

What do you think some of the challenges to remote delivery would be?	Values, n (%)
Lack of appropriate internet access	22 (21)
Reduced privacy	11 (10)
Reduced security	8 (8)
Interruption to communications	19 (18)
Lack of personal presence	17 (16)
Difficulties developing a therapeutic relationship	14 (13)
Challenges delivering or receiving therapeutic modality	14 (13)
Issues of safety in case of negative response	1 (1)

Qualitative Results

Thematic analysis of the qualitative data isolated three main themes: (1) being forced into change, (2) daring to deliver mental health services using digital health, and (3) future possibilities offered by digital health. In each theme, participants' responses encompassed considerations for service providers, organizations, and clients. What followed was an elaboration of these themes, together with tables of subthemes and supporting quotes.

Theme 1: Forced Into Change

Service Providers

Providers' willingness to use digital health was integral to delivering trauma therapy to MMs, veterans, and PSP amid physical distancing restrictions (Textbox 2). Personal experience, skill level, and readiness to adopt technology were reasons for both mental health service provider excitement around and reluctance to accept the use of digital health to deliver trauma therapy.

XSL•FO
RenderX

Textbox 2. Forced into change—service provider subthemes.

Addressing Personal Attitudes and Mindset

- *I think we're the bigger problem then patients are in some cases. I think we're the larger barrier to being able to shift to a digital-based solution. Because I was concerned about it, I was anxious.* [Focus group 3; P18]
- *But it was provider willingness, I think some of it was technology-based, some of it was just the unwillingness of certain providers and I don't know why - was it you're missing something? was it them - how am I licensed to do this? Some college issues? Personal professional issues? Or maybe it's just fear of the unknown - 'well I don't understand this technology so I'm just going to stick with pen and paper.'* [Interview, P1]

Managing Increased Workflow Demands

- *It definitely was a transition for us on a clinician level. There is more to manage, there's more things to review. So I think for some people it kind of slowed down the process because we needed to [...] talk about what it meant to do these virtual sessions, [...] what it meant to [...] find a place in your home where you can do this trauma work that's going to be sensitive, that's going to be appropriate, [...] the logistics.* [Focus group 3; P14]

Accessing Client Information

- *[F]or somebody with a head injury, [...] they can't get a neuro psych assessment; that's just been put on hold indefinitely. Again, you know, not being able to get this kind of multidisciplinary information to inform the supports that I'm giving or the treatment direction that I'm giving you know what part is post-concussion syndrome and what part is mental health and how am I going to tease that out? Well, I don't have the information. So we are kind of making do.* [Focus group 5; P33]

Exchanging Therapeutic Material With Clients

- *[H]ow do we get the worksheets to the person? How do we get the person the materials that are part of the treatment? How do we have a conversation where I can't see the worksheets? For example, sending emails - so I have to remember I got to make sure that I send the email for this week's homework or I gotta remember on my end to make sure that I send out the workbook for the person to do cognitive processing therapy, and then it might take a week or two for it to actually get to them, given how the mail system is right now.* [Focus group 3; P14]

Making Reliable Clinical Observations: The Body Tells the Story

- *We have been trained our whole careers [...] to vet all emotion out of everything we write, every report has to be completely neutral. [...] [I]f we don't have the body to tell us the story, if we don't have that ability to observe, we're really limited [...]. Tone of voice and all those things are really helpful, but more seasoned Cops can become really good at [...] hiding all real emotion.* [Focus group 2; P9]

Facilitating the Therapeutic Alliance

- *[T]he physical connection and the closeness of actually being with someone is a huge part to allowing them to trust you, to feel your empathy, to allow you into what they're going through, and doing it screen to screen, especially when there are glitches and stuff that happens, just is not as effective and I have noticed that our members are not on board with it.* [Focus group 2; P13]
- *From my point of view why it doesn't work again, is just that there's something real with the energy you share with someone when you're with them. And this takes that away. So from my point of view it takes so much longer to build trust, to get people to open up. So I think from my point of view that's what I saw, they were just not as willing to be open and honest of what they're going through... they're just not as comfortable having that conversation online versus face to face. And I understand that, because I really feel like that connection is not there... members do not want to have these meetings online, they just don't.* [Focus group 2; P13]

Conducting Risk Assessments and Safety Planning

- *I would find it very difficult to do [...] a real thorough assessment if someone was actively suicidal in a virtual setting. [...] I would find it very difficult to see how that can be conducted wholesomely, also without seeing – because agitation can be seen in the feet. There's so much that I would take under consideration when I assess a patient who is suicidal.* [Focus group 1; P3]
- *Where is your client geographically? If you lose contact with them and they're in crisis how are you going to get them help? So really thinking through, what are the clinician guidelines about knowing where people are?* [Focus group 5; P33]

Workload responsibilities reportedly increased with the shift to remote delivery of mental health service provision. The proficient use of technologies was challenging for some and resulted in increased work time and administrative duties. Providing clients with therapy resources or exchanging confidential materials required creative problem-solving. One participant described needing to be a clinical service *technology wizard* since the start of the COVID-19 pandemic. Participants also noted that the additional time and effort required to implement practice change was often underestimated or ignored by employers.

Communication with colleagues around client care was impacted by a shift to digital health. Participants described having fewer opportunities to confer and collaborate with their colleagues. They also indicated that client information was not as readily accessible, thereby making assessments and case histories more difficult to complete. Furthermore, securely and remotely accessing and transferring client files required additional consideration.

The delivery of trauma therapy also requires careful adjustment and adaptation. Participants noted that the sense of engagement

XSL•FO
RenderX

and energy were notably different between in-person and digital health sessions. They also noted that assessing and monitoring nonverbal cues in a virtual setting is challenging. As MMs, veterans, and PSP typically suppress emotions, Mental health service providers rely on nonverbal facial expressions and body language to determine what a client may be feeling. One participant noted that "the body is what tells us the story." The inability to "see below a client's head and neck" can impede an SP's ability to pick up on mental health indicators (eg, overall appearance, hygiene, and psychomotor agitation). Participants also felt that digital health negatively impacted the therapeutic alliance and, at times, resulted in the clinician feeling distant or cut off from their client.

Assessing risk over digital health was also consistently identified as a topic requiring consideration. Participants questioned how to best manage client disclosures of suicidal ideation, intent of harm to self or others, or domestic violence. Dealing with such situations was identified as the primary reason that using digital health was uncomfortable and anxiety provoking.

Organizations

Although organizations and service providers wanted to provide the best mental health services possible to MMs, veterans, and PSP with OSIs during the COVID-19 pandemic, they encountered numerous barriers to service provision (Textbox 3).

Organizations and regulatory bodies for regulated health professionals were unprepared for a sudden shift to digital health service delivery. For example, some participants noted that adopting digital health was prohibited by organizational mandates just days before the COVID-19 lockdown. As a result, policies, practices, systems, and resources were viewed as being limited or inappropriate. Mental health service providers indicated that these realities impeded their use of digital health when responding to client needs.

Security and privacy concerns were strongly highlighted. Participants noted that access to organization-specific computer hardware, software, and systems was lacking, making it necessary for them to use personal devices in both work and home environments. Given this reality, organizational security and privacy policies were often either overly restrictive or not restrictive enough to address the unique factors associated with offering trauma therapy from home.

Textbox 3. Forced into change—organization subthemes.

Changing Organizational Practices

- *Some agencies are so stuck in the way they've always done it. (...)[S]o this is a whole new chapter, truly.* [Focus group 2; P9]
- *We don't have the latitude in the military on our defense wide network of deviating from the rules, because everything is sort of being scrutinized, and so we've got clinicians who want to be able to communicate because the patients are asking to communicate on email. [...] [T]rying to find a way to do it within the bounds of the rules has been a challenge for us.* [Focus group 3; P21]

Exploring Digital Health Platforms to Meet Client Needs

- *[W]hile everyone is keen to use Zoom, [...] we're not able to use it on our network because of the concerns over security. So it raises issues for us in terms of what platforms we can actually use on our network; [...] the military but also to our federal partners [...] are bound by the same security parameters so that's a real impediment for us.* [Focus group 3; P21]
- *[T]he policies that have been in place and the privacy concerns that have existed since we started have been really huge barriers to our delivery of care. Our population usually has cognitive deficits, has difficulty with technology and just problem solving at its most basic. For us, one of the policies for example is to encrypt all emails. People could not figure out how to unencrypt them and access our Zoom groups or our Zoom meetings if we encrypted them.* [Focus group 1; P5]

Managing Security Concerns

- *I agree that to just start up with any platform and start talking to patients without any real understanding of the privacy implications is a real risk.* [Focus group 1; P3]
- *Videotaping that has always been something that is part of our sessions for therapy purposes. That has been turned off by [agency] as a status-quo for all our sessions on Zoom, so even when we went to leadership to turn that off so we could Zoom videotape, and the patient gave consent, we couldn't do that because the legal team wouldn't allow us.* [Focus group 1; P5]

Clients

The service provider participants thoughtfully considered ways to enable clients to access mental health services and overcome barriers to digital health use (Textbox 4). These participants spoke of the importance of ensuring that clients have the appropriate technology to support digital health delivery, understand how to use the technology and appreciate the risks and benefits of using digital health, and provide consent before engaging in trauma therapy. Participants highlighted that digital health also presumes a certain level of socioeconomic prosperity (ie, access to a computer, high-speed internet, and a private or secure location), which may not be attainable for all clients. As some clients may be uncomfortable with the possibility of a session being monitored by an employer or third party, study participants emphasized the importance of reassuring clients of the confidential nature of therapy sessions.

XSL•FO
RenderX

Textbox 4. Forced into change—client subthemes.

Managing Inequalities

- *If I'm out of work and or don't have [...] high speed internet, and I only have a phone or [...] a tablet that's maybe outdated, [...] my experience engaging with you as a provider is going to be very different [...] interrupted or lower quality or I'm going to be so preoccupied. [...] [T]hat's really a socioeconomic issue.* [Interview; P1]
- *Immediately there were some challenges with some of our clients, knowing how to use technology being able to navigate the digital world, particularly for older clients and then the technology itself whether that was the specific one that we were using or other ones. [...] Other issues were related to people not having appropriate technologies.* [Focus group 5; P29]

Addressing Confidentiality and Privacy Concerns

- *[Y]ou actually have to tell clients that they shouldn't take the call sitting against an airplane hangar while there's people passing in front of it. That they actually need some privacy and security in a place where they won't be interrupted in order to do this work. So because we're not containing the space, we need to educate our clients about how to create that space for themselves.* [Focus group 5; P33]
- *[I]n terms of privacy, if you're now consulting by video call [...], you could potentially see their background and you could sort of see into their homes, [...] people walking by or dogs or cats or children.* [Focus group 1; P4]
- *I have most of my conversations as conversations because they don't want anything in writing ever. [...] [A]nything that's important they want to know it's staying between you and them.* [Focus group 2; P11]

Building Digital Relationships

- *For me, my experience is when I'm working with folks like that, if I have a relationship with them already talking on the phone is, is really effective. We don't do video conferencing. But if I don't have a relationship with them and then I want to phone and talk to them about an issue, it's not as effective.* [Focus group 4; P23]
- *[T]hey have so many trust issues, and they have anxiety about going in to see a clinician in the first place that, now you throw the computer in there, they're unsure of what's being recorded, [...] who has access to it [...]. I think it just increases their anxiety about opening up and sharing.* [Focus group 2; P13]

Facilitating Coregulation

- *It's that self-soothing piece because you're not gonna be in the room to be able to support me. [...] [I]t's really giving people the tools [...]. They're being activated in a way that you can't necessarily see, but also could stay activated afterwards [...]. [T]hey might process that more quickly if they're in the room with you.* [Focus group 4; P23]
- *[C]ops are reluctant to talk about emotions at all; [...] they don't want to take their cork out of the bottle if there's not somebody to catch them.* [Focus group 2; P11]

Accommodating for Cognitive and Trauma Challenges

- *[W]ith members who have been dealing with PTSD [posttraumatic stress disorder], just practically, sitting and staring at a screen for an hour, would be very difficult [...]. Bright lights. Some guys can't even watch TV, so I think even practically that would be a big thing.* [Focus group 2; P13]
- *There are absolutely people who have a huge adversity to technology, and it's now linked to their trauma, [...] [T]hey're hearing it and are helpless to fix the situation.* [Focus group 2; P9]
- *[S]ome members, [...] although they feel it's not going to be as productive as face to face, would still do it; [other] members would just be like, "no thanks, I would rather just deal with it." [...] [I]t would be case to case.* [Focus group 2; P13]

Conducting trauma therapy using digital health, although workable for many clients, was not viewed by the participants as appropriate for all MMs, veterans, and PSP. Some participants felt it may be more challenging to build therapeutic alliances virtually rather than in person with certain clients. Participants emphasized that care must be taken in selecting the best format of service delivery for each individual client, especially MMs, veterans, and PSP, who may experience issues of trust and attachment.

Participants expressed that clients with cognitive dysfunction, who have experienced a brain injury, and who are highly emotionally dysregulated and in frequent need of active coregulation may not be suitable for digital health delivery of trauma therapy. Furthermore, those clients for whom technology is a trigger (eg, police officers who may be required to watch graphic web-based content) may also not benefit from digital health-delivered trauma therapies.

Theme 2: Daring to Deliver

Service Providers

The use of digital health afforded service providers unexpected benefits (Textbox 5). Most participants noted that digital health service delivery *saved time* because of a lack of commuting and fewer workday interruptions, which enabled them to respond to more clients. Participants noted that the efficiency experienced was also impacted by changes in their attitudes toward digital health and by gaining familiarity with the technology and software platforms.

XSL•FO
RenderX

Digital health offered new intervention possibilities and options for providing outcome-oriented services. Some mental health service providers reported adapting psychoeducational approaches, developing new ways of conducting exposure-based interventions, and facilitating client connections to support groups. Participants noted the importance of allowing service providers to develop creative solutions, with 1 participant stating, "Stay nimble...keep looking at how to make things better...leverage this." Furthermore, as service providers acclimated to digital health service delivery, they were better able to accommodate for the lack of access to nonverbal cues. Provision of therapy to a client in their home environment afforded service providers a glimpse into the client's living situation, offering valuable therapeutic information.

Textbox 5. Daring to deliver—service providers.

Increasing Efficiency Through the Use of Digital Health

- *Pre-COVID there were lots of barriers to virtual care and [...] working from home, but we have proven during this crisis that it can be done and we can be as efficient or almost as efficient working from home and taking care of our patient than we would be sitting in our offices with them in front of us.* [Focus group 3; P17]
- *[I]t became, I guess, time management wise, actually a lot easier. She could sit at our console, get a video screen, [...] and do a team's call with the one, the two or even a group session.* [Focus group 4; P25]
- *The fact that I don't have walk-ins has allowed me to do a lot more with more patients. That I'm not interrupted by people knocking on my door, I can finish on time, I can focus better in my sessions, it's easier, and I can still have patients communicate with me through my secretary so that I can get back to them when I have time. Putting that barrier actually helped me focus more on my work.* [Focus group 1; P5]

Accepting and Gaining Familiarity of Digital Health

- *There's been this real kind of shift in my thinking from 'this is second rate. This is not as good as the in person' to really try to dig in and find what's the unique benefit of a video platform in terms of if we're going to be doing this for a while, how do we lean into the technology and what it gives in terms of immediacy, of consolidation and implementation?* [Focus group 5; P33]
- *I'll FaceTime somebody now instead of phoning them because we've gotten used to looking at people on screens and so I feel that that's actually been a third benefit for us here is the use of FaceTime, as opposed to just a phone call. So you can actually read some body language and although it's through a camera, look at people in the eye.* [Focus group 4; P22]

Creatively Delivering Interventions

- *So for me I was still in the office, facilitating exposures, where I could. For some folks, though, that I was working with that were immuno-compromised. Well, we couldn't do face to face live exposures and so we took it online. And so what that looked like was we started with the members sending me a list of videos that she needed me to record on my phone [...] and I walked around headquarters with my phone trying to be as steady as I could and recorded them, [...] So those were really, really successful and she was really happy with them and there was really good feedback.* [Focus group 4; P22]

Enhancing Insight Into a Client's Life Circumstances

- *[Digital health provides] a glimpse into their world. You see what their environments look like; [...] whether they're in a state of chaos, just based on their room or the noise in the background. So some of those little things give you a bit of a sense of where they're at and what's going on for them. So I think that is [...] beneficial.* [Focus group 5; P29]

Organizations

Adaptations at the organizational level were made to enhance digital health service delivery (Textbox 6). Participants reported that organizations that seemed to be able to adapt to digital health most effectively embraced new opportunities for resource sharing and interagency collaboration. For example, connections with colleagues and national and international experts were also made more possible as a result of digital health. The shift to digital health also enabled organizations to expand their service offerings, build greater capacity, and facilitate more equitable mental health service provision, particularly with geographic limitations being removed by digital health. This move to more equitable mental health service provision was deemed to be extremely beneficial and relevant to MMs and PSP who are often required to work in remote and isolated locations.

Textbox 6. Daring to deliver—organizations.

Sharing Agency Resources

- *[Digital health has been] positive in terms of helping other agencies create care options, [...] start reintegration programs. Zoom has been tremendous at connecting agencies. [...] I have done a lot of teaching over Zoom since all of this has started. So, that in terms of teaching [...] or connecting agencies [...] that has been awesome! [...] [I]t's one level removed from the care to the individual, but I think it's still so relevant.* [Focus group 2; P9]

Providing More Equitable Services

- *The reality of that lived experience is the occupational injury clinics and the therapeutic resources are concentrated in the metropolitan environments. So, prior to COVID, we took a stance that the support we were able to provide was equitable but not equal. [...] [A]n unintended benefit is that the playing field is now level in that everybody has the exact same type of interaction.* [Focus group 1; P2]

XSL•FO
RenderX

Clients

The effectiveness of digital health for use with MMs, veterans, and PSP was associated with several factors (Textbox 7). Participants identified the benefit of proactively reaching out to clients to provide regular support and information regarding how and when they might be able to access digital health therapy.

Overall, the participants perceived that the use of digital health reduced barriers to access and enhanced client comfort during appointments. The participants reported that the no-show rates for client sessions were unchanged in the shift to digital health and that missed appointments decreased. The strength of the therapeutic relationship impacted the ease with which service providers were able to transition from in-person to digital health delivery, with transitions being easier when a therapeutic relationship had been established through in-person sessions. Participants perceived that clients demonstrated a *forbearance flexibility* with the shift to digital health and an appreciation for the steps taken to maintain their safety and mental health during the pandemic.

Textbox 7. Daring to deliver—clients.

Continuing Access to Appointments

- *I think for trauma-focused treatments my no-shows have gone down.* [Focus group 3; P14]
- *Very often now people are in their own space, they're comfortable, there's actually more intimacy and there can be more safety.* [Focus group 5; P33]
- *[W]e have not seen fewer clients or had fewer sessions. They're very comparable. And we have not seen any change in our no-show rate for appointments, [...] So clients are attending, [...] [and] getting just as many sessions.* [Focus group 3; P15]

Creating Connections

- *My colleague and I ended up doing some groups, partly so people felt connected, and yet we were teaching skills and mindfulness [...]. So I think that there's a real benefit to that, even just to increase access.* [Focus group 3; P16]
- *[We're] really staying in close contact. Right from that first week, clinicians [...] phoned every client to say 'we know this is a difficult time, [...], here are the options available to you.' So that proactive outreach, I think, was a really important piece of keeping our clients engaged.* [Focus group 3; P15]
- *We started a sentinel program in the unit so we had people responsible to check in on their people routinely.* [Focus group 1; P6]

Adapting to Use of Digital Health for Mental Health Service Delivery

- *[W]hereas it took longer to establish a therapeutic relationship at the beginning of COVID, I'm finding that [...] it's pretty seamless now. So people are adapting. I'm adapting. People are adapting.* [Focus group 5; P33]
- *The feedback I'm getting from the ships is that yeah, this is different, but what we're providing is good and sufficient. I mean they understand the constraints, they very much appreciate that we're not putting them at physical risk by insisting that we send a team in. And so there's that patience, forbearance flexibility that's required on the other end to make this work. But our observations is that the folks are A: professional enough to deal with that, but B: they realize the need for the support. So they take the compromised approach rather than just say, don't bother. So there's an adaptation that's occurred on everybody's part right now, and it's not just technology.* [Focus group 4; P25]

Theme 3: Looking Forward

Service Providers

Viewing current circumstances as a catalyst for change, participants made recommendations to enhance mental health service delivery based on their experience of digital health use (Textbox 8). Suitability of clients for trauma therapy using digital health was discussed, with safety concerns and risk factors being primary considerations. Participants noted that using digital health to deliver mental health services could work very well for clients experiencing anxiety, depression, or fear of stigma or finding it difficult to get to mental health appointments. However, they also noted that not all of their clients fared well using digital health, and they often had to meet with clients in person, despite COVID-19 restrictions, as the client had deteriorated.

To address this concern, service providers identified that a continuum of mental health service delivery, from in person through the web, may be more effective in these populations as clients could choose to receive treatment in person, digitally, or a combination of both. Participants further discussed the possibility that a hybrid model could be used as a way to build rapport between service providers and clients before progressing to a web-based format. Considerations associated specifically with providing trauma therapy using digital health with MMs, veterans, and PSP were also noted.

Managing screen fatigue—the sense of fatigue caused by staring at a computer screen—was also considered critical to the ongoing delivery of mental health services using digital health. Service providers emphasized the importance of self-care, social support, and adequate breaks to do so. It is possible that service providers may need to see fewer clients per day as a result of decreasing billable hours. Fatigue may also be prevalent among clients, necessitating that agencies refrain from overwhelming clients with too much information or on-screen sessions.

The importance of training service providers in the delivery of trauma therapy in a virtual context was emphasized by participants. This may include the use of virtual platforms to deliver therapy and adaptations of therapies to a virtual format, including CPT and eye movement desensitization and

XSL•FO
RenderX

reprocessing. Practical issues regarding training were also raised, such as when, where, and how this type of training might be provided.

Textbox 8. Looking forward—service providers.

Catalyzing Change

- *[T]he greatest travesty that could come out of this [...] whole pandemic is to have us revert back to old ways without being able to solidify some of the real values that have come out of it.* [Focus group 3; P21]
- *We have some opportunity right now in this environment to [...] highlight some of these important issues on how we can move them forward.* [Focus group 1; P6]

Determining Client Suitability for Digital Health

- *I had to see half of my people in person, because they were [...] deteriorating at home. And so that was sort of an adjustment that I had to make following some of our Provincial recommendations and guidelines from our college.* [Focus group 5; P30]
- *There are lots of treatment programs. [...] [How do] you know which one to use? How do you know if it's actually being effective in your situation?* [Focus group 4; P24]

Managing Screen Fatigue

- *On a really busy clinical day and an exceptional day, I could see eight people in a day and I sort of rolled into the telehealth system with that schedule in place. I'd be wiped out for a couple of days after that. And so I've found the busiest day that I can really put in now is about five. And so there is this kind of screen fatigue that happens.* [Focus group 5; P33]

Providing Continuing Education

- *Not to say that digital health can't be used, but as you said, I think it's a whole other way of being [...] [I]t's almost as if we're going to have to retrain all of our clinicians [...] to make sure that people have those skills and those competencies is huge because for me there's a level of fear and uncomfortableness.* [Focus group 1; P35]
- *I had no virtual care training [...] So that's a big, big training piece that clinicians should undertake. Now how does that get instituted? How does that get overseen? Who pays for that? This is all time consuming, and clinicians are all so busy to take that on. I think a lot of them are just learning as they go.* [Focus group 1; P3]

Organizations

Many organizations and regulatory bodies for regulated health professionals have rapidly developed or enhanced existing policies and practices to support the use of digital health (Textbox 9). Participants noted, however, that policies still require adaptations as the provision of digital health continues to evolve, especially for organizations that had previously been uncomfortable or avoided using digital health. Participants also noted the intentionality required to deliver quality mental health sessions using digital health and strongly suggested eliminating distractions during digital health appointments such as phone calls, emails, and personal interruptions, which were found to be more difficult in a virtual setting than in person. Administrative support and resources were also cited as being essential elements for both clients and service providers in problem-solving technological challenges associated with digital health.

Textbox 9. Looking forward—organizations.

Aligning Practice and Policy

- *I know in our group there's been such an impetus to move forward that we're just trying to catch up with policy implementation. We just don't have the time to get it all done and written and out there. Policy needs to catch up with the process, and that's a challenge.* [Focus group 1; P3]
- *For organizations, do they now have a policy that supports this sort of virtual online world? And then it's about ensuring that is the appropriate equipment in place? The technological support if somebody is having a challenge accessing remotely that there's someone they can call that will help walk them through because there's nothing more frustrating than a technical problem, when maybe you're trying to reach out or get support or something like that, right? You just sort of say, well I just give up on that, right.* [Focus group 2; P10]

Reducing Work Space Distractions

- *Number one, eliminating distractions. Like, even as I'm sitting here, I'm looking out the window. I've been checking my phone. I've been plugging away at a draft email. And so it's harder to eliminate the distractions when we have a venue like this, as opposed to being in the room. So we need to make sure that we're very clear with folks that we might be doing this for whatever reason, that you need to work extra hard to eliminate the distractions in the room.* [Focus group 4; P22]

Providing Administrative Supports and Resources

- *But the idea of having dedicated administrative support to send out the emails, to help clients troubleshoot and feel more comfortable, to walk to a clinician's office and help them troubleshoot and feel more comfortable. The sooner we can work through some of those barriers, the more comfortable, people will feel because if they continue to struggle, it's easier just to give up and go back to what you know.* [Focus group 3; P15]

XSL•FO
RenderX

Clients

Service providers perceived that the use of digital health enhanced service options and reduced barriers for clients accessing mental health services (Textbox 10). They indicated that this was particularly beneficial for clients residing in rural areas where mental health services may be sparse. They also felt that the flexibility of service delivery contributed to a person-centered approach, where a client who may not be ready to access in-person support is still provided service. In addition, in their experiences delivering mental health care using digital health, participants emphasized the need to match generational norms with digital health, with younger generations tending to be more comfortable with digital health delivery.

Textbox 10. Looking forward—clients.

Enhancing Service Options

- *Rural communities or officers that are in two person detachments not having any access to care- like there is no psychologist in a 400 kilometer radius [...] I even look at post-shooting, how many times have we had [...] two officers [...] in a small community and they have nothing, and the communities kind of turn on them. [...] [W]hen we're looking at smaller communities, this is a game changer for actually having people access what they need.* [Focus group 2; P9]

Reducing Barriers

- *[F]or some patients with social anxiety, for example, internet-based sessions are more accessible [...] [I]f they feel demotivated, [...] it's easier to log in through the internet then it is to come to the clinic.* [Focus group 1; P5]
- *Our goal is always to try to remove all of those barriers. [...] [S]ome people may feel more comfortable starting with text messaging or in person and other people might say the last thing I want to do is actually go in person and talk to somebody, I'm not ready for that yet, but being able to work up. [...] [S]hould we return to a [...] normal world, how do we take the best of what we're learning and give more options to the people we're working with? [...] So this pandemic has really forced organizations to pivot really quickly. [...] [and] take a lot more risks than they would have before because they haven't had any other choices.* [Focus group 2; P10]

Responding to Generational Differences

- *I've had two members with teenagers that needed to see a clinician, and both member's kids chose to do it online when they had the option to go in and see the psychologist face to face. [...] [T]hey just felt more comfortable that way. [...] [O]nline's fine.* [Focus group 2; P13]

Discussion

Principal Findings

This study aimed to explore the perspectives of MMs, PSP, and veteran organizational leaders and policymakers and decision-makers; subject matter experts and researchers; and mental health service providers supporting MMs, veterans, and PSP experiencing OSIs amid the transition to digital health because of the COVID-19 pandemic. In light of the impact of the pandemic on health care services and as mental health service provider comfort with and acceptance of technology is a critical component of successful implementation of new initiatives and treatment modalities [29], such an exploration is timely. This study is especially important in that it begins to address the knowledge gap regarding the experiences of those who work within the Canadian context.

Similar themes and a general consensus emerged across focus groups regarding the use of digital health for mental health service delivery to MMs, veterans, and PSP. Almost all participants reported using digital health for mental health service delivery amid the COVID-19 pandemic and suggested that digital health could be adopted as a standard delivery mode for trauma therapy [30]. Similar to research on the use of digital health in the civilian population, some participants identified opportunities and benefits from the widespread adoption of digital health, including more equitable access to mental health services, especially in geographically remote locations; reduction in mental health barriers and stigma; and the ability to develop novel and creative mental health solutions to ongoing challenges [31]. Specifically, as MMs and PSP often work in rural settings and geographically isolated locations, the ability to offer evidence-based mental health treatment to these workers who would otherwise have no access to these supports should command significant attention. Many participants felt that returning to in-person delivery would be a *step back* and that such a return would fail to capitalize on the lessons learned and the work accomplished during the COVID-19 pandemic.

Although digital health provides a number of significant benefits and has proven to be feasible, attention needs to be given to certain issues before its widespread adoption. For instance, systemic changes are needed to effectively facilitate the transition from in-person to digital health service delivery. Of primary concern are the changes needed to update policies and procedures concerning privacy and security. Critical attention needs to be given to platform selection based not only on bottom-line cost but also on functionality and evidence of overall user experience and satisfaction. Such procurement considerations may be at odds with the current legislated policy [32]. Infrastructure, hardware, software, and connectivity at the organizational, clinical, and client levels all need to be explicitly supported. In addition, the workflow, appointments, electronic medical records, liability, and billing codes in place pre–COVID-19 are all based on in-person service delivery and are not always easy to adapt to digital health trauma therapy. Regulatory and reimbursement hurdles faced by individuals and institutions implementing telehealth have been previously published [33,34].

Active support for service providers throughout the transition to the use of digital health for mental health service provision is critical. At the time the focus groups for this study were

XSL•FO
RenderX

conducted, participants expressed a degree of openness to digital health delivery. Some mental health service providers participants remained highly skeptical of digital health; however, they further highlighted the critical need for digital health specific training for service providers working with MMs, veterans, and PSP. With the abrupt shift to digital health, mental health service providers found themselves outstripping policies and hastily putting into place ad hoc procedures. They sometimes experienced uncertainty and perceived changes as potentially conflicting with best practices and client interests. Research before the COVID-19 pandemic illustrated that mental health service providers who experienced technological challenges reported feeling more fatigued and experiencing professional self-doubt and a loss of confidence [8,35]. Generational differences may also be an important factor to consider.

Population-specific training would better prepare mental health service providers to deliver quality mental health care to MMs, veterans, and PSP experiencing OSIs. Training is needed to effectively engage with clients when one cannot rely as much on nonverbal cues, as is the case with digital health. The perceived reduction in observed interpersonal cues was frequently identified as a challenge by mental health service providers in Britain using digital health during COVID-19 [31]. As nonverbal cues are a key component of trauma therapy, the reduction or removal of cues represents a potential loss of valuable information in the treatment process [31]. This is particularly significant when delivering trauma therapy to clients with complex mental health needs and at an increased risk of harm to self or others. Such concerns become even more salient when clients are highly trained professionals who have knowledge of weaponry and lethal force, which can put them at greater risk for completed suicide.

Mental health service providers were acutely aware of the need to be attentive to subtle verbal and nonverbal cues to establish the therapeutic relationship, deliver trauma therapies, and assess both treatment effectiveness and risk of self-harm. Geller [33] had argued that when mental health service providers are "fully in the moment" and attuned with their clients, this sense of mutual safety and strong alliance invites clients to engage in the necessary therapeutic work by increasing their ability to emotionally regulate. Although not fully known at this time, this attunement and subsequent emotional regulation may be lessened in the context of digital health [36]. Research into *Zoom fatigue* indicates that videoconferencing is more mentally taxing, the level of communication is not completely synchronous, and greater cognitive energy is required [35,37]. Cognitive demand may be intensified by the lack of nonverbal cues, which would typically assist communication during in-person encounters [37]. Within the context of intensive trauma therapy, the cognitive endurance required may be even greater than what would be required for an in-person therapeutic encounter.

The appropriateness of digital health for clients is also dependent on cognitive functioning, the source and type of trauma, and technical literacy. Given that cognitive impairment (including memory, attention, concentration, and executive processes) is fairly common in trauma-induced mental health disorders, it is unclear whether some elements of digital health could be specifically problematic or challenging to certain MMs, veterans, and PSP [38,39]. Jones et al [19] determined that the use of digital health for trauma therapy in these populations showed comparable results to that of in-person therapy. Similar results have been found regarding the feasibility of telehealth for serious mental illnesses (ie, schizophrenia, schizoaffective disorder, and bipolar disorder) [37,38]. Such evidence suggests that digital health, or at least the use of a hybrid approach (blend of digital health and in-person services), could be effective for MMs, veterans, and PSP. It is possible that some clients may benefit from in-person delivery as rapport is developed, with the delivery of therapy shifting to digital health platforms once the therapeutic alliance is well established. There may also be certain aspects of rehabilitation and recovery that may be more or less suited to digital health. For example, clients who require acute stabilization may be better suited for in-person service delivery and monitoring compared to those with more stable symptoms. A continuum of service offerings and flexible and adaptable delivery mechanisms on the part of mental health service providers may best meet the needs of MMs, veterans, and PSP experiencing OSIs and seeking mental health treatment. Delivering safe, effective, and timely trauma therapy to MMs, veterans, and PSP by mental health service providers who are comfortable with both in-person and digital health environments would significantly broaden service delivery options and likely result in increased access to mental health care.

Recommendations and Future Research

As digital health is a novel method for delivering trauma therapy, it requires further systematic research and should not currently be considered synonymous with in-person therapy. The aforementioned considerations should be explored, and modifications should be made to ensure safe and effective mental health care to widely implement digital health delivery of mental health services. Research into digital health is imperative to define the best approaches for effective digital health-delivered care and treatment of these and other patient populations. Special consideration should be given to the fit of technology within the organization, environment, culture, and other relevant contexts [39].

Key recommendations from this study:

- Specific infrastructure, technologies, policies, and procedures are needed if digital health is to be permanently integrated as a predominant mode of mental health service delivery.
- Mental health service providers need access to administrative, technological, colleague support, and recognition and mitigation of new job demands in a digital health environment.
- Training is required if mental health service providers are to effectively, comfortably, and competently deliver trauma therapies to MMs, veterans, and PSP in a digital health environment.
- A hybrid model may bridge services from in-person to digital health service provision and allow for the delivery of flexible, personalized care.
- Proactive outreach is needed to support clients with OSIs.

XSL•FO
RenderX

- Clients also need to be educated about protocols, procedures, potential benefits, and challenges associated with digital health service delivery.

In addition, if digital health is to be broadly used as a medium for mental health delivery for MMs, veterans, and PSP, further research is needed in a number of areas. First, further exploration regarding which trauma therapies are most conducive to digital health environments is required. To date, most of the research on the delivery of digital health trauma interventions has focused on CPT and prolonged exposure, with limited studies on eye movement desensitization and reprocessing [19]. Research is also needed to determine the indications and contraindications to the use of digital health. Significant research is needed regarding the management and mitigation of risk when delivering digital health trauma therapy. Finally, research should explore the potential of digital health to be used to deliver novel and creative trauma therapy. In this study, some participants highlighted the possibility of delivering therapy when a client was walking, whereas others noted that trauma therapy could be provided via smartphones.

Limitations

This study had some limitations. First, this study used a small convenience sample, which drew on pre-existing relationships with national, provincial, and municipal military, veteran, and PSP partners. Second, the ability to recruit was affected by the COVID-19 pandemic as two of the populations of this study (ie, active-duty MMs and PSP) were frontline SPs during the pandemic and significant work demands limited their ability to engage in the study. Third, there were limited frontline participants (potentially because of the impact of the pandemic on their work), resulting in more perspectives being gathered from policy makers and SPs than from persons with lived experience. Consequently, the views of client experiences described in this study are mostly from the perspective of clinicians or mental health service providers. Finally, data may have been biased by focus group participants being potentially swayed by the views of others or participants not feeling comfortable disclosing their personal views in a group.

Conclusions

This study explored the use of digital health for the provision of mental health services during the COVID-19 pandemic from the perspective of stakeholders, including mental health service providers and policy makers, providing trauma therapy to MMs, veterans, and PSP. The findings suggest that digital health is a viable component of a model of care for MMs, veterans, and PSP, which is inclusive of in-person and virtual modes of trauma-focused service delivery. The findings offer considerations for whom and at what point in treatment digital health is appropriate; clarification of training, support, resources, and guidelines necessary for service providers to be successful in the digital delivery of trauma therapy; and a better understanding of the factors influencing service provider perceptions and acceptance of digital health. These results can inform the implementation and uptake of mental health interventions via digital health for MMs, veterans, and PSP and may equally apply to other trauma-affected populations. As the COVID-19 pandemic continues, remote service delivery methods for trauma therapy will be increasingly needed to support the mental health and well-being of MMs, veterans, and PSP who continue to serve and respond to the needs of communities.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Focus group (World Cafe) or interview questions.
[DOCX File , 14 KB - formative_v5i8e26369_app1.docx]

References

1. Liu S, Yang L, Zhang C, Xiang Y, Liu Z, Hu S, et al. Online mental health services in China during the COVID-19 outbreak. Lancet Psychiatry 2020 Apr;7(4):17-18 [FREE Full text] [doi: 10.1016/s2215-0366(20)30077-8]
2. Halladay J, Bélanger SA. Fast facts on COVID-19 and the Canadian Armed Forces. J Mil Veteran Fam Health 2020 Nov 01;6(S2):7-8 [FREE Full text] [doi: 10.3138/jmvfh-6.s2-0004]
3. Camp II RD. Leadership in the COVID-19 environment: coping with uncertainty to support PSP mental health. J Community Safe Well Being 2020 Jul 15;5(2):37-38. [doi: 10.35502/jcswb.140]
4. Ricciardelli R, Czarnuch S, Carleton R, Gacek J, Shewmake J. Canadian Public Safety Personnel and Occupational Stressors: how PSP interpret stressors on duty. Int J Environ Res Public Health 2020 Jul 01;17(13):4736 [FREE Full text] [doi: 10.3390/ijerph17134736] [Medline: 32630259]
5. Luxton D, Pruitt L, O'Brien K, Kramer G. An evaluation of the feasibility and safety of a home-based telemental health treatment for posttraumatic stress in the U.S. Military. Telemed J E Health 2015 Nov;21(11):880-886 [FREE Full text] [doi: 10.1089/tmj.2014.0235] [Medline: 26103565]
6. Ashwick R, Turgoose D, Murphy D. Exploring the acceptability of delivering Cognitive Processing Therapy (CPT) to UK veterans with PTSD over Skype: a qualitative study. Eur J Psychotraumatol 2019 Feb 04;10(1):1573128 [FREE Full text] [doi: 10.1080/20008198.2019.1573128] [Medline: 30774784]

XSL•FO
RenderX

7. Acierno R, Gros D, Ruggiero K, Hernandez-Tejada BM, Knapp RG, Lejuez CW, et al. Behavioral activation and therapeutic exposure for posttraumatic stress disorder: a noninferiority trial of treatment delivered in person versus home-based telehealth. Depress Anxiety 2016 May;33(5):415-423 [FREE Full text] [doi: 10.1002/da.22476] [Medline: 26864655]
8. Wade VA, Eliott JA, Hiller JE. Clinician acceptance is the key factor for sustainable telehealth services. Qual Health Res 2014 May;24(5):682-694. [doi: 10.1177/1049732314528809] [Medline: 24685708]
9. McClellan M, Florell D, Palmer J, Kidder C. Clinician telehealth attitudes in a rural community mental health center setting. J Rur Ment Health 2020 Jan;44(1):62-73 [FREE Full text] [doi: 10.1037/rmh0000127]
10. Jang-Jaccard J, Nepal S, Alem L, Li J. Barriers for delivering telehealth in rural australia: a review based on Australian trials and studies. Telemed J E Health 2014 May;20(5):496-504. [doi: 10.1089/tmj.2013.0189] [Medline: 24801522]
11. Barney A, Buckelew S, Mesheriakova V, Raymond-Flesch M. The COVID-19 pandemic and rapid implementation of adolescent and young adult telemedicine: challenges and opportunities for innovation. J Adolesc Health 2020 Aug;67(2):164-171 [FREE Full text] [doi: 10.1016/j.jadohealth.2020.05.006] [Medline: 32410810]
12. Roncero C, García-Ullán L, de la Iglesia-Larrad JI, Martín CH, Andrés P, Ojeda A, et al. The response of the mental health network of the Salamanca area to the COVID-19 pandemic: the role of the telemedicine. Psychiatry Res 2020 Sep;291:113252 [FREE Full text] [doi: 10.1016/j.psychres.2020.113252] [Medline: 32623263]
13. Taylor C, Fitzsimmons-Craft E, Graham A. Digital technology can revolutionize mental health services delivery: the COVID-19 crisis as a catalyst for change. Int J Eat Disord 2020 Jul;53(7):1155-1157 [FREE Full text] [doi: 10.1002/eat.23300] [Medline: 32449523]
14. Waller G, Pugh M, Mulkens S, Moore E, Mountford VA, Carter J, et al. Cognitive-behavioral therapy in the time of coronavirus: clinician tips for working with eating disorders via telehealth when face-to-face meetings are not possible. Int J Eat Disord 2020 Jul;53(7):1132-1141 [FREE Full text] [doi: 10.1002/eat.23289] [Medline: 32383530]
15. COVID-19 practice guidance: Telehealth in occupational therapy practice. College of Occupational Therapists of British Columbia. 2020. URL: https://cotbc.org/wp-content/uploads/COVID-19-Practice-Guidance-Telehealth-in-Occupational-Therapy-Practice-Dec-2020-Update.pdf [accessed 2021-02-27]
16. COVID-19 : Information for social workers. Alberta College of Social Workers. URL: https://acsw.in1touch.org/site/COVID-19?nav=sidebar [accessed 2021-02-27]
17. Advice to the Profession: COVID-19: Virtual care. College of Physicians and Surgeons of Alberta. 2020. URL: http://www.cpsa.ca/wp-content/uploads/2020/03/AP_COVID-19-Virtual-Care.pdf [accessed 2021-02-27]
18. Joint Task Force for the Development of Telepsychology Guidelines for Psychologists. Guidelines for the practice of telepsychology. Am Psychol 2013 Dec;68(9):791-800. [doi: 10.1037/a0035001] [Medline: 24341643]
19. Jones C, Miguel-Cruz A, Smith-MacDonald L, Cruikshank E, Baghoori D, Chohan AK, et al. Virtual trauma-focused therapy for military members, veterans, and public safety personnel with posttraumatic stress injury: systematic scoping review. JMIR Mhealth Uhealth 2020 Sep 21;8(9):e22079 [FREE Full text] [doi: 10.2196/22079] [Medline: 32955456]
20. Barkin S, Schlundt D, Smith P. Community-engaged research perspectives: then and now. Acad Pediatr 2013 Mar;13(2):93-97. [doi: 10.1016/j.acap.2012.12.006] [Medline: 23498079]
21. Esmail L, Moore E, Rein A. Evaluating patient and stakeholder engagement in research: moving from theory to practice. J Comp Eff Res 2015 Mar;4(2):133-145 [FREE Full text] [doi: 10.2217/cer.14.79] [Medline: 25825842]
22. Denzin N, Lincoln Y. Introduction: The discipline and practice of qualitative research. In: Denzin NK, Lincoln YS, editors. The SAGE Handbook of Qualitative Research. Thousand Oaks, CA: SAGE Publications Ltd; 2005:1-32.
23. Israel B, Eng E, Schulz A, Parker E, Satcher D. Methods for Community-Based Participatory Research for Health. San Francisco, CA: John Wiley & Sons; 2012:1-524.
24. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: Building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
25. Pope C, Mays N, Popay J. Synthesising Qualitative And Quantitative Health Evidence: A Guide To Methods. Berkshire, England: McGraw Hill Education; 2007:A.
26. Liu L, Miguel Cruz A, Rios Rincon A, Buttar V, Ranson Q, Goertzen D. What factors determine therapists' acceptance of new technologies for rehabilitation – a study using the Unified Theory of Acceptance and Use of Technology (UTAUT). Disabil Rehabil 2015;37(5):447-455 [FREE Full text] [doi: 10.3109/09638288.2014.923529] [Medline: 24901351]
27. Ben-Zeev D. The digital mental health genie is out of the bottle. Psychiatr Serv 2020 Dec 01;71(12):1212-1213 [FREE Full text] [doi: 10.1176/appi.ps.202000306] [Medline: 32576123]
28. McBeath A, du Plock S, Bager-Charleson S. The challenges and experiences of psychotherapists working remotely during the coronavirus pandemic. Couns Psychother Res 2020 Jun 23:394-405 [FREE Full text] [doi: 10.1002/capr.12326] [Medline: 32837329]
29. Key policies and directives for procurement. Government of Canada. 2017. URL: https://buyandsell.gc.ca/policy-and-guidelines/Policy-and-Legal-Framework/Key-Policies-and-Directives#10 [accessed 2021-02-27]
30. Bajowala S, Milosch J, Bansal C. Telemedicine pays: billing and coding update. Curr Allergy Asthma Rep 2020 Jul 27;20(10):60 [FREE Full text] [doi: 10.1007/s11882-020-00956-y] [Medline: 32715353]

XSL•FO
RenderX

31. Torous J, Hsin H. Empowering the digital therapeutic relationship: virtual clinics for digital health interventions. NPJ Digit Med 2018:16 [FREE Full text] [doi: 10.1038/s41746-018-0028-2]
32. Aafjes-van Doorn K, Békés V, Prout T. Grappling with our therapeutic relationship and professional self-doubt during COVID-19: will we use video therapy again? Counsel Psychol Q 2020 Jun 03:1-12 [FREE Full text] [doi: 10.1080/09515070.2020.1773404]
33. Geller S. Cultivating online therapeutic presence: strengthening therapeutic relationships in teletherapy sessions. Counsel Psychol Q 2020 Jul 26:1-17 [FREE Full text] [doi: 10.1080/09515070.2020.1787348]
34. Wiederhold B. Connecting through technology during the coronavirus disease 2019 pandemic: avoiding "Zoom Fatigue". Cyberpsychol Behav Soc Netw 2020 Jul;23(7):437-438 [FREE Full text] [doi: 10.1089/cyber.2020.29188.bkw] [Medline: 32551981]
35. Soble J, Critchfield E, O'Rourke JJ. Neuropsychological evaluation in traumatic brain injury. Phys Med Rehabil Clin N Am 2017 May;28(2):339-350 [FREE Full text] [doi: 10.1016/j.pmr.2016.12.009] [Medline: 28390517]
36. Doneva S. Mild traumatic brain injury in military service personnel: key issues and considerations. J Mil Veteran Fam Health 2018 Oct;4(2):121-135 [FREE Full text] [doi: 10.3138/jmvfh.2017-0012]
37. Santesteban-Echarri O, Piskulic D, Nyman R, Addington J. Telehealth interventions for schizophrenia-spectrum disorders and clinical high-risk for psychosis individuals: a scoping review. J Telemed Telecare 2020;26(1-2):14-20. [doi: 10.1177/1357633X18794100] [Medline: 30134781]
38. Naslund J, Marsch L, McHugo G, Bartels S. Emerging mHealth and eHealth interventions for serious mental illness: a review of the literature. J Ment Health 2015;24(5):321-332 [FREE Full text] [doi: 10.3109/09638237.2015.1019054] [Medline: 26017625]
39. Ammenwerth E, Iller C, Mahler C. IT-adoption and the interaction of task, technology and individuals: a fit framework and a case study. BMC Med Inform Decis Mak 2006 Jan 09;6:3 [FREE Full text] [doi: 10.1186/1472-6947-6-3] [Medline: 16401336]

Abbreviations

CPT: cognitive processing therapy
MM: military member
OSI: operational stress injury
PSP: public safety personnel
REDCap: Research Electronic Data Capture

Edited by G Eysenbach; submitted 02.03.21; peer-reviewed by S Chokshi, M Galea; comments to author 28.06.21; revised version received 06.07.21; accepted 06.07.21; published 12.08.21.

Please cite as:
Smith-MacDonald L, Jones C, Sevigny P, White A, Laidlaw A, Voth M, Mikolas C, Heber A, Greenshaw AJ, Brémault-Phillips S
The Experience of Key Stakeholders During the Implementation and Use of Trauma Therapy via Digital Health for Military, Veteran, and Public Safety Personnel: Qualitative Thematic Analysis
JMIR Form Res 2021;5(8):e26369
URL: https://formative.jmir.org/2021/8/e26369
doi:10.2196/26369
PMID:34387549

©Lorraine Smith-MacDonald, Chelsea Jones, Phillip Sevigny, Allison White, Alexa Laidlaw, Melissa Voth, Cynthia Mikolas, Alexandra Heber, Andrew J Greenshaw, Suzette Brémault-Phillips. Originally published in JMIR Formative Research (https://formative.jmir.org), 12.08.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Digital Patient-Reported Outcome Measures for Monitoring of Patients on Cancer Treatment: Cross-sectional Questionnaire Study

Mayuran Ananth Sivanandan[1], MBBS, MA; Catherine Sharma[1], MA; Pippa Bullard[1], BN; Judith Christian[1], MD

Department of Oncology and Radiotherapy, Nottingham University Hospitals NHS Trust, Nottingham, United Kingdom

Corresponding Author:
Mayuran Ananth Sivanandan, MBBS, MA
Department of Oncology and Radiotherapy, Nottingham University Hospitals NHS Trust
Nottingham City Hospital
Hucknall Road
Nottingham, NG5 1PB
United Kingdom
Phone: 44 07859877635
Email: ananthsiva@doctors.org.uk

Abstract

Background: Oncology has been facing increasing outpatient activity associated with higher cancer incidence, better survival rates, and more treatment options. Innovative technological solutions could help deal with this increasing demand. Using digital patient-reported outcome measures (PROMs) to identify patients who need a face-to-face (FTF) appointment is a potential approach.

Objective: This study aims to assess the feasibility of digital PROM questionnaires to enable remote symptom monitoring for patients undergoing cancer treatment and their ability to highlight the requirement for an FTF appointment.

Methods: This study was performed at a tertiary oncology center between December 2018 and February 2019. The Common Terminology Criteria for Adverse Events were adapted into patient-friendly language to form the basis of treatment-specific digital questionnaires covering specific cancer drugs and radiotherapy treatments. These treatment-specific digital PROM questionnaires were scored by both patients and their clinicians during FTF appointments. Patients and clinicians did not see each other's scored PROMs. Agreement between patients and clinicians was assessed using descriptive statistics. Patient and staff feedback was also obtained.

Results: In total, 90 patients participated in the study across 10 different treatment pathways. By comparing paired patient and clinician responses, the sensitivity of the patient-completed questionnaires in correctly highlighting the need for FTF review was 94% (44/47), and all patients with severe or grade 3+ symptoms were identified (6/6, 100%). Patient-completed PROMs appropriately revealed that 29% (26/90) of the participating patients did not need FTF review based on their symptoms alone. Certain oncological treatment pathways, such as immunotherapy, were found to have a larger proportion of patients with minimal symptoms than others, such as conventional chemotherapy. Patient and staff feedback showed high approval of digital PROMs and their potential for use in remote monitoring.

Conclusions: Digital PROM questionnaires can feasibly highlight the need for FTF review in oncology clinics for treatment. Their use with specific treatments could safely reduce the requirement for FTF care, and future work should evaluate their application in the remote monitoring of patients.

(JMIR Form Res 2021;5(8):e18502) doi:10.2196/18502

KEYWORDS

patient-reported outcome measures; patient-reported outcomes; remote monitoring; toxicity; outpatients; digital technology; digital health; mobile health; oncology; chemotherapy; immunotherapy; radiotherapy

XSL•FO
RenderX

Introduction

Background

Oncology is a predominantly outpatient specialty; hence, the increases in outpatient activity are of particular relevance. There has been an increase in National Health Service (NHS) outpatient appointments in England from 63.2 million to 118.6 million in the 10-year period ending between 2016 and 2017 [1] and projected significant increases in the demand for oncology services in both the United States [2] and Europe [3]. Growing service pressures on oncology outpatient activities are specifically driven by increased cancer incidence [4], improved survival rates [5], and an expanded treatment repertoire [6]. Current pressures on outpatient services have been stated to negatively affect patient and clinician experience [7]. Furthermore, global workforce shortages are increasing and are predicted to increase further [2,3,8]. Therefore, the outpatient system will struggle to continue to offer the capacity to deal with the increasing demand in its current traditional form.

Consequently, improving the efficiency of oncology care is paramount [2]; for example, the UK NHS' Long Term Plan advocates a fundamental remodeling of outpatients working with technology to help drive a reduction in face-to-face (FTF) outpatient appointments of up to a third in the coming 5 years [9]. This is particularly relevant given the impact of the COVID-19 pandemic. An application of technology that will help achieve this ambitious target is to allow alternative consultation methods outside a traditional FTF encounter to review patients. An example is remote monitoring, where technology can allow patients' health to be checked at a distance by clinical staff, such as through the completion of symptom-related questionnaires incorporating patient-reported outcome measures (PROMs) [7,10].

The clinical benefits of PROMs being used as a part of the care of patients with cancer have been shown to include increased awareness of symptoms by patients and clinicians, streamlining of consultations, improved interprofessional communication [11], and improved health care outcomes for patients, including quality of life and survival [12]. Furthermore, their use is associated with patient-centered care and improved patient self-efficacy [13]. The use of PROMs and digital technology has been advocated in cancer strategy reports by the NHS [14] and the Independent Cancer Taskforce [15].

The strategy of using PROMs in monitoring patients remotely has been applied successfully in gastroenterology in patients with inflammatory bowel disease on immunosuppressive treatment [16]. A similar strategy would be equally attractive in oncology, where a large proportion of follow-up activities involve regular attendance to monitor patients on treatments [17], including both radiotherapy (RT) and systemic treatments. In the research setting, the use of electronic PROMs to allow regular reporting of chemotherapy side effects by patients on cancer treatment has been evaluated in the context of randomized controlled trials. These studies have indicated several improved patient outcomes, such as improved quality of life and reduced hospitalization rates through improved symptom management [12,18-20]. However, the data for actually replacing routine FTF outpatient follow-up of patients on oncological treatment with remote monitoring with PROMs in the standard setting are sparse [21].

For patients on cancer treatments, the National Cancer Institute Common Terminology Criteria for Adverse Events (CTCAE) [22] is the standard tool used by clinicians to grade and record treatment-related adverse events (Textbox 1), and this is typically done in *on treatment* outpatient clinics. Many adverse events are based on a patient's subjective experience, and this has led to individual groups rephrasing CTCAE, which are designed for clinicians, into a patient-understandable language to generate a PROM that directly captures the patient perspective and maintains the clinical usefulness of CTCAE [23-25]. The National Cancer Institute has developed its own PROM based on CTCAE (patient-reported outcomes Common Terminology Criteria for Adverse Events [PRO-CTCAE]) for use in patients in cancer clinical trials [26]; however, it does not currently map onto the severity grades of CTCAE that are used for clinical decision-making.

Textbox 1. Common Terminology Criteria for Adverse Events grading for adverse events.

Common Terminology Criteria for Adverse Events (CTCAE) Grade and Description Adapted From National Cancer Institute (2017): CTCAE (Version 5.0) [22]

- Grade 0: no symptom
- Grade 1: mild symptom not needing intervention
- Grade 2: moderate symptom where intervention is indicated
- Grade 3: severe symptom that requires hospitalization

Additional benefits of using a PROM in gauging oncological treatment–related adverse events are that evidence suggests that clinicians can underreport their severity [27] and that the recording of toxicity by clinicians in routine practice can be suboptimal [28]. Trials that have evaluated the utility of patient-modified CTCAE as a PROM have examined its use in addition to existing FTF hospital appointments [12,25,29], and it has not been assessed as a tool to help determine whether an FTF appointment is actually needed.

Our oncology department represents the largest oncological facility in the East Midlands [30] in one of the largest hospital trusts in England [31]. Locally, the department's outpatient activity has consistently increased on an annual basis, with a growth of approximately 2500 appointments per year on average over the last 6 years. A major driver for this increase was found

XSL•FO
RenderX

to be *on treatment* appointments. Therefore, strategies to reduce footfall within the oncology outpatient department would be beneficial.

Objective

In this context, it was hypothesized that remotely completed questionnaires based on patient-modified CTCAE could serve as a triage tool to ascertain the need for a patient to attend an FTF appointment. It was believed that PRO-CTCAE would not be appropriate in this setting as it does not map onto the CTCAE severity grades that are clinically relevant to help determine the need for an FTF appointment; moreover, PRO-CTCAE is advocated to be used only for symptoms that occurred in the previous 7 days [26]. We, therefore, performed a feasibility study at our center to assess if digital PROM questionnaires based on patient-modified CTCAE could be used for symptom monitoring in oncology *on treatment* clinics and compared paired patient and clinician questionnaires to identify whether these questionnaires could accurately highlight the requirement for an FTF appointment.

Methods

Overview

A cross-sectional study was undertaken at Nottingham University Hospitals NHS Trust (United Kingdom) to evaluate the use of digital PROM questionnaires between December 2018 and February 2019. A multi-professional team led this project, and input was sought from clinicians, information technology staff, quality improvement specialists, and patient representatives. The technology partner for this project was DrDoctor (London), who provided an electronic portal to allow the completion of questionnaires.

Patient Groups

Specific oncological treatment pathways for this feasibility study were chosen to cover the breadth of both radical RT and systemic drug pathways. For the systemic drug pathways, the study included several patient groups who were considered less likely to have significant side effects on treatment, such as single-agent immunotherapy patients and patients on oral targeted drugs. A similar theme was chosen for the RT group; therefore, patients with adjuvant breast and radical prostate RT were targeted. Nevertheless, it was also decided to test in some groups, such as metastatic prostate cancer patients on chemotherapy and patients with radical RT for head and neck cancer, where the opportunities for FTF reduction in care might be less obvious.

PROM Development

It was decided that treatment-specific PROMs would be designed to assess treatment-related symptoms and side effects. The symptoms that needed assessment, and therefore, inclusion in each treatment-specific questionnaire, were decided by a project-team clinician through review of the appropriate treatment-specific literature (eg, summary of product characteristics) and trusted UK cancer information websites [32]. Subsequently, appropriate questions were developed by adapting relevant items of the CTCAE [22] and World Health Organization Performance Status (PS) for relevant questionnaires pertaining to systemic treatment, into a patient-friendly language in a similar approach to previous groups [25,27,29]. Responses to CTCAE items were based on grades on an ordinal scale of 0 (not present), 1 (mild), 2 (moderate), and 3 (severe) and PS on a scale from 0 to 4.

It was recognized that certain symptoms, such as fever or palpitations, were more appropriate for a binary question (yes-no) alone rather than an ordinal-scale question, and this approach was used where required. It was also deemed that the questionnaires should determine the presence of emotional concerns in patients. There was no appropriate CTCAE item to capture this; therefore, a binary question about emotional concerns was created by clinicians and added to all questionnaires. A collaborative approach with site-specific oncological teams was implemented with a review of relevant treatment questionnaires before use. They made comments and suggested amendments that were enacted before the questionnaires were used by patients in this study.

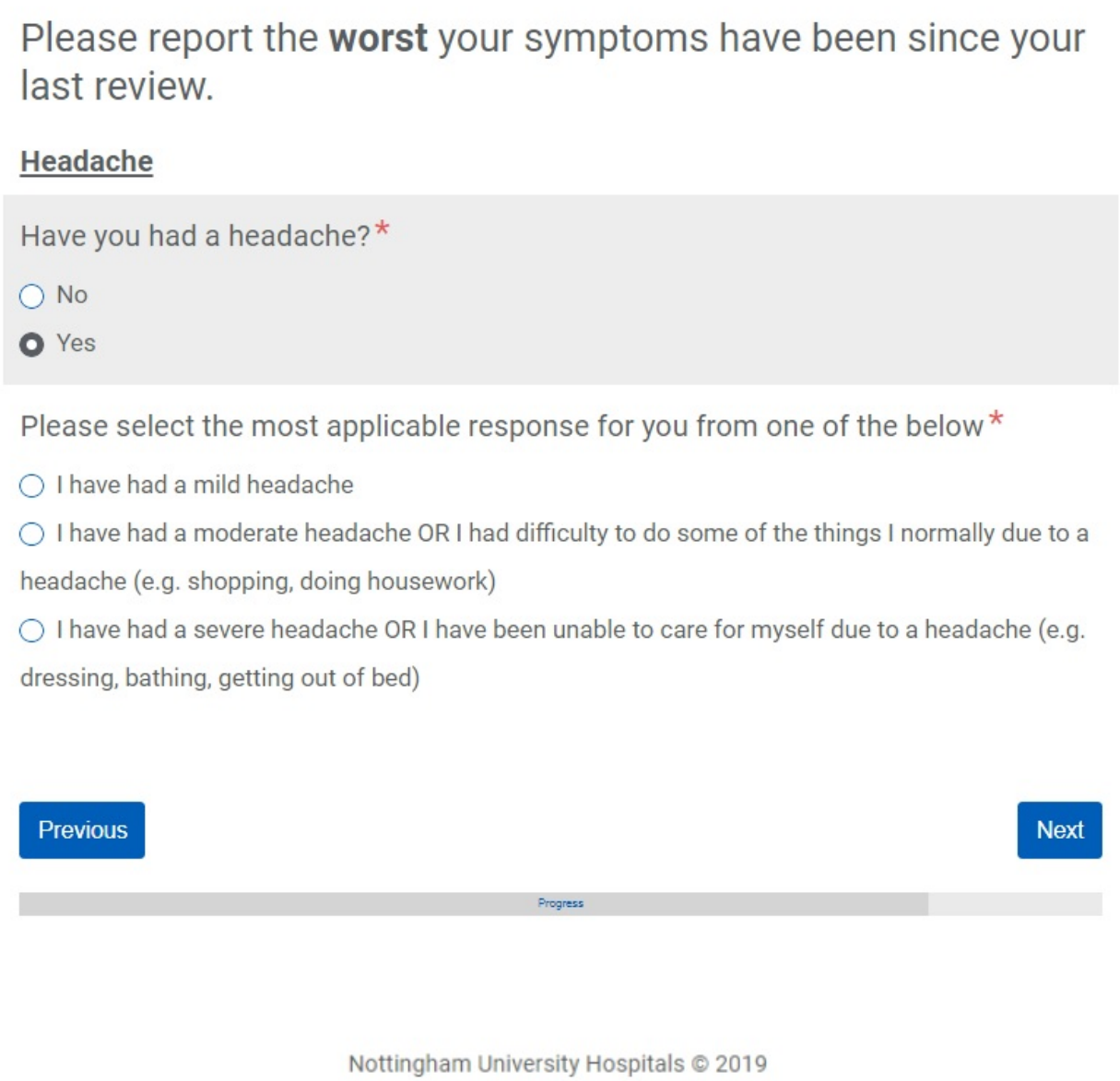
Digital Interface Development

The questionnaires were converted into a digital format by a member of the information technology team using the internet-based Formstack system (Formstack LLC) and subsequently uploaded to the DrDoctor portal, which is a cloud-based platform. The design of the electronic questionnaires was based on the work of previous research groups whose electronic questionnaire design was found to be acceptable to patients [18,29]. Apart from the PS question, symptom occurrence had to be indicated by answering a yes-no question, and if yes was selected, the corresponding graded responses would appear for a patient to mark as appropriate. The authors felt this would minimize the amount of reading for patients and thus the burden on their time. Each question had to be answered before moving to the next to ensure that all questions were completed. Questionnaires were designed to be simple to reduce break-off rates [33], and a progress bar was placed at the bottom of each page to increase the likelihood of completing the survey [34]. An example of a question from a digital PROM questionnaire is shown in Figure 1.

The DrDoctor portal is password-protected, and a member of the study team allocated the appropriate treatment-specific questionnaires for patients and clinicians to complete during the study. Completed questionnaires contained no patient-identifiable data and were assigned a letter to allow corresponding patient and clinician questionnaires to be analyzed for concordance.

XSL•FO
RenderX

Figure 1. An example of a question from a digital patient-reported outcome measure questionnaire.



Study Design

Patients eligible for this study were recommended by their clinical teams; they had to be aged at least 18 years, able to understand written English, and have specific cancer diagnoses currently receiving specific treatments (Table 1). Patients had to provide verbal informed consent, and patients unable to complete the questionnaires were excluded from the study. Eligible patients were approached by a member of the study team to complete a treatment-specific digital PROM questionnaire in the oncology outpatient department and RT review clinics at Nottingham University Hospitals NHS Trust. Patients completed the digital PROM questionnaire before their FTF appointment unless the time pressures of their FTF appointment required completion after their FTF appointment. This was deemed acceptable, as a previous study showed no significant difference if patients completed their questionnaires before or after seeing their clinician [27]. The patients completed the questionnaire electronically on a tablet device in a private room in the outpatient department with a member of the study team. Clinicians were asked to complete a corresponding digital PROM questionnaire following a participating patient's FTF appointment. CTCAE was common knowledge to all clinicians before this study, but comprehensive knowledge of the precise CTCAE symptom grades was not required as questionnaire responses were designed to equate to the appropriate CTCAE grading. The process of asking both patients and clinicians to score symptoms blind of each other was a new process needed for this study.

XSL•FO
RenderX

Table 1. Patient characteristics (N=90).

Treatment pathway	Tumor site	Treatment intent	Patients, n (%)
Capecitabine chemotherapy with or without oxaliplatin chemotherapy	Colorectal	Adjuvant	11 (12)
Abiraterone and enzalutamide	Prostate	Palliative	10 (11)
Breast radiotherapy	Breast	Adjuvant	10 (11)
Head and neck radiotherapy	Head and neck	Radical	10 (11)
Single-agent immunotherapy (nivolumab or pembrolizumab)	Melanoma	Palliative and adjuvant	10 (11)
Prostate radiotherapy	Prostate	Radical	10 (11)
Pazopanib	Renal	Palliative	10 (11)
Imatinib	Gastrointestinal stromal tumor	Palliative and adjuvant	8 (9)
Stereotactic ablative radiotherapy	Lung	Radical	6 (7)
Docetaxel chemotherapy	Prostate	Palliative	5 (6)

The rationale for clinicians completing a corresponding electronic questionnaire was to enable the comparison of paired responses between patients and clinicians. The current standard outpatient pathway for the assessment of treatment-related side effects is dependent on a clinician's interpretation of a patient's symptoms; therefore, comparison of paired patient and clinician questionnaires would enable assessment of the feasibility and accuracy of a patient-completed PROM on its own to triage the need for further assessment. Patients and clinicians did not see each other's PROM results, and the results were not used for clinical decision-making. This was done so that the suitability of our designed questionnaires could be assessed before consideration for routine clinical use. The current standard of care for FTF appointments was maintained for all patients to ensure patient safety. Our method is similar to that of other groups who have performed similar interventions [25,27].

Participants were asked to complete a feedback form enquiring about the usability of the digital questionnaire, thoroughness of the questionnaire, and acceptance of future use on a 10-point Likert scale. Participating study clinicians were asked to complete a similar feedback form after the completion of the study. The authors wanted to assess not only the user experience with the digital interface but also the content of the questionnaires. Hence, the authors designed a bespoke feedback form to assess both because it was not possible to use a pre-existing tool, such as the System Usability Scale, which is solely focused on usability. It was decided that the feedback form would comprise 3 questions to maximize response rates. A 10-point Likert scale was chosen to enable sufficient distinction between positive and negative responses and generate quantitative data for analysis [35]. Examples of feedback form questions can be found in Multimedia Appendices 1 and 2.

This study was deemed to not present a risk to patient safety or patient data protection by the trust's chief clinical information officer. As this study formed part of a local service improvement project, no further formal ethics review was deemed necessary in keeping with appropriate guidelines [36].

Data Analysis

The criteria in a completed questionnaire that were deemed to indicate the need for an FTF review were defined as the presence of any of the following: any grade 2 or higher response to a CTCAE-based question, having any symptom assessed with a binary question, or a PS in the range of 3-4. Using these criteria, the concordance between paired patient and clinician questionnaires for containing an FTF indicator was analyzed; the specific FTF indicator did not need to match in the paired questionnaires. Concordance was assessed by cross-tabulating the presence of an FTF indicator in paired patient- and clinician-completed questionnaires (Table 2).

Table 2. Cross tabulation of patient- and clinician-completed patient-reported outcome measure questionnaires by the presence of a face-to-face indicator.

Presence of an FTF[a] indicator in patient-completed PROM[b]	Presence of an FTF indicator in clinician-completed PROM	
	Yes	No
Yes	True positive	False positive
No	False negative	True negative

[a]FTF: face-to-face.

[b]PROM: patient-reported outcome measure.

As the current standard of care comprises clinician interpretation of patient symptoms, the clinician-completed PROM represented the *standard*, and the patient-completed PROM represented the *test variable*. Sensitivity was calculated as true positive/(true positive+false negative) and specificity as true negative/(true negative+false positive). A similar method of cross-tabulation was performed to assess the presence of any severe binary

XSL•FO
RenderX

symptoms or CTCAE grade 3 or higher symptoms in paired patient and clinician questionnaires.

The concordance of grading of common individual symptoms between paired patient and clinician questionnaires was assessed. Concordance was analyzed using descriptive statistics without the use of the Cohen κ statistic, as it was deemed to be the most accurate technique considering the predicted asymmetrical scoring differences in the ordinal data in line with recommendations from similarly conducted studies [25,27].

The Likert scale data from patient and staff feedback surveys were analyzed using descriptive statistics.

Results

In total, 90 patients participated in the study across 10 different oncology treatment pathways, as shown in Table 1. The concordance between paired patient and clinician questionnaires for the presence of an indicator for FTF review is shown in Table 3.

Table 3. Concordance between paired patient-reported outcome measure questionnaires in highlighting the need for face-to-face review (N=90).

Presence of an FTF[a] indicator in patient-completed PROM[b] (n)	Presence of an FTF indicator in clinician-completed PROM (n)	
	Yes (n=47)	No (n=43)
Yes (61)	44	17
No (29)	3	26

[a]FTF: face-to-face.

[b]PROM: patient-reported outcome measure.

Thus, the sensitivity of the patient-completed questionnaires in correctly highlighting the need for FTF review was 94% (44/47) and specificity was 60% (26/43). False-negative patient-completed PROMs (ie, a patient questionnaire not indicating the need for FTF review but the clinician questionnaire indicating so) was 3% (3/90) of the total. Further analysis showed that these were all for symptoms that the clinician determined were of moderate severity (grade 2). Therefore, acknowledging these false negatives, 97% (87/90) of patient questionnaires flagged in a clinically appropriate manner.

All questionnaires were completed by participants in their entirety except for PS data being unavailable for 5 patients; 4 out of 5 of these patients had patient-completed questionnaires that already contained indicators for FTF review, with the remaining patient having corresponding patient- and clinician-completed questionnaires displaying no significant symptoms. Hence, the missing PS data were not considered likely to affect the above analysis.

Furthermore, 29% (26/90) of the paired questionnaires were concordant for the absence of any FTF indicators. This figure equates to the percentage of patients who were correctly identified by questionnaires not to need an FTF review and, therefore, the potential for FTF appointment reduction. Stratification by treatment pathway demonstrated that this percentage of potential FTF reduction by questionnaire varied considerably across the pathways from 0% (0/10) in patients receiving head and neck radical RT and 0% (0/5) in those receiving prostate chemotherapy to up to 70% (7/10) in those receiving single-agent immunotherapy, as shown in Table 4.

Table 4. Concordant questionnaires that contained no indicators for face-to-face review stratified by treatment pathway (N=90).

Treatment pathway (number of patients)	Concordant questionnaires per pathway without indicators for FTF[a] review, n (%)
Immunotherapy (n=10)	7 (70)
Lung SABR[b] (n=6)	3 (50)
Abiraterone and enzalutamide (n=10)	4 (40)
Pazopanib (n=10)	4 (40)
Imatinib (n=8)	3 (38)
Breast RT[c] (n=10)	3 (30)
Prostate RT (n=10)	1 (10)
Colorectal chemotherapy (n=11)	1 (9)
Head and neck RT (n=10)	0 (0)
Prostate chemotherapy (n=5)	0 (0)

[a]FTF: face-to-face.

[b]SABR: stereotactic ablative radiotherapy.

[c]RT: radiotherapy.

Regarding concordance between paired patient and clinician questionnaires for the presence of a severe or grade 3+ symptom or higher (Table 5), the sensitivity of patient questionnaires was 100% (6/6) and specificity was 87% (73/84).

XSL•FO
RenderX

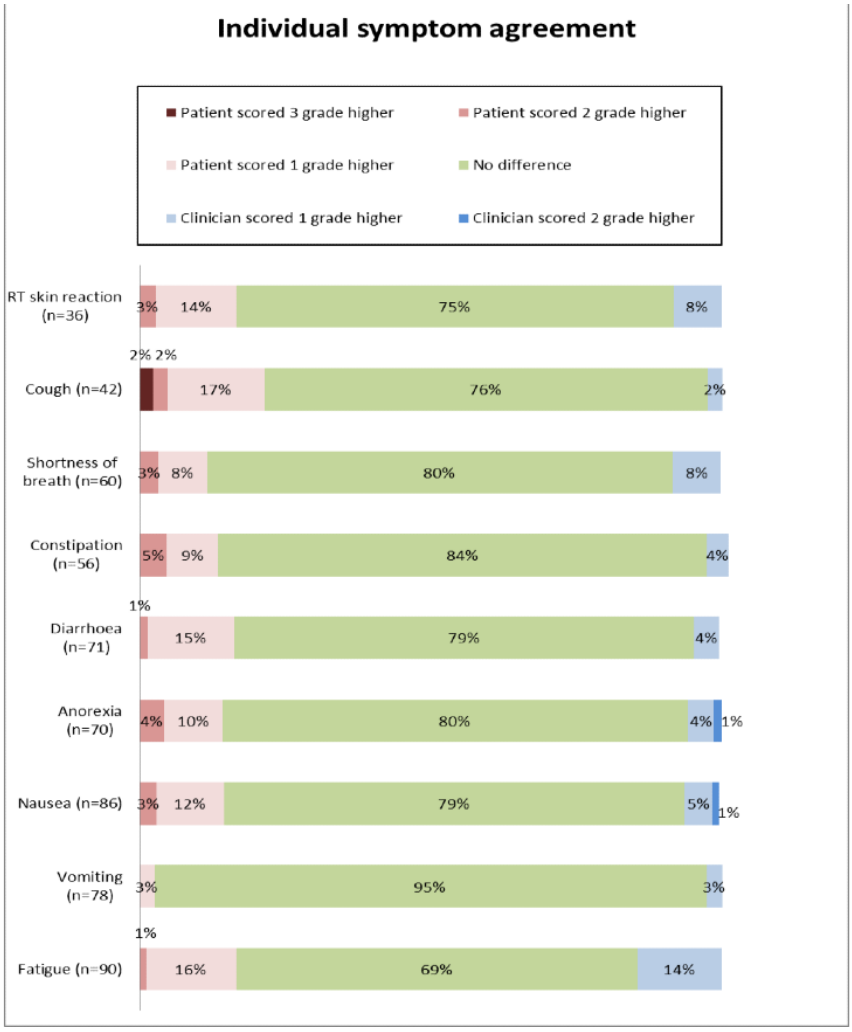
For frequently appearing symptoms in the different treatment-specific questionnaires (fatigue, vomiting, nausea, anorexia, diarrhea, constipation, shortness of breath, cough, and RT skin reaction), the exact agreement between patients and clinicians ranged from 69% (62/90) agreement for fatigue to 95% (74/78) for vomiting (Figure 2 and Multimedia Appendix 3). When there were individual symptom discrepancies between patients and clinicians, they were mostly within 1 grading point, and patients were more likely to assign greater severity to symptoms.

Table 5. The concordance between paired questionnaires for the presence of a severe symptom (N=90).

Presence of a severe or grade 3+ symptom in patient-completed PROM[a] (n)	Presence of a severe or grade 3+ symptom in clinician-completed PROM (n)	
	Yes (n=6)	No (n=84)
Yes (17)	6	11
No (73)	0	73

[a]PROM: patient-reported outcome measure.

Figure 2. Agreement of common individual symptoms between patients and clinicians. RT: radiotherapy.



Of the 90 patients, 77 (86%) completed the feedback form. On a 10-point Likert scale, the mean patient approval score was 9.3 for usability, 9.1 for questionnaire thoroughness, and 8.8 for acceptance of questionnaires to supplement outpatient care. There were missing data for 3 patients who did not provide an answer for the latter question.

In addition, 48% (10/21) of clinicians completed a feedback form. On a 10-point Likert scale, the mean clinician approval score was 9.4 for usability, 9.8 for questionnaire thoroughness, and 9.6 for future use of the questionnaires to supplement outpatient care.

Discussion

Principal Findings

By comparing patient and clinician questionnaires, this study has shown that acute toxicity questionnaires based on

XSL•FO
RenderX

patient-modified CTCAE can act as a triage tool to help highlight the need for FTF review in oncology *treatment* clinics. Patient questionnaires successfully detected all patients with severe symptoms. Our results indicate that the use of patient questionnaires to enable remote monitoring in certain treatment pathways could significantly reduce the need for FTF outpatient reviews. Patients and staff provided positive feedback on questionnaire usability and content and accepted its use to assist symptom monitoring. Our study thus contributes to the existing literature regarding the use of PROMs in routine outpatient settings [37-39], particularly the way that PROMs can usefully aid clinical decision-making and guide the need for FTF review in *on treatment* oncology clinics.

For common individual graded symptoms, the agreement between patients and clinicians was good; when there were differences, they were usually small, with patients more likely to indicate greater severity than clinicians, comparable with previous studies [25,27]. These individual differences rarely affected those patients who needed FTF review, with our results showing high sensitivity of our questionnaires, incorporating the presence of our predefined FTF indicators, to detect patients who needed FTF review and patients with potentially severe symptoms. This suggests there would be a low risk that patients who would potentially need clinical intervention would be missed. The tendency for some patients to rate symptoms more severely than clinicians explains the lower specificity of the questionnaires to determine the need for FTF review.

This study confirms the potential benefit of PROM questionnaires in acting as a triage tool for determining FTF review. Our results indicate that a significant proportion of participating patients (26/90, 29%) were correctly determined not to need an FTF appointment from their questionnaire results alone. There was a further proportion of patients, comprising 19% (17/90) of the cohort, in which the patient-completed questionnaires indicated a need for FTF review, but the corresponding clinician questionnaires suggested that this was not needed. This suggests that subsequent review of patients through a telephone or video consultation could be beneficial as a method to increase the specificity of patient questionnaires.

The study has also highlighted that the use of PROM questionnaires for the purpose of FTF reduction could be especially advantageous in certain follow-up treatment pathways. Of note, a large proportion of the pathways that seem particularly suitable for remote monitoring based on our results are the newer oncological systemic treatments, such as immunotherapy and tyrosine kinase inhibitor treatments. Patients can be on these treatments for many months and potentially years unlike traditional chemotherapy drugs where the course of treatment is usually a few months. Therefore, the benefits of appropriate FTF reduction to various stakeholders would be particularly discernible for patients with reduced hospital visits, leading to decreased burden on their time and finances, for clinicians with more productive use of their time, and for managers with more effective clinic use [9]. Moreover, PROMs have been shown to have broader clinical benefits for patients and clinicians [11], suggesting that more widespread use of digital PROMs would have additional health care benefits outside the primary scope of our study.

Technological solutions are being espoused to help with outpatient working [7,9], and our study demonstrates both patient and staff acceptance of our particular digital strategy. This helps justify that such an approach would work if it were to be implemented into routine oncological practice with both strong patient and staff willingness to drive its success. Digital PROMs are only one of the many technological tools that can help make oncology work more efficiently. Video consultations to enable remote review of patients have been shown to be safe and effective when used appropriately [40], and their use has expanded rapidly in response to the COVID-19 pandemic [41]. Other technological solutions that seek to improve the efficiency of a number of aspects of oncology work include artificial intelligence applied to radiomics, such as breast screening interpretation [42], and streamlining RT workflows, such as through auto-contouring during RT outlining and voxel-based dose prediction approaches to refine the treatment planning process [43]. Thus, digital technology, including electronic PROMs, looks set to have a significant impact on oncology practice.

Limitations

The questionnaires in this study were largely based on CTCAE, which has the limitation of not being formally validated [27]. However, they form the standard for adverse event reporting in oncology [22], and in line with previous studies [27,29], modification of terminology into patient-understandable language enables patient reporting of symptom severity while mapping onto an established grading system that is well known and widely used by clinicians. Furthermore, the UK Oncology Nursing Society triage tool was used for the emergency assessment of chemotherapy toxicity in our study center [44] and the UK Oncology Nursing Society tool is based on CTCAE criteria; thus, this was felt to additionally aid acceptance of the digital PROMs used by clinicians in this study. Currently, the National Cancer Institute's patient-reported outcome tool PRO-CTCAE does not map exactly onto the recognized CTCAE grading system; therefore, it would be difficult to use it as a remote monitoring tool to determine the need for FTF assessment. Furthermore, CTCAE has been applied to other specialties outside of oncology, such as in trials pertaining to hypertension and HIV [45], making it generalizable to other medical specialties.

In this study, clinicians were asked to complete the PROM questionnaires based on information gathered from routine FTF appointments. The questions asked in these FTF consultations were up to the clinician's discretion as per their routine practice. Therefore, there is the possibility that clinicians may have completed PROM questionnaires with insufficient information. This limitation reflects the standard clinical practice for *on treatment* reviews, which would have a less *systematic* approach than a PROM questionnaire.

Potential FTF reduction using digital questionnaires was estimated through the absence of predefined indicators for FTF review. The authors recognize that patients may want to see their medical team in an FTF appointment for reasons other than these indicators. Actual FTF reduction will, therefore, likely be lower in practice, but these figures demonstrate the

XSL•FO
RenderX

large opportunity for follow-up reduction if an appointment is deemed unnecessary from both the patient and clinician perspectives. If such questionnaires are to be applied to routine care, questionnaires should be designed to allow patients to explicitly state their request for an FTF review to enable a patient-centered approach to care.

It can be stated that our study is limited by the fact that statistical analyses, such as Cohen κ statistics, were not used to formally assess agreement between patients and clinicians. However, previous studies have criticized Cohen κ statistics in this setting because of the asymmetry across the scoring differences and have advocated descriptive statistics, as used in this study, as being sufficient for determining interrater concordance in this particular situation.

Although the patient feedback form response rate was high at 86% (77/90), it was recognized that the staff feedback form response rate was significantly lower at 48% (10/21). This is partly explained as patients were asked to complete this directly after their FTF appointment following an in-person request from a member of the project team, whereas clinicians were asked to do this via email after the study was completed. Hence, nonresponse bias may affect the strength of the conclusions that can be drawn from the staff feedback data.

Future Work

Moving forward, we have organized patient focus groups to provide detailed qualitative feedback about patient understanding and acceptance of the designed questionnaires. These will occur before a planned pilot study to use *remotely completed* digital PROMs in selected oncology treatment pathways to assess their ability to reduce the need for FTF care. We are also considering remote monitoring for patients who have completed their cancer treatment via PROMs as part of their long-term follow-up. These PROM questionnaires would target symptoms suggestive of recurrence as well as the consequences of their cancer treatment.

Conclusions

This study demonstrates the potential efficacy and utility of PROM questionnaires to facilitate remote monitoring of patients undergoing oncology treatments to reduce the need for FTF care. They have a high approval rating from both patients and clinicians. Significantly, they appeared to correctly identify patients with severe adverse treatment effects. From our data, a treatment strategy using digital PROMs in our oncology center alone, which has approximately 30,000 follow-up attendances per year, could safely reduce the need for thousands of FTF appointments. The use of remote monitoring via PROMs could lead to a more patient-centered model of care with a reduced need for hospital visits with resultant benefits to patients, clinicians, and the wider health system.

Acknowledgments

This work was supported by the East Midlands Academic Health Science Network and Nottingham Hospitals Charity. They played no role in the study design, data collection, analysis, or writing of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Patient feedback questions.
[DOCX File , 13 KB - formative_v5i8e18502_app1.docx]

Multimedia Appendix 2
Clinician feedback questions.
[DOCX File , 13 KB - formative_v5i8e18502_app2.docx]

Multimedia Appendix 3
Raw data set for individual common symptom agreement between patients and clinicians.
[XLSX File (Microsoft Excel File), 12 KB - formative_v5i8e18502_app3.xlsx]

References

1. Hospital Outpatient Activity 2016-17. United Kingdom: NHS Digital; 2017.
2. Patlak M, Levit L. Ensuring Quality Cancer Care Through the Oncology Workforce: Sustaining Care in the 21st Century: Workshop Summary Internet. Washington, DC: The National Academies Press; 2009.
3. Seruga B, Sullivan YR, Fundytus A, Hopman XW, Ocana A, Joffe YJ, et al. Medical oncology workload in Europe: one continent, several worlds. Clin Oncol (R Coll Radiol) 2020 Jan;32(1):e19-e26. [doi: 10.1016/j.clon.2019.06.017] [Medline: 31324474]

XSL•FO
RenderX

4. Cancer Incidence Statistics. Cancer Research UK. URL: https://www.cancerresearchuk.org/health-professional/cancer-statistics/incidence#heading-Zero [accessed 2019-03-09]
5. Cancer Survival Statistics. Cancer Research UK. URL: https://www.cancerresearchuk.org/health-professional/cancer-statistics/survival [accessed 2019-03-09]
6. Palumbo M, Kavan P, Miller W, Panasci L, Assouline S, Johnson N, et al. Systemic cancer therapy: achievements and challenges that lie ahead. Front Pharmacol 2013;4:57 [FREE Full text] [doi: 10.3389/fphar.2013.00057] [Medline: 23675348]
7. Outpatients: the Future – Adding Value Through Sustainability. London, UK: Royal College of Physicians; 2018.
8. Clinical Oncology UK Workforce Census Report 2018. The Royal College of Radiologists. URL: https://www.rcr.ac.uk/publication/clinical-oncology-uk-workforce-census-2018-report [accessed 2019-03-09]
9. NHS. The NHS long term plan. 2019. URL: https://www.longtermplan.nhs.uk/ [accessed 2019-03-09]
10. Teagle A, Gilbert D. Remote follow-up strategies after cancer treatment: a lot of opportunities. Clin Oncol (R Coll Radiol) 2014 Oct;26(10):622-624. [doi: 10.1016/j.clon.2014.05.009] [Medline: 24969682]
11. Yang LY, Manhas DS, Howard AF, Olson RA. Patient-reported outcome use in oncology: a systematic review of the impact on patient-clinician communication. Support Care Cancer 2018 Jan;26(1):41-60. [doi: 10.1007/s00520-017-3865-7] [Medline: 28849277]
12. Basch E, Deal AM, Kris MG, Scher HI, Hudis CA, Sabbatini P, et al. Symptom monitoring with patient-reported outcomes during routine cancer treatment: a randomized controlled trial. J Clin Oncol 2016 Feb 20;34(6):557-565 [FREE Full text] [doi: 10.1200/JCO.2015.63.0830] [Medline: 26644527]
13. Basch E, Stover AM, Schrag D, Chung A, Jansen J, Henson S, et al. Clinical utility and user perceptions of a digital system for electronic patient-reported symptom monitoring during routine cancer care: findings from the pro-tect trial. JCO Clin Cancer Inform 2020 Oct;4:947-957 [FREE Full text] [doi: 10.1200/CCI.20.00081] [Medline: 33112661]
14. Achieving World-class Cancer Outcomes: Taking the Strategy Forward. NHS. 2016. URL: https://www.england.nhs.uk/wp-content/uploads/2016/05/cancer-strategy.pdf [accessed 2019-03-09]
15. Achieving World-Class Cancer Outcomes: A Strategy for England 2015-2020. Cancer Research UK. 2015. URL: https://www.cancerresearchuk.org/sites/default/files/achieving_world-class_cancer_outcomes_-_a_strategy_for_england_2015-2020.pdf [accessed 2019-03-09]
16. Aziz A, Reynolds R, Ansari A. Process and systems: a population-based model of care for people with inflammatory bowel disease - patient-reported outcomes. Future Healthc J 2019 Feb;6(1):30-35 [FREE Full text] [doi: 10.7861/futurehosp.6-1-30] [Medline: 31098583]
17. Introduction Pack for New Registrars in Clinical Oncology. Oncology Registrars' Forum. 2016. URL: https://www.rcr.ac.uk/sites/default/files/25._introduction_pack_for_new_registrars_in_clinical_oncology_updated_sb_27.4.16.pdf [accessed 2019-03-09]
18. Maguire R, Fox PA, McCann L, Miaskowski C, Kotronoulas G, Miller M, et al. The eSMART study protocol: a randomised controlled trial to evaluate electronic symptom management using the advanced symptom management system (ASyMS) remote technology for patients with cancer. BMJ Open 2017 Jun 6;7(5):e015016 [FREE Full text] [doi: 10.1136/bmjopen-2016-015016] [Medline: 28592577]
19. Final Report Summary - Esmart (Randomised Controlled Trial to Evaluate Electronic Symptom Management Using the Advanced Symptom Management System (ASYMS) Remote Technology for Patients With Cancers). European Commission. URL: https://cordis.europa.eu/project/id/602289/reporting [accessed 2021-01-26]
20. Absolom K, Warrington L, Hudson E, Hewison J, Morris C, Holch P, et al. Phase III randomized controlled trial of erapid: ehealth intervention during chemotherapy. J Clin Oncol 2021 Mar 1;39(7):734-747. [doi: 10.1200/JCO.20.02015] [Medline: 33417506]
21. Scheibe M, Herrmann A, Schmitt J, Einhart N, Sedlmayr B, Kowalski C. Implementation of patient-reported outcome assessment in routine cancer care: a systematic review of multicentric programs in Europe. Z Evid Fortbild Qual Gesundhwes 2020 Nov;156-157:11-23. [doi: 10.1016/j.zefq.2020.08.001] [Medline: 32980282]
22. Common Terminology Criteria for Adverse Events (CTCAE) v5. National Cancer Institute. 2017. URL: https://ctep.cancer.gov/protocoldevelopment/electronic_applications/docs/CTCAE_v5_Quick_Reference_5x7.pdf [accessed 2019-03-09]
23. Basch E, Artz D, Dulko D, Scher K, Sabbatini P, Hensley M, et al. Patient online self-reporting of toxicity symptoms during chemotherapy. J Clin Oncol 2005 May 20;23(15):3552-3561. [doi: 10.1200/JCO.2005.04.275] [Medline: 15908666]
24. Holch P, Warrington L, Potrata B, Ziegler L, Hector C, Keding A, et al. Asking the right questions to get the right answers: using cognitive interviews to review the acceptability, comprehension and clinical meaningfulness of patient self-report adverse event items in oncology patients. Acta Oncol 2016;55(9-10):1220-1226. [doi: 10.1080/0284186X.2016.1213878] [Medline: 27551774]
25. Christodoulou M, McCloskey P, Stones N, Bayman N, Burt P, Chittalia A, et al. Investigation of a patient reported outcome tool to assess radiotherapy-related toxicity prospectively in patients with lung cancer. Radiother Oncol 2014 Aug;112(2):244-249. [doi: 10.1016/j.radonc.2014.07.008] [Medline: 25107555]
26. PRO-CTCAE. National Cancer Institute. 2019. URL: https://healthcaredelivery.cancer.gov/pro-ctcae/ [accessed 2020-01-25]

XSL•FO
RenderX

27. Basch E, Iasonos A, McDonough T, Barz A, Culkin A, Kris MG, et al. Patient versus clinician symptom reporting using the National Cancer Institute Common Terminology Criteria for Adverse Events: results of a questionnaire-based study. Lancet Oncol 2006 Nov;7(11):903-909. [doi: 10.1016/S1470-2045(06)70910-X] [Medline: 17081915]
28. For Better, for Worse? A Report by the National Confidential Enquiry Into Patient Outcome and Death. National Confidential Enquiry into Patient Outcome and Death (NCEPOD). 2008. URL: https://www.ncepod.org.uk/2008report3/Downloads/SACT_report.pdf [accessed 2019-03-09]
29. Absolom K, Holch P, Warrington L, Samy F, Hulme C, Hewison J, eRAPID systemic treatment work group. Electronic patient self-Reporting of Adverse-events: patient Information and aDvice (eRAPID): a randomised controlled trial in systemic cancer treatment. BMC Cancer 2017 May 8;17(1):318 [FREE Full text] [doi: 10.1186/s12885-017-3303-8] [Medline: 28482877]
30. Oncology Services. Nottingham University Hospitals NHS Trust. 2019. URL: https://www.nuh.nhs.uk/oncology-services- [accessed 2019-01-06]
31. Overview - Nottingham University Hospitals NHS Trust. National Health Service. 2018. URL: https://www.nhs.uk/Services/Trusts/Overview/DefaultView.aspx?id=1368# [accessed 2019-09-15]
32. Macmillan. URL: https://www.macmillan.org.uk/ [accessed 2021-01-26]
33. Mavletova A, Couper M. A meta-analysis of breakoff rates in mobile web surveys. In: Mobile Research Methods: Opportunities and Challenges of Mobile Research Methodologies. London, UK: Ubiquity Press; 2015:81-98.
34. Yan T, Conrad F, Tourangeau R, Couper M. Should I stay or should I go: the effects of progress feedback, promised task duration, and length of questionnaire on completing web surveys. International Journal of Public Opinion Research 2010 Dec 15;23(2):131-147. [doi: 10.1093/ijpor/edq046]
35. Boynton P, Greenhalgh T. Selecting, designing, and developing your questionnaire. Br Med J 2004 May 29;328(7451):1312-1315 [FREE Full text] [doi: 10.1136/bmj.328.7451.1312] [Medline: 15166072]
36. BMJ Quality and Safety. URL: https://qualitysafety.bmj.com/pages/wp-content/uploads/sites/44/2018/08/Policy-on-Ethic-Reviews2018.pdf [accessed 2019-01-19]
37. Nilsson E, Orwelius L, Kristenson M. Patient-reported outcomes in the Swedish National Quality Registers. J Intern Med 2016 Feb;279(2):141-153 [FREE Full text] [doi: 10.1111/joim.12409] [Medline: 26306802]
38. Value Based Healthcare: Learning From Practice. King's Health Partners. URL: https://www.kingshealthpartners.org/assets/000/002/178/KHP_VBHC_Lessons_from_practice_-_Aneurin_Bevan_University_Health_Board_v3_original.pdf?1530185305 [accessed 2021-01-26]
39. Howell D, Molloy S, Wilkinson K, Green E, Orchard K, Wang K, et al. Patient-reported outcomes in routine cancer clinical practice: a scoping review of use, impact on health outcomes, and implementation factors. Ann Oncol 2015 Sep;26(9):1846-1858 [FREE Full text] [doi: 10.1093/annonc/mdv181] [Medline: 25888610]
40. O'Cathail M, Sivanandan MA, Diver C, Patel P, Christian J. The use of patient-facing teleconsultations in the national health service: scoping review. JMIR Med Inform 2020 Mar 16;8(3):e15380 [FREE Full text] [doi: 10.2196/15380] [Medline: 32175911]
41. Car J, Koh G, Foong P, Wang C. Video consultations in primary and specialist care during the covid-19 pandemic and beyond. Br Med J 2020 Oct 20;371:m3945. [doi: 10.1136/bmj.m3945] [Medline: 33082127]
42. Dembrower K, Wåhlin E, Liu Y, Salim M, Smith K, Lindholm P, et al. Effect of artificial intelligence-based triaging of breast cancer screening mammograms on cancer detection and radiologist workload: a retrospective simulation study. Lancet Digit Health 2020 Sep;2(9):e468-e474. [doi: 10.1016/S2589-7500(20)30185-0]
43. Rattan R, Kataria T, Banerjee S, Goyal S, Gupta D, Pandita A, et al. Artificial intelligence in oncology, its scope and future prospects with specific reference to radiation oncology. BJR Open 2019;1(1):20180031 [FREE Full text] [doi: 10.1259/bjro.20180031] [Medline: 33178922]
44. Oncology/Haematology 24 Hour Triage Rapid Assessment and Access Toolkit. The UK Oncology Nursing Society. 2016. URL: https://www.ukons.org/site/assets/files/1134/oncology_haematology_24_hour_triage.pdf [accessed 2021-01-26]
45. Dose Limiting Toxicity (DLT) and Common Toxicity Criteria (CTC) / Common Terminology Criteria for Adverse Events (CTCAE). On Biostatistics and Clinical Trials. 2015. URL: http://onbiostatistics.blogspot.com/2015/01/dose-limiting-toxicity-dlt-and-ctc-ctcae.html [accessed 2019-06-10]

Abbreviations

CTCAE: Common Terminology Criteria for Adverse Events
FTF: face-to-face
NHS: National Health Service
PRO-CTCAE: patient-reported outcomes Common Terminology Criteria for Adverse Events
PROM: patient-reported outcome measure
PS: performance status
RT: radiotherapy

XSL•FO
RenderX

Edited by G Eysenbach; submitted 03.03.20; peer-reviewed by L McCann, S Nabhani-Gebara; comments to author 25.09.20; revised version received 06.03.21; accepted 13.04.21; published 13.08.21.

Please cite as:
Sivanandan MA, Sharma C, Bullard P, Christian J
Digital Patient-Reported Outcome Measures for Monitoring of Patients on Cancer Treatment: Cross-sectional Questionnaire Study
JMIR Form Res 2021;5(8):e18502
URL: https://formative.jmir.org/2021/8/e18502
doi:10.2196/18502
PMID:34398785

©Mayuran Ananth Sivanandan, Catherine Sharma, Pippa Bullard, Judith Christian. Originally published in JMIR Formative Research (https://formative.jmir.org), 13.08.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Development of a Mobile Health App (TOGETHERCare) to Reduce Cancer Care Partner Burden: Product Design Study

Ingrid Oakley-Girvan[1], PhD, MPH; Sharon Watkins Davis[1], MPA; Allison Kurian[2], MD, MSc; Lisa G Rosas[2], PhD, MPH; Jena Daniels[1], BSc; Oxana Gronskaya Palesh[2], PhD, MPH; Rachel J Mesia[2], PhD; Arif H Kamal[3], MD, MBA, MHS; Michelle Longmire[1], MD; Vasu Divi[2], MD

[1]Medable Inc, Palo Alto, CA, United States
[2]Stanford University, Stanford, CA, United States
[3]Duke University, Durham, NC, United States

Corresponding Author:
Ingrid Oakley-Girvan, PhD, MPH
Medable Inc
525 University Avenue
Suite A 70
Palo Alto, CA, 94301
United States
Phone: 1 408 656 2948
Email: ingrid@medable.com

Abstract

Background: Approximately 6.1 million adults in the United States serve as care partners for cancer survivors. Studies have demonstrated that engaging cancer survivors and their care partners through technology-enabled structured symptom collection has several benefits. Given the high utilization of mobile technologies, even among underserved populations and in low resource areas, mobile apps may provide a meaningful access point for all stakeholders for symptom management.

Objective: We aimed to develop a mobile app incorporating user preferences to enable cancer survivors' care partners to monitor the survivors' health and to provide care partner resources.

Methods: An iterative information gathering process was conducted that included (1) discussions with 138 stakeholders to identify challenges and gaps in survivor home care; (2) semistructured interviews with clinicians (n=3), cancer survivors (n=3), and care partners (n=3) to identify specific needs; and (3) a 28-day feasibility field test with seven care partners.

Results: Health professionals noted the importance of identifying early symptoms of adverse events. Survivors requested modules on medication, diet, self-care, reminders, and a version in Spanish. Care partners preferred to focus primarily on the patient's health and not their own. The app was developed incorporating quality-of-life surveys and symptom reporting, as well as resources on home survivor care. Early user testing demonstrated ease of use and app feasibility.

Conclusions: TOGETHERCare, a novel mobile app, was developed with user input to track the care partner's health and report on survivor symptoms during home care. The following two clinical benefits emerged: (1) reduced anxiety among care partners who use the app and (2) the potential for identifying survivor symptoms noted by the care partner, which might prevent adverse events.

Trial Registration: ClinicalTrials.gov NCT04018677; https://clinicaltrials.gov/ct2/show/NCT04018677

(JMIR Form Res 2021;5(8):e22608) doi:10.2196/22608

KEYWORDS

cancer; oncology; mHealth; caregiver; cancer survivor; mobile app; smartphone; feasibility; caregiver burden; symptom reporting

Introduction

Approximately 6.1 million adults in the United States serve as care partners for cancer survivors [1]. Informal care partners, defined as unpaid spouses, relatives, and friends of the survivor, are essential partners with oncology teams in the delivery of complex cancer care services at home [2]. Care partners assist with activities of daily living, medication administration, wound care, transportation, meals, finances, advocacy, and emotional

XSL•FO
RenderX

support. Care partners frequently attend medical visits with the survivor, often keeping track of physician instructions and medication changes [3]. Based on data collected through the 2015-2017 Behavioral Risk Factor Surveillance System (BRFSS) run by the Centers for Disease Control and Prevention (CDC), 24% of adults aged 45 years or older were care partners for relatives or friends [4]. Approximately 25% of those caring for cancer survivors spend more than 40 hours a week providing these services to their family or friends [5].

Studies have demonstrated that engaging cancer survivors and their care partners through technology-enabled structured symptom collection has several benefits [6-9]. For example, in a randomized trial, Basch et al found an increase in health-related quality of life and a decrease in emergency room visits and hospitalizations for survivors who were provided with a tablet computer–based symptom reporting system [10]. Studies have also found that survivor symptom self-reporting (patient-reported outcomes [PROs]) resulted in an increase in survival compared with usual care [8,11-13].

Complementary to patient reports, care partners bring a different and important perspective (observer-reported outcomes [ObsROs]) when reporting on survivor symptoms and may notice symptoms the survivor does not. Further, data demonstrate that these perspectives are feasible to collect. One study demonstrated that a series of systematic questions presented to care partners of children in palliative care were easy to complete and identified symptoms underdiagnosed by medical teams. Reporting by children aged 7 years or over and their care partners were consistent for common symptoms, but care partners reported irritability and nervousness more frequently than children [14]. In adult prostate cancer survivors, survivors who had a care partner were more likely to discuss pain at doctor visits than those without a care partner. This study concluded that tools encouraging early symptom reporting could lead to enhanced symptom and disease management [15]. Care partners' assessments of symptoms in survivors were similar with survivor self-reports, indicating that the care partner could serve as a proxy [16,17]. There is also evidence that using an electronic symptom reporting system decreases the emotional distress of care partners [18,19].

Given the high utilization of mobile technologies, even among underserved populations and in low resource areas [20-24], mobile apps may provide a meaningful access point for all stakeholders for symptom management [25-27]. A Deloitte survey of US health care consumers found the following three key areas of consumer engagement: (1) consumers want to partner with clinical teams on their health care and management; (2) consumers are increasingly trusting and using online information; and (3) consumers, particularly those with chronic conditions, are increasingly utilizing health technologies [28], providing additional reasons to consider this avenue of engagement. Mobile health technologies show promise as solutions for health care needs across the cancer continuum [29] and have the potential to improve health care outcomes by providing consumers with a platform that can address all three of these key areas.

The development of most mHealth apps does not involve user input [30], despite evidence that incorporating feedback from appropriate stakeholders, including care partners, into smartphone app development can result in a more successful mobile tool [31,32]. The purpose of this study was to design and develop a mobile app in collaboration with users and other stakeholders for informal care partners to remotely monitor cancer survivors' health and for providing care partner resources. While additional work is required to confirm the clinical effectiveness for specific outcomes, this paper documents the development process of this app (TOGETHERCare) and the preliminary results of early user testing.

Methods

Overview

This study was conducted in three sequential phases guided by the Technology Acceptance Model and user-centered design principles. The objective of phase 1 was to understand current care gaps and needs in cancer care through interviews with health care providers, survivors, and care partners. In phase 2, we gathered design and content specifications for the planned app through semistructured interviews with care partners, survivors, and clinicians. Finally, in phase 3, we created the first version of the TOGETHERCare mobile app for informal care partners to remotely monitor cancer survivors' health and their own health and for providing care partner resources. We collected feedback from a small beta testing group of care partners who used the app for 28 days in two geographically different academic cancer centers (Duke and Stanford). This study has been registered on clinicaltrials.gov (NCT04018677). This iterative information gathering process was conducted to address three main questions. Table 1 lists the questions and our methods for addressing the questions.

Table 1. Main questions and methods for addressing the questions.

Main questions	Methods to address the questions
Phase 1. What are the needs and gaps for cancer care partners?	Discussions with 138 stakeholders including 32 care partners
Phase 2. What features would users like to see in such an app?	Semistructured interviews with three physicians, three cancer survivors, and three care partners
Phase 3. Is an mHealth[a] app for cancer care partners feasible for them to use?	28-day beta iOS (Apple) user testing with feasibility and acceptability feedback from seven care partners in two geographically diverse academic cancer centers (Duke University and Stanford University)

[a]mHealth: mobile health.

XSL•FO
RenderX

Phase 1. Stakeholder Discussions

We talked to stakeholders who would have insights into the needs of cancer patients, including providers, pharmaceutical company scientists, advocates, social workers, medical directors, researchers, and care partners. Stakeholders were identified first through Project Team Advisory Group (PTAG) members (N=14) and Patient Advocacy Council (PAC) members (N=4). PTAG members included survivor advocates, clinicians, and researchers with direct experience working with cancer survivors. PAC members included cancer survivors and care partners. A snowball technique was used to identify additional stakeholders, so that each informant identified additional people to interview until saturation of concepts was achieved.

The purpose of the interviews was to identify gaps in or barriers to cancer care, care partners, workflows, and the potential value of possible app components to organizations and individuals. These interviews were open-ended and did not use interview guides. Responses to the interviews were categorized by the research team [33].

Phase 2. Semistructured Interviews

From a convenience sample from the Stanford Cancer Institute and Monterey-Salinas California health care systems, two interviewers conducted nine interviews using Institutional Review Board (IRB)-approved semistructured guides. The interviews were conducted with three physicians who work with cancer survivors, three cancer survivors, and three care partners currently caring for cancer survivors. Three interviews with Spanish-speaking survivors and care partners were conducted with a medical interpreter translating for the interviewer. Notes from the interviews were transcribed into a prepared template. The results were compiled, and response concepts were coded by two members of the core team.

Phase 3. Beta Test

We conducted a beta test with cancer care partners to assess feasibility. Informed consent was collected through the app. The app included the following tabs: “Profile” for care partner’s name and demographics; “Activities” for informed consent, Health Insurance Portability and Accountability Act authorization, PRO surveys related to care partners’ health, survivor demographics, and surveys related to survivors’ health, including a targeted symptom list and a preappointment concerns survey; and “Resources” with links to resources related to care partners’ health, caregiving tasks, and local referral information.

The aim of the beta test was to test the consent and enrollment process and gather feedback from real care partners. Beta testing was also used to catch software or interoperability bugs before extensive usability testing. Cancer survivors (n=6) were recruited by staff at Stanford and Duke Universities and asked to consent to the project and identify their care partners. After providing informed consent, the care partners (n=7) (one survivor named two care partners) were enrolled in the study and instructed on how to download the app and enter data. Stanford University recruited survivors from survivor support groups, and Duke University recruited survivors from the palliative care clinic.

We conducted semistructured interviews with most beta participants at day 7 and all participants at day 28 to see what they liked and did not like about using the app. Interviews were zoom calls, typically 30 to 45 min in length, and the same script was followed in each interview, except for the first question (“Tell us about your care partner situation”) that we did not ask at the day 28 interview. Interviews provided much more in-depth information about care partners’ experiences using the app than we would have obtained using generic app satisfaction scales such as the System Usability Scale (SUS) [34,35] and mHealth App Usability Questionnaire (MAUQ) [36]. Interviews were recorded and transcribed.

Transcripts of the interviews were coded to identify themes. The thematic coding was done based on the principles of content analysis, where textual data are identified, analyzed, and grouped to form meaningful categories [37]. The level of analysis was entire sentences, as the interviews were semistructured and coding by word would give undue weight to a single respondent who had longer responses. Sentences from the interviews were coded based on the main concept contained in them, and the most commonly occurring themes across interviews were compiled.

Statistical methods for analyzing the in-app surveys included frequency counts and percentages to determine the completion rate of each survey, the number of surveys that were completed during the 4-week beta test, and the time required to complete each survey.

All procedures and study materials were independently approved by IRBs at Stanford and Duke Universities (clinicaltrials.gov NCT04018677). Because this was an unrandomized trial, no CONSORT checklist was filed. All procedures were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all participants included in the study.

Results

Phase 1. Stakeholder Discussions

The responses from the 138 stakeholders resulted in 262 responses that fell into different categories as follows: burnout/stress (n=108), cancer types/comorbid conditions (n=65), revenue/business (n=40), office value (n=22), insurance (n=12), and media/events attended by respondents (n=15). The major categories that emerged across all areas included the following: (1) survivors who come in for crisis visits are older, have comorbidities, have care partners at home who are burned out, live geographically far away, or have difficulty accessing appropriate resources and services; (2) buyers of new applications are risk sensitive or want a tool supported by validation studies proving reduction of clinical burdens and crisis visits when care partners are supported and less stressed; and (3) fitting into the electronic health record (EHR) system is important.

Responses by the 108 stakeholders who commented on burnout/stress are detailed in Table 2. Scheduling, keeping

XSL•FO
RenderX

upcoming visits organized, or feeling overwhelmed with too much information related to different doctors, medications, and appointments were mentioned 27 times by most stakeholders across job functions.

The 32 care partners who were interviewed wanted an app to help them (1) be better organized and feel prepared to take care of their loved ones, including organizational support for scheduling visits, keeping up on the treatment and following treatment protocols, or knowing what signs/symptoms to look out for in order to identify adverse events earlier; (2) be more knowledgeable about various side effects, insurance options, treatment options, and other resources available to the survivor and care partner; and (3) have a sense of support and understanding of challenges from the clinical team, friends, family, and coworkers.

Table 2. Stakeholder discussion responses related to burnout and stress.

Stakeholder	Primary cause of stress/burnout for informants						
	Confusion navigating the health care system/insurance	Schedule/organization/too much info (meds, doctors, and appointments)	Socioeconomic costs, missing work	Distance from the doctor's office/transportation	Elderly survivor	Emotional burden/depression	Other
Clinician/pharma (n=16)	2	3	3	3	1	3	1
Researcher (n=3)	0	1	0	1	0	0	1
Care partner (n=9)	1	4	2	0	0	1	1
Social worker/counselor (n=23)	3	7	3	4	1	3	2
Nurse (n=9)	2	3	2	1	1	0	0
Attorney (n=2)	1	1	0	0	0	0	0
Chief medical officer (n=1)	0	0	0	0	0	1	0
Clinician (n=17)	0	3	4	3	3	4	0
Patient advocate (n=4)	0	0	2	1	1	0	0
Chief executive officer (n=7)	0	2	1	1	1	1	1
Medical doctor (n=17)	2	3	3	2	2	4	1
Total (n=108)[a]	11	27	20	16	10	17	7

[a]108 of the 132 stakeholder comments were classified into the burnout/stress category.

Phase 2. Semistructured Interviews

The demographics for the semistructured interviewees were quite varied. Physicians (n=3) were between 39 and 66 years of age and had been providing health care for those aged 14 to 39 years. There were two medical oncologists and one surgeon. Survivors (n=3) ranged in age from 43 to 74 years and had been diagnosed between 8 months and 19 years prior. Cancer diagnoses included breast cancer, melanoma, ovarian cancer, colon cancer, and thyroid cancer (one survivor had multiple diagnoses). Care partners (n=3) ranged in age from 22 to 63 years and had been care partners for a range of less than 1 year to 6 years. We interviewed two white non-Hispanic females, one white non-Hispanic male, one Asian male, three Latino females, and two Latino males.

For many of the coded interview concepts, there was general agreement across the physicians, cancer survivors, and care partners. Although all three groups agreed that there is currently no systematic way for specialists to keep in touch with survivors once they have moved to community care, survivorship care plans (SCPs) would be useful. The SCP provides treatment history, management of side effects, and information on who to contact. However, they currently do not receive or prepare an SCP. All three groups concurred that the survivor had to initiate either a visit or call to the specialist. All three groups agreed that they have smartphones and that an app including the ability to communicate between the different groups, along with other modules, such as guidance on assisting with daily medical tasks and activities of daily living, would be useful.

There were also differences between the three groups of semistructured interviewees in concepts coded for responses to four of the interview questions. Care partners and survivors had different kinds of questions they would like to ask physicians, compared with questions clinical staff felt they heard frequently from care partners. Cancer survivors indicated no concerns about the app, and one care partner mentioned a concern about keeping medical information private, but clinical staff were concerned about the added workload and whether the app would prove useful unless all members of the care team participated in the effort. Clinical staff had specific ideas for smartphone app modules, including communication tips for survivors and care partners to make better use of their time with the clinical team; a dashboard or status bar that would track the survivor through the care process, from the initial treatment through all treatments provided by different specialties; and addition of PROs or ObsROs. Survivors and care partners were interested in modules on medication, diet, self-care, reminders, and a version in Spanish.

XSL•FO
RenderX

Phase 3. Beta Test Qualitative Analysis

Evidence from the beta test semistructured qualitative surveys indicated that care partners (n=7) found the app to be useful, and would continue to use it, as well as recommend it to others. They also suggested that adding features to the app, such as the ability to search for specialized information and insert open-ended notes, would greatly enhance the functionality of the app. Care partners were focused on survivor health, and were not too interested in responding to health questions about themselves; instead, they felt that most of the questions should focus on the survivor. Moreover, they requested more feedback from the daily surveys, such as an explanation of what their survivors' health measurements indicated in terms of the survivors' current health statuses. The four main themes that emerged from the beta test qualitative interviews are shown in Table 3.

Table 3. Beta test qualitative analysis themes with illustrative comments.

Theme	Concept	Selected comments
The app is useful	All care partners found the app to be useful in multiple ways, including reducing their anxiety by focusing on the fixed number of survey questions, as well as serving as a learning tool and raising their level of awareness.	*Finds it very helpful. Helps her not to worry and eases her mind. Helps keep things in perspective - she focuses on set number of questions, not a huge list.* *I thought the measurement of mental health was helpful. It might help avoid depression. Well, I think it's easy to use and to understand.*
Add functionality to the app	Care partners suggested that the app should provide specialized information and contain an open-ended notes section.	*Especially if more types of information could be added to this app that I can use other than general Information about caregiving.* *Would be nice to have a box where you can put notes in - especially about the survivor.*
Care partners are focused on survivor health	Care partners remarked that the surveys contained too many questions pertaining to the care partners rather than focusing on survivors.	*I want to focus more on the cancer patient because that person is the one who needs help, more than the care partner.* *Actually, I would rather focus more on the cancer patient myself.*
Care partners need more feedback	Care partners noted that they would have liked feedback about their survey responses.	*It would have helped me to know that someone on the end is monitoring my responses or I could receive a response back when I did things on the app.* *It would be better if you could receive some kind of report back of what the measurement means, especially about patient's health and stress level, depression, exercise.*

Phase 3. Beta Test Quantitative Data Analysis

Seven care partners participated in the beta test. The completion rates below refer to the percentage of surveys for which care partners answered all the questions. During the beta test, care partners were told that the demographics section was optional, so that was the least frequently completed survey (Stanford, 25%; Duke, 33%). The Patient Health Questionnaire-4 [38] survey about the care partner, which measures depression and anxiety in four items, was completed only 31% of the time by both groups combined, with the care partners infrequently answering any of the questions. Demographics about the survivor were completed 90% of the time, and "My loved one's health" (about the survivor) and the preappointment survey were completed 91% of the time. All other surveys were completed between 97% and 100% of the time by both groups combined.

Within the 4-week test period, care partners were expected to answer each survey a certain number of times. Table 4 lists the survey name, the frequency with which the survey came up on the app, the number of times it was available during the 28-day beta test time frame, and the median and mean numbers of care partner beta testers (n=7) who started the survey. Care partners started surveys close to the frequency at which they were available, but not all the survey questions were completed (participants could skip any questions they did not wish to answer). The "number started" refers to the number of surveys for which at least one question was answered, regardless of whether all questions in that survey were completed. One-time surveys took on average from 32 seconds (care partner eligibility) to 211 seconds (e-consent document and signature [consent previously explained and reviewed by the clinical team]) to complete. Repeated surveys were completed, on average, in 2 minutes or less. The Patient Health Questionnaire-4 (PHQ-4) had an average completion time of 8 seconds, and the "My loved one's health" was completed on average in 129 seconds.

XSL•FO
RenderX

Table 4. Surveys included in the care partner app by availability frequency and mean number started by beta testers (n=7) during the beta testing period (all surveys were completed by the care partner).

Survey name	Requested Frequency	Number of times available	Median number started	Mean number started
Care partner consent	Once	1	1	1.29
Care partner eligibility	Once	1	1	1.14
Daily (about care partner sleep and mood)	Daily	25	9	13.86
Demographics (about care partner)	Once	1	1	1.00
EQ5D[a] [39] (about care partner)	Weekly	5	5	4.86
My loved one (survivor demographics)	Once	1	1	1.14
My loved one's health (about survivor)	Weekly	5	5	4.57
PROMIS[b] [40] (about care partner)	Biweekly	4	3	2.86
PHQ-4[c] [38] (about care partner)	Biweekly	4	3	2.71
Preappointment (care partner concerns re: survivor appointment)	Weekly	5	5	4.71
Weekly (about care partner health)	Weekly	5	5	4.86

[a]EQ5D: EuroQol 5-Dimension.

[b]PROMIS: Patient-Reported Outcomes Measurement Information System.

[c]PHQ-4: Patient Health Questionnaire-4.

Discussion

This paper describes the user-involved development process of TOGETHERCare, a smartphone app for care partners. This is one of few studies regarding app development for cancer care partners that utilized a rigorous development approach involving users [31,41]. Engaging users in the design of an mHealth app facilitates app adoption and usage [42].

As the US population ages, more care is being delivered at home. Limits on rehabilitation and nursing home payments can result in survivors being discharged before they are ready [43,44]. Starting in 2015, more money was spent nationally on home care than care provided in nursing homes [45]. This increase in caring for survivors at home has raised the burden on informal care partners, such as family and friends. This heavy burden can affect mental and physical health, and the care partner's health can appreciably impact the survivor [2,46-51]. Clinical benefits have been associated with PROs [10-13]. While studies on care partner reporting are limited [19], there is evidence that care partners can identify early symptoms [14-16], and the use of a symptom reporting system may reduce caregiver distress [18].

TOGETHERCare is a mobile app that provides for care partner symptom reporting for themselves and the survivors. The following three phases were completed in the development of the app: (1) stakeholder discussions, (2) semistructured interviews, and (3) beta testing of the app. These phases are essential to ensure that the app is developed incorporating user preferences to increase its value [30].

In our stakeholder discussions, all stakeholders felt that an app for informal care partners would be beneficial. Stakeholders frequently mentioned that care partners feel overwhelmed with too much information and financial considerations, and have emotional issues including stress and depression. Clinical staff, care partners, and survivors included in the semistructured interviews all agreed that an app designed to help care partners would be welcome. Results of early user beta testing showed that TOGETHERCare is feasible to use, with users able to complete surveys and commenting that the app was useful and helpful. Our qualitative interviews with testers revealed that care partners are primarily focused on their survivor's health, not their own, and the quantitative analysis indicated the need to reduce the number of surveys about care partner's health. Several care partners mentioned that the survivor-focused surveys helped to reduce their anxiety and bring relief by reducing the universe of things they had to worry about. Similar findings were observed by Chih et al, who examined an online symptom-reporting system for advanced cancer survivors [18].

This study has some limitations. While an open-ended interview with stakeholders allowed us to receive perspectives that might otherwise not have been discussed, having a more structured interview may have allowed us to examine opinions in a more standardized way. Limitations in this study also include the small convenience sample of care partners involved in the semistructured interviews and the beta testing. However, small sample sizes for exploratory studies are common, and other published research studies on app development and acceptability have used between 5 and 11 users for feasibility and usability testing [52-54]. The beta test version of the app did not include visualizations that are expected to be developed in future versions of the app, although mockups of visualizations were presented during the qualitative interviews, and feedback was obtained. Survey data completed by care partners were not presented within the EHR to clinicians because we were not testing the feedback component at this stage. Future work in a larger clinical sample will include providing survey data completed by care partners to clinicians within a work flow they are already using. Care partners in our study were required

XSL•FO
RenderX

to have an iPhone, as the beta test version was developed for the iOS platform, with a future intention to include Android phones.

Next steps include testing the app with a larger population, providing data recorded by care partners about their survivors' symptoms to the clinical team within the EHR, and ultimately testing the app impact on specific outcomes in a randomized controlled trial. Further evaluation in a randomized clinical trial is needed to provide evidence that the app would result in fewer hospitalizations and emergency room visits and potentially extend survival.

TOGETHERCare, a novel mobile app, was developed together with care partners, health care professionals, and survivors to track the health of care partners and report on a targeted list of survivor symptoms during home care. The following two clinical benefits have emerged: (1) the reduction of anxiety among care partners who use the app and (2) the potential for the identification of symptoms exhibited by the survivor and noted by the care partner, which, if caught early, might prevent unnecessary emergency room or hospital care.

Acknowledgments

The Medable Patient Advisory Council provided substantial input and feedback for the app (TJ Sharpe, Richie Kahn, Jennifer McNary, and Joan Venticinque). Miriam Green also provided input and edits from a survivor's perspective. Research reported in this publication was partially supported by the National Cancer Institute of the National Institutes of Health under contract number HHSN261201700030C. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest

IOG, SWD, JD, and ML are employed by Medable Inc, which developed the TOGETHERCare app with funding from the National Institutes of Health.

References

1. Applebaum AJ, Breitbart W. Care for the cancer caregiver: a systematic review. Palliat Support Care 2013 Jun;11(3):231-252 [FREE Full text] [doi: 10.1017/S1478951512000594] [Medline: 23046977]
2. Bevans M, Sternberg EM. Caregiving burden, stress, and health effects among family caregivers of adult cancer patients. JAMA 2012 Jan 25;307(4):398-403 [FREE Full text] [doi: 10.1001/jama.2012.29] [Medline: 22274687]
3. Kasper JD, Steinbach U, Andrews J. Caregiver role appraisal and caregiver tasks as factors in ending caregiving. J Aging Health 1994 Aug;6(3):397-414. [doi: 10.1177/089826439400600307] [Medline: 10135717]
4. Caregiving for Family and Friends — A Public Health Issue. Centers for Disease Control Prevention. URL: https://www.cdc.gov/aging/caregiving/caregiver-brief.html [accessed 2021-07-21]
5. Caregiving in the U.S. 2020. The National Alliance for Caregiving. URL: https://www.caregiving.org/caregiving-in-the-us-2020/ [accessed 2020-06-12]
6. Bouazza YB, Chiairi I, El Kharbouchi O, De Backer L, Vanhoutte G, Janssens A, et al. Patient-reported outcome measures (PROMs) in the management of lung cancer: A systematic review. Lung Cancer 2017 Nov;113:140-151. [doi: 10.1016/j.lungcan.2017.09.011] [Medline: 29110842]
7. Fiteni F, Cuenant A, Favier M, Cousin C, Houede N. Clinical Relevance of Routine Monitoring of Patient-reported Outcomes Clinician-reported Outcomes in Oncology. In Vivo 2019;33(1):17-21 [FREE Full text] [doi: 10.21873/invivo.11433] [Medline: 30587597]
8. Aapro M, Bossi P, Dasari A, Fallowfield L, Gascón P, Geller M, et al. Digital health for optimal supportive care in oncology: benefits, limits, and future perspectives. Support Care Cancer 2020 Oct 12;28(10):4589-4612 [FREE Full text] [doi: 10.1007/s00520-020-05539-1] [Medline: 32533435]
9. Matsuda A, Yamada Y, Ishizuka N, Matsushima E, Kobayashi K, Ohkubo T, et al. Effectiveness of a Self-Monitoring Quality of Life Intervention For Patients with Cancer Receiving Palliative Care: A Randomized Controlled Clinical Trial. Asian Pac J Cancer Prev 2019 Sep 01;20(9):2795-2802 [FREE Full text] [doi: 10.31557/APJCP.2019.20.9.2795] [Medline: 31554379]
10. Basch E, Deal AM, Kris MG, Scher HI, Hudis CA, Sabbatini P, et al. Symptom Monitoring With Patient-Reported Outcomes During Routine Cancer Treatment: A Randomized Controlled Trial. J Clin Oncol 2016 Feb 20;34(6):557-565 [FREE Full text] [doi: 10.1200/JCO.2015.63.0830] [Medline: 26644527]
11. Basch E, Deal AM, Dueck AC, Scher HI, Kris MG, Hudis C, et al. Overall Survival Results of a Trial Assessing Patient-Reported Outcomes for Symptom Monitoring During Routine Cancer Treatment. JAMA 2017 Jul 11;318(2):197-198 [FREE Full text] [doi: 10.1001/jama.2017.7156] [Medline: 28586821]
12. Basch E, Deal AM, Dueck AC, Bennett A, Atkinson T, Scher H, et al. Overall survival results of a randomized trial assessing patient-reported outcomes for symptom monitoring during routine cancer treatment. JCO 2017 Jun 20;35(18_suppl):LBA2-LBA2 [FREE Full text] [doi: 10.1200/JCO.2017.35.18_SUPPL.LBA2]

XSL•FO
RenderX

13. Denis F, Basch E, Septans A, Bennouna J, Urban T, Dueck AC, et al. Two-Year Survival Comparing Web-Based Symptom Monitoring vs Routine Surveillance Following Treatment for Lung Cancer. JAMA 2019 Jan 22;321(3):306-307 [FREE Full text] [doi: 10.1001/jama.2018.18085] [Medline: 30667494]
14. Madden K, Magno Charone M, Mills S, Dibaj S, Williams JL, Liu D, et al. Systematic Symptom Reporting by Pediatric Palliative Care Patients with Cancer: A Preliminary Report. J Palliat Med 2019 Aug;22(8):894-901 [FREE Full text] [doi: 10.1089/jpm.2018.0545] [Medline: 30759031]
15. Drudge-Coates L, Oh WK, Tombal B, Delacruz A, Tomlinson B, Ripley AV, et al. Recognizing Symptom Burden in Advanced Prostate Cancer: A Global Patient and Caregiver Survey. Clin Genitourin Cancer 2018 Apr;16(2):e411-e419 [FREE Full text] [doi: 10.1016/j.clgc.2017.09.015] [Medline: 29111175]
16. Armstrong TS, Wefel JS, Gning I, Acquaye A, Vera-Bolanos E, Gilbert MR, et al. Congruence of primary brain tumor patient and caregiver symptom report. Cancer 2012 Oct 15;118(20):5026-5037 [FREE Full text] [doi: 10.1002/cncr.27483] [Medline: 22415423]
17. Kurtz ME, Kurtz JC, Given CC, Given B. Concordance of cancer patient and caregiver symptom reports. Cancer Pract 1996;4(4):185-190. [Medline: 8900759]
18. Chih M, DuBenske LL, Hawkins RP, Brown RL, Dinauer SK, Cleary JF, et al. Communicating advanced cancer patients' symptoms via the Internet: a pooled analysis of two randomized trials examining caregiver preparedness, physical burden, and negative mood. Palliat Med 2013 Jun;27(6):533-543 [FREE Full text] [doi: 10.1177/0269216312457213] [Medline: 22988042]
19. Wishart LR, Brown B, Nund RL, Fotinos E, Hutchison AR, Ward EC, et al. A prospective study monitoring carer distress during (chemo)radiotherapy for head and neck cancer via an electronic platform. J Med Radiat Sci 2021 Jun;68(2):139-148 [FREE Full text] [doi: 10.1002/jmrs.448] [Medline: 33169922]
20. Smith A. Americans and Their Cell Phones. Pew Research Center. 2011. URL: https://www.pewresearch.org/internet/Reports/2011/Cell-Phones.aspx [accessed 2020-06-11]
21. Price M, Williamson D, McCandless R, Mueller M, Gregoski M, Brunner-Jackson B, et al. Hispanic migrant farm workers' attitudes toward mobile phone-based telehealth for management of chronic health conditions. J Med Internet Res 2013 Mar 30;15(4):e76 [FREE Full text] [doi: 10.2196/jmir.2500] [Medline: 23624105]
22. Dang CM, Estrada S, Bresee C, Phillips EH. Exploring potential use of internet, E-mail, and instant text messaging to promote breast health and mammogram use among immigrant Hispanic women in Los Angeles County. Am Surg 2013 Oct;79(10):997-1000. [Medline: 24160786]
23. Purnell JQ, Griffith J, Eddens KS, Kreuter MW. Mobile technology, cancer prevention, and health status among diverse, low-income adults. Am J Health Promot 2014;28(6):397-402 [FREE Full text] [doi: 10.4278/ajhp.120816-ARB-396] [Medline: 24200336]
24. Anderson-Lewis C, Darville G, Mercado RE, Howell S, Di Maggio S. mHealth Technology Use and Implications in Historically Underserved and Minority Populations in the United States: Systematic Literature Review. JMIR Mhealth Uhealth 2018 Jun 18;6(6):e128 [FREE Full text] [doi: 10.2196/mhealth.8383] [Medline: 29914860]
25. Mobile Fact Sheet. Pew Research Center. 2021. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2020-06-10]
26. Chi N, Demiris G. The Roles of Telehealth Tools in Supporting Family Caregivers: Current Evidence, Opportunities, and Limitations. J Gerontol Nurs 2017 Feb 01;43(2):3-5. [doi: 10.3928/00989134-20170111-04] [Medline: 28128391]
27. Bae WK, Kwon J, Lee HW, Lee S, Song E, Shim H, et al. Feasibility and accessibility of electronic patient-reported outcome measures using a smartphone during routine chemotherapy: a pilot study. Support Care Cancer 2018 Nov;26(11):3721-3728. [doi: 10.1007/s00520-018-4232-z] [Medline: 29732483]
28. Betts D, Korenda L. A consumer-centered future of health. Deloitte Center for Health Solutions. 2019. URL: https://www2.deloitte.com/content/dam/Deloitte/nl/Documents/life-sciences-health-care/deloitte-nl-healthcare-consumer-centered-future-of-health.pdf
29. Davis SW, Oakley-Girvan I. mHealth Education Applications Along the Cancer Continuum. J Cancer Educ 2015 Jun;30(2):388-394. [doi: 10.1007/s13187-014-0761-4] [Medline: 25482319]
30. Davis SW, Oakley-Girvan I. Achieving value in mobile health applications for cancer survivors. J Cancer Surviv 2017 Aug;11(4):498-504. [doi: 10.1007/s11764-017-0608-1] [Medline: 28342093]
31. Wittenberg E, Xu J, Goldsmith J, Mendoza Y. Caregiver Communication About Cancer: Development of a mhealth resource to support family caregiver communication burden. Psychooncology 2019 Feb;28(2):365-371 [FREE Full text] [doi: 10.1002/pon.4950] [Medline: 30471236]
32. Harder H, Holroyd P, Burkinshaw L, Watten P, Zammit C, Harris PR, et al. A user-centred approach to developing bWell, a mobile app for arm and shoulder exercises after breast cancer treatment. J Cancer Surviv 2017 Dec;11(6):732-742 [FREE Full text] [doi: 10.1007/s11764-017-0630-3] [Medline: 28741202]
33. Glaser B, Strauss A. Discovery of Grounded Theory: Strategies for Qualitative Research. London, United Kingdom: Routledge; 2017.
34. Bangor A, Kortum PT, Miller JT. An Empirical Evaluation of the System Usability Scale. International Journal of Human-Computer Interaction 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]

XSL•FO
RenderX

35. System Usability Scale (SUS). usability.gov. URL: https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html [accessed 2020-06-01]
36. Zhou L, Bao J, Setiawan IMA, Saptono A, Parmanto B. The mHealth App Usability Questionnaire (MAUQ): Development and Validation Study. JMIR Mhealth Uhealth 2019 Apr 11;7(4):e11500 [FREE Full text] [doi: 10.2196/11500] [Medline: 30973342]
37. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
38. Kroenke K, Spitzer RL, Williams JBW, Löwe B. An ultra-brief screening scale for anxiety and depression: the PHQ-4. Psychosomatics 2009;50(6):613-621. [doi: 10.1176/appi.psy.50.6.613] [Medline: 19996233]
39. Stolk E, Ludwig K, Rand K, van Hout B, Ramos-Goñi JM. Overview, Update, and Lessons Learned From the International EQ-5D-5L Valuation Work: Version 2 of the EQ-5D-5L Valuation Protocol. Value Health 2019 Jan;22(1):23-30 [FREE Full text] [doi: 10.1016/j.jval.2018.05.010] [Medline: 30661630]
40. Sikorskii A, Victorson D, O'Connor P, Hankin V, Safikhani A, Crane T, et al. PROMIS and legacy measures compared in a supportive care intervention for breast cancer patients and caregivers: Experience from a randomized trial. Psychooncology 2018 Sep;27(9):2265-2273 [FREE Full text] [doi: 10.1002/pon.4825] [Medline: 29956396]
41. Badr H, Carmack CL, Diefenbach MA. Psychosocial interventions for patients and caregivers in the age of new communication technologies: opportunities and challenges in cancer care. J Health Commun 2015;20(3):328-342 [FREE Full text] [doi: 10.1080/10810730.2014.965369] [Medline: 25629218]
42. van Gemert-Pijnen JEWC, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011 Dec 05;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]
43. Medicare's Limited Nursing Home Coverage. ElderLawAnswers. 2021. URL: https://www.elderlawanswers.com/medicares-limited-nursing-home-coverage-6705 [accessed 2021-06-30]
44. Who pays for home care for cancer patients? American Cancer Society. URL: https://www.cancer.org/treatment/finding-and-paying-for-treatment/choosing-your-treatment-team/home-care-agencies/who-pays.html [accessed 2020-03-04]
45. Rogers K. As the US population ages, the need for home health-care workers skyrockets. CNBC. 2018. URL: https://www.cnbc.com/2018/05/31/as-the-us-population-ages-the-need-for-home-health-care-workers-skyrockets.html [accessed 2020-03-04]
46. Ochoa CY, Buchanan Lunsford N, Lee Smith J. Impact of informal cancer caregiving across the cancer experience: A systematic literature review of quality of life. Palliat Support Care 2020 Apr;18(2):220-240. [doi: 10.1017/S1478951519000622] [Medline: 31588882]
47. Caruso R, Nanni MG, Riba MB, Sabato S, Grassi L. The burden of psychosocial morbidity related to cancer: patient and family issues. Int Rev Psychiatry 2017 Oct;29(5):389-402. [doi: 10.1080/09540261.2017.1288090] [Medline: 28753076]
48. Lee S, Colditz GA, Berkman LF, Kawachi I. Caregiving and risk of coronary heart disease in U.S. women. American Journal of Preventive Medicine 2003 Feb;24(2):113-119. [doi: 10.1016/s0749-3797(02)00582-2] [Medline: 12568816]
49. Perpiñá-Galvañ J, Orts-Beneito N, Fernández-Alcántara M, García-Sanjuán S, García-Caro MP, Cabañero-Martínez MJ. Level of Burden and Health-Related Quality of Life in Caregivers of Palliative Care Patients. Int J Environ Res Public Health 2019 Nov 29;16(23):4806-4806 [FREE Full text] [doi: 10.3390/ijerph16234806] [Medline: 31795461]
50. Riffin C, Van Ness PH, Wolff JL, Fried T. Multifactorial Examination of Caregiver Burden in a National Sample of Family and Unpaid Caregivers. J Am Geriatr Soc 2019 Feb;67(2):277-283 [FREE Full text] [doi: 10.1111/jgs.15664] [Medline: 30452088]
51. Schulz R, Beach SR. Caregiving as a risk factor for mortality: the Caregiver Health Effects Study. JAMA 1999 Dec 15;282(23):2215-2219. [doi: 10.1001/jama.282.23.2215] [Medline: 10605972]
52. Nguyen AD, Frensham LJ, Wong MX, Meslin SM, Martin P, Lau AY, et al. mHealth App Patient Testing and Review of Educational Materials Designed for Self-Management of Gout Patients: Descriptive Qualitative Studies. JMIR Mhealth Uhealth 2018 Oct 15;6(10):e182 [FREE Full text] [doi: 10.2196/mhealth.9811] [Medline: 30322835]
53. Gordon JS, Armin JS, Cunningham JK, Muramoto ML, Christiansen SM, Jacobs TA. Lessons learned in the development and evaluation of RxCoach™, an mHealth app to increase tobacco cessation medication adherence. Patient Educ Couns 2017 Apr;100(4):720-727 [FREE Full text] [doi: 10.1016/j.pec.2016.11.003] [Medline: 27839891]
54. Brown-Johnson CG, Berrean B, Cataldo JK. Development and usability evaluation of the mHealth Tool for Lung Cancer (mHealth TLC): a virtual world health game for lung cancer patients. Patient Educ Couns 2015 Apr;98(4):506-511 [FREE Full text] [doi: 10.1016/j.pec.2014.12.006] [Medline: 25620075]

Abbreviations

EHR: electronic health record
IRB: Institutional Review Board
ObsRO: observer-reported outcome
PAC: Patient Advocacy Council

XSL•FO
RenderX

PRO: patient-reported outcome
PTAG: Project Team Advisory Group
SCP: survivorship care plan

Edited by G Eysenbach; submitted 20.07.20; peer-reviewed by M Bass, D Roosan, J Li; comments to author 20.08.20; revised version received 13.10.20; accepted 18.06.21; published 13.08.21.

Please cite as:
Oakley-Girvan I, Davis SW, Kurian A, Rosas LG, Daniels J, Palesh OG, Mesia RJ, Kamal AH, Longmire M, Divi V
Development of a Mobile Health App (TOGETHERCare) to Reduce Cancer Care Partner Burden: Product Design Study
JMIR Form Res 2021;5(8):e22608
URL: https://formative.jmir.org/2021/8/e22608
doi:10.2196/22608
PMID:34398787

©Ingrid Oakley-Girvan, Sharon Watkins Davis, Allison Kurian, Lisa G Rosas, Jena Daniels, Oxana Gronskaya Palesh, Rachel J Mesia, Arif H Kamal, Michelle Longmire, Vasu Divi. Originally published in JMIR Formative Research (https://formative.jmir.org), 13.08.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Understanding Physicians' Preferences for Telemedicine During the COVID-19 Pandemic: Cross-sectional Study

Sarah Nies[1], BA; Shae Patel[1], BSc; Melissa Shafer[1], BSc; Laura Longman[1], BSc; Iman Sharif[1], MPH, MD; Paulo Pina[1], MPH, MD

Family Health Centers at New York University Langone, Brooklyn, NY, United States

Corresponding Author:
Paulo Pina, MPH, MD
Family Health Centers at New York University Langone
5800 3rd Avenue
Brooklyn, NY, 11220
United States
Phone: 1 2153474369
Email: paulo.pina@nyulangone.org

Abstract

Background: In contrast to the current broad dissemination of telemedicine across medical specialties, previous research focused on the effectiveness of telemedicine in special populations and for behavioral health encounters, demonstrating that both physician and patient factors impact the efficacious use of telemedicine.

Objective: We aim to evaluate physician perceptions of the appropriateness of telemedicine for patients attending the primary care practices of a federally qualified health center in New York City.

Methods: We used an anonymous cross-sectional survey including closed- and open-ended questions. We used chi-square to test whether providers from certain specialties were more likely to state they would use telemedicine in the future. We used *t* tests to compare age between those who would versus would not use telemedicine. We then used logistic regression to test whether age and specialty were both correlated with the desire to use telemedicine in the future. We used thematic content analysis to describe the reasons providers felt they would not want to use telemedicine in the future and to describe the situations for which they felt telemedicine would be appropriate.

Results: Of 272 health care providers who were sent the electronic survey, 157 (58%) responded within the 2-week survey time frame. The mean age of providers was 45 (range 28-75) years. Overall, 80% (126/157) stated they would use telemedicine in the future. Compared to the family medicine, internal medicine, behavioral health, dental, and obstetrics and gynecology specialties, providers from pediatrics, med-peds, subspecialties, and surgery (protelemedicine specialties) were more likely to believe telemedicine would be useful post pandemic (61/67 [91%] vs 65/90 [72%]; $P<.001$). Providers who reported they would use telemedicine in the future were younger (mean age 44, range 42-46 years vs mean age 50, range 46-55 years; $P=.048$). In the regression analysis, both protelemedicine specialties and age were significantly associated with odds of reporting they would use telemedicine in the future (prospecialties: odds ratio 5.2, 95% CI 1.7-16.2; younger age: odds ratio 1.05, 95% CI 1.01-1.08). Providers who did not want to use telemedicine in the future cited concerns about inadequate patient care, lack of physical patient interaction, technology issues, and lack of necessity. Providers who felt telemedicine would be useful cited the following situations: follow-up visits, medication refills, urgent care, patient convenience, and specific conditions such as behavioral health, dermatology visits, and chronic care management.

Conclusions: The majority of health providers in this resource-limited setting in a federally qualified health center believed that telemedicine would be useful for providing care after the pandemic is over.

(JMIR Form Res 2021;5(8):e26565) doi:10.2196/26565

KEYWORDS

telemedicine; federally qualified health care center; primary care; COVID-19; telehealth; physician; doctor; preference; perspective; dissemination; appropriate; cross-sectional; survey

XSL•FO
RenderX

Introduction

Telemedicine instantly became the preferred, and for many *only*, mechanism for health care delivery in New York City during the COVID-19 pandemic [1]. Health care institutions quickly established a variety of strategies to deliver telemedicine services using audio-video or audio only platforms compatible with the Health Insurance Portability and Accountability Act (HIPAA) to provide patient's access to their providers [2-4].

In contrast to the current broad dissemination of telemedicine across medical specialties, prior research focused on the effectiveness of telemedicine mostly in specific populations [5,6] and for behavioral health encounters [7-10]. Research shows that optimal and efficacious use of telemedicine requires willingness of both the physician and patient to engage on these nontraditional platforms [9]. Physician satisfaction and preference for telemedicine, however, has not been studied as abundantly, especially after the emergence of COVID-19 [10,11]. Although physician personality (ie, judging vs perceiving) and preference for telemedicine demonstrate some correlation, there are few studies on the association between physician age or specialty with physician preference to use telemedicine for clinical practice [11].

This study qualitatively and quantitatively evaluates physician preferences regarding the use of telemedicine for patients in a large federally qualified health system in Brooklyn, New York. We hypothesized that younger physicians and physicians who provide behavioral health services would be more likely to cite telemedicine as an appropriate and preferred modality of care post pandemic. This hypothesis was formed on the assumption that younger physicians would be more familiar with nontraditional technology platforms and the assumption that behavioral health care service does not require physical assessments. With our qualitative data, it is also our hope to explore and to identify any reservations or shortcomings they may have, in efforts to provide insight into the use of telemedicine as an efficient means of providing quality care in the future.

Methods

Design

We devised a unique and anonymous cross-sectional electronic survey for this project to collect qualitative and quantitative data. We surveyed health care providers working for a large federally qualified health care system based in Brooklyn, New York that is comprised of 8 primary care practices (medicine, pediatrics, obstetrics and gynecology [OB/GYN], behavioral health), 6 dental clinics, 9 community medicine sites, and 52 school-based health centers. The study was categorized as exempt research by the New York University (NYU) Institutional Review Board.

Beginning in March 2020, providers had the option of using either Webex, Doximity, or MyChart to deliver telemedicine visits to patients. All three of these telemedicine modalities are compliant with HIPAA [3,4]. Webex appointments were scheduled by practice registration staff, and patients were sent emails with instructions on how to log in to the appointment. To use Doximity, providers individually signed up for the service and downloaded the app. They could then send a text message to a patient's cell phone number asking them to join a video call, or they could directly call the patient's phone number to conduct an audio-only visit. Anyone enrolled in the patient portal, Mychart, could access the visit through that application. All providers were encouraged to conduct audio-video visits over audio-only visits if possible.

Survey Procedures

Based on previous studies, we created a brief survey and emailed a survey web link to all providers in May 2020, approximately 2 months into the COVID-19 pandemic in Brooklyn, New York. Providers provided consent and received several reminders to complete the survey over a 2-week period.

In the survey, physicians were asked which telemedicine platforms they used since the beginning of the pandemic in March 2020. They then were asked in the survey to indicate "yes" or "no" to whether they would like to use any of the telemedicine platforms routinely for patient care if the pandemic was over. If the physician responded "yes," they were asked which of the platforms they would want to use going forward. If the physician responded "no," they were prompted to explain "why not" and if there were exceptions as to when telemedicine would be useful in patient care. To assess survey consistency, Conbrach alpha was calculated with an acceptable score of .63. To further assess and identify common themes of physician preference, another free writing prompt within the survey asked the physicians to identify what specific patient care situations they felt telemedicine would be most helpful to use.

To assess physician age preference and specialty preference, the survey asked the physician to fill in their age and specialty. The listed specialties included pediatrics, family medicine, internal medicine, med-peds, behavioral health, dental, OB/GYN, and other. If "other" was selected, the participant was asked to describe the specialty.

Analyses

We tabulated descriptive statistics for all survey participants. Missing data for age was imputed at the mean value. We used chi-square to test whether providers from certain specialties were more likely to state they would use telemedicine in the future. We used *t* tests to compare age between those who would and those who would not use telemedicine. We then used logistic regression to test whether age and specialty were both correlated with desire to use telemedicine in the future. For this analysis, we combined specialties that were more likely to state they would use telemedicine in the future into one binary variable, "pro-telemedicine specialties."

To better understand the reasons why providers would or would not use telemedicine in the future, we used thematic content analysis to describe the themes from the open-ended responses. All open-ended responses were read and coded separately by two of the authors who then compared notes and, after discussion with the senior author (IS), came to a consensus of thematic groupings.

XSL•FO
RenderX

Results

Of 272 health care providers who were sent the electronic survey, 157 (58%) responded within the 2-week survey time frame. Demographics and survey responses are shown in Table 1.

There was a statistically significant difference in preference to use telemedicine in the future by specialty. Compared to the family medicine, internal medicine, behavioral health, dental, and OB/GYN specialties, providers from pediatrics, med-peds, medical subspecialties, and surgery (protelemedicine specialties) were more likely to believe telemedicine would be useful post pandemic (61/67 [91%] vs 65/90 [72%]; $P<.001$). Furthermore, med-peds, pediatrics, and medical subspecialties had the highest percentage of "will use telemedicine in the future" with 100% (2/2), 94% (34/36), and 85% (17/20), respectively. OB/GYN and internal medicine had the lowest percentage of "will use telemedicine in the future" with OB/GYN at 50% (4/8) and internal medicine at 63% (17/27).

Providers who reported they would use telemedicine in the future were younger (mean age 44, range 42-46 years vs mean age 50, range 46-55 years; $P=.048$). In the regression analysis, both protelemedicine specialties and younger age were significantly associated with increasing odds of reporting that they would use telemedicine in the future (protelemedicine specialties: odds ratio 5.2, 95% CI 1.7-16.2; younger age: odds ratio 1.05, 95% CI 1.01-1.08).

Table 1. Descriptive statistics for the survey sample.

Variable	Participants (N=157)
Age (years), mean (SD)	45.7 (11.5)
Specialty, n (%)	
Pediatrics	36 (22.9)
Family medicine	26 (16.5)
Internal medicine	27 (17.1)
Med-peds	2 (1.3)
Behavioral health	9 (5.7)
Dental	20 (12.7)
OB/GYN[a]	8 (5.0)
Surgery	9 (5.7)
Medical subspecialties	20 (12.7)
Modalities deemed effective (multiple responses allowed; n=309), n (%)	
Telephone	70 (22.7)
Doximity audio + video	82 (26.5)
Doximity audio only	32 (10.4)
Webex audio + video	59 (19.1)
Webex audio only	18 (5.8)
Mychart audio + video	29 (9.4)
None Effective	1 (0.3)
Did not use	18 (5.8)
Would use in the future (multiple response allowed; n=126), n (%)	
Telephone	36 (28.6)
Doximity audio + video	73 (57.9)
Doximity audio only	13 (10.4)
Webex audio + video	41 (32.5)
Webex audio only	6 (4.8)
Mychart audio + video	50 (39.7)
Other (Zoom, Facetime, Whatsapp etc.)	4 (3.2)

[a]OB/GYN: obstetrics and gynecology

XSL•FO
RenderX

There were 22 open-ended responses to the question "why not?" for respondents who said they would not want to use telemedicine once the pandemic is over, citing concerns about inadequate patient care, lack of physical patient interaction, technology issues, and lack of necessity. The responses were divided into thematic grouping, as seen in Table 2. Concerns included the inability "to perform physical examination" and technology issues such as the "time [required] for the provider to connect." Other statements were focused on the necessary use of telemedicine during a pandemic, as one of the participants said "I will like to have this technology as an option for certain circumstances but not as a routine way to provide patient care."

Table 2. Themes for open-ended responses for why providers would not use telemedicine once the pandemic is over.

Thematic category	Responses (n=22), n
Patient care (diagnosis, vitals, physical exam, labs)	9
Lack of physical patient interaction	6
Technology issues	4
Necessity (use *only* during a pandemic)	2

There were 151 open-ended responses to the question "For which situations do you think telemedicine would be useful?" The responses were categorized under the following themes: follow-up visits, medication refills, urgent care, patient convenience, psychiatric complaints, dermatology complaints, and chronic care management (Table 3). Some situations were specialty specific:

> *MFM [maternal fetal-medicine] consultations, other consultations that do not require a physical exam. Follow up prenatal visits that do not require a physical exam (i.e. lab review)*

Other suggestions were patient population specific, for example:

> *Patients with mobility issues. Patients who can't come in easily for different reasons. Patients who frequently No-show. Patients who can't get transportation easily. Patients who have caregivers with them who can be together.*

Other suggestions fit more general clinical management such as:

> *Follow up [visits] to discuss test results; check in for medications refill requests; all other "I want to speak to my doctor" situations should be routinely "web" appointments and should be billable and compensated as they all take time and effort*

> *Non-annual visits for routine follow up that do not require a physical exam. examples- responses to medication initiation/titration, lab results, medication refills, Diabetes f/u that are less than 3-4 months.*

Table 3. Thematic coding results for situations believed to be useful for telemedicine after the pandemic is over.

Thematic category	Responses (n=151), n (%)
Follow-up/lab result visits	47 (31)
Medication refill visits	26 (17)
Urgent care/acute symptom triage	24 (16)
Patient convenience (eg, no transportation, includes older adults)	17 (11)
Psych complaints	14 (9)
Dermatology complaints	13 (9)
Chronic care management (patient can self-report HgA_{1c}[a], BP[b], lifestyle, symptoms)	10 (7)

[a]HgA_{1c}: glycated hemoglobin.

[b]BP: blood pressure.

Discussion

The majority of health providers in this resource-limited setting of a federally qualified health center believed that telemedicine would be useful for providing care after the pandemic is over.

As previous studies have demonstrated [12,13], we also found that older providers are also less likely to prefer telemedicine. Our data show that health care providers older than 60 years were more likely to discontinue use of telemedicine post COVID-19 compared to those younger than 60 years. The reasons for this preference among physicians older than 60 years are unclear and can be an area for further research to identify specific barriers.

Our findings also suggest there is a significant difference among specialty providers in relation to telemedicine preference. The specialty with the highest number of providers willing to continue telemedicine use post COVID-19 was pediatrics. This comes as no surprise considering there is a substantial body of research outlining the benefits and considerations for using telemedicine in the pediatric setting [14,15]. Responses from the qualitative data set indicated *ease of follow-up* as the most common reason for continued use of virtual visits among pediatricians. Conversely, internal medicine had the highest

XSL•FO
RenderX

number of providers unwilling to continue its use in the future compared to other specialties. This was surprising because of the focus on history, imaging, and laboratory findings involved in the exam and diagnostic process of internal medicine primary care.

Most past research has focused on patient preference and outcomes of care [9,12,16-21] rather than provider inclination. Studies suggest that telemedicine has been accepted more by patients than by providers [22], with providers citing technological barriers to care provision [13,23-25]. Interestingly, the most commonly cited reason for not continuing use among our cohort, as revealed in the open-ended question responses, was lack of fundamental patient interaction required for health care, such as vitals and certain physical exams. In medicine, providers pride themselves on creating therapeutic relationships based on sitting in the same room with a person. Although there is an intuitive feeling of what a therapeutic relationship feels like, few studies have examined whether and how physical senses (eg, touch or eye gaze) enhance the therapeutic relationship [25]. Without the ability to interact with a patient physically and apply their nuanced senses, physicians in our cohort were less likely to prefer telemedicine as compared to in-person patient interactions. In addition, although the management of several major chronic conditions such as diabetes, heart disease, and chronic obstructive pulmonary disease have been shown to be adequately treated via telemedicine, it would be useful to have follow-up studies that specifically identify which patient populations and diseases physicians found this type of interaction particularly critical for [18,24].

Although the response rate in this survey was strong for a 2-week time frame, the study represents a snapshot in time with data from one system; the preferences of providers in this setting may not be generalizable to other institutions and settings, and they may also evolve over time. Our sample size was small across all specialties and different preferences may be discovered in a larger population. Our data was also limited to physicians within the NYU health care system, and results may reflect the biases of urban communities. Another possibility is that surveyees had differing interpretations of our question regarding willingness to continue telemedicine use in the future. It is possible that opinions would change if it was more precisely worded to be a supplement to a practice rather than an exclusive option as often required during the COVID-19 pandemic. The qualitative data in this study can guide researchers and practice leaders to work with providers to optimize the use of telemedicine in health care going forward, presenting a small silver lining to the COVID-19 pandemic.

Telemedicine as a method to provide patient care has a wide array of implications that can drastically shape the future of health care. Patients have expressed high satisfaction rates when engaging in technology-based health care interactions [9]. Understanding the reservations of medical professionals, based on age and specialty, can lead to improvements that address their concerns and expand the use of telemedicine in practice. The thematic issues described in survey responses of this study can be expanded upon for future research to help deliver more efficient and advantageous telemedicine delivery.

Conflicts of Interest

None declared.

References

1. AJMC Staff. A timeline of COVID-19 developments in 2020. AJMC. 2021 Jan 01. URL: https://www.ajmc.com/view/a-timeline-of-covid19-developments-in-2020 [accessed 2020-10-25]
2. Using telehealth to expand access to essential health services during the COVID-19 pandemic. Centers for Disease Control and Prevention. 2020 Jun 10. URL: https://www.cdc.gov/coronavirus/2019-ncov/hcp/telehealth.html [accessed 2020-10-25]
3. Telehealth services. Doxy.me. URL: https://doxy.me/en/features/ [accessed 2020-10-25]
4. Exceptional telehealth experiences. Webex by Cisco. URL: https://www.webex.com/industries/healthcare.html [accessed 2020-10-25]
5. Fact sheet: telehealth. American Hospital Association. URL: https://www.aha.org/factsheet/telehealth [accessed 2020-10-25]
6. Kruse CS, Mileski M, Moreno J. Mobile health solutions for the aging population: a systematic narrative analysis. J Telemed Telecare 2017 May;23(4):439-451 [FREE Full text] [doi: 10.1177/1357633X16649790] [Medline: 27255207]
7. Cartwright M, Hirani SP, Rixon L, Beynon M, Doll H, Bower P, Whole Systems Demonstrator Evaluation Team. Effect of telehealth on quality of life and psychological outcomes over 12 months (Whole Systems Demonstrator telehealth questionnaire study): nested study of patient reported outcomes in a pragmatic, cluster randomised controlled trial. BMJ 2013 Feb 26;346:f653 [FREE Full text] [doi: 10.1136/bmj.f653] [Medline: 23444424]
8. Choi NG, Hegel MT, Marti N, Marinucci ML, Sirrianni L, Bruce ML. Telehealth problem-solving therapy for depressed low-income homebound older adults. Am J Geriatr Psychiatry 2014 Mar;22(3):263-271 [FREE Full text] [doi: 10.1097/JGP.0b013e318266b356] [Medline: 23567376]
9. Emerson JF, Welch M, Rossman WE, Carek S, Ludden T, Templin M, et al. A multidisciplinary intervention utilizing virtual communication tools to reduce health disparities: a pilot randomized controlled trial. Int J Environ Res Public Health 2015 Dec 22;13(1):ijerph13010031 [FREE Full text] [doi: 10.3390/ijerph13010031] [Medline: 26703661]

XSL•FO
RenderX

10. Breen P, Murphy K, Browne G, Molloy F, Reid V, Doherty C, et al. Formative evaluation of a telemedicine model for delivering clinical neurophysiology services part I: utility, technical performance and service provider perspective. BMC Med Inform Decis Mak 2010 Sep 15;10:48 [FREE Full text] [doi: 10.1186/1472-6947-10-48] [Medline: 20843309]
11. Miner H, Fatehi A, Ring D, Reichenberg JS. Clinician telemedicine perceptions during the COVID-19 pandemic. Telemed J E Health 2021 May;27(5):508-512. [doi: 10.1089/tmj.2020.0295] [Medline: 32946364]
12. Mayworm AM, Lever N, Gloff N, Cox J, Willis K, Hoover SA. School-based telepsychiatry in an urban setting: efficiency and satisfaction with care. Telemed J E Health 2020 Apr;26(4):446-454. [doi: 10.1089/tmj.2019.0038] [Medline: 31120378]
13. Scott Kruse C, Karem P, Shifflett K, Vegi L, Ravi K, Brooks M. Evaluating barriers to adopting telemedicine worldwide: a systematic review. J Telemed Telecare 2018 Jan;24(1):4-12 [FREE Full text] [doi: 10.1177/1357633X16674087] [Medline: 29320966]
14. Narasimha S, Madathil KC, Agnisarman S, Rogers H, Welch B, Ashok A, et al. Designing telemedicine systems for geriatric patients: a review of the usability studies. Telemed J E Health 2017 Jun;23(6):459-472. [doi: 10.1089/tmj.2016.0178] [Medline: 27875667]
15. Burke BL, Hall RW, Section of Telehealth Care. Telemedicine: pediatric applications. Pediatrics 2015 Jul;136(1):e293-e308 [FREE Full text] [doi: 10.1542/peds.2015-1517] [Medline: 26122813]
16. Utidjian L, Abramson E. Pediatric telehealth: opportunities and challenges. Pediatr Clin North Am 2016 Apr;63(2):367-378. [doi: 10.1016/j.pcl.2015.11.006] [Medline: 27017042]
17. Soegaard Ballester JM, Scott MF, Owei L, Neylan C, Hanson CW, Morris JB. Patient preference for time-saving telehealth postoperative visits after routine surgery in an urban setting. Surgery 2018 Apr;163(4):672-679. [doi: 10.1016/j.surg.2017.08.015] [Medline: 29398042]
18. Batbaatar E, Dorjdagva J, Luvsannyam A, Savino MM, Amenta P. Determinants of patient satisfaction: a systematic review. Perspect Public Health 2017 Mar;137(2):89-101. [doi: 10.1177/1757913916634136] [Medline: 27004489]
19. Ekeland AG, Bowes A, Flottorp S. Effectiveness of telemedicine: a systematic review of reviews. Int J Med Inform 2010 Nov;79(11):736-771. [doi: 10.1016/j.ijmedinf.2010.08.006] [Medline: 20884286]
20. Polinski JM, Barker T, Gagliano N, Sussman A, Brennan TA, Shrank WH. Patients' satisfaction with and preference for telehealth visits. J Gen Intern Med 2016 Mar;31(3):269-275 [FREE Full text] [doi: 10.1007/s11606-015-3489-x] [Medline: 26269131]
21. Tasneem S, Kim A, Bagheri A, Lebret J. Telemedicine video visits for patients receiving palliative care: a qualitative study. Am J Hosp Palliat Care 2019 Sep;36(9):789-794. [doi: 10.1177/1049909119846843] [Medline: 31064195]
22. Polisena J, Coyle D, Coyle K, McGill S. Home telehealth for chronic disease management: a systematic review and an analysis of economic evaluations. Int J Technol Assess Health Care 2009 Jul;25(3):339-349. [doi: 10.1017/S0266462309990201] [Medline: 19619353]
23. Bashshur RL, Howell JD, Krupinski EA, Harms KM, Bashshur N, Doarn CR. The empirical foundations of telemedicine interventions in primary care. Telemed J E Health 2016 May;22(5):342-375 [FREE Full text] [doi: 10.1089/tmj.2016.0045] [Medline: 27128779]
24. Brooks E, Turvey C, Augusterfer EF. Provider barriers to telemental health: obstacles overcome, obstacles remaining. Telemed J E Health 2013 Jun;19(6):433-437. [doi: 10.1089/tmj.2013.0068] [Medline: 23590176]
25. Kerr F, Wiechula R, Feo R, Schultz T, Kitson A. Neurophysiology of human touch and eye gaze in therapeutic relationships and healing: a scoping review. JBI Database System Rev Implement Rep 2019 Feb;17(2):209-247 [FREE Full text] [doi: 10.11124/JBISRIR-2017-003549] [Medline: 30730854]

Abbreviations

HIPAA: Health Insurance Portability and Accountability Act
MFM: maternal fetal-medicine
NYU: New York University
OB/GYN: obstetrics and gynecology

Edited by G Eysenbach; submitted 16.12.20; peer-reviewed by C Doarn, D Alden; comments to author 10.01.21; revised version received 30.01.21; accepted 31.05.21; published 13.08.21.

Please cite as:
Nies S, Patel S, Shafer M, Longman L, Sharif I, Pina P
Understanding Physicians' Preferences for Telemedicine During the COVID-19 Pandemic: Cross-sectional Study
JMIR Form Res 2021;5(8):e26565
URL: https://formative.jmir.org/2021/8/e26565
doi:10.2196/26565
PMID:34227993

XSL•FO
RenderX

©Sarah Nies, Shae Patel, Melissa Shafer, Laura Longman, Iman Sharif, Paulo Pina. Originally published in JMIR Formative Research (https://formative.jmir.org), 13.08.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Copyright as a Barrier to Music Therapy Telehealth Interventions: Qualitative Interview Study

Amanda Reid[1*], PhD, JD; Alex Kresovich[1*], MA

Hussman School of Journalism and Media, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[*]all authors contributed equally

Corresponding Author:
Amanda Reid, PhD, JD
Hussman School of Journalism and Media
University of North Carolina at Chapel Hill
393 Carroll Hall
Campus Box 3365
Chapel Hill, NC, 27599
United States
Phone: 1 9044828702
Email: areid@unc.edu

Abstract

Background: Music therapy is a multifaceted discipline that harnesses the power of music to treat a wide range of patient populations. A therapist who plays music in a private room for a patient is not subject to copyright restrictions on public performances. However, in the wake of the COVID-19 pandemic, music therapy is no longer strictly confined to the face-to-face setting. This study explores music therapists' perceptions of copyright law with respect to their ability to provide mediated services to their clients.

Objective: The objectives of our study were two-fold. The first was to investigate whether concerns about copyright law are hampering the diffusion of telehealth innovations, and the second was whether these concerns are causing music therapists to avoid therapeutically beneficial telehealth interventions.

Methods: Semistructured interviews were conducted with credentialed music therapists (n=18) in the United States between May 2020 and June 2020. With participants' consent, we used video conference technology to record and transcribe the in-depth interviews. The median interview length was 45 (SD 16.37) minutes. This theoretically informed study employed thematic analysis of the interview data.

Results: The COVID-19 pandemic accelerated the adoption of telehealth interventions to facilitate therapy outside of private face-to-face environments: environments where music therapy practices are largely shielded from copyright infringement concerns. Five main themes emerged, including therapists' uncertainty about permissible uses of music and therapists' erring on the side of caution causing lost opportunities for care. Our interview data suggest music therapists have altered telehealth interventions in suboptimal ways to avoid copyright liability in a physically distanced environment.

Conclusions: Some music therapists "drag their feet" on offering therapeutically appropriate telehealth services to clients because of copyright concerns. Our findings suggest innovative mediated therapies were shied away from or abandoned. These findings offer a novel contribution to the public health literature by highlighting copyright law as an unexpected and unwelcome barrier to the diffusion of music therapy practices in technology-mediated settings.

(JMIR Form Res 2021;5(8):e28383) doi:10.2196/28383

KEYWORDS

telehealth; music therapy; diffusion of innovations; COVID-19; copyright law; fair use

Introduction

Music therapy is an evidence-based discipline that uses music to treat a wide range of physiological and psychological conditions and a variety of patient populations [1-10]. The clinical use of musical interventions to achieve therapeutic goals is an increasingly popular telehealth service [11], and the practice benefits millions of youth and adults annually [12]. Telehealth is a vital tool, which is often used to reach

XSL•FO
RenderX

underserved patients in rural and urban settings [13]. Preliminary research suggests music therapy via telehealth improves access to care and community music engagement [14,15]. Research also suggests web-based services are promising avenues to increase mental health awareness and treatment options [16,17]. For those practicing via mediated technologies, there are several freely available and fee-based platforms to choose from, including Doxy.me (Doxy.me LLC, Rochester, NY), Microsoft Teams (Microsoft Corporation, Redmond, WA), Skype for Business (Microsoft Corporation), WebEx (Webex by Cisco, Milpitas, CA), and Zoom (Zoom Video Communications Inc, San Jose, CA). These platforms offer real-time video and audio communication between patients and providers [18].

Depending on the client's needs, a music therapist can create a treatment plan that includes singing, performing, creating, or listening to music. The varied uses of music can include active music listening, song lyric analysis, improvised music playing, and songwriting. It is well established that patient-preferred music yields superior results across various therapeutic interventions [19,20]. Many of the different modes of music therapy intervention involve playing or performing patient-preferred music — music that is often subject to copyright protection [21].

For copyrighted music, rightsholders have several exclusive rights; an exclusive right means the right to exclude others from specified uses of a copyrighted work [22]. Absent an exemption or limitation (eg, fair use [22]), a rightsholder has the exclusive right to make copies of the work, create derivatives of the work, perform the work publicly, and distribute copies of the work publicly. Notably, not all performances of a copyrighted work are proscribed; purely private performances of copyrighted music are not within the rightsholder's exclusive rights [23]. A legal opinion letter advised the American Music Therapy Association (AMTA) that the use of copyrighted music in face-to-face therapy is not infringing: "A hospital room is private. Hence, one-on-one patient-therapist work within a hospital room is not the kind of setting where performing music or playing music in the form of a sound recording will result in copyright infringement" [24]. In other words, playing copyrighted music is freely permitted when it is not in a public setting. A therapist who plays music in a private room for a patient does not infringe the rightsholder's exclusive right to perform the work publicly. However, music therapy is no longer strictly confined to the face-to-face setting.

The COVID-19 pandemic catalyzed the rapid adoption of telehealth services to safely deliver care while limiting the risk of exposure to contagion inherent in a face-to-face setting. To facilitate more comprehensive access to physically distanced health care options, the federal government relaxed Health Insurance Portability and Accountability Act (HIPAA) compliance guidelines for telehealth [25]. To further encourage telehealth services, the AMTA issued a statement early in the pandemic clarifying that it "supports the use of telehealth as a means to provide music therapy interventions when beneficial to clients" [26]. Despite the advantages of telehealth, music therapists may be less comfortable using copyrighted music for fear of infringement when using mediated technologies for therapy [27]. The lack of clarity on copyright law in the telehealth space threatens the interventions of those who need physical and mental health care via music therapy. Copyright enforcement efforts can come directly from rightsholders or via automated tools that use algorithms to identify infringing content [28]. The public health literature identifies barriers to real-time telehealth interventions, including a lack of access to basic internet services, inadequate internet speed for audio and video quality, resistance to technology (by patients and providers), and health insurance reimbursement issues [29,30]. However, this literature fails to account for another barrier to telehealth: the threat of copyright enforcement against music therapy interventions that are conducted by mediated communication technologies.

While there is scant literature discussing copyright as a barrier to telehealth, one researcher has captured a snapshot of music therapists' concerns about copyright. To study telehealth music therapy during the COVID-19 crisis, Carvajal [27] conducted a content analysis of Facebook posts collected from a music therapy group, *Music Therapists Unite*. These posts were collected between January 2020 and April 2020. Three of the music therapists' posts illustrate uncertainty about the permissibility of therapeutic uses of music under copyright law. One therapist queried the group:

> *Copyright Question: If I made a recording of myself singing more modern songs and it is a private playlist that only those with the link can see, would that be okay?*

Another asked:

> *I did a Facebook live group called household rhythms where we jammed out using items found around the room to popular songs. I got flagged on Facebook as a copyright [infringer] for one of the songs. It was my understanding if it was for Educational (sic) purposes it was okay to provide recorded music. Can someone offer me some insight as they have asked me to do it again weekly, and I don't want to get them or myself in trouble.*

And lastly, a music therapist asked the Facebook group:

> *If I wanted to use pre-recorded music for an intervention, can I play that through my phone/speakers at home without getting a copyright strike on Youtube (sic)?*

These questions reflect the uncertainty and unease that music therapists have when using copyrighted music to serve clients in virtual spaces. These questions also reflect therapists' desire to stay out of trouble — copyright trouble either at the hands of rightsholders or algorithmic enforcement.

We hypothesize that such copyright worries may cause music therapists to deviate from preferred treatments. Copyright concerns may not only affect treatment modes but also discourage the delivery of telehealth services altogether. In other words, there is the potential for demonstrable "delays, deformations, and failure to execute mission" [31] in the offering of remote music therapy because of copyright uncertainty. Anxiety and uncertainty about copyright risks further exacerbate existing treatment gaps to those in need [32,33].

XSL•FO
RenderX

According to Rogers' [34] Diffusion of Innovations theory, an innovation needs wide adoption to self-sustain. And, individuals need clear information about the advantages and disadvantages of adoption. However, the excerpts from Carvajal's [27] content analysis of the music therapist Facebook group suggest that copyright concerns may hamper telehealth innovations. The early adopters of music therapy telehealth are entering uncharted legal territory. It is unclear whether it is permissible to use copyrighted music in technology-mediated music therapy sessions. We hypothesize that legal uncertainty — rather than uncertainty about telehealth's practical applications — may be an obstacle to the adoption of this innovation in health care delivery.

Methods

Design Strategy

This research adopted a qualitative design using an interpretive description approach, which is appropriate for the study of applied health disciplines [35,36]. Ethical review was performed by the Office of Human Research Ethics at the University of North Carolina at Chapel Hill, and it was determined to be exempt from further review (exemption reference ID 261072). Our interviews were conducted in the United States during the summer of 2020.

The aim of this study was to interrogate the effects of legal uncertainty on the adoption and diffusion of telehealth innovations. Researchers have long recognized that knowledge is contingent upon human practices and social context [37,38]. Meaning is constructed as humans engage with and interpret their sociocultural contexts [39]. The flexibility of interpretive description design permits identification of commonalities of experiences, while also recognizing individual variation [40,41]. The unique experiences of participants are situated within broader patterns of the phenomenon of study [42].

The epistemological perspective of this project views knowledge as socially constructed, and this perspective aligned positively with the phenomena of study. US copyright law is an instrumental doctrine designed to encourage the production and dissemination of creative works.

The lead researcher's ontological and epistemological assumptions influenced the research strategies of this project. Specifically, the philosophical basis of this study was guided by the lead researcher's legal training, coupled with her expertise in US copyright law. Consistent with a constructionist perspective, when situating oneself within the research process, it must be acknowledged that a researcher does not approach a study objectively. This reflexivity acknowledges that the lead researcher's expertise introduced the risk of bias.

To enhance trustworthiness and rigor, our findings were confirmed through 2 methods of triangulation [43,44]. First, participants with a range of experience contributed to the data sources of this study. This data triangulation contributed to a robust and comprehensive understanding of phenomena. Second, an independent communication researcher collaborated in the thematic analysis of the data and provided an unbiased perspective. Both researchers equally contributed to the interpretive thematic analysis of the data. This analytical triangulation contributed independent and credible corroboration of the thematic findings. Moreover, verbatim quotes are included below to offer a clear audit trail of evidence to support these thematic findings.

Participants

The inclusion criterion was credentialed music therapists working in the United States. Participants were identified using a combination of purposive and snowball sampling approaches. A purposive sampling approach was used to ensure that those recruited represented clinical faculty as well as practitioners and those employed at hospital facilities as well as those self-employed. Participants were initially recruited from a list of music therapists provided to the lead author by a representative at the AMTA. The lead researcher used referrals from these initial interviewees' networks and then recruited more interviewees via snowball sampling. Finally, some interviewees were recruited using contact information obtained using internet searches for qualified participants. Through these various recruitment and sampling methods, 18 credentialed music therapists met the criterion for an interview.

Data Collection

All 18 interviews were conducted by the lead researcher between May 19, 2020 and June 19, 2020 and were recorded using Zoom video conferencing. These in-depth, semistructured interview sessions produced over 13 hours of interview data, and each interview averaged 45 minutes in length (SD 16.37 minutes), ranging from 23 minutes to 97 minutes. At the beginning of each interview session, an institutional review board consent script was presented to the participants visually (ie, shared on the computer screen) and aurally (ie, read aloud). Participation was completely voluntary, and consent could be withdrawn at any time. With participants' consent, the one-on-one interviews were digitally recorded, and the interviews were transcribed by the video conference software. At the end of each interview, participants were offered a US $50 electronic gift card.

The digital recordings and transcripts were uploaded to a secure university-issued computer and then deleted from the video conference platform. The lead researcher anonymized the transcripts, and then a graduate assistant reviewed the transcripts against the audio-only recordings to ensure accuracy. Any questions about grammar or spelling were resolved by the lead researcher. Only the lead author knows the identity of the participants.

Interview sessions began with a topic guide exploring participants' general experiences as music therapists and the types of therapeutic services they provided. Follow-up questions asked in what particular ways they used music with their clients and why they might use copyrighted music. Participants were then asked how their practices were affected by the COVID-19 pandemic and whether their comfort with using copyrighted music was altered in telehealth settings. An interview schedule initially guided the questioning, and ad hoc follow-up questions were used to further explore salient points.

XSL•FO
RenderX

Data Analysis

The anonymized and proofed interview transcripts were analyzed using the coding approach by MacQueen et al [45]. Focusing on our hypothesis about the effect of legal uncertainty on telehealth services, we collaborated on codebook development, and we each independently highlighted significant quotes from the interview transcripts [46]. In the coding process, we identified inductive categories through thematic analysis [47]. Through collaboration and iterative refinement, we identified coding patterns reflecting the data's main themes about telehealth services [48]. Verbatim quotes are included because they richly capture music therapists' attitudes about copyright. These thick descriptions not only serve to increase conformability in the data analysis process but also reduce the researchers' influence in the data analysis. Moreover, quotes are attributed to the various participants to illustrate that these emergent themes were consistent from more than one interviewee. With these findings, this study aims to expand the public health literature by revealing an unexpected barrier to the diffusion of innovative telehealth practices.

Results

Overview

In total, 18 credentialed music therapists met the criterion for in-depth, semistructured interviews. These participants averaged over 18 years of experience in the field. When asked to self-describe their level of expertise most (14/18, 78%) identified as expert; 3 identified as intermediate (17%), and 1 identified as novice (5%). Most of the participants were female (15/18, 83%). Participants' varied work settings and client population groups are reflected in Table 1. No other demographic data were solicited from participants.

Table 1. Expertise and experience of music therapist (MT) participants (n=18).

MT number	Self-described expertise	Years of MT experience	Work settings and client populations
1	Intermediate	4	Emergency care, intensive care, detox
2	Expert	20	Children
3	Expert	13	Hospice care, cancer care
4	Expert	12	Adolescents from limited-resource communities, mental health
5	Expert	10	Cancer care, dementia care
6	Intermediate	26	Children, adolescents, elder care, mental health
7	Expert	12	Children, adolescents, elder care
8	Intermediate	8	Cancer care, burn care, bereavement
9	Expert	25	Children, mental health
10	Expert	38	Mental health
11	Expert	16	Elder care, hospice care, children, mental health
12	Expert	32	Mental health, cancer care
13	Expert	23	Mental health, substance use disorder, dementia care
14	Novice	>1	Children, adolescents, rehab patients
15	Expert	30	Mental health, elder care, hospice care, children
16	Expert	20	Dementia care, elder care, hospice care
17	Expert	19	Children, adolescents, hospice care
18	Expert	30	Children, adolescents, cancer care

Thematic Analysis

Overview

Analyses of the interview data suggest there are physically distanced clients who are not getting the telehealth services they need — fueled in part by copyright concerns. Our results suggest that risk aversion to copyright liability in technology-mediated interventions is causing some therapists to alter and avoid preferred treatments. Concerns about copyright liability even caused a few therapists to forgo telehealth services altogether because licensing was impractical, unaffordable, or both. Our analysis generated 5 main themes. These themes and representative quotes are listed in Textbox 1.

XSL•FO
RenderX

Textbox 1. Thematic analysis of copyright's consequences for music therapy telehealth, including representative interview statements for each of the 5 main themes.

1. Adaptation to COVID-19 pandemic protocols accelerated adoption of technology-mediated interventions:
   - "Some of the conversations that I've been having with colleagues these days is that the clinical need has changed for the people that we serve, because of the pandemic. If you're not face-to-face with your folks, then the clinical need has changed." [MT #5]
   - "Teletherapy can also involve downloadables or things that get put on a platform, like YouTube or something like that. And so, the relationship is still happening via email or other platforms of conversation, but the work can happen in these sorts of asynchronous models. And that's where I think potentially copyright gets involved…" [MT #4]
   - "We have had to record and do everything like this [Zoom]." [MT #2]
2. Uncertainty about permissible uses of copyrighted music in technology-mediated interventions:
   - "[Music therapists] could also be videotaping, and all of the sudden you're not just doing live performance, you are talking about potentially a sync license, or mechanical license on top of it. And that gets very complicated very quickly." [MT #10]
   - "For music therapists, it's kind of a gray area; you're essentially playing other people's music and you're getting paid to provide the service." [MT #5]
   - "Eeek! It [copyright] makes my hair stand on end. I just can't feel safe. I can't ever get a straight answer." [MT #16]
3. Concerns about copyright grew as therapists attempted to adopt technology-mediated interventions:
   - "And that was when we really realized like, oh my goodness, there's a huge gap here, and we don't know whether we can do this legally using copyrighted music." [MT #17]
   - "As much as I would love to create music therapy sessions and put them on the internet for other people to use, that's where we run into a lot of copyright issues." [MT #3]
   - "So, as we moved more into creating online content . . . it was so hard to get the answers about what was okay and what was not okay." [MT #2]
4. Aversion to copyright risks prompted erring on the side of caution:
   - "I've just been really careful to avoid using previously composed music in this [pandemic] moment." [MT #3]
   - "The problem is there's just like so many gray areas. It's hard to determine sometimes what is okay and what is not. And so, I'm always just super cautious about it. I just take a step back and say, I'm not even really going to push that boundary. I'm not going to go there at all, and I'm going to, you know, stay safe over here." [MT #7]
   - "I know one of my colleagues just recently was thinking about [copyright] because she wanted to do a country music sing along on our closed-circuit TV. But she really wasn't sure if that was appropriate and sort of abandoned the idea." [MT #8]
   - "It's causing me to not present things that I think could potentially be valuable. So, I might not present the idea of creating a live DJ, or a live performance, because I'm worried about something getting taken down [from YouTube]." [MT #4]
5. Erring on the side of caution triggered lost opportunities of care:
   - "That's absolutely happening… There are practice limitations that are happening. Yeah." [MT #17]
   - "Lack of clarity [on the law] creates those types of environments where I feel like I'm not sure I can suggest something that feels clinically valuable." [MT #4]
   - "Familiar music helps to connect people with their past and with the people that are important to them. We can't deliver that to them because doing a recording of that would be a violation of copyright law. So, our clients are not getting music that they would have before. They're not able to access that because of the virus combined with copyright." [MT #11]

Theme 1: Adaptation to COVID-19 Pandemic Protocols Accelerated Adoption of Technology-Mediated Interventions

The COVID-19 pandemic wrought profound changes to therapists' conventional delivery of music therapy interventions to clients:

> *COVID put the brakes on face-to-face therapy in outpatient settings.* [MT #13]

Another therapist said:

> *I do grief counseling with young children, which has been totally shut down now due to COVID.* [MT #8]

In sum:

> *COVID makes it infinitely more complicated because we can't do the work that we've been able to do.* [MT #3]

The physical distancing protocols prompted by the pandemic forced many outpatient music therapists to shift to mediated platforms to continue to provide services to their clients, as illustrated by this quote:

XSL•FO
RenderX

> *My preference is to do live as much as possible so that we can interact and I can adapt things in the moment. But in some [rural] places, broadband access isn't there. So, the live interaction isn't possible, just because of internet access. And in marginalized communities (in places where they don't have a lot of [broadband] access), recordings are easier to use to have a smooth musical experience — without a lot of freezing and hiccups and things like that.* [MT#11]

The COVID-19 pandemic forced many outpatient music therapists to shift to mediated platforms to continue to provide services to their clients:

> *We have had to record and do everything like this [Zoom].* [MT #2]

Another therapist confirmed the need to shift to mediated services because:

> *Now, with COVID, I don't have any face-to-face clients.* [MT #16]

Nearly all therapists were affected by the pandemic:

> *Some of the conversations that I've been having with colleagues these days is that the clinical need has changed for the people that we serve, because of the pandemic. If you're not face-to-face with your folks, then the clinical need has changed.* [MT #5]

Some music therapists reported difficulties as they attempted to adapt to the requirements of physical distancing. In addition to video conference platforms, therapists tried to experiment with alternative methods for delivering therapy, but often with dissatisfying results:

> *There are so many music therapists who are trying to figure it out, but there isn't a clear and easy way to do that.* [MT #11]

For some, therapy continued — albeit in different modes:

> *Teletherapy can also involve downloadables or things that get put on a platform, like YouTube or something like that. And so, the relationship is still happening via email or other platforms of conversation, but the work can happen in these sorts of asynchronous models. And that's where I think potentially copyright gets involved…* [MT #4]

Theme 2: Uncertainty About Permissible Uses of Copyrighted Music in Technology-Mediated Interventions

Copyright restrictions were cited as a reason it was not clear and easy to use platform-mediated options for telehealth. In virtual spaces, music therapists might need more than just public performance licenses:

> *…they could also be videotaping, and all of the sudden you're not just doing live performance, you are talking about potentially a sync license or mechanical license on top of it. And that gets very complicated very quickly.* [MT #10]

Frustration at the law's complexity and uncertainty was expressed by several interview participants. Despite attempts to research copyright law, one therapist perceived fair use to be "incredibly nebulous" [MT #2]. Another therapist concluded that it was ultimately "unknowable" which uses of copyrighted music were safe and which were not [MT #6]. As one therapist stated:

> *For music therapists, it's kind of a gray area; you're essentially playing other people's music and you're getting paid to provide the service.* [MT #5]

Another therapist expressed the uncertainty more colorfully:

> *Eeek! It [copyright] makes my hair stand on end. I just can't feel safe. I can't ever get a straight answer… I'm just flummoxed because it's so confusing.* [MT# 16]

In sum, interviewees lacked clarity on which uses of copyrighted music in a technology-mediated delivery format were permissible.

Theme 3: Concerns About Copyright Grew as Therapists Attempted to Adopt Technology-Mediated Interventions

It was in the process of exploring technology-mediated therapy options that music therapists reported encountering copyright concerns:

> *And that was when we really realized like, oh my goodness, there's a huge gap here, and we don't know whether we can do this legally using copyrighted music.* [MT #17]

In virtual spaces, copyright questions were complex and daunting:

> *Well, we just haven't been able to devote the time and effort and resources to do that. You can talk to some of these [licensing] companies, but they only deal with certain kinds of licenses. That's the other complicating factor.* [MT#10]

Another therapist noted:

> *As much as I would love to create music therapy sessions and put them on the internet for other people to use, that's where we run into a lot of copyright issues.* [MT #3]

"So, as we moved more into creating online content," one therapist bemoaned that "it was so hard to get the answers about what was okay and what was not okay" [MT #2]. Novel copyright questions were met with unsatisfyingly vague and unclear answers.

In the private, face-to-face settings, therapists uniformly reported feeling comfortable using copyrighted music:

> *When I'm face to face with the patient I never have said, I'm not going to play that because of copyright. In my mind, I'm not getting paid because I play that music. I'm getting paid for the clinical service.* [MT #5]

Another therapist emphasized this point:

XSL•FO
RenderX

> *When I'm face to face with people, I'm not worried about [copyright] at all… Let me be clear, 99.99% of the time I don't worry about copyright; I think it's covered by fair use. We're not monetizing anything, and we're not performing anything in public.* [MT #3]

However, when therapy was no longer confined to private face-to-face settings, therapists reported copyright concerns. In response to this uncertainty, some interviewees admitted that they shy away from using copyrighted music in technology-mediated therapy sessions:

> *I've just been really careful to avoid using previously composed music in this [pandemic] moment.* [MT #3]

When asked if there ever was an instance when a copyrighted song would have been more therapeutically appropriate, but because of copyright concerns, it was not used, an interviewee emphatically answered, "Yes, that has happened" [MT #3]. Therapists reported giving wide berth to avoid potential liability:

> *The problem is there's just like so many gray areas. It's hard to determine sometimes what is okay and what is not. And so, I'm always just super cautious about it. I just take a step back and say, I'm not even really going to push that boundary. I'm not going to go there at all, and I'm going to, you know, stay safe over here.* [MT #7]

In the face of copyright uncertainty, another therapist said she would avoid "anything that could put you at jeopardy," and "since you don't know the answer black and white, go the opposite direction…" [MT #2].

But when therapists entirely avoid copyrighted music, it can be therapeutically problematic:

> *If all our research says patient-preferred music [yield superior results], and then suddenly we can't use patient-preferred music because of copyright, then that's a huge problem.* [MT #5]

One therapist explained the importance of using patient-preferred music:

> *The problem with [using non-copyrighted music] is that some people are going to be more engaged if you can find that music that they're really familiar with — that they're really comfortable with. And most of our referrals are children that are not doing great. It's the kids in the hospital that are having a tougher time. So, you're already talking about a kid that's scared, that's anxious, that's maybe anticipating a poke or some kind of difficult procedure. So, using that familiar music is a way of making that connection — building that rapport — and then helping them to start regulating themselves. Honestly, using copyrighted material for a lot of music therapists is really integral to how they practice.* [MT #17]

In other words, the inability to use a patient's favorite song hampers a therapist's ability to connect and build rapport with the patient.

Theme 4: Aversion to Copyright Risks Prompted Erring on the Side of Caution

In response to these concerns, some interviewees admitted that they shied away from using copyrighted music in technology-mediated therapy sessions. One therapist candidly acknowledged, "I'm dragging my feet" on therapeutic activities because of copyright uncertainty [MT #4]. This copyright uncertainty:

> *…creates a situation where I don't feel comfortable making therapeutically necessary suggestions that I would otherwise make.* [MT #4]

Another therapist noted:

> *I know one of my colleagues just recently was thinking about [copyright] because she wanted to do a country music sing along on our closed-circuit TV. But she really wasn't sure if that was appropriate and sort of abandoned the idea.* [MT #8]

Therapists acknowledged that copyright concerns are exacerbating lost opportunities for care:

> *It's causing me to not present things that I think could potentially be valuable. So, I might not present the idea of creating a live DJ or a live performance because I'm worried about something getting taken down [from YouTube].* [MT #4]

Automated algorithmic enforcement on platforms like YouTube and Facebook was identified as a concern. Nevertheless, some therapists openly doubted whether famous musicians would directly bring an enforcement action:

> *I don't think Ozzy Osbourne is going to come knock on somebody's door because you use their song in a heartbeat recording.* [MT #5]

Another noted:

> *I cannot imagine a world where a [grieving] family in that situation would get in trouble.* [MT # 17]

Another therapist said that when she has asked musicians for permission to use songs "just for the therapy session," "almost always they say yes . . . usually they have said, 'Go for it. No fees'" [MT #13]. In one instance, a therapist wrote to a musician explaining how the song "My Shot" from the play "Hamilton," was used in therapy to help "clients reflect on what is their shot, what steps they are taking not to throw it away, and where in their lives to they need to rise up" [MT #13]. And shortly thereafter, the therapist "got a really nice note back from Lin Manuel [Miranda] or his publicist (whoever writes his things) that said, 'This is awesome. I wish you and your patients the best'" [MT #13]. While therapists did not think most musicians would object to therapeutic uses of their music, it was better to be safe than sorry:

> *I highly doubt that Taylor Swift would come after a child with cancer. But I suppose that there's a chance that that could happen, or that she could come after the hospital as well. And I don't want to get anybody in trouble. I don't want to have a legal battle for one 30-minute session that we're doing. So, I'm going to*

XSL•FO
RenderX

> *step away from that and not make that a problem in the first place.* [MT #7]

Despite perceiving a low likelihood of enforcement actions by rightsholders, it was seen as safer to err on the side of caution.

Theme 5: Erring on the Side of Caution Fueled Unmet Needs for Therapy

Some therapists acknowledged a demand for therapy that was not being satisfied:

> *I'm going to say, probably at least 60 clients [in nursing homes] that we would be reaching that we're not reaching right now because of that [copyright].* [MT # 11]

The therapist explained the unmet need of nursing home patients — patients who are "basically in solitary confinement" and their "functioning is declining" without social interaction and connection with others [MT #11]. Music therapy could potentially help these patients because therapists find that music helps with "socialization and engagement and staying in the here and now" [MT #11]. However, copyright liability was a reason therapists did not feel comfortable offering recorded music therapy sessions. As the therapist put it:

> *Familiar music helps to connect people with their past and with the people that are important to them. We can't deliver that to them because doing a recording of that would be a violation of copyright law. So, our clients are not getting music that they would have before. They're not able to access that because of the virus combined with copyright.* [MT #11]

Thus, to avoid copyright concerns, some therapists avoided mediated services altogether — because there were no adequate alternatives to patients' familiar (and copyrighted) music.

Other therapists confirmed that lost opportunities of care were fueled by copyright concerns:

> *…lack of clarity [on the law] creates those types of environments where I feel like I'm not sure I can suggest something that feels clinically valuable.* [MT #4]

When asked if patients had therapeutic needs that were unmet, one interviewee confirmed, "That's absolutely happening":

> *I've been telling my staff that we have to honor these things [copyright]. And then they push back. And then I push back. And they push back. And, you know, we're like making a diamond. So yeah, absolutely. There are practice limitations that are happening. Yeah.* [MT #17]

In other words, our interviews revealed that supervisors and facility administrators are discouraging certain therapeutic practices because it is unclear whether these practices risk copyright liability. Thus, concerns about copyright liability are manifesting as "practice limitations." Therapeutic practices can be constrained not only by a therapists' own concerns but also by a therapist's supervisor's concerns.

Discussion

Principal Findings

Our primary objective was to obtain greater insight into whether copyright concerns affected music therapists' delivery of computer-mediated therapy. While telehealth services are not new, widespread experimentation with telehealth services was hastened by COVID-19's physical distancing requirements. The diffusion of telehealth services offers a range of salutary benefits [30,49]. Nevertheless, the sudden shift to mediated interventions caused considerable confusion in the music therapy community. Research suggests therapists scrambled to answer novel questions about telehealth technology affordances, privacy regulations, and billing protocols [27]. In addition to these questions, our interviews identified uncertainty about copyright law as a contributing factor to music therapists' hesitance to embrace innovative interventions.

A copyright holder's exclusive rights to public performances, reproductions, and derivatives of copyrighted music present a unique and underexamined barrier for telehealth services. As our findings suggest, copyright law is not typically a concern for music therapists providing private, face-to-face treatment. A copyright holder cannot restrict purely private performances of copyrighted materials [21,23]. But when therapy transitions to remote delivery via mediated technologies, copyright law is not so clear. Exemptions to patent protection have received notable attention during the pandemic, especially by US policymakers [50]. However, new copyright exemptions have received little attention [51]. These findings can help to start those important discussions.

The complexity of US copyright law is reputed to rival the complexity of the US tax code. No court case has evaluated the permissibility of therapeutic uses of copyrighted music. No court case has clarified whether therapeutic uses of music are fair uses of music. Music therapy's novel uses of copyrighted works put remote delivery of music therapy within a gray zone of US copyright law.

In this gray area of the law, music therapists reported unease. Attempts to research copyright law failed to yield clear, definitive answers, further amplifying music therapists' confusion. Courts decide fair use questions on a case-by-case basis, and fair use is notoriously difficult to assess in novel settings. As legal scholars note, "The fair use doctrine is famous for its uncertainty" [52].

Not only are questions about fair uses untested in the courts but rightsholders have also failed to offer an efficient mechanism for therapists to license their varied therapeutic uses of music [23]. With neither clear legal precedent on fair use nor an easy way to license their uses, music therapists may be inclined to err on the side of caution to avoid liability. Supporting a risk-averse approach, the AMTA's COVID-19 Task Force Advisory noted, "A conservative approach simply avoids music covered under copyright in any setting" [53].

Synthesis of Findings With Prior Work

Emerging research confirms that technology-based music therapy holds tantalizing promise [54]. In the mental health

XSL•FO
RenderX

field, the widespread adoption of telehealth has made it the "new normal" [18]. Technology's perceived usefulness and ease of use are important factors in the successful adoption of new technology [55]. However, the legal dimension to technology's use has been underexplored in the public health literature.

The value of our work lies in providing new perspectives on barriers to adopting innovative telehealth interventions. The extensive literature on diffusion of innovation has failed to incorporate legal uncertainty as a barrier. Our participants' thick descriptions offer unique insights on the effects of gray areas of the law.

This study builds on prior inquiry on the effects of copyright law on music therapists. This study draws from a rich corpus of 13 hours of data to answer new research questions, identify distinct themes, and incorporate unique verbatim quotes. Prior research explored how copyright concerns prompt music therapists to act as gatekeepers who discourage patients from sharing therapeutic artifacts on social media [21]. This prior work examined the secondary effects of copyright law, where copyright uncertainty co-opted music therapists to act as proxies for rightsholders. On the other hand, this present work examines the primary effects of copyright uncertainty on music therapists who offer services via telehealth. Rather than discouraging patients' activities, this study highlights that copyright uncertainty discourages music therapists' own telehealth activities. We acknowledge the value in sharing health information on social media [56]; however, the results from this study are arguably more urgent because these findings suggest there are lost opportunities of care to those in need.

Strengths and Limitations

This exploratory study involved 18 music therapists working in the United States. A key strength is the depth of experience of the participants; these participants had over 300 years of combined experience in the field. This study provides an important foundation for exploring music therapists' views about copyright implications for telehealth. Our findings do not claim generalizability; rather, our aim was a more in-depth understanding of this social phenomenon. To complement the findings in this study, survey analysis should assess a wider pool of music therapists. The views expressed by these participants may not be typical of music therapists. Future research should extend this study through quantitative approaches and should assess how pervasive the concerns identified herein are among a wider sample of music therapists.

Future research should also interrogate whether music therapists in other countries have had similar experiences. Differences in copyright laws in other countries, differences in musical preferences (eg, patient preferences for noncopyrighted or public domain songs), and differences in music therapists' experiences with telehealth technologies may impair the generalizability of these findings outside of the United States. Moreover, the COVID-19 experience has been unprecedented; it has been generations since the West has experienced a pandemic. Future research should assess how music therapists in other countries adapted during other, recent epidemics, like SARS (severe acute respiratory syndrome) [57] and MERS (Middle East respiratory syndrome) [58]. Future research should assess whether concerns highlighted herein are isolated to Western societies with copyright laws similar to the United States.

Other limitations also exist. The ratio of female to male participants was high, with only 3 of 18 (17%) participants being male. Novice therapists were underrepresented, with only 1 of 18 (6%) self-describing as novice. Lastly, demographic characteristics of the interviewees (eg, age, ethnicity, and religion) were not collected. These limitations in the data are acknowledged.

It is also acknowledged that the lead researcher's prior research on copyright law had the potential to bias the data analysis. To mitigate these risks, trustworthiness measures were adopted, including collaboration with an independent researcher and inclusion of extensive verbatim quotes. The lead researcher's law training, while a potential limitation, positively contributed to the quality of the interview data; it enabled thoughtful and appropriate follow-up questions in the semistructured interviews. This exploratory project treads novel ground with music therapists and how copyright concerns are a barrier to telehealth innovations. The quality of these data provides valuable insights not only about music therapists but also may offer insights about other professions that incorporate copyrighted works, like filmmakers [59] or remixers [60,61].

Conclusions

The lack of clarity on copyright law in the telehealth space has undermined the robust diffusion of therapeutic uses of music. This gray area of the law has prompted some therapists to err on the side of caution. This avoidance is suboptimal in the wake of the shift to technology-mediated interventions and calls for expanded access to underserved populations. Our in-depth interviews confirmed that some music therapists were indeed eschewing their traditional practice modes when using new telehealth innovations because of copyright restrictions. To address the unmet needs for therapy, Congress could amend the copyright statute to include a statutory exemption for therapeutic uses of music [23].

Our interviews revealed that music therapists' risk aversion has served as a barrier to the diffusion of much-needed telehealth services. These findings offer a unique contribution to the public health literature by (1) highlighting copyright law as an impediment in a technology-mediated setting and (2) revealing there are lost opportunities of care because music therapists are concerned about copyright liability. Thus, legal uncertainty is an underappreciated and unwelcome barrier to the diffusion of telehealth innovations.

Acknowledgments

This research was supported by a seed grant from the UNC Hussman School of Journalism and Media.

XSL•FO
RenderX

Conflicts of Interest

None declared.

References

1. Hartling L, Newton AS, Liang Y, Jou H, Hewson K, Klassen TP, et al. Music to reduce pain and distress in the pediatric emergency department: a randomized clinical trial. JAMA Pediatr 2013 Oct;167(9):826-835. [doi: 10.1001/jamapediatrics.2013.200] [Medline: 23857075]
2. Gallagher LM, Lagman R, Rybicki L. Outcomes of music therapy interventions on symptom management in palliative medicine patients. Am J Hosp Palliat Care 2018 Mar;35(2):250-257. [doi: 10.1177/1049909117696723] [Medline: 28274132]
3. İçel S, Başoğul C. Effects of progressive muscle relaxation training with music therapy on sleep and anger of patients at Community Mental Health Center. Complement Ther Clin Pract 2021 May;43:101338. [doi: 10.1016/j.ctcp.2021.101338] [Medline: 33639518]
4. Yinger OS, Gooding L. Music therapy and music medicine for children and adolescents. Child Adolesc Psychiatr Clin N Am 2014 Jul;23(3):535-553. [doi: 10.1016/j.chc.2013.03.003] [Medline: 24975624]
5. Palumbo A. Integrated rehabilitation post-stroke: combined music therapy and occupational therapy group for physical, emotional, and social well-being. Arch Phys Med Rehabil 2019 Oct;100(10):e91-e92 [FREE Full text] [doi: 10.1016/j.apmr.2019.08.266]
6. Gómez-Romero M, Jiménez-Palomares M, Rodríguez-Mansilla J, Flores-Nieto A, Garrido-Ardila EM, González López-Arza MV. Benefits of music therapy on behaviour disorders in subjects diagnosed with dementia: a systematic review. Neurologia 2017 May;32(4):253-263 [FREE Full text] [doi: 10.1016/j.nrl.2014.11.001] [Medline: 25553932]
7. Bieleninik A, Ghetti C, Gold C. Music therapy for preterm infants and their parents: a meta-analysis. Pediatrics 2016 Sep;138(3):1-17 [FREE Full text] [doi: 10.1542/peds.2016-0971] [Medline: 27561729]
8. Hadley S, Yancy G, editors. Therapeutic Uses of Rap and Hip-hop. New York, NY: Routledge; Sep 1, 2011.
9. Warren M. The impact of music therapy on mental health. National Alliance on Mental Illness. 2016 Dec 19. URL: https://www.nami.org/Blogs/NAMI-Blog/December-2016/The-Impact-of-Music-Therapy-on-Mental-Health [accessed 2021-07-31]
10. Trombetti A, Hars M, Herrmann FR, Kressig RW, Ferrari S, Rizzoli R. Effect of music-based multitask training on gait, balance, and fall risk in elderly people: a randomized controlled trial. Arch Intern Med 2011 Mar 28;171(6):525-533. [doi: 10.1001/archinternmed.2010.446] [Medline: 21098340]
11. What is music therapy. American Music Therapy Association. 2005. URL: https://www.musictherapy.org/about/musictherapy/ [accessed 2021-07-31]
12. 2017 AMTA member survey and workforce analysis. American Music Therapy Association. URL: https://musictherapy.org/assets/1/7/17WorkforceAnalysis.pdf [accessed 2021-07-31]
13. Patel SY, Huskamp HA, Busch AB, Mehrotra A. Telemental health and US rural-urban differences in specialty mental health use, 2010-2017. Am J Public Health 2020 Sep;110(9):1308-1314. [doi: 10.2105/AJPH.2020.305657] [Medline: 32673109]
14. Vaudreuil R, Langston DG, Magee WL, Betts D, Kass S, Levy C. Implementing music therapy through telehealth: considerations for military populations. Disabil Rehabil Assist Technol 2020 Jul 01;1:1-10. [doi: 10.1080/17483107.2020.1775312] [Medline: 32608282]
15. Spooner H, Lee JB, Langston DG, Sonke J, Myers KJ, Levy CE. Using distance technology to deliver the creative arts therapies to veterans: case studies in art, dance/movement and music therapy. Arts Psychother 2019 Feb;62:12-18. [doi: 10.1016/j.aip.2018.11.012]
16. Eccles H, Nannarone M, Lashewicz B, Attridge M, Marchand A, Aiken A, et al. Perceived effectiveness and motivations for the use of web-based mental health programs: qualitative study. J Med Internet Res 2020 Jul 31;22(7):e16961 [FREE Full text] [doi: 10.2196/16961] [Medline: 32735216]
17. Liu Z, Qiao D, Xu Y, Zhao W, Yang Y, Wen D, et al. The efficacy of computerized cognitive behavioral therapy for depressive and anxiety symptoms in patients with COVID-19: randomized controlled trial. J Med Internet Res 2021 May 14;23(5):e26883 [FREE Full text] [doi: 10.2196/26883] [Medline: 33900931]
18. Martinez-Martin N, Dasgupta I, Carter A, Chandler JA, Kellmeyer P, Kreitmair K, et al. Ethics of digital mental health during COVID-19: crisis and ppportunities. JMIR Ment Health 2020 Dec 22;7(12):e23776 [FREE Full text] [doi: 10.2196/23776] [Medline: 33156811]
19. Silverman M, Letwin L, Nuehring L. Patient preferred live music with adult medical patients: a systematic review to determine implications for clinical practice and future research. Arts Psychother 2016 Jul;49:1-7 [FREE Full text] [doi: 10.1016/j.aip.2016.05.004]
20. Najafi Ghezeljeh T, Mohades Ardebili F, Rafii F, Haghani H. The effects of music intervention on background pain and anxiety in burn patients: randomized controlled clinical trial. J Burn Care Res 2016;37(4):226-234. [doi: 10.1097/BCR.0000000000000266] [Medline: 26132048]
21. Reid A, Miño P. When therapy goes public: copyright gatekeepers and sharing therapeutic artifacts on social media. Intl J Commu 2021;15:950-969 [FREE Full text] [doi: 10.1016/b978-1-84334-693-7.50014-3]

XSL•FO
RenderX

22. Copyright Act, 17 U.S.C. § 106. Exclusive rights in copyrighted works. 1976. URL: https://www.govinfo.gov/content/pkg/USCODE-2010-title17/pdf/USCODE-2010-title17-chap1-sec106.pdf [accessed 2021-07-31]
23. Reid A. Social utility of music: a case for a copyright exemption for therapeutic uses. Cornell J Law and Public Policy 2020;30(1):1-43 [FREE Full text]
24. Hazard JW. Re: Use of Music in Hospital Settings. American Music Therapy Association. 2001 May 08. URL: http://www.musictherapy.org/assets/1/7/copyrightletter.pdf [accessed 2021-07-31]
25. The Office for Civil Rights. Notification of enforcement discretion for telehealth remote communications during the COVID-19 nationwide public health emergency. U.S. Department of Health & Human Services. 2021 Jan 20. URL: https://www.hhs.gov/hipaa/for-professionals/special-topics/emergency-preparedness/notification-enforcement-discretion-telehealth/index.html [accessed 2021-07-31]
26. APAC/AIAC COVID-19 response: FAQs. American Music Therapy Association. 2020 Mar 25. URL: https://www.musictherapy.org/assets/1/7/APAC_AIAC_COVID-19_Response_FAQs.pdf [accessed 2021-07-31]
27. Carvajal MA. Telehealth music therapy: considerations and changes during the Covid-19 crisis (Thesis submitted to the College of Music in partial fulfillment of the requirements for the degree of Master of Music). The Florida State University. 2020. URL: https://www.proquest.com/openview/60d3f525bd74c4eb24f16fe8c6b8a7ab/1 [accessed 2021-07-31]
28. Montagnani ML. Virtues and perils of algorithmic enforcement and content regulation in the EU: a toolkit for a balanced algorithmic copyright enforcement. Case Western Reserve J Law, Tech & the Internet 2020;11:1-49 [FREE Full text]
29. Almathami HKY, Win KT, Vlahu-Gjorgievska E. Barriers and facilitators that influence telemedicine-based, real-time, online consultation at patients' homes: systematic literature review. J Med Internet Res 2020 Feb 20;22(2):e16407 [FREE Full text] [doi: 10.2196/16407] [Medline: 32130131]
30. Sasangohar F, Bradshaw MR, Carlson MM, Flack JN, Fowler JC, Freeland D, et al. Adapting an outpatient psychiatric clinic to telehealth during the COVID-19 pandemic: a practice perspective. J Med Internet Res 2020 Oct 01;22(10):e22523 [FREE Full text] [doi: 10.2196/22523] [Medline: 32936768]
31. Aufderheide P, Milosevic T, Bello B. The impact of copyright permissions culture on the US visual arts community: the consequences of fear of fair use. New Media Soc 2016 Jul 09;18(9):2012-2027. [doi: 10.1177/1461444815575018]
32. Kazdin AE. Addressing the treatment gap: a key challenge for extending evidence-based psychosocial interventions. Behav Res Ther 2017 Jan;88:7-18. [doi: 10.1016/j.brat.2016.06.004] [Medline: 28110678]
33. Sabatello M, Burke TB, McDonald KE, Appelbaum PS. Disability, ethics, and health care in the COVID-19 pandemic. Am J Public Health 2020 Oct;110(10):1523-1527. [doi: 10.2105/AJPH.2020.305837] [Medline: 32816541]
34. Rogers EM. Diffusion of Innovations. New York, NY: Simon and Schuster; 2003.
35. Hunt MR. Strengths and challenges in the use of interpretive description: reflections arising from a study of the moral experience of health professionals in humanitarian work. Qual Health Res 2009 Sep;19(9):1284-1292. [doi: 10.1177/1049732309344612] [Medline: 19690208]
36. Burnett E, Corlett J. Understanding risk perceptions and responses of the public and health care professionals toward Clostridium difficile: a qualitative interpretive description study. Am J Infect Control 2017 Feb 01;45(2):133-138. [doi: 10.1016/j.ajic.2016.07.020] [Medline: 27789069]
37. Crotty MJ. The Foundations of Social Research: Meaning and Perspective in the Research Process. New York, NY: SAGE Publications Ltd; 1998.
38. Mullen EJ. Evidence-based knowledge in the context of social practice. Scand J Public Health 2014 Mar;42(13 Suppl):59-73. [doi: 10.1177/1403494813516714] [Medline: 24553855]
39. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
40. Thorne S. Interpretive Description. Walnut Creek, CA: Left Coast Press; 2008.
41. Berterö C. Developing qualitative methods - or "same old wine in a new bottle". Int J Qual Stud Health Well-being 2015;10:27679 [FREE Full text] [doi: 10.3402/qhw.v10.27679] [Medline: 25977121]
42. Thorne S, Kirkham S, O'Flynn-Magee K. The analytic challenge in interpretive description. Int J Qual Methods 2016 Nov 29;3(1):1-11 [FREE Full text] [doi: 10.1177/160940690400300101]
43. Natow R. The use of triangulation in qualitative studies employing elite interviews. Qual Res 2019 Feb 19;20(2):160-173 [FREE Full text] [doi: 10.1177/1468794119830077]
44. Rooshenas L, Paramasivan S, Jepson M, Donovan JL. Intensive triangulation of qualitative research and quantitative data to improve recruitment to randomized trials: the QuinteT approach. Qual Health Res 2019 Apr;29(5):672-679. [doi: 10.1177/1049732319828693] [Medline: 30791819]
45. MacQueen K, McLellan E, Kay K, Milstein B. Codebook development for team-based qualitative analysis. CAM Journal 2016 Sep 14;10(2):31-36 [FREE Full text] [doi: 10.1177/1525822x980100020301]
46. Boyatzis RE. Transforming Qualitative Information: Thematic Analysis and Code Development. New York, NY: SAGE Publications, Inc; 1998.
47. Luker K. Salsa Dancing into the Social Sciences: Research in an Age of Info-glut. Cambridge, MA: Harvard University Press; 2010.
48. Saldana J. The Coding Manual for Qualitative Researchers. New York, NY: SAGE Publications Ltd; 2015.

XSL•FO
RenderX

49. Olayiwola JN, Magaña C, Harmon A, Nair S, Esposito E, Harsh C, et al. Telehealth as a bright spot of the COVID-19 pandemic: recommendations from the virtual frontlines ("Frontweb"). JMIR Public Health Surveill 2020 Jun 25;6(2):e19045 [FREE Full text] [doi: 10.2196/19045] [Medline: 32479413]
50. Maxmen A. In shock move, US backs waiving patents on COVID vaccines. Nature 2021 May 06:1. [doi: 10.1038/d41586-021-01224-3] [Medline: 33958787]
51. Hilty RM, Batista PHD, Carls S, Kim D, Lamping M, Slowinski PR. Covid-19 and the role of intellectual property: position statement of the Max Planck Institute for innovation and competition of 7 May 2021. Max Planck Institute for Innovation and Competition. 2021. URL: https://www.ip.mpg.de/fileadmin/ipmpg/content/stellungnahmen/2021_05_25_Position_statement_Covid_IP_waiver.pdf [accessed 2021-07-31]
52. Urban JM, Falzone A. Demystifying fair use: the gift of the Center for Social Media statements of best practices. J Copyright Soc of the USA 2010 Mar;57(3):337-350 [FREE Full text] [doi: 10.4324/9781315637549-8]
53. Advisory from the AMTA COVID19 Task Force. Copyright and licensing: maintaining compliance in online/remote environments. American Music Therapy Association. 2020 Jun. URL: https://www.musictherapy.org/assets/1/7/Advisory_Copyright_from_the_AMTA_COVID19_Task_Force_06_20_20.pdf [accessed 2021-07-31]
54. Potts C, Bond R, Ryan A, Mulvenna M, McCauley C, Laird E, et al. Ecological momentary assessment within a digital health intervention for reminiscence in persons with dementia and caregivers: user engagement study. JMIR Mhealth Uhealth 2020 Jul 06;8(7):e17120 [FREE Full text] [doi: 10.2196/17120] [Medline: 32420890]
55. Davis F. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 1989 Sep;13(3):319-340 [FREE Full text] [doi: 10.2307/249008]
56. Moorhead SA, Hazlett DE, Harrison L, Carroll JK, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. J Med Internet Res 2013 Apr 23;15(4):e85 [FREE Full text] [doi: 10.2196/jmir.1933] [Medline: 23615206]
57. Peiris JSM, Yuen KY, Osterhaus ADME, Stöhr K. The severe acute respiratory syndrome. N Engl J Med 2003 Dec 18;349(25):2431-2441. [doi: 10.1056/NEJMra032498] [Medline: 14681510]
58. Zumla A, Hui DS, Perlman S. Middle East respiratory syndrome. The Lancet 2015 Sep 05;386(9997):995-1007 [FREE Full text] [doi: 10.1016/S0140-6736(15)60454-8] [Medline: 26049252]
59. Aufderheide P, Sinnreich A. Documentarians, fair use, and free expression: changes in copyright attitudes and actions with access to best practices. Inf Commun Soc 2015 May 27;19(2):178-187 [FREE Full text] [doi: 10.1080/1369118x.2015.1050050]
60. Pappalardo K, Aufderheide P. Romantic remixers: hidden tropes of romantic authorship in creators' attitudes about reuse. Cultural Sci J 2020;12(1):1-12. [doi: 10.5334/csci.117]
61. Fiesler C, Bruckman AS. Remixers' understandings of fair use online. 2014 Presented at: 17th ACM conference on Computer supported cooperative work & social computing; February 15-19, 2014; Baltimore, MD URL: https://dl.acm.org/doi/10.1145/2531602.2531695 [doi: 10.1145/2531602.2531695]

Abbreviations

AMTA: American Music Therapy Association
HIPAA: Health Insurance Portability and Accountability Act
MERS: Middle East respiratory syndrome
SARS: severe acute respiratory syndrome

Edited by G Eysenbach; submitted 03.03.21; peer-reviewed by D Schmidt, P Phiri; comments to author 05.05.21; revised version received 19.05.21; accepted 12.07.21; published 13.08.21.

Please cite as:
Reid A, Kresovich A
Copyright as a Barrier to Music Therapy Telehealth Interventions: Qualitative Interview Study
JMIR Form Res 2021;5(8):e28383
URL: https://formative.jmir.org/2021/8/e28383
doi:10.2196/28383
PMID:34319241

©Amanda Reid, Alex Kresovich. Originally published in JMIR Formative Research (https://formative.jmir.org), 13.08.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium,

XSL•FO
RenderX

provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Efficiency, Usability, and Outcomes of Proctored Next-Level Exams for Proficiency Testing in Primary Care Education: Observational Study

Birgitte Schoenmakers[1*], MD, PhD; Johan Wens[2*], MD, PhD

[1]Department of Public Health and Primary Care, KU Leuven, Leuven, Belgium

[2]Department of Primary and Interdisciplinary Care, University of Antwerp, Antwerpen, Belgium

[*]all authors contributed equally

Corresponding Author:
Birgitte Schoenmakers, MD, PhD
Department of Public Health and Primary Care
KU Leuven
Kapucijnenvoer 33/7001
Leuven, 3000
Belgium
Phone: 32 495235639
Email: birgitte.schoenmakers@kuleuven.be

Abstract

Background: The COVID-19 pandemic has affected education and assessment programs and has resulted in complex planning. Therefore, we organized the proficiency test for admission to the Family Medicine program as a proctored exam. To prevent fraud, we developed a web-based supervisor app for tracking and tracing candidates' behaviors.

Objective: We aimed to assess the efficiency and usability of the proctored exam procedure and to analyze the procedure's impact on exam scores.

Methods: The application operated on the following three levels to register events: the recording of actions, analyses of behavior, and live supervision. Each suspicious event was given a score. To assess efficiency, we logged the technical issues and the interventions. To test usability, we counted the number of suspicious students and behaviors. To analyze the impact that the supervisor app had on students' exam outcomes, we compared the scores of the proctored group and those of the on-campus group. Candidates were free to register for off-campus participation or on-campus participation.

Results: Of the 593 candidates who subscribed to the exam, 472 (79.6%) used the supervisor app and 121 (20.4%) were on campus. The test results of both groups were comparable. We registered 15 technical issues that occurred off campus. Further, 2 candidates experienced a negative impact on their exams due to technical issues. The application detected 22 candidates with a suspicion rating of >1. Suspicion ratings mainly increased due to background noise. All events occurred without fraudulent intent.

Conclusions: This pilot observational study demonstrated that a supervisor app that records and registers behavior was able to detect suspicious events without having an impact on exams. Background noise was the most critical event. There was no fraud detected. A supervisor app that registers and records behavior to prevent fraud during exams was efficient and did not affect exam outcomes. In future research, a controlled study design should be used to compare the cost-benefit balance between the complex interventions of the supervisor app and candidates' awareness of being monitored via a safe browser plug-in for exams.

(JMIR Form Res 2021;5(8):e23834) doi:10.2196/23834

KEYWORDS

primary care; education; graduate; medical education; testing; assessment; app; COVID-19; efficiency; accuracy

XSL•FO
RenderX

Introduction

The COVID-19 pandemic has heavily affected education and assessment programs. Flemish (Belgian) universities were, as those in many other countries, not sufficiently prepared to fluently switch from analogous teaching and testing to digital teaching and testing. Moreover, the COVID-19 measures have also necessitated physical distancing before and during exams, which has resulted in a lower number of candidates participating in exam sessions. This situation has resulted in complex planning for managing human resources, locations, and adequate equipment.

The challenge of reorganization is even more impressive in interuniversity collaborations for education. The Advanced Master of Family Medicine program in Flanders, Belgium, is formally organized and is offered by 4 Flemish universities. This collaboration comprises a common administration, common courses and examinations, common residencies, and separate registers of residents for the university of their choice. In the three phases of the advanced master program, we have over 900 residents in training and in education. The planning of examinations is therefore a complex logistic and administrative matter. To address this logistic challenge, we built (more than 1 decade ago) an intelligent, comprehensive, and interactive assessment platform. From other authors, we have learned that (medical) students are in favor of computerized exams, and these exams may enhance learning experiences and effects without affecting final test outcomes [1-4]. The platform we developed provides an interface for summative and formative knowledge testing (in 6 questions formats) and Objective Structured Clinical Examination–performance and proficiency testing.

In general, it is not desirable to postpone or reschedule high-stakes tests or assessments [5,6]. Such exams are often organized apart from the regular exam schedule. The exam regulations and planning process of the proficiency test for admission to the Advanced Master of Family Medicine program in Flanders are also different from those of regular exams. Further, the same proficiency test is organized for the four Flemish universities. The test involves a 3-stage procedure that starts with an administrative stage. This is followed by the actual exam, and the test is finalized with a jury exam for candidates who failed the exam in stage 2. The actual exam is a machine-assisted test that is conducted on the digital assessment platform. This whole procedure has been conducted since 2016, and this format has proven to be reliable, acceptable, and feasible [7].

The combination of planning an off-schedule exam and offering exams on campus (ie, the ruling policy) while adhering to the original exam format has forced us to opt for a creative solution. We decided to prioritize the format of this high-stakes test and organize—in addition to the original on-campus exam—a proctored off-campus exam. Web-based exams (or remote exams) for high-stakes assessments are considered user-friendly and cost-friendly and are flexible in terms of planning, but their validity is questionable if fraud cannot be ruled out [3,8-10]. The development and administration processes of proctored exams are still under construction and have been boosted during the COVID-19 pandemic [5]. The most commonly applied antifraud measures are direct face visualization and visual verification through a webcam [11-13]. However, screening all of the webcam images of large groups of students is time consuming, and image quality is highly dependent on webcam technology, a stable internet connection, and background lighting [11,12].

To minimize the occurrence of fraudulent events during the exam, we developed a web-based supervisor app for tracking and tracing candidates' behaviors during the exam. The technology we have built and applied goes beyond that of traditional proctored exam systems that focus on tracing sounds and images. In addition to real-time supervision (image and sound), all student behavior is recorded and logged for posttest reassessment in case of suspiciousness. In this paper, we report on the efficiency and usability of our antifraud measures and the impact that a proctored exam has on exam outcomes.

Methods

Study Design

Through collaboration with the developers of the assessment platform and through discussions with the coordinators and exam supervisors (called the expert group) of the Advanced Master Education program, we determined the criteria and conditions for designing the supervisor app. The application operated on the following three levels: the recording of actions, analyses of behavior, and live supervision. First, during the exam, the system recorded data from three sources—the computer screen, the camera, and the microphone. These recordings were immediately encrypted and saved on a secured server. Second, the supervisor app used a pattern recognition algorithm for response, clicking behavior, and time stamp analyses (Textbox 1). The application also analyzed individual behavior as well as correlations across (collaborating) candidates. Each event was given a score. Every suspicious event increased the final suspicion rating by 0.1. Based on the consensus of the expert group, a candidate with a suspicion rating of 0.5 or higher was considered suspicious. The tracking data were stored in the users' browsers and sent to the server after 20 nonsuspicious events occurred or when suspicious events occurred. If the suspicion rating rose above 0.5, the exam submission was flagged, and a report on the flagged behavior was downloaded for assessment. Suspicious events were defined as switching to another browser, returning to the exam page, closing the exam page, disconnecting from the internet, and making sounds and noises.

XSL•FO
RenderX

Textbox 1. Types of events that were tracked and traced by the supervisor app.

Events

- Blur page (suspicious)
  - This means that the user switched to another browser tab or another program.
- Focus page (suspicious)
  - This means that the user returned their focus to the page.
  - Focus page events can be used (after blur page events) to determine how long the user had been gone.
- Close page (suspicious)
  - This means that the user closed the current tab or opened a different page in the same tab.
- Load page
  - This event occurs when the user loads a page.
  - Load page events can be used (after close page events) to determine how long the user had been gone.
- Keyboard events
  - Logged with a basic key logger
- Connected to the internet
- Disconnected from the internet (suspicious)
  - The user can stop their internet connection to stop video recordings. This data were sent to the server when the user reconnected to the internet.
- Mouse clicks
  - Includes click location coordinates
- Sound detected (suspicious)
  - The app tracks any noise above certain volume levels with a duration of >1 ms.
- Administrator started watching
  - This is marked in blue on the tracking page.
- Administrator started talking
  - This is marked in blue on the tracking page.
- Finalization of the test component
  - This is needed because when a user completes the test component, a close page event is triggered.

The third level of supervision allowed for optional human oversight during the examination. The human monitor could immediately join the live feed of each candidate to obtain more information or to send a warning to candidates in private. In case the supervisor app crashed, the affected candidates were able to switch to the Safe Exam Browser (Eidgenössische Technische Hochschule Zürich), which allows only the exam interface to remain accessible on a user's computer.

In addition to the technical solutions, we expected the application to have a preventive effect on candidates' behaviors. Candidates were comprehensively briefed in advance on how to install the application, on how to test it, and on the specific features of the supervising technology. We used an animation video and a written instruction form for the briefing.

A voluntary panel that consisted of experienced teachers and exam supervisors tested the supervisor app in 2 sessions. These test participants were asked to behave in a suspicious manner (ie, talking, making noise, turning away from the screen, using the internet, typing, clicking, and using the mouse). After the first session, we made the following adjustments to the application and the procedure: we isolated the suspicious sound events from the other suspicious events, and we increased the overall suspicion rating from 0.5 to 1.

The comparison group was composed of the on-campus candidates who were supervised under the usual circumstances and conditions. To assess efficiency, we logged technical issues and the interventions that were used during the exam. Technical issues were defined as those reported by students, and

XSL•FO
RenderX

interventions referred to the active responses of the team. To analyze the impact that the supervisor app had on students' exam outcomes, we compared the scores of the proctored group to those of the on-campus group. To test usability, we counted the number of suspicious students and behaviors.

All students who applied for admission to the advanced master program were included in this study. To take the exam, candidates were free to register for off-campus participation or on-campus participation. On campus (n=4), a human supervisor was present, and candidates used the campus' gear. For off-campus monitoring, we had an experienced supervising team of 6 staff members. A team of 2 developers was also fully available during the exam. The off-campus supervisors were able to send notifications or warnings to candidates who were behaving suspiciously, and these supervisors intervened in cases involving technical issues.

The actual exam and the associated procedures were set up in the same manner as those that were conducted before the COVID-19 pandemic; all candidates completed the same exams, and candidates who failed the machine-assisted test (the actual exam) were invited to a jury exam after 1 week. Candidates who were flagged as those exhibiting suspicious behaviors during the exam or in postexam analyses were also invited to the jury exam.

Ethical Approval

Ethical approval was not required for this study. Consent for the administration of the proctored exam was obtained from the Permanent Education Commission of the Faculty of Medicine of KU Leuven.

Data Set

The complete data set is available upon simple request and can be sent as a link to a Google Drive directory.

Results

A total of 593 candidates subscribed to the exam. Of these candidates, 472 (79.6%) used the supervisor app for off-campus exams and 121 (20.4%) were present on campus (Table 1). The test results of both the off-campus and on-campus groups did not differ significantly (*P*=.15).

Table 1. Participation and test results of off-campus students and on-campus students.

Characteristics	Off-campus students	On-campus students	*P* value[a]
Candidates (N=593), n (%)	472 (79.6)	121 (20.4)	N/A[b]
Candidates at each university, n (%)			
Leuven University	227 (84.1)	43 (15.9)	N/A
Antwerp University	29 (35.8)	52 (64.2)	N/A
Brussels University	13 (50)	13 (50)	N/A
Gent University	203 (94)	13 (6)	N/A
Average test results (out of 100)	72	72.8	.15

[a]The *P* value was determined via a one-tailed pooled *t* test.

[b]N/A: not applicable.

Overall, we registered and solved 15 technical issues that occurred during the off-campus exam (Table 2). Of these issues, 8 were the result of software problems (in particular, loading a reading text in a new tab). Further, 2 candidates experienced a negative impact on their exam performance due to technical issues. The development team made one of these candidates switch to the Safe Exam Browser to complete the exam. Further, based on the postexam analyses and after deliberation, the coordinator exempted the other candidate of the jury exam.

XSL•FO
RenderX

Table 2. Comparison of exam procedures and outcomes of the off-campus exam and on-campus exam.

Characteristics	Off-campus exam	On-campus exam
Technical issues, n	15	1
Technical issues with an impact on the exam, n	2	0
Type of issue, n		
Internet failure	2	0
Hardware issue	4	1
Camera crash	1	0
Software issue	8	0
Average suspicion score	0.4	N/A[a]
Median suspicion score	0.3	N/A
Suspicious candidates, n (%)		
Detected by the application	22 (4)	N/A
Flagged by supervisors	2 (<1)	N/A
Flagged due to noncritical events (≤1 event)	455 (96)	N/A
Flagged without events	15 (3)	N/A
Flagged due to noise events	472 (100)	N/A
Interventions during the exam, n	8	1
Technical interventions	2[b]	0
Warnings to candidates	1[c]	0

[a]N/A: not applicable.

[b]Technical interventions were needed for 2 individuals.

[c]Warnings were given to 1 group due to background noise.

Of the 593 candidates, the application detected 22 (3.7%) with a suspicion rating of >1. All cases were the result of 1 or more noise events (background noise). These students received an immediate but nonoffensive request to stop all background noise without further consequences or impacts on the progression of their exams. Additionally, the supervisor sent a request to all students to reduce background noise. All other noncritical events included leaving the webpage, closing a page, or typing text. Live monitoring and a postexam review of the recordings confirmed that all of these events occurred unintentionally and without fraudulent intent. In the on-campus group, we did not detect suspicious or fraudulent students.

The monitoring supervisors flagged two candidates who were typing more than the expected amount (in a multiple-choice exam). After the review of the recordings, it was found that these two candidates were using the “control find” function to search for words in the reading text. During the exam, the supervisors intervened 8 times due to technical issues, warned 2 candidates to stop talking to themselves, and sent a group message to ask candidates to reduce background noise.

Discussion

This pilot observational study demonstrated that a supervisor app that records and registers behavior was accurate because it was able to detect all suspicious events without having an impact on exam performance. The efficiency of the proctored supervisor system seemed low, as shown by the availability of a 12-person bystander team, which monitored and solved a relatively low number of technical issues.

Students have shown certain preferences for different exam formats depending on how much confidence they have in their techniques, the organization, exam procedures, and exam outcomes and how confident they are with technological aspects and issues [9,12,14,15]. The unbalanced number of registrations for off-campus and on-campus exams among the universities was striking, but this can be explained by the availability of infrastructure (classes and information technology gear). We indeed found that the candidates from larger universities were more likely to register for an off-campus exam. The psychometrics of the exam were comparable between the off-campus and on-campus groups. Therefore, we can conclude that there was probably no bias induced by the voluntary option for participating in off-campus exams or on-campus exams. Other authors have compared the scores of proctored exams to scores of traditional exams that were conducted in the past several years and found no meaningful differences [6,8,14].

The major weakness of proctored exam systems lies in the occurrence of technical or technological failures and in a lack of user experience [2,4,12,16]. The number of technical issues was low and was mainly related to software issues (technology issues). The complexity of the application (the registration of behavior, recording, etc), in combination with a multicomponent

XSL•FO
RenderX

exam with different questions types, necessitated very performant technological equipment. In one case, a technical issue led to the interruption of the exam of the affected candidate. The supervisors and the development team solved all other issues within an acceptable time span. However, these issues revealed that the system was vulnerable to the low quality of the personal gear of the candidates. Therefore, education institutes should guarantee a fair and safe exam environment for every candidate by offering high-quality infrastructure and logistics [4,12,13].

During the entire exam, a team of 12 persons was constantly monitoring and maintaining contact with the candidates. Although we did not perform a cost-benefit analysis, we assume that the whole procedure does not save costs, since the on-campus exam opportunity also remained available [4].

On average, candidates induced 3 events, of which the noise and sound event was the most prevalent and critical. The development team increased the threshold sound level during the exam, since the application immediately (in the first few minutes) flagged all candidates. Sound is indeed difficult to avoid and control (we noticed children crying, birds singing, candidates talking out loud, doors being slammed, street work being conducted, etc) [11,12].

A very small number of candidates was flagged as suspicious by the application, but due to the live monitoring of the exam, they were exempted from suspicion. The supervisors reviewed the recordings of two candidates who were flagged by the system and via the live monitoring of the exam. The records did not reveal any fraudulent behavior. During the exam, the team was constantly watching the students, and the system continuously logged behaviors. Additionally, students were very aware of the presence of the supervisors. They were also aware of the fact that they were being watched and screened. Further, the exam was performed within a restricted time span; therefore, students did not have time to cheat. Consequently, students were probably very reluctant to cheat, and this might have increased their stress levels [10,13].

Prior to the exam, we thoroughly instructed the candidates on the registration of suspicious events. This single intervention might have been decisive in stopping candidates from committing fraud during this high-stakes exam [2,4,10]. A simple recording of sounds and images combined with a safe web browser that blocked all other webpages might have been just as efficient as the supervisor app [2,4]. A less comprehensive registration and recording process for suspicious events might have also reduced the number of technical issues and allowed the supervisors to focus on the live monitoring of the exam [8].

During the exams, supervisors limited the number of interventions and warnings to avoid distracting the candidates. Technical issues were solved, and supervisors sent a nonoffensive warning to only two candidates who were talking to themselves. After the exam, we only received reports of technical issues; we received no complaints regarding supervisors' interventions or interruptions caused by the supervisors.

A weakness of the intervention is that we did not question the candidates about their experiences. This group of graduation candidates was, at the time of writing this report, very hard to reach. However, the staff and the formal faculty administration did not receive comments or objections from students in the period following the exam. Other reports have shown a relatively high satisfaction rate among master program candidates, and complaints have typically been about technical issues [16]. A second weakness is the arbitrary, expert-based threshold for suspicious behavior, which was set to 0.5. However, during the exam, it appeared that this threshold was set too low, since most students immediately obtained high scores mainly due to background noise.

A major strength of this intervention is the fact that we were able to use the application for a large group of candidates taking a high-stakes exam [2]. Additionally, we compared the exam outcomes of the intervention group with those of a control group, which was composed of candidates who preferred to take the exam on campus. We could not completely rule out bias, but since exam outcomes were comparable between both groups, we believe that the risk of bias was low. Based on other authors' papers, we also knew that candidates' preferences for computerized exams or paper-based exams do not influence exam outcomes [15].

Conclusion

A sophisticated supervisor app that registered behaviors and recorded sounds and images to prevent fraud during exams proved to be efficient and did not affect exam outcomes. However, live monitoring and having a team on standby for solving technical issues resulted in a high amount human workload. In future research, a controlled study design should be used to compare the cost-benefit balance between the complex interventions of the supervisor app and candidates' awareness of being monitored via a safe browser plug-in for exams.

Acknowledgments

We would like to acknowledge An Stockmans, Guy Gielis, Bert Aertgeerts, Roy Remmen, Dirk Devroey, Aram Hovsepyan, Chavdar Gushterov, Peter Decat, and Madieke Maseman for their intellectual and logistic support in the development and unrolling of this plug-in.

Conflicts of Interest

None declared.

XSL•FO
RenderX

References

1. Ogilvie RW, Trusk TC, Blue AV. Students' attitudes towards computer testing in a basic science course. Med Educ 1999 Nov;33(11):828-831. [doi: 10.1046/j.1365-2923.1999.00517.x] [Medline: 10583791]
2. Bengtsson L. Take-home exams in higher education: A systematic review. Educ Sci (Basel) 2019 Nov 06;9(4):267. [doi: 10.3390/educsci9040267]
3. Karay Y, Schauber SK, Stosch C, Schüttpelz-Brauns K. Computer versus paper--does it make any difference in test performance? Teach Learn Med 2015;27(1):57-62. [doi: 10.1080/10401334.2014.979175] [Medline: 25584472]
4. Boevé AJ, Meijer RR, Albers CJ, Beetsma Y, Bosker RJ. Introducing computer-based testing in high-stakes exams in higher education: Results of a field experiment. PLoS One 2015 Dec 07;10(12):e0143616. [doi: 10.1371/journal.pone.0143616] [Medline: 26641632]
5. Munshi F, Alsughayyer A, Alhaidar S, Alarfaj M. An online clinical examination for fellowship certification during the COVID-19 pandemic. Med Educ 2020 Oct;54(10):954-955 [FREE Full text] [doi: 10.1111/medu.14267] [Medline: 32501565]
6. Sam AH, Reid MD, Amin A. High-stakes, remote-access, open-book examinations. Med Educ 2020 Aug;54(8):767-768 [FREE Full text] [doi: 10.1111/medu.14247] [Medline: 32421858]
7. Schoenmakers B, Wens J. Proficiency testing for admission to the postgraduate family medicine education. J Family Med Prim Care 2018;7(1):58-63 [FREE Full text] [doi: 10.4103/jfmpc.jfmpc_163_17] [Medline: 29915734]
8. Rios JA, Liu OL. Online proctored versus unproctored low-stakes internet test administration: Is there differential test-taking behavior and performance? American Journal of Distance Education 2017 Feb 02;31(4):1-14. [doi: 10.1080/08923647.2017.1258628]
9. Alessio HM, Malay N, Maurer K, Bailer AJ, Rubin B. Interaction of proctoring and student major on online test performance. International Review of Research in Open and Distributed Learning 2018 Nov;19(5):165-185. [doi: 10.19173/irrodl.v19i5.3698]
10. Migut G, Koelma D, Snoek CGM, Brouwer N. Cheat me not: Automated proctoring of digital exams on bring-your-own-device. 2018 Jul 02 Presented at: ITiCSE '18: 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education; July 2-4, 2018; Larnaca, Cyprus p. 388. [doi: 10.1145/3197091.3205813]
11. Asep HSG, Bandung Y. A design of continuous user verification for online exam proctoring on m-learning. 2019 Presented at: 2019 International Conference on Electrical Engineering and Informatics (ICEEI); July 9-10, 2019; Bandung, Indonesia. [doi: 10.1109/iceei47359.2019.8988786]
12. Baume M. Online proctored exams: Where and how are they used? Basics, practical scenarios and technical solutions for online proctoring at european universities and educational institutions. 2019 Presented at: 13th International Technology, Education and Development Conference; March 11-13, 2019; Valencia, Spain p. 5216-5225. [doi: 10.21125/inted.2019.1301]
13. Spence D, Ward R, Wooden S, Browne M, Song H, Hawkins R, et al. Use of resources and method of proctoring during the NBCRNA Continued Professional Certification Assessment: Analysis of outcomes. J Nurs Regul 2019 Oct;10(3):37-46 [FREE Full text] [doi: 10.1016/s2155-8256(19)30147-4]
14. Daffin Jr LW, Jones AA. Comparing student performance on proctored and non-proctored exams in online psychology courses. Online Learning Journal 2018 Mar 01;22(1):131-145. [doi: 10.24059/olj.v22i1.1079]
15. Hochlehnert A, Brass K, Moeltner A, Juenger J. Does medical students' preference of test format (computer-based vs. paper-based) have an influence on performance? BMC Med Educ 2011 Oct 25;11:89 [FREE Full text] [doi: 10.1186/1472-6920-11-89] [Medline: 22026970]
16. Milone AS, Cortese AM, Balestrieri RL, Pittenger AL. The impact of proctored online exams on the educational experience. Curr Pharm Teach Learn 2017;9(1):108-114. [doi: 10.1016/j.cptl.2016.08.037] [Medline: 29180142]

Edited by G Eysenbach; submitted 25.08.20; peer-reviewed by H Salminen, I Zamberg; comments to author 23.09.20; revised version received 06.10.20; accepted 16.05.21; published 16.08.21.

Please cite as:
Schoenmakers B, Wens J
Efficiency, Usability, and Outcomes of Proctored Next-Level Exams for Proficiency Testing in Primary Care Education: Observational Study
JMIR Form Res 2021;5(8):e23834
URL: https://formative.jmir.org/2021/8/e23834
doi:10.2196/23834
PMID:34398786

©Birgitte Schoenmakers, Johan Wens. Originally published in JMIR Formative Research (https://formative.jmir.org), 16.08.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium,

XSL•FO
RenderX

provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Development and Evaluation of Acceptability and Feasibility of a Web-Based Intervention for Patients With Bipolar Disorder in Iran: Implementation Study

Agaah Ashrafi[1], MD, MPH; Maryam Tabatabaee[1], MD, MPH; Vandad Sharifi[1], MD

Department of Psychiatry, Roozbeh Hospital, Tehran University of Medical Sciences, Tehran, Iran

Corresponding Author:
Vandad Sharifi, MD
Department of Psychiatry, Roozbeh Hospital
Tehran University of Medical Sciences
South Kargar Avenue
Roozbeh Hospital
Tehran
Iran
Phone: 98 9123875869
Fax: 98 2155419113
Email: vsharifi@tums.ac.ir

Abstract

Background: Psychoeducation for bipolar disorder has a significant impact on symptoms and treatment adherence. In Iran, as a low-resource setting, infrastructural barriers, such as inadequate mental health professionals, difficulties in transportation, and costs of care, may hinder optimum delivery of this evidence-based intervention to patients.

Objective: This study sought to explore the acceptability and feasibility of a web-based intervention for bipolar patients in Iran.

Methods: A website has been developed as a platform for providing psychoeducational content about bipolar disorder. Patients were chosen via a convenient sampling method in 2018-2019. The main component of the intervention included streaming 7 weekly video clips after attending a single in-person meeting, as well as a medication self-monitoring application. Information was collected about the feasibility and acceptability of the intervention.

Results: We invited 45 patients from the day center and the outpatient clinic of Roozbeh psychiatric hospital and some private clinics in Tehran. Of the 23 patients (51%) who attended the first in-person session and provided informed consent, 14 patients dropped out during the study. While 9 patients completed the course (attended 4 or more online sessions), only 5 watched all the video sessions. The rate of adherence to the intervention and frequency of exposure to the website were much higher for those recruited from the private and outpatient clinics.

Conclusions: This web-based intervention can be feasible and acceptable only for a subgroup of patients with specific educational status and socioeconomic level.

(JMIR Form Res 2021;5(8):e23360) doi:10.2196/23360

KEYWORDS

bipolar disorder; psychoeducation; web-based intervention; feasibility; acceptability

Introduction

Bipolar disorder (BD) is a substantial public health problem, with lifetime prevalence of around 2.4% in the general population according to world mental health survey in 2011 [1] and 1% among the Iranian population [2]. Patients frequently present relapsing episodes and usually undergo subsyndromal symptoms, cognitive problems, functional impairment, and repeated hospitalizations which cause significant burden both for their families and the whole community. Medications is considered as a gold-standard treatment modality, however, psychosocial interventions, such as psychoeducation (PE), self-help, and psychotherapy (individual, couple, and family), have shown promising treatment outcomes for BD [3].

PE is considered an effective modality in the treatment and management of BD [4-6]. In a recent systematic review conducted by Demissie et al [5] in 2018, it was concluded that

XSL•FO
RenderX

PE would enhance patients' adherence to treatment, knowledge and attitude toward the condition and their quality of life as well as it would minimize relapse rates and hospitalizations. In Iran, several studies which investigated the effectiveness of PE interventions showed that PE significantly decrease relapse and re-hospitalization rate [7]. Faridhosseini and colleagues (2017), investigated the effectiveness of a culturally adapted structured PE program for BD. They also concluded that PE for patients with BD could improve their quality of life and minimize relapse risk [8].

In contrast to the vast literature regarding the efficacy of PE programs, there are still some limitations in the application and delivery of these services, specifically in low resource settings such as Iran. Traditionally, these services are provided face-to-face, which is not always possible especially for patients who have low treatment commitments or major problems for attending the face-to-face sessions, including long distance, limitation in affording needed time, energy and/or money, and limited access to health professionals.

Novel web-based technologies offer an opportunity to employ standardized psychological treatments that can deal with some of the aforementioned limitations. Considering the fact that these technologies are widely accessible 24-hours a day, the timing of interventions can be tailored to the patients' needs and availability [9].

In some high-income countries such as the United Kingdom, health-care policymakers have shown interests about employing the power of the internet to allow patients to take more responsibility in their illness management. "Beating Bipolar" is a web-based PE intervention developed in the United Kingdom and has proven to be as a promising PE treatment modality for BD patients; this intervention focuses on illness awareness, adherence to treatment, early detection of recurrence and lifestyle regularity [9-11].

Taken together, to fill the gap between availability and demand, we need to think of a treatment platform which is more accessible and feasible. In developing countries such as Iran, the obstacles, including distance, and limited number of mental-health staffs in remote areas, have posed serious problems regarding management of this group of patients and this calls for employing novel treatment approaches such as online PE. Here we tend to offer a preliminary web-based platform for PE and to evaluate its acceptability and feasibility for a group of patients with BD in Iran.

Methods

Developing the Web-Based Intervention

After reviewing available widely used and studied web-based intervention platforms in developed countries [9,12-14], we developed a preliminary design and proposed it to a professional website developer firm in Tehran. The website included 2 main parts: (1) 7 visually animated video clips (modules) with PE-related content that were simply narrated and had a maximum length of 20 minutes; the content was based on evidence-based knowledge on BD and our team's previous experience with face-to face PE programs [15]; (2) weekly medication tables for patients to self-record their medications to improve their feelings of agency and mastery and emphasize the importance of self-monitoring and adherence to medications. It is worth mentioning that to respect the privacy of patients, each user was allocated a personal portal to register their progress and data.

Following the first face-to-face session, the 7 internet-based modules were provided. Table 1 shows the topics that were covered.

Table 1. Contents of the web-based psychoeducation course for bipolar disorder.

Session/week	Content
1	In-person orientation meeting
2	Bipolar disorder and definitions
3	Etiology and risk factors
4	Bipolar disorder, medications and treatments
5	Alarm signs and relapse prevention
6	Adaptation to bipolar disorder
7	Problem solving, an essential skill for patients
8	Other topics such as marriage, exercise, driving, diet, occupation, and more for bipolar patients and summary

Topics, as mentioned in Table 1, were incorporated into the visual content (educational video clips) and written material (downloadable pdf files). To persuade and encourage patients to use the website during the course, we sent them a reminder text message at the beginning of each week. Moreover, a module to record questions from patients was incorporated in the website, and it was answered regularly by one of the authors. After developing a preliminary version, we randomly chose 3 patients of different ages and genders to participate and give us some feedback regarding their experience in a pilot course using this website.

Implementation and Evaluation of the Web-Based Intervention

Participants

We recruited subjects from the day center and outpatient clinic of Roozbeh Psychiatric Hospital and some private clinics in

XSL•FO
RenderX

Tehran. Our inclusion criteria were patients with a DSM-5 BD, diagnosed by a psychiatrist, age 15-65 years old, being able to access a computer and the internet, a minimum educational level of the 6th grade in primary school, no previously receipt or completion of any PE courses, and full or partial remission at the time of the study as indicated by their psychiatrist. We excluded patients with severe visual impairment, intellectual disabilities, or neurodevelopmental disorders.

Procedures

The first session of each course (1st module) was held in person to familiarize each patient with other participants, the research team, and the overall process of the project and to obtain informed consent. Patients were informed regarding the importance of PE in the management of psychiatric conditions and more specifically BD. In each round of the interventions, we invited patients via a phone call, and we reminded them of the session time via text message 1 day before the meeting. In this first session, personal username and password were allocated to each patient to fill in their data. After this session, we invited the patients to participate in our study for 7 consecutive weeks and to expose themselves to the aforementioned PE video clips on the website. Figure 1 demonstrates the front page of the website where the patients were requested to enter their own username and password to access their individual profile. The Farsi language was selected for the website so a language barrier would not become an issue for our patients who all were Iranian.

Figure 2 shows the main page of the website, which showed 8 round circles at the top of the page as a manifestation of the course progress and the week the patients were in at any time, 2 main sections in the middle of the page, and some ancillary tabs on the right side.

Figure 1. The front page of the irbiplar.ir website.

Figure 2. The main page of the irbipolar.ir website.

One of the 2 main sections linked to the weekly streaming PE video clips (Figure 3; Multimedia Appendix 1), and the other section showed a weekly timetable of medications to take and allowed them to indicate their pattern of actual medication administration during the week (Figure 4).

XSL•FO
RenderX

Figure 3. Page with the psychoeducation video clips.

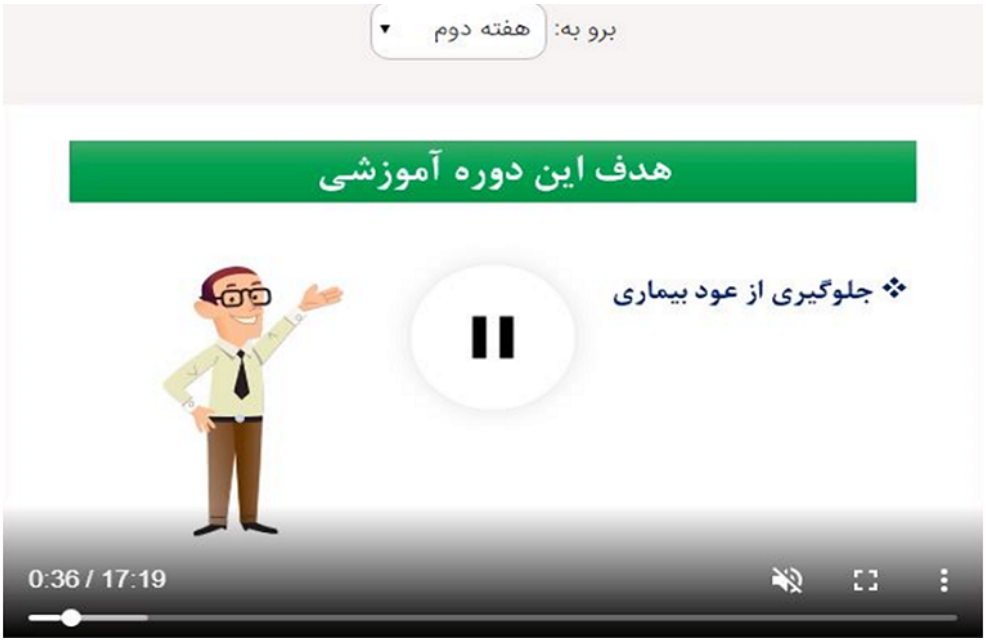


Figure 4. The weekly table for recording the pattern of adherence to medication.

Initially, this section was supposed to be a reminder to regulate the patients' patterns of taking medication as well as give them a sense of mastery over their treatment plan; however, we thought a reminder might be negatively viewed by the patients as being controlled and might even decrease their cooperation. Therefore, this section remained optional, although strongly suggested, for the patients to use.

It is worth mentioning that during the intervention period, patients received other mental health services as well, which included medications and if indicated, psychotherapy and rehabilitation services provided by the hospital. For cases for whom we suspected a relapse, patients were referred for a medication review, and afterwards if the patient was still interested in participating in the study, we asked him or her to update his or her medication table on the website and continue the course. If the patient was not still interested, he or she was excluded from the study.

Assessments

Our primary objective was to determine the level of feasibility and acceptability of this intervention. Feasibility outcomes were assessed via the recruitment rate, adherence to the website training course protocol, participant drop-out rate (participants who did not attend ≥4 weeks during their 8-week course). In addition, we conducted qualitative interviews with the team members and patients at the end of study to discuss obstacles and facilitators of the implementation platform and how to address them. Acceptability was assessed via the frequency and pattern of website use by participants, which was designed to

XSL•FO
RenderX

be reported as an Excel file by the programmer of the website, as well as patients' satisfaction rate, including whether the program was user-friendly, comprehensibility of the content, usefulness of the content, and usefulness of reminders. These variables were rated on a Likert scale, ranging from 1 to 5, for each item by the participants in terms of an interview conducted by 1 of the authors. SPSS v20 and Excel 2016 were used for statistical calculations and analyses as needed.

Ethical Considerations

A written informed consent form was completed by each patient. Patients were assured that the data are confidential. The current design was approved by the Research Ethics Committee of Tehran University of Medical Sciences (IR.TUMS.MEDICINE.REC.1397.579).

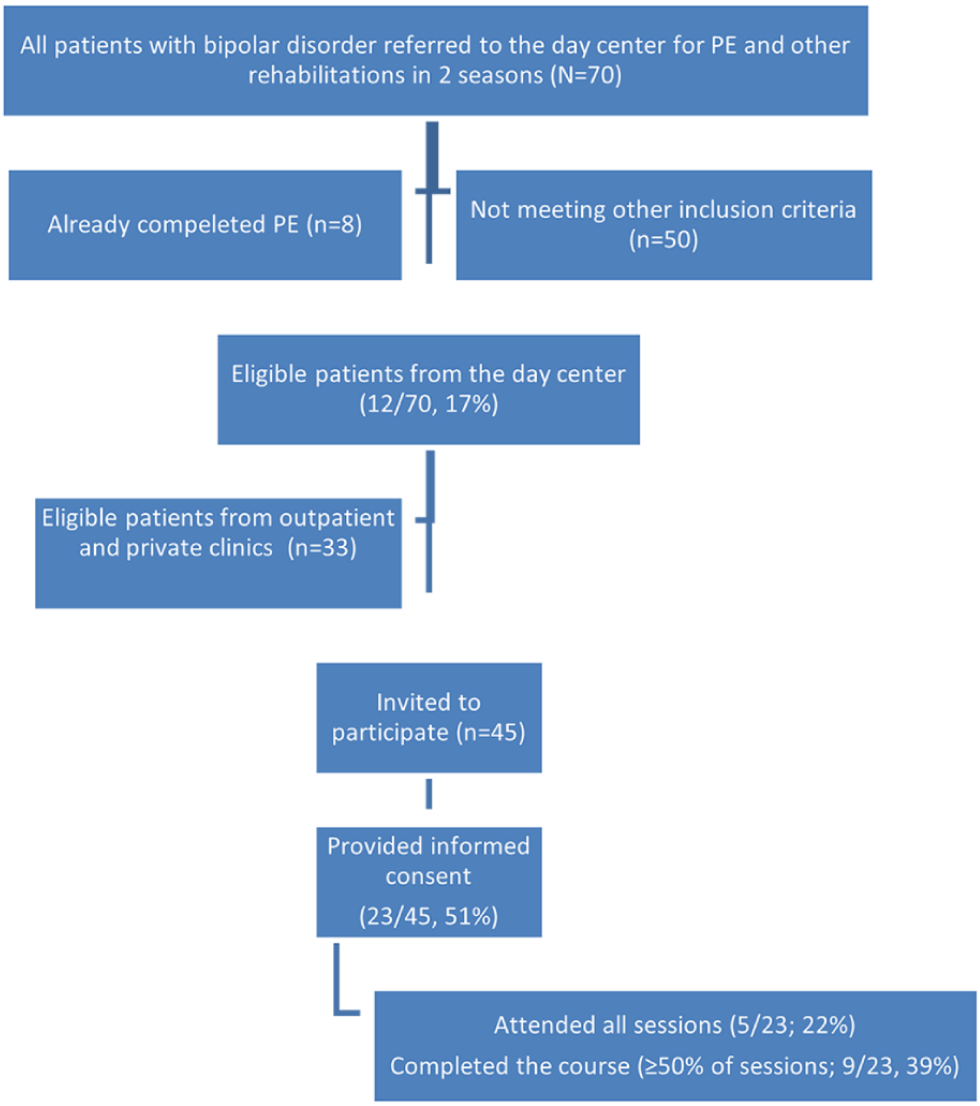
Results

Patient Recruitment Process

Among 70 bipolar patients referred to Roozbeh Psychiatric Hospital in the second half of 2018, 12 were eligible to be invited to our study. Of these, 2 attended the first in-person session, and 8 could not participate because they had already completed the same course in previous face-to-face sessions. The remaining 50 patients of the 70 patients referred were excluded for not meeting the inclusion criteria, such as not having enough interest or lack of familiarity with the internet. Then, we invited an additional list of 33 patients from the hospital outpatient clinics and private clinics in Tehran in January 2019.

Finally, among the 45 invited patients, 23 attended the first in-person session (51%) and agreed to participate in the study. Figure 5 shows the sampling process.

Figure 5. The flow of subjects in the study. PE: psychoeducation.



XSL•FO
RenderX

Feasibility Outcomes

Recruitment Rate

Among the patients referred to the day center, the recruitment rate was 3% (2/70). It was 63% (21/33) for the outpatient and private clinics.

Adherence Rate

We defined a dropout as being absent in 4 or more of the 8 sessions of the whole course. There was a 100% dropout rate during the first round of the study, which was held for the patients referred to the day center of Roozbeh Hospital, as none of the patients attended online meetings. Among all patients, 14 patients dropped out during the study. Of the 9 patients who completed the course, only 5 (22%) attended all the sessions, and 4 (17%) patients were absent in some sessions, but they completed the course. Furthermore, the level of commitment of all patients to participate in each week is illustrated in Table 2.

Table 2. Adherence to the web-based psychoeducation modules.

Module	Number of patients who completed the module
1	23
2	12
3	9
4	9
5	8
6	7
7	6
8	6

Dropout Rate

As mentioned, among 23 patients who entered the study, 14 dropped out, which accounts for 61% of the total sample who attended the first in-person meeting. Of these 14 participants, 50% (7/14) provided reasons that could be considered as a "lack of enough motivation." Table 3 shows the possible reasons for nonadherence (dropout) that were obtained from interviews with the participants.

Table 3. Dropout reasons (n=14).

Dropout reasons	n (%)
Relapse	2 (14)
Difficulty using web-based modules	3 (21)
Not having access to a computer during certain weeks of the course	2 (14)
Other reasons[a]	7 (50)

[a]Other reasons include those that could be considered a lack of motivation such as a lack of spare time and forgetfulness.

Acceptability Outcomes

Frequency and Pattern of Website Use

This includes the number of patient logins to the website and the entire time that was spent on the website by patients. As shown in Table 4, an average frequency of 1-2 logins per week, each lasting about 21 minutes, by patients who completed the course indicates the relative acceptability of this course, while for noncompleters, this variable is meaningfully lower (Table 4). While 14 (14/23, 60%) watched at least one PE video, only 5 (5/23, 22%) patients watched all of the videos.

Table 4. Frequency and pattern of participant use of the website over the entire 7 weeks of the study.

Parameter	Patients who dropped out (n=14)	Patients who completed the course (n=9)
Total logins, n	40	117
Average logins per participant, mean (SD)	2.85 (1.75)	13 (3.85)
Total duration (minutes)	233	1333
Average duration per participant (minutes), mean (SD)	16.65 (12.23)	148 (29.1)

Participants' Experiences

To assess participants' experiences using the online platform, we interviewed 13 participants who at least watched one PE video to gain answers to satisfaction questions, with responses rated from 1 to 5 on a Likert scale; the results are presented in Table 5. The average range of satisfaction with each item shows the relative acceptability of the intervention.

XSL•FO
RenderX

Table 5. Participants' experiences using the website (score range: 1-5).

Characteristic	Rating, mean (SD)	Mean rating divided by the total possible score of 5, %
User-friendly	4.2 (1.12)	84
Simplicity and usefulness of the contents	4.5 (0.86)	90
Usefulness of the Q&A panel	3 (1.37)	60
Usefulness of reminders	4 (1.20)	80

Discussion

Generally, our findings demonstrated this intervention was feasible and acceptable only for some groups of patients with specific educational and socioeconomic status who met the requirements of the procedure.

Principal Findings

In recent years, novel communication and information platforms have represented a promising prospect to offer psychological intervention via web-based tools and to tackle some of the drawbacks of face-to-face sessions. In comparison with in-person sessions, they are readily accessible and minimize the delivery timeline. Furthermore, the timing of the therapy can be arranged according to the specific need and availability of each user.

In developed countries, telemedicine and more specifically, telepsychiatric platforms (websites, applications) are becoming increasingly used in the treatment plan of bipolar patients [16-18]. In Iran, however, as a low-resource setting, this area is still young, and we do not have a strong evidence base regarding the efficacy and feasibility of web-based interventions for the management of psychiatric conditions.

Currently, PE, as one of the most studied psychotherapy modalities, has been revealed as a practical approach to the management and treatment of BD [19-21]. Considering the interest of many patients in searching the internet for their conditions and treatments, the vast amount of invalid and low-quality content that is currently available could put them at higher risk. Hence, it appears that providing a platform with evidence-based materials supported by responsive, knowledgeable staff to provide answers to patients anytime from anywhere via the platform is necessary. Furthermore, the low number of mental health professionals in many remote geographical zones with an increasing number of patients calls for adopting a novel approach to address the specific needs of this patient group in our country.

According to the results of our study, we can conclude that our web-based PE intervention platform is not feasible and acceptable for patients who were referred to the day center of Roozbeh Psychiatric Hospital. We speculate that for this first group of patients who may have a more severe illness and poorer educational and socioeconomic background that makes it likely they are clients of a nonprivate university hospital, the poor attendance is due to a lack of motivation or knowledge about the importance of these interventions in the management of their disorder as well as lower skill level or access to these types of technology-based services. However, this platform was sufficiently accepted and feasibly implemented for another group of patients who were recruited from private clinics or outpatient clinics in Roozbeh Hospital who were already more actively seeking the needed care and had better socioeconomic backgrounds.

Obstacles and Facilitators to Designing the Website

Influential Factors

After holding a number of sessions, the research team reached consensus about influential factors in creating the website and sorted these factors into 2 categories: factors related to the individuals involved in the study and factors related to the intervention.

Factors Related to the Individuals Involved in the Study

For the web design team, the main obstacle was that our research team did not have any previous experience in designing a website, so we had to outsource this stage to a reliable web design firm. Since the web design firm did not have any professional knowledge regarding our subject, we had to hold a number of sessions to align our expectations and knowledge with their experience in designing an online platform.

Regarding the research team, a limited number of professional staff and insufficient team members for implementing different stages of the study and follow-ups meant one person carried all these responsibilities, which could increase the number of errors.

For patients, first, we should state that a noticeable proportion of patients did not have enough motivation or skills to participate in the current study, and some patients thought that participating in this study may lead to a relapse of their symptoms. Several patients also preferred face-to-face sessions and did not identify enough with the proposition of online platforms.

Factors Related to the Intervention

First, we can point to the novel feature of this intervention in our setting, which can result in some trial and error. Furthermore, we faced some unfortunate incidences such as filtering the website in the middle of the study, which happened accidently by the government. The main limitation was the necessity to have 1 face-to-face session at the initiation of the study for patients. This obstacle manifested as the participation gap between the potential 45 volunteers at the preliminary invitation and the ~50% attendance (n=23) at the first in-patient session. Another limitation was specifying a 1-week period for each module; according to some patients, they had a problem attending some modules due to a lack of spare time, sickness, or travelling, and sometimes, they lost their motivation in continuing the course when they were absent for 1 module. Another issue was that watching a video clip or answering some questions did not take more than 1 or 2 hours per week, and this

XSL•FO
RenderX

limited involvement during the week could lead to decreased motivation to continue participation. We could have tackled this issue by providing daily tasks for participants to complete and revive their motivation on a regular basis.

Our results are consistent with the outcome of the qualitative study conducted by Poole and colleagues [18]; they analyzed the feasibility, acceptability, and impact of an internet-based PE platform, "beating bipolar." They found it feasible to deliver and acceptable for patients to use via a computer. This intervention had a satisfying impact on insight concerning illness, health behavior, personal habits, and positive attitude toward medication [18]. However, in their study, thematic analysis was employed to describe the participants' experience. In the current study, we used a more quantitative approach. Hidalgo-Mazzei et al [12] designed a simple smartphone application to provide PE contents for patients with BD. Consistent with our results, this type of online intervention was an efficient approach to improve self-management by patients with BD. However, their small sample size (51 patients) made generalization difficult, as we can also say about this study [12].

Conclusion

Altogether, we can claim that this intervention was feasible and acceptable only for some groups of patients with specific educational and socioeconomic status meeting the requirements of the procedure, for example, access to and knowing how to use the internet. Through the last two decades, the opportunity to provide PE contents through internet-based platforms have been investigated. According to various studies, those platforms have shown good to acceptable retention rates [12-14,17,18,22,23]. However, all these earlier studies were designed and implemented in developed countries. In countries such as Iran, this area is still undeveloped and in its infancy; therefore, implementing these novel interventions should be followed cautiously. Results of the current study confirm that using a web-based PE platform is acceptable and feasible for a specific subgroup of patients, and we may need to revise the intervention to tailor it to the needs and features of other patient groups. However, further studies in a larger sample pool should be conducted to reach more conclusive results.

Acknowledgments

We thank Dr. Sina Rezayi for assistance with designing and programming the website.

We also show our gratitude to Dr. Mani Monajemi's comments on an earlier version and Dr. Zahra Mirsepasi's comments on a later version of this manuscript that were great help in improving the manuscript.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors or any kind of sponsors, and no specific financial resources were linked or applied to this study.

Authors' Contributions

AA primarily contributed to conceptualization, data curation, formal analysis, investigation, methodology, software, and writing the original draft. MT also had a great role in the process of conceptualization, methodology, resources, software, supervision, validation, and the writing, review, and editing of the prefinal version of the manuscript. VS was also a great consultant during the process of conceptualization, methodology, supervision, validation, and writing, review, and editing of the final version of manuscript. All authors have approved the final article.

Conflicts of Interest

This study took place at Roozbeh Hospital (Community Psychiatry Subdivision, day center), which is a university hospital, and the authors of this manuscript who contributed to the design and development of the entire study including the website were active members of this department, and 2 of them are still active members.

Multimedia Appendix 1
Example of a psychoeducational video clip which was embeded in website in the second week of the course.
[MP4 File (MP4 Video), 89417 KB - formative_v5i8e23360_app1.mp4]

References

1. Merikangas KR, Jin R, He J, Kessler RC, Lee S, Sampson NA, et al. Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative. Arch Gen Psychiatry 2011 Mar;68(3):241-251 [FREE Full text] [doi: 10.1001/archgenpsychiatry.2011.12] [Medline: 21383262]
2. Sharifi V, Amin-Esmaeili M, Hajebi A, Motevalian A, Radgoodarzi R, Hefazi M, et al. Twelve-month prevalence and correlates of psychiatric disorders in Iran: the Iranian Mental Health Survey, 2011. Arch Iran Med 2015 Feb;18(2):76-84 [FREE Full text] [Medline: 25644794]
3. American Psychiatric Association. Practice guideline for the treatment of patients with bipolar disorder (revision). Am J Psychiatry 2002 Apr;159(4 Suppl):1-50. [Medline: 11958165]

XSL•FO
RenderX

4. Colom F, Vieta E, Martinez-Aran A, Reinares M, Goikolea JM, Benabarre A, et al. A randomized trial on the efficacy of group psychoeducation in the prophylaxis of recurrences in bipolar patients whose disease is in remission. Arch Gen Psychiatry 2003 Apr;60(4):402-407. [doi: 10.1001/archpsyc.60.4.402] [Medline: 12695318]
5. Demissie M, Hanlon C, Birhane R, Ng L, Medhin G, Fekadu A. Psychological interventions for bipolar disorder in low- and middle-income countries: systematic review. BJPsych Open 2018 Sep;4(5):375-384 [FREE Full text] [doi: 10.1192/bjo.2018.46] [Medline: 30202599]
6. Colom F, Vieta E, Reinares M, Martínez-Arán A, Torrent C, Goikolea JM, et al. Psychoeducation efficacy in bipolar disorders: beyond compliance enhancement. J Clin Psychiatry 2003 Sep;64(9):1101-1105. [doi: 10.4088/jcp.v64n0917] [Medline: 14628987]
7. Mottaghipour Y, Tabatabaee M. Family and Patient Psychoeducation for Severe Mental Disorder in Iran: A Review. Iran J Psychiatry 2019 Jan;14(1):84-108 [FREE Full text] [Medline: 31114622]
8. Faridhosseini F, Baniasadi M, Fayyazi Bordbar MR, Pourgholami M, Ahrari S, Asgharipour N. Effectiveness of Psychoeducational Group Training on Quality of Life and Recurrence of Patients with Bipolar Disorder. Iran J Psychiatry 2017 Jan;12(1):21-28 [FREE Full text] [Medline: 28496498]
9. González-Ortega I, Ugarte A, Ruiz de Azúa S, Núñez N, Zubia M, Ponce S, Faria, et al. Online psycho-education to the treatment of bipolar disorder: protocol of a randomized controlled trial. BMC Psychiatry 2016 Dec 22;16(1):452 [FREE Full text] [doi: 10.1186/s12888-016-1159-0] [Medline: 28007034]
10. Atherton H, Majeed A. An information revolution: time for the NHS to step up to the challenge. J R Soc Med 2011 Jun;104(6):228-230 [FREE Full text] [doi: 10.1258/jrsm.2011.110062] [Medline: 21659396]
11. High Quality Care for All: NHS Next Stage Review Final Report. National Health Service (NHS) UK. 2008. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/228836/7432.pdf [accessed 2021-07-01]
12. Hidalgo-Mazzei D, Mateu A, Reinares M, Undurraga J, Bonnín CM, Sánchez-Moreno J, et al. Self-monitoring and psychoeducation in bipolar patients with a smart-phone application (SIMPLe) project: design, development and studies protocols. BMC Psychiatry 2015 Mar 20;15:52 [FREE Full text] [doi: 10.1186/s12888-015-0437-6] [Medline: 25884824]
13. Lauder S, Chester A, Castle D, Dodd S, Gliddon E, Berk L, et al. A randomized head to head trial of MoodSwings.net.au: an Internet based self-help program for bipolar disorder. J Affect Disord 2015 Jan 15;171:13-21. [doi: 10.1016/j.jad.2014.08.008] [Medline: 25282145]
14. Parker G, Fletcher K, Blanch B, Greenfield L. Take-up and profile of individuals accessing a web-based bipolar self-test screening measure. J Affect Disord 2012 Apr;138(1-2):117-122. [doi: 10.1016/j.jad.2011.12.013] [Medline: 22284017]
15. Tabatabaee M, Mottaghipour Y, Zarghami F. Patient education manual: Bipolar disorder. Tehran: Gisa Publisher; 2014.
16. Faurholt-Jepsen M, Vinberg M, Frost M, Christensen EM, Bardram JE, Kessing LV. Smartphone data as an electronic biomarker of illness activity in bipolar disorder. Bipolar Disord 2015 Nov;17(7):715-728. [doi: 10.1111/bdi.12332] [Medline: 26395972]
17. Lauder S, Chester A, Castle D, Dodd S, Berk L, Klein B, et al. Development of an online intervention for bipolar disorder. www.moodswings.net.au. Psychol Health Med 2013;18(2):155-165. [doi: 10.1080/13548506.2012.689840] [Medline: 22712771]
18. Poole R, Simpson SA, Smith DJ. Internet-based psychoeducation for bipolar disorder: a qualitative analysis of feasibility, acceptability and impact. BMC Psychiatry 2012 Sep 13;12:139 [FREE Full text] [doi: 10.1186/1471-244X-12-139] [Medline: 22971042]
19. Batista TA, Von Werne Baes C, Juruena MF. Efficacy of psychoeducation in bipolar patients: Systematic review of randomized trials. Psychology & Neuroscience 2011 Jul;4(3):409-416. [doi: 10.3922/j.psns.2011.3.014]
20. Kanba S, Kato T, Terao T, Yamada K, Committee for Treatment Guidelines of Mood Disorders, Japanese Society of Mood Disorders, 2012. Guideline for treatment of bipolar disorder by the Japanese Society of Mood Disorders, 2012. Psychiatry Clin Neurosci 2013 Jul;67(5):285-300 [FREE Full text] [doi: 10.1111/pcn.12060] [Medline: 23773266]
21. Yatham LN, Kennedy SH, Parikh SV, Schaffer A, Beaulieu S, Alda M, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) and International Society for Bipolar Disorders (ISBD) collaborative update of CANMAT guidelines for the management of patients with bipolar disorder: update 2013. Bipolar Disord 2013 Feb;15(1):1-44. [doi: 10.1111/bdi.12025] [Medline: 23237061]
22. Nicholas J, Proudfoot J, Parker G, Gillis I, Burckhardt R, Manicavasagar V, et al. The ins and outs of an online bipolar education program: a study of program attrition. J Med Internet Res 2010 Dec 19;12(5):e57 [FREE Full text] [doi: 10.2196/jmir.1450] [Medline: 21169169]
23. Taylor-Rodgers E, Batterham PJ. Evaluation of an online psychoeducation intervention to promote mental health help seeking attitudes and intentions among young adults: randomised controlled trial. J Affect Disord 2014 Oct;168:65-71. [doi: 10.1016/j.jad.2014.06.047] [Medline: 25038293]

Abbreviations

BD: bipolar disorder

XSL•FO
RenderX

PE: psychoeducation

Edited by G Eysenbach; submitted 10.08.20; peer-reviewed by N Mohammad Gholi Mezerji, M K.; comments to author 24.09.20; revised version received 05.05.21; accepted 31.05.21; published 17.08.21.

Please cite as:
Ashrafi A, Tabatabaee M, Sharifi V
Development and Evaluation of Acceptability and Feasibility of a Web-Based Intervention for Patients With Bipolar Disorder in Iran: Implementation Study
JMIR Form Res 2021;5(8):e23360
URL: https://formative.jmir.org/2021/8/e23360
doi:10.2196/23360
PMID:34402794

©Agaah Ashrafi, Maryam Tabatabaee, Vandad Sharifi. Originally published in JMIR Formative Research (https://formative.jmir.org), 17.08.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Development of a Mobile App to Support Self-management of Anxiety and Depression in African American Women: Usability Study

Terika McCall[1,2,3], MBA, MPH, PhD; Muhammad Osama Ali[3], MPS; Fei Yu[3,4,5], MPS, PhD; Paul Fontelo[6], MPH, MD; Saif Khairat[3,7,8], MPH, PhD

[1]Center for Medical Informatics, Yale School of Medicine, New Haven, CT, United States
[2]Division of Health Informatics, Department of Biostatistics, Yale School of Public Health, New Haven, CT, United States
[3]Carolina Health Informatics Program, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States
[4]School of Information & Library Science, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States
[5]Health Sciences Library, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States
[6]National Library of Medicine, Bethesda, MD, United States
[7]School of Nursing, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States
[8]Cecil G. Sheps Center for Health Services Research, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

Corresponding Author:
Terika McCall, MBA, MPH, PhD
Center for Medical Informatics
Yale School of Medicine
300 George Street
Suite 501
New Haven, CT, 06511
United States
Phone: 1 203 737 5379
Email: terika.mccall@yale.edu

Abstract

Background: Anxiety and depressive disorders are the most common mental health conditions among African American women. Despite the need for mental health care, African American women significantly underuse mental health services. Previous mobile health studies revealed significant improvements in anxiety or depressive symptoms after intervention. The use of mobile apps offers the potential to eliminate or mitigate barriers for African American women who are seeking access to mental health services and resources.

Objective: This study aims to evaluate the usability of the prototype of an app that is designed for supporting the self-management of anxiety and depression in African American women.

Methods: Individual usability testing sessions were conducted with 15 participants in Chapel Hill, North Carolina. Cognitive walkthrough and think-aloud protocols were used to evaluate the user interface. Eye-tracking glasses were used to record participants' visual focus and gaze path as they performed the tasks. The Questionnaire for User Interface Satisfaction was administered after each session to assess the participants' acceptance of the app.

Results: Participants rated the usability of the prototype positively and provided recommendations for improvement. The average of the mean scores for usability assessments (ie, overall reactions to the software, screen, terminology and app information, learning, and app capabilities) ranged from 7.2 to 8.8 on a scale of 0-9 (low to high rating) for user tasks. Most participants were able to complete each task with limited or no assistance. Design recommendations included improving the user interface by adding graphics and color, adding a tutorial for first-time users, curating a list of Black women therapists within the app, adding details about tracking anxiety and depression in the checkup graphs, informing users that they can use the talk-to-text feature for journal entries to reduce burden, relabeling the mental health information icon, monitoring for crisis support, and improving clickthrough sequencing.

Conclusions: This study provides a better understanding of user experience with an app tailored to support the management of anxiety and depression for African American women, which is an underserved group. As African American women have high

XSL•FO
RenderX

rates of smartphone ownership, there is a great opportunity to use mobile technology to provide access to needed mental health services and resources. Future work will include incorporating feedback from usability testing and focus group sessions to refine and develop the app further. The updated app will undergo iterative usability testing before launching the pilot study to evaluate the feasibility and acceptability of the prototype.

(JMIR Form Res 2021;5(8):e24393) doi:10.2196/24393

KEYWORDS

African Americans; women; mental health; anxiety; depression; telemedicine; mHealth; mobile applications; digital health; user-centered design; mobile phone

Introduction

Background

Approximately 1 in 4 African American women in the United States has experienced mental illness [1]. Furthermore, anxiety and depressive disorders are the most common mental health conditions among African American women [2]. However, African American women significantly underuse mental health services compared to their White counterparts (12.8% vs 28.7%, respectively) [1]. Historically, mental illness has been underreported in the African American community [3,4]; therefore, the true burden may be significantly higher than the reported prevalence estimates.

Data from the 2019 *National Survey on Drug Use and Health* [1] showed that approximately 30% of non-Hispanic Black women who reported experiencing mental illness in the past year did not receive mental health treatment during that time. Barriers such as the stigma of mental illness, limited access to treatment, lack of or inadequate health insurance, mistrust of providers, and limited health literacy (LHL) all prevent traditionally marginalized populations from seeking care [5-7]. A recent survey of 395 African American women [8] revealed that the most common reasons for not seeking mental health treatment or counseling when needed were attributed to cost, not knowing where to go to access services, lack of time, and stigma. The use of mobile apps may help to eliminate or mitigate barriers by providing information on affordable options for mental health care, facilitating connections with preferred therapists, eliminating travel time using remote services, and reducing potential stigma by providing a discreet way to receive care in a preferred setting (eg, in the privacy of their home).

Numerous interventions have successfully used apps to help participants in reducing their anxiety or depressive symptoms [9-12]. Two meta-analyses of randomized controlled trials exploring the use of smartphone mental health interventions to reduce anxiety or depressive symptoms [13,14] revealed that participants experienced a significant reduction in anxiety or depressive symptoms postintervention. However, the majority of published studies were conducted with a predominantly White sample, which may affect the generalizability of the results to other racial and ethnic groups. Furthermore, a study by Sarkar et al [15], which investigated the usability of commercially available apps for depression, found that, “while patients express interest in using technologies for self-management, current tools are not consistently usable for diverse patients.” The results of a recent systematic review [16] found only 3 studies focused on culturally informed telehealth interventions for managing anxiety and depression in African American adults [17-19] and only 2 included women [17,18]. Moreover, findings from focus group interviews revealed that African American women desired a culturally informed mental health app that can address their specific needs and preferences (eg, information to find a Black woman therapist) [8]. Previous studies have shown that African American women are comfortable with participating in mobile health (mHealth) research and interventions [20,21], and 80% of African American women own smartphones [22]. On average, they spend 19 hours per week on smartphone apps [22]. Therefore, there is a great opportunity to use apps to increase access to culturally informed resources and services for supporting the management of anxiety and depression in African American women, which is a significantly underserved population.

Recent years have witnessed a growing awareness of the effectiveness of using apps for psychological interventions. This interest includes studies on both the mental health benefits and usability of apps [23]; however, scientific evidence does not support the effectiveness of most mental health apps in the market [24]. Furthermore, many apps do not incorporate any evidence-based practices or clinical expertise regarding their usability [25]. An analysis of the user reviews for mental health apps taken from the App Store and Google Play revealed that the major issues with the apps were bugs and a poor user interface design [26]. Poor usability of mental health apps contributes to low engagement [26,27].

Objective

The purpose of this study is to evaluate the usability of a prototype for an app designed for supporting the self-management of anxiety and depression in African American women. Specifically, participants evaluated the user interface on how well it helped them to complete basic tasks (eg, finding information about therapists). The findings will guide the further development of the app.

Methods

Prototype Development

The efficacious components of apps for anxiety and depression have been widely cited in the literature. Previous studies have highlighted the need for educational, psychotherapy, self-tracking, and personal development components in a mental health app that is designed to help users manage anxiety or depression [24,28-31]. Users who engaged in self-tracking and goal-setting experienced reduced depressive symptoms [30]. Furthermore, apps that offer guidance, such as personalized

XSL•FO
RenderX

feedback and supportive messages, have been shown to have a positive effect on mental health outcomes [13,32].

The mental health app evaluated in this study was developed by a multidisciplinary team at the University of North Carolina (UNC) at Chapel Hill with expertise in counseling psychology, user experience and user interface design, mobile app development, and health informatics. The initial prototype, a native app for Android devices, included basic features informed by a review of the literature and a survey of mental health and wellness apps available in the App Store and Google Play. The primary features (Textbox 1) included a guided thought journal, information about anxiety and depression (including facts about the prevalence of anxiety and depression among African American women), self-assessments for depression (using questions from the Patient Health Questionnaire 9-item scale [33]) and anxiety (using questions from the Generalized Anxiety Disorder 7-item scale [34]), mood rating, graphs to track trends in depression and anxiety severity and mood rating history, culturally informed resources (eg, links to the Therapy for Black Girls therapist directory and podcast), and a self-care planner. Figure 1 shows a screenshot of the app's home screen.

Textbox 1. Description of primary app features.

Journal

- A guided thought journal that allows the user to record their thoughts and feelings, and if applicable, prompts them to think about future actions and displays a supportive message.

Info

- Provides information about anxiety and depression (including facts about the prevalence of anxiety and depression among African American women). The user is also presented with information about symptoms, causes, treatments, and tips for managing anxiety or overcoming depression.

Checkup

- Allows the user to complete self-assessments to screen for the presence and severity of depression using questions from the Patient Health Questionnaire 9-item scale, and anxiety using questions from the Generalized Anxiety Disorder 7-item scale.

Mood

- Displays a Likert scale of images (emoticons) for users to rate how they feel.

Graphs

- Allows the user to track trends in their anxiety and depression severity, and mood rating history.

Resources

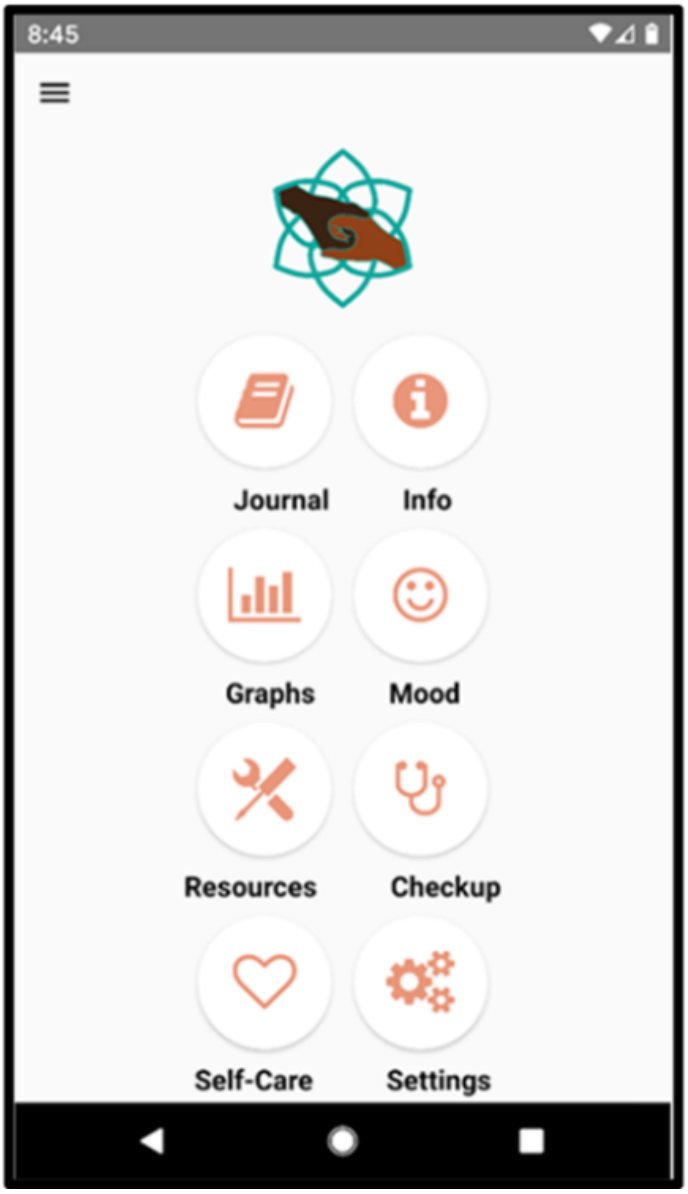
- Presents a library of culturally informed resources that links users to mental health information, therapists' directories (eg, Therapy for Black Girls), mental health and wellness podcasts (eg, Balanced Black Girl), financial assistance, and suicide crisis information.

Self-care planner

- Allows the user to create a self-care plan and checks in with the user after the chosen end date to see if the activities were completed.

XSL•FO
RenderX

Figure 1. Screenshot of the app home screen.



Recruitment

The usability study was exempted from a full review by the UNC at Chapel Hill's institutional review board. Prior research showed that 5 participants could reveal about 85% of the problems in a formative usability study [35,36]. In addition, findings from a notable study by Virzi [37] showed that 80% of usability problems were detected in 4 or 5 participants. A total of 15 participants were recruited to test the usability of the app. Participants were recruited via posts on social media (eg, Facebook and Twitter), a recruitment listing on Research for Me @UNC, and flyers posted in the Durham and Chapel Hill communities inviting women (18 years or older) who identified as Black or African American or multiracial (ie, Black or African American and another race) and had a history of anxiety or depression to participate in the study. However, study participation did not require a clinical diagnosis of an anxiety or depressive disorder. Each participant received a US $25 gift card for their completion of the study.

Procedures

In February and March 2020, individual usability testing sessions were held for each participant at the UNC at Chapel Hill School of Nursing's Biobehavioral Lab. Before the start of usability testing, a researcher went through consent forms with participants and obtained their signatures. In addition, participants were informed that the study would last approximately 1 hour.

First, the participants received a brief overview of the study aims and a description of the Tobii eye-tracking glasses and software (Tobii Pro AB) [38]. The Tobii Pro 2 glasses were calibrated for each participant before beginning the usability testing. Participants were provided with an Android mobile phone (ie, Google Pixel 2), and were assigned a persona (ie, a fictional character created to represent a target user) and scenario (ie, a fictitious story about a target user and their motivation for using the app; Multimedia Appendix 1). Then, they were asked to perform a series of 4 tasks in the app and think aloud as they completed them. To reduce participants' workload while testing

XSL•FO
RenderX

most of the app's features, the study sample was split and participants were assigned either scenario 1 or scenario 2 (Textbox 2). The instructions for each task included fictitious information for the participants to enter so that they could complete the task. No personal information was solicited or involved during the interaction between the participants and the app.

Textbox 2. List of tasks for scenarios (differences between the scenarios have been italicized).

Scenario 1 tasks

- Find out your levels of *anxiety* for the past 6 weeks.
- Find information on how to *manage anxiety.*
- *Add a new entry to your journal.*
- Locate a therapist to schedule an appointment.

Scenario 2 tasks

- Find out your levels of *depression* for the past 6 weeks.
- Find information on how to *overcome depression.*
- Create a *self-care* plan.
- Locate a therapist to schedule an appointment.

The cognitive walkthrough method [39] was used to evaluate the user interface design on how well it supported users in learning to complete tasks. Specifically, this method was used to "evaluate the ease with which users can perform a task with little or no formal instruction or informal coaching" [39]. Participants were told to speak aloud their thoughts and actions so that they could be recorded using the Tobii software. The researcher read a persona, scenario, and the first task to the participant and then instructed them to begin. This process was repeated for tasks 2-4. The Tobii software recorded the videos of participants' interactions with the app while they completed the tasks, including taps on the phone screen, eye movements, and the amount of time spent on each task.

Measures

The usability of the app was measured by participants' ability to complete tasks efficiently (measured by time and number of taps to complete each task) and their satisfaction with the user interface (measured by scores in each domain of the Questionnaire for User Interface Satisfaction [QUIS]; [40]).

Benchmarks for Cognitive Walkthrough Tasks

Benchmarks comprising a list of actions that should be performed to complete each task efficiently were created by the research team. Each task was divided into steps that should be taken to complete the actions along the *happy path* (ie, the most efficient sequence of steps to produce the desired outcome). The benchmarks were used to evaluate the participants' actions while completing the tasks (Multimedia Appendix 2).

The Questionnaire for User Interface Satisfaction

After the usability testing was completed, each participant was given a hardcopy of an adapted version of the QUIS [40] for further assessment. In the QUIS, the words *system* and *computer* were replaced with *app* in sections of the instrument. The QUIS "measures the user's subjective rating of the human-computer interface" [40]. The five domains of the QUIS covered (1) overall reaction to the software (eg, How easy was the app to use? How stimulating was the app?), (2) screen (eg, How easy was it to read the characters on the screen? How clear was the organization of information on the screen?), (3) terminology and app information (eg, How consistent was the use of terms throughout the app? How clear were the messages on the screen that prompted input from the user?), (4) learning (eg, How easy was it to explore new features through trial and error? How often could tasks be performed in a straightforward manner?), and (5) app capabilities (eg, How fast was the app? How easy was it to correct your mistakes?). Response options for each question were displayed on a Likert scale ranging from low to high (scores of 0-9). All 27 questions were weighted equally and collapsed into 5 mean scores, one for each domain and for each individual participant. The minimum and maximum mean scores, average of means, and SDs were calculated for each domain. Participants were informed to select *N/A* (not applicable) for the survey items that were not applicable. The QUIS also included 2 qualitative questions that asked participants to list the most positive and negative aspects of the app.

Statistical Analysis

Quantitative Data Analysis

Descriptive statistics were calculated as means, SDs, and ranges for continuous variables (eg, age) and as frequencies and percentages for categorical variables (eg, education level) for sample characteristics. To measure efficiency, the mean, SD, and range were calculated for the time required to complete each task and the number of taps needed to complete each task. To measure user interface satisfaction, the means, SDs, and ranges for scores in each QUIS domain were calculated. Statistical analyses were conducted using SPSS (version 26, IBM Corp) [41].

Qualitative Data Analysis

The results of the cognitive walkthrough sessions were summarized qualitatively. The Tobii eye-tracking software produced heat maps that revealed the focus of the participants'

XSL•FO
RenderX

visual attention on the screens. Furthermore, the most positive and negative aspects of the app reported on the QUIS were summarized. Specifically, common issues with the app prototype that were identified were discussed, and positive aspects were highlighted.

Results

Participants

A total of 15 participants tested the usability of the app using the cognitive walkthrough and think-aloud methods. Participants were in the age range of 20-66 years (mean age 29.8 years, SD 12.4 years), and all identified as either Black or African American or multiracial (ie, Black or African American and another race) and female. Most participants (13/15, 87%) obtained a bachelor's degree or higher and indicated that they used mobile apps 4 or more times per day (14/15, 93%). Table 1 summarizes the characteristics of the study participants.

Table 1. Characteristics of study participants (N=15).

Variables	Values
Age (years), mean (SD)	29.8 (12.4)
Education, n (%)	
Less than a bachelor's degree	2 (13)
Bachelor's degree or higher	13 (87)
Mobile app use, n (%)	
1-3 times per day	1 (7)
4 or more times per day	14 (93)

Cognitive Walkthrough

Scenario 1 Tasks

After the persona and scenario were read to the participants (n=8), they were instructed to think aloud as they completed each of the 4 tasks for scenario 1. Table 2 provides a summary of the results for the cognitive walkthrough for scenario 1 tasks. Tasks included finding the recorded levels of anxiety for the past 6 weeks, finding information on how to manage anxiety, adding a new entry in the guided thought journal, and locating a therapist to schedule an appointment.

Table 2. Summary of results for the cognitive walkthrough for scenario 1 tasks (n=8).

Task	Completed, n (%)	Benchmark		Participant outcomes	
		Time	Taps, n	Time, range	Taps, range
Find out your levels of anxiety for the past 6 weeks	8 (100)	13 seconds	3	20 seconds to 1 minute 41 seconds	3-6
Find information on how to manage anxiety	8 (100)	25 seconds	6	48 seconds to 4 minutes 24 seconds	6-37
Add a new entry to your journal	5 (63)	3 minutes 10 seconds	15	2 minutes 48 seconds to 6 minutes 10 seconds	17-19
Locate a therapist to schedule an appointment	8 (100)	56 seconds	8	1 minute 42 seconds to 3 minutes 21 seconds	10-35

Most participants were able to complete each task fully with limited or no assistance. On average, participants took longer to complete the tasks than the benchmark times. Adding a new entry to the journal proved to be the most cumbersome task as it required participants to be thoughtful when they entered text and selected their feelings from a list. Two participants had difficulty with locating the information on how to manage anxiety, searching in the resources section of the app instead of information. To locate a therapist, participants were required to tap a button that was linked to a website with a directory of primarily Black women therapists. Although all participants were able to complete the task of locating a therapist, many reported that the interface of the website was not mobile friendly and required a lot of scrolling to find out if the therapist was accepting new clients. The actions completed, number of taps, and the amount of time spent on each task were recorded for each participant in Multimedia Appendix 3.

Scenario 2 Tasks

Following the protocol, the persona and scenario were read to the participants (n=7). Next, they were instructed to think aloud as they completed each of the 4 tasks. Table 3 provides a summary of the results for the cognitive walkthrough for scenario 2 tasks. Tasks included finding the recorded levels of depression for the past 6 weeks, finding information on how to overcome depression, creating a plan for self-care, and locating a therapist to schedule an appointment.

XSL•FO
RenderX

Table 3. Summary of results for the cognitive walkthrough for scenario 2 tasks (n=7).

Task	Completed, n (%)	Benchmark		Participant outcomes	
		Time	Taps, n	Time, range	Taps, range
Find out your levels of depression for the past 6 weeks	7 (100)	14 seconds	3	27 seconds to 48 seconds	3-6
Find information on how to overcome depression	7 (100)	20 seconds	6	28 seconds to 2 minutes 14 seconds	6-14
Create a self-care plan	7 (100)	1 minute 17 seconds	15	1 minute 17 seconds to 2 minutes 47 seconds	16-23
Locate a therapist to schedule an appointment	7 (100)	49 seconds	8	1 minute 7 seconds to 2 minutes 9 seconds	9-13

All participants were able to determine their depression levels for the past 6 weeks and reported the most recent level of depression recorded in the graph. One participant had difficulty with locating the information on how to overcome depression and searched in the resources section of the app instead of the information feature. The participants liked the ability to track their self-care. The time to complete a self-care plan varied according to the typing speed of the participants. The actions completed, number of taps, and the amount of time spent on each task were recorded for each participant in Multimedia Appendix 4.

QUIS Scores

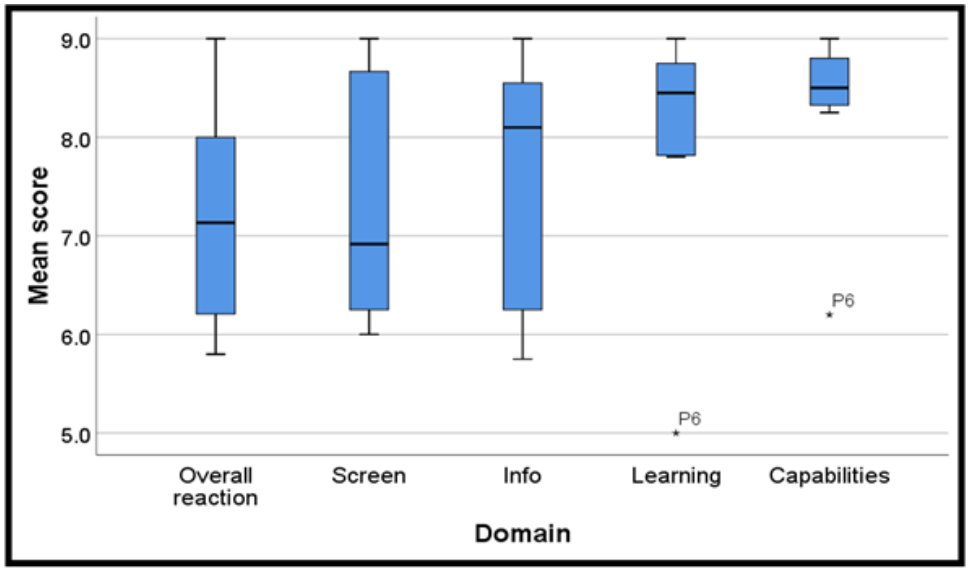
Scenario 1 Tasks

The average of mean scores for each domain ranged from 7.2 to 8.3 on a scale of 0-9 (low to high rating; Table 4). The average of mean scores for the overall reaction to the software was 7.2 (SD 1.1), that for screen was 7.3 (SD 1.3), that for terminology and app information was 7.6 (SD 1.3), that for learning was 8.0 (SD 1.3), and that for app capabilities was 8.3 (SD 0.9). Figure 2 displays a boxplot of the mean scores for the 5 domains of the QUIS for scenario 1 tasks. There was 1 outlier (participant #6) with a score of 5.0 for the learning domain and 6.2 for app capabilities.

Table 4. Summary of mean scores for the 5 domains of the Questionnaire for User Interface Satisfaction for scenario 1 tasks (n=8).

Domain	Mean scores, range	Average of mean scores, mean (SD)
Overall reaction to the software	5.8-9.0	7.2 (1.1)
Screen	6.0-9.0	7.3 (1.3)
Terminology and app information	5.8-9.0	7.6 (1.3)
Learning	5.0-9.0	8.0 (1.3)
App capabilities	6.2-9.0	8.3 (0.9)

Figure 2. Boxplots for the 5 domains of the Questionnaire for User Interface Satisfaction for scenario 1 tasks. P6: participant #6.



XSL•FO
RenderX

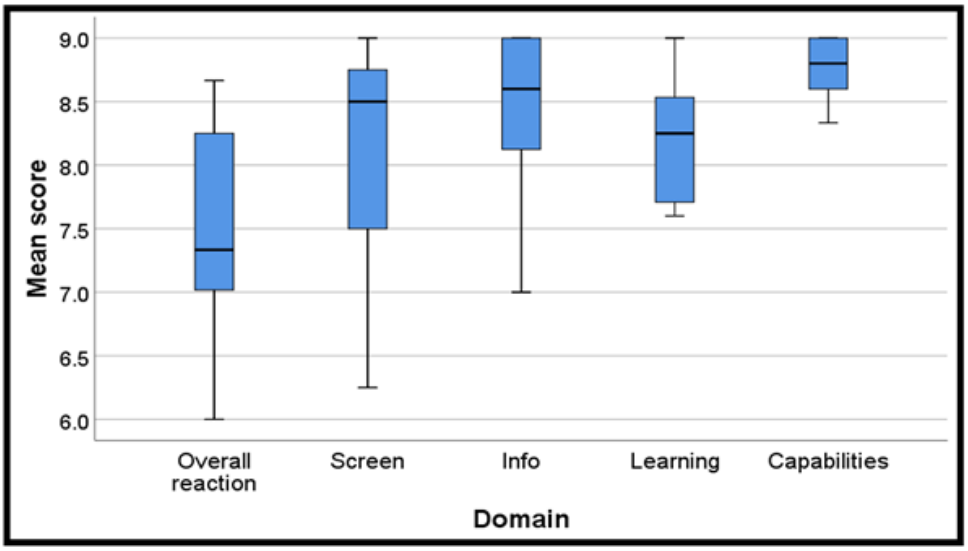
Scenario 2 Tasks

The average of mean scores for each domain ranged from 7.5 to 8.8 on a scale of 0-9 (low to high rating; Table 5). The average of mean scores for the overall reaction to the software was 7.5 (SD 1.0), that for screen was 8.0 (SD 1.0), that for terminology and app information was 8.4 (SD 0.8), that for learning was 8.2 (SD 0.5), and that for app capabilities was 8.8 (SD 0.3). Figure 3 displays a boxplot of the mean scores for the 5 domains of the QUIS for scenario 2 tasks.

Table 5. Summary of mean scores for the 5 domains of the Questionnaire for User Interface Satisfaction for scenario 2 tasks (n=7).

Domain	Mean score, range	Average of mean scores, mean (SD)
Overall reaction to the software	6.0-8.7	7.5 (1.0)
Screen	6.3-9.0	8.0 (1.0)
Terminology and app information	7.0-9.0	8.4 (0.8)
Learning	7.6-9.0	8.2 (0.5)
App capabilities	8.3-9.0	8.8 (0.3)

Figure 3. Boxplots for the 5 domains of the Questionnaire for User Interface Satisfaction for scenario 2 tasks.



Qualitative Results

Table 6 shows the participants' feedback on the app design. Overall, participants felt that the app was easy to use, organized well, and had fast processing speed. General recommendations for improvement included adding more graphics and color, including instructions for data entry textboxes, a tutorial for first-time users, and crisis support monitoring. Regarding content, participants thought the information provided was of high quality, they also liked that links were provided to helpful outside resources (including a link to Black women therapists with their availability), and that graphs and text provided information to track their anxiety and depression severity. Participants recommended that a curated list of Black women therapists should be provided within the app to avoid linking externally to websites that are not mobile friendly (eg, required too much scrolling).

XSL•FO
RenderX

Table 6. Feedback on app design.

Themes	Positive feedback	Recommendations
Overall reactions to the app	• Easy to use • Organized well • Fast processing speed • Every screen had a header • App saves information entered and indicates the completion of task • Nice logo	• Improve user interface by adding more graphics and color • Add instructions under start and end date fields to explain what to enter in the textboxes • Add a tutorial for first-time use • Monitor for crisis support
Content	• Quality of information • Linked to helpful outside resources • List of Black women therapists with their availability • Graphs and text provide information to track anxiety and depression • Tips to help manage anxiety	• Curate a list of Black women therapists in the app to avoid linking externally to websites that are not mobile friendly
Features	• Tracking anxiety and depression • Self-care planning • Guided thought journal	• Add footnotes to anxiety and depression history graphs that it can be used to track trends • Informing users that they can use the talk-to-text feature for journal entries to reduce burden
Navigation and error prevention	• Icons were helpful for guidance on navigating the app • Drop-down menu clearly lists all features • Great prompts in the guided thought journal • System uses highlighting to indicate the selection made • System alerts when there is a mistake or missing information	• Relabel information icon to make clear it contains mental health information • Improve clickthrough sequencing

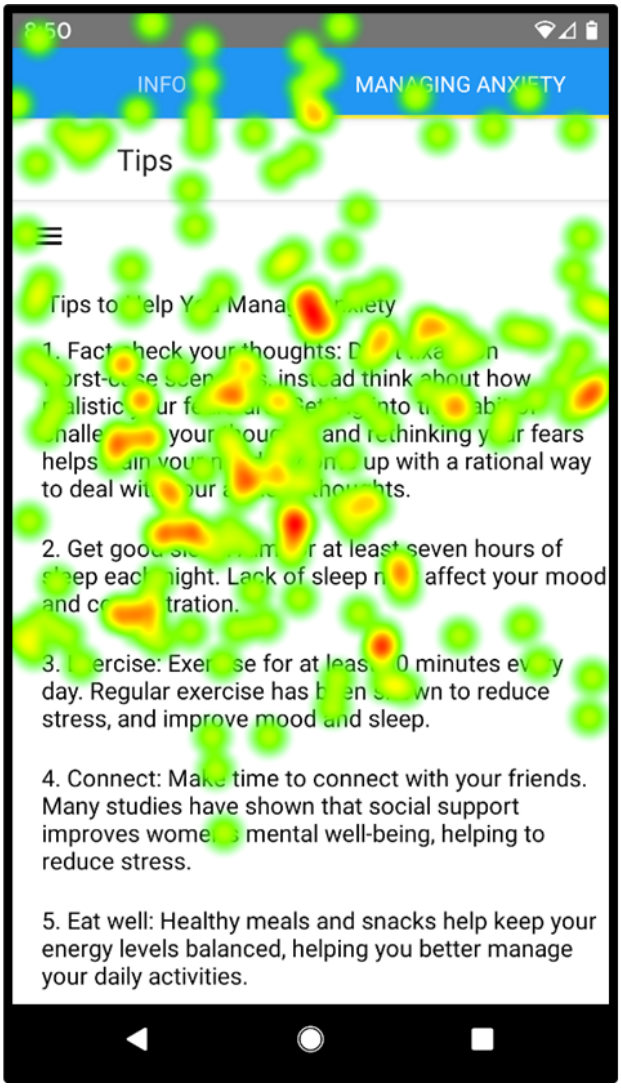
Anxiety and depression checkups, self-care planner, and guided thought journal features were rated positively. Recommendations for improvement included adding a footnote to anxiety and depression history graphs to indicate that it can be used to track trends and adding instructions to the journal feature for informing users that they can use the talk-to-text feature for journal entries to reduce burden. Regarding navigation and error prevention, participants thought that the icons were helpful for guidance on navigating the app and liked that the drop-down menu clearly lists all features. They also noted the great prompts in the guided thought journal. Participants recommended that the information feature be relabeled to elucidate that it contains mental health information and that clickthrough sequencing be improved.

Heat Maps

The eye-tracking software showed that participants primarily focused on the left and middle areas of the screen while looking for information. The heat map for the managing anxiety tips screen (Figure 4) revealed that participants spent the most time viewing the title, left side, and middle of the screen (the areas of highest intensity). Similarly, Figure 5 shows that on the heat map of the journal entry summary screen, the titles, left side of the screen, and *Save* button were the areas where participants focused on most (the areas of highest intensity).

XSL•FO
RenderX

Figure 4. Heat map of the managing anxiety tips screen showing areas that participants focused on most.



XSL•FO
RenderX

Figure 5. Heat map of the journal entry summary screen showing areas that participants focused on most.

Discussion

Principal Findings

To our knowledge, this is one of the first mobile apps specifically designed to support the self-management of anxiety and depression in African American women, irrespective of physical health conditions or special circumstances. This is an important distinction because there are studies that focus on culturally informed mental health interventions for Black women; however, the participants had a particular physical health condition (eg, HIV positive) [18] or special circumstances (eg, caregivers of patients with dementia) [17]. Our results demonstrate that the participants were mostly satisfied with the user interface of the app prototype. Moreover, the average of means scores for the overall reaction to the software, screen, terminology, app information, learning, and app capabilities were high. They ranged from 7.2 to 8.3 on a scale of 0-9 (low to high rating) for scenario 1 tasks (Table 4), and 7.5 to 8.8 for scenario 2 tasks (Table 5). Most participants were able to complete each task fully with limited or no assistance. However, the variance in time to complete each task was primarily attributed to the amount of time it took for participants to talk through their thoughts and actions, differences in typing speed, or using an alternative method to search (eg, using a link from the hamburger menu instead of tapping the feature icon).

Furthermore, a few participants were unsure if they needed to first record their anxiety levels in the app before locating it in the graph and confused about where to find information about anxiety and depression, looking at the resources or self-care features of the app instead of the information feature. Participants expressed that they thought the information feature was the place to find information about the app and not about anxiety. Moreover, although the majority of the participants had no problem finding the button within the app that was linked to the website with the therapist directory, some participants expressed that the website itself was not mobile friendly and required a lot of scrolling to find out if the therapist was accepting new clients.

The aforementioned usability testing results produced the following considerations for the development of a mental health app to support the self-management of anxiety and depression in African American women: (1) the app should be intuitive and easy to use, (2) the app should include a feature to self-monitor mental health (eg, depression severity monitoring), (3) the app should allow users to learn coping skills (eg, tips on how to overcome depression), (4) the app should connect users

XSL•FO
RenderX

with needed resources (eg, therapist), and (5) the app should provide users with the option to plan activities for self-care. In addition, the interface should be visually appealing and have clear labeling. These recommendations are consistent with those found in the literature that highlight the need for educational, psychotherapy, and personal development components in a mental health app designed to help users manage anxiety and depression [28,29]. Having a tutorial on how to use the app features and find content would also help improve user satisfaction and increase engagement. Furthermore, culturally tailoring the content and resources helps personalize the app to meet the specific needs and preferences of African American women.

The eye-tracking software demonstrated that the participants primarily focused on the left side and middle of the screen when looking for information. This is consistent with a previous study that showed that people spend 80% of their time looking at the left side of the screen [42]. Developers should consider placing important information on the left side or middle of the screen to make it easier for users to find. As external sites that the app links to may not be mobile friendly or have good usability, consider either only linking to mobile-friendly sites or placing all important content within the app. In addition, navigation buttons should be large and spaced well so that users with longer fingernails do not have difficulty tapping them. According to Fitts Law, the longer the distance to the target and the smaller the size of the target, the longer it takes to complete the movement [43]. Feedback from the usability testing sessions informed the current design and development of the app, in terms of both anesthetics (eg, layout, color scheme) and information architecture (eg, renaming the information feature).

Downloading an app does not necessarily mean that the user will continue to use it consistently in the long term. Past research conducted on the usage patterns of mental health apps showed that despite the high amount of initial installations and daily active minutes spent, only a small percentage of users continued to use the apps for longer than a couple of weeks [44]. There are several challenges in using apps for mental health care. Factors that affect the uptake include both the health and digital literacy of users. Approximately one-third of adults in the United States have LHL [45]. African American adults have a higher prevalence of LHL than their White counterparts [45]. Individuals with LHL are less likely to use digital devices for health-related purposes, which can hinder the use of smartphone mental health app interventions [46]. To promote adoption among users with LHL, smartphone health interventions should incorporate features that target a person's own health literacy needs and technical skills [47].

Furthermore, poor usability in which an app may appear to be filled with bugs if a function or more does not seem to work, also affects the long-term use of the app. Another challenge is the lack of user engagement within apps that can cause app use to decline [27]. A previous case study on building a highly rated mental health app showed that users tend to prefer apps that focus on self-development and change rather than interventions from external sources [28]. The app tested in this usability study will promote engagement by prompting users to enter their mood daily. Users are also encouraged to record their thoughts and feelings in the guided thought journal for reflection and planning future actions. Push notifications will also remind users to complete the anxiety (Generalized Anxiety Disorder 7-item scale) and depression (Patient Health Questionnaire 9-item scale) checkups every 2 weeks to track progress. In addition, the app checks in with the user after the chosen end date to confirm whether activities in their self-care plan were completed. Tailored feedback is incorporated into some components of the app, such as the guided thought journal, mood tracking, and self-care plan.

When designing the user interface of the app itself, several variables should be considered. The experience of the user with app usage, their background, how they might function under stress while using the app, and the environment in which the app is being used are just some of the important factors that should be considered to ensure a positive user experience. A good mental health app should consider users' need for support, sociocultural factors, and personal development goals. For example, an app focusing on providing anxiety management should educate its users about anxiety, provide culturally informed self-help options to manage their anxiety through low-intensity techniques, and give users the ability to track their progress and connect with preferred providers for additional support if necessary. Self-care apps should assume that their users are independent and provide support to them when needed [27]. A *one-size-fits-all* approach to designing mHealth interventions may lead to more options but continued disparity in receiving mental health care. The inequitable design of digital health tools further perpetuates the exclusion of underserved populations from mental health care [48]. Incorporating recommendations from intended users and knowledge of their technology use and behaviors can help mitigate potential intervention-generated inequalities [49].

One caveat is that the use of apps to receive mental health support may not be appropriate for everyone. If used as an adjunct to therapy, clients should be screened to determine whether the use of this modality is appropriate for treatment [50]. A survey of African American women revealed that video calls were an acceptable modality to communicate with a professional to receive help in managing anxiety or depression [20], whereas text messaging was not [51]. Therefore, an option to communicate with a mental health professional via video calls within the app may be a useful feature to include. The extent to which a mental health app is used may vary. It can be used for self-management only, peer support, the primary modality to receive mental health care from a professional (eg, telecounseling), or as an adjunct to in-person or other methods for remote counseling.

Strengths and Limitations

The main strengths of this usability study are its rigorous design and use of the cognitive walkthrough and think-aloud method, eye-tracking technology-assisted usability evaluation, and administration of the QUIS to capture user performance, physiological data, and qualitative feedback completely. In addition, 15 participants were recruited to participate in the usability testing of the app, thus providing a more than adequate sample size.

XSL•FO
RenderX

One of the main limitations was that our participants were mostly younger women (under 50 years old) with at least a bachelor's degree. This may limit the generalizability of the findings to older African American women and those with less than a bachelor's degree. However, previously reported statistics revealed that younger Black women had a higher prevalence of lifetime anxiety (eg, generalized anxiety disorder) and mood disorders (eg, major depressive disorder) than older Black women (50 years or older) [2]. Although access to mental health services and resources may be less of an issue for African American women with at least a bachelor's degree, access to culturally informed resources and a professional that meets the ethnicity and gender preferences of the patient may remain an issue as less than 5% of active psychologists in the United States are African American women [52].

In addition, the geographical restriction in recruiting participants may have resulted in the opinions and perceptions of the participants not reflecting those of a nationally representative sample. Another limitation was that this study did not focus on the efficacy of the app to reduce anxiety and depressive symptoms, as this would require a randomized controlled trial and significant resources to be done properly. Despite these limitations, the study yielded useful information that provides guidance for designing an app to help African American women and other populations in managing anxiety and depression.

Conclusions and Future Directions

African American women have high rates of smartphone ownership (80%) [22], and there is a great opportunity to use mobile technology to provide mental health resources and services to them. Poor usability severely affects the engagement and effectiveness of mHealth interventions. Therefore, usability testing should be incorporated in the design and development of mental health apps to increase adoption, engagement, and user satisfaction. This study contributes to an improved understanding of users' experiences with an app tailored to support the self-management of anxiety and depression in African American women, which is an underserved population. It is recommended that future researchers and app designers consider the proposed content, features, and considerations highlighted by this study while developing a mental health app for this population to enhance the user experience. We are planning to further develop the app by incorporating feedback from this study, and the new version will undergo iterative usability testing before the launch of the pilot study to evaluate the feasibility and acceptability of the prototype.

Acknowledgments

The authors would like to acknowledge the members of Computer Science + Social Good at the UNC at Chapel Hill for assisting with the development of the app. In addition, the authors also thank Dr Craig Locatis, Fang Liu, Dr Clinton Bolton III, and Malvika Pillai for their support in the development of the app. The first author (TM) was supported by funding from the National Library of Medicine's Institutional Training Grant for Research Training in Biomedical Informatics and Data Science at the Carolina Health Informatics Program (T15LM012500) and Yale Center for Medical Informatics (T15LM007056). This research was also supported by the Intramural Research Program of the National Library of Medicine, National Institutes of Health.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Usability testing scenarios and tasks.
[PDF File (Adobe PDF File), 120 KB - formative_v5i8e24393_app1.pdf]

Multimedia Appendix 2
Cognitive walkthrough benchmarks.
[PDF File (Adobe PDF File), 1449 KB - formative_v5i8e24393_app2.pdf]

Multimedia Appendix 3
Usability testing participant performance versus benchmark for scenario 1 tasks.
[PDF File (Adobe PDF File), 1229 KB - formative_v5i8e24393_app3.pdf]

Multimedia Appendix 4
Usability testing participant performance versus benchmark for scenario 2 tasks.
[PDF File (Adobe PDF File), 1095 KB - formative_v5i8e24393_app4.pdf]

References

1. National Survey on Drug Use and Health 2019 (NSDUH-2019-DS0001) [data file]. U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, and Center for Behavioral Health Statistics and Quality. 2020. URL: https://datafiles.samhsa.gov [accessed 2021-04-06]

XSL•FO
RenderX

2. Lacey KK, Parnell R, Mouzon DM, Matusko N, Head D, Abelson JM, et al. The mental health of US Black women: the roles of social context and severe intimate partner violence. BMJ Open 2015 Oct 19;5(10):e008415 [FREE Full text] [doi: 10.1136/bmjopen-2015-008415] [Medline: 26482770]
3. Mental Health: Culture, Race, and Ethnicity. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2001.
4. Takayanagi Y, Spira AP, Roth KB, Gallo JJ, Eaton WW, Mojtabai R. Accuracy of reports of lifetime mental and physical disorders: results from the Baltimore Epidemiological Catchment Area study. JAMA Psychiatry 2014 Mar;71(3):273-280 [FREE Full text] [doi: 10.1001/jamapsychiatry.2013.3579] [Medline: 24402003]
5. Thompson VLS, Bazile A, Akbar M. African Americans' perceptions of psychotherapy and psychotherapists. Prof Psychol Res Pr 2004;35(1):19-26. [doi: 10.1037/0735-7028.35.1.19]
6. Hines-Martin V, Malone M, Kim S, Brown-Piper A. Barriers to mental health care access in an African American population. Ment Health Nurs 2003;24(3):237-256. [doi: 10.1080/01612840305281]
7. Merritt-Davis OB, Keshavan MS. Pathways to Care for African Americans With Early Psychosis. Psychiatr Serv 2006 Jul;57(7):1043-1044. [doi: 10.1176/ps.2006.57.7.1043]
8. McCall T. Mhealth for Mental Health: Culturally-tailored Interventions for Managing Anxiety and Depression in African American Women [doctoral Dissertation]. ProQuest. 2020. URL: https://search.proquest.com/docview/2436913713 [accessed 2020-08-27]
9. Arean PA, Hallgren KA, Jordan JT, Gazzaley A, Atkins DC, Heagerty PJ, et al. The use and effectiveness of mobile apps for depression: results from a fully remote clinical trial. J Med Internet Res 2016 Dec 20;18(12):e330 [FREE Full text] [doi: 10.2196/jmir.6482] [Medline: 27998876]
10. Clarke J, Proudfoot J, Birch MR, Whitton AE, Parker G, Manicavasagar V, et al. Effects of mental health self-efficacy on outcomes of a mobile phone and web intervention for mild-to-moderate depression, anxiety and stress: secondary analysis of a randomised controlled trial. BMC Psychiatry 2014 Sep 26;14:272 [FREE Full text] [doi: 10.1186/s12888-014-0272-1] [Medline: 25252853]
11. Proudfoot J, Clarke J, Birch MR, Whitton AE, Parker G, Manicavasagar V, et al. Impact of a mobile phone and web program on symptom and functional outcomes for people with mild-to-moderate depression, anxiety and stress: a randomised controlled trial. BMC Psychiatry 2013 Nov 18;13:312 [FREE Full text] [doi: 10.1186/1471-244X-13-312] [Medline: 24237617]
12. Moberg C, Niles A, Beermann D. Guided self-help works: randomized waitlist controlled trial of pacifica, a mobile app integrating cognitive behavioral therapy and mindfulness for stress, anxiety, and depression. J Med Internet Res 2019 Jun 8;21(6):e12556 [FREE Full text] [doi: 10.2196/12556] [Medline: 31199319]
13. Firth J, Torous J, Nicholas J, Carney R, Pratap A, Rosenbaum S, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. World Psychiatry 2017 Oct;16(3):287-298 [FREE Full text] [doi: 10.1002/wps.20472] [Medline: 28941113]
14. Firth J, Torous J, Nicholas J, Carney R, Rosenbaum S, Sarris J. Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. J Affect Disord 2017 Aug 15;218:15-22 [FREE Full text] [doi: 10.1016/j.jad.2017.04.046] [Medline: 28456072]
15. Sarkar U, Gourley GI, Lyles CR, Tieu L, Clarity C, Newmark L, et al. Usability of commercially available mobile applications for diverse patients. J Gen Intern Med 2016 Dec;31(12):1417-1426. [doi: 10.1007/s11606-016-3771-6] [Medline: 27418347]
16. McCall T, Bolton CS, Carlson R, Khairat S. A systematic review of telehealth interventions for managing anxiety and depression in African American adults. Mhealth 2021;7:31 [FREE Full text] [doi: 10.21037/mhealth-20-114] [Medline: 33898600]
17. Glueckauf RL, Davis WS, Willis F, Sharma D, Gustafson DJ, Hayes J, et al. Telephone-based, cognitive-behavioral therapy for African American dementia caregivers with depression: initial findings. Rehabil Psychol 2012 May;57(2):124-139. [doi: 10.1037/a0028688] [Medline: 22686551]
18. Himelhoch S, Mohr D, Maxfield J, Clayton S, Weber E, Medoff D, et al. Feasibility of telephone-based cognitive behavioral therapy targeting major depression among urban dwelling African-American people with co-occurring HIV. Psychol Health Med 2011 Mar;16(2):156-165 [FREE Full text] [doi: 10.1080/13548506.2010.534641] [Medline: 21328144]
19. Hightow-Weidman LB, Muessig KE, Pike EC, LeGrand S, Baltierra N, Rucker AJ, et al. Healthmpowerment.Org: building community through a mobile-optimized, online health promotion intervention. Health Educ Behav 2015 Aug;42(4):493-499. [doi: 10.1177/1090198114562043] [Medline: 25588932]
20. McCall T, Schwartz T, Khairat S. Acceptability of Telemedicine to Help African American Women Manage Anxiety and Depression. Stud Health Technol Inform 2019 Aug 21;264:699-703. [doi: 10.3233/SHTI190313] [Medline: 31438014]
21. James DCS, Harville C, Whitehead N, Stellefson M, Dodani S, Sears C. Willingness of African American Women to Participate in e-Health/m-Health Research. Telemed J E Health 2016 Mar;22(3):191-197. [doi: 10.1089/tmj.2015.0071] [Medline: 26313323]
22. Reaching Black Women Across Media Platforms. The Nielsen Company. 2017 Dec 11. URL: www.nielsen.com/us/en/insights/news/2017/reaching-black-women-across-media-platforms.html [accessed 2018-11-14]

XSL•FO
RenderX

23. Miralles I, Granell C, Díaz-Sanahuja L, Van Woensel W, Bretón-López J, Mira A, et al. Smartphone Apps for the Treatment of Mental Disorders: Systematic Review. JMIR Mhealth Uhealth 2020 Apr 02;8(4):e14897 [FREE Full text] [doi: 10.2196/14897] [Medline: 32238332]
24. Van Ameringen M, Turna J, Khalesi Z, Pullia K, Patterson B. There is an app for that! The current state of mobile applications (apps) for DSM-5 obsessive-compulsive disorder, posttraumatic stress disorder, anxiety and mood disorders. Depress Anxiety 2017 Jun;34(6):526-539. [doi: 10.1002/da.22657] [Medline: 28569409]
25. Shen N, Levitan MJ, Johnson A, Bender JL, Hamilton-Page M, Jadad AAR, et al. Finding a depression app: a review and content analysis of the depression app marketplace. JMIR Mhealth Uhealth 2015;3(1):e16 [FREE Full text] [doi: 10.2196/mhealth.3713] [Medline: 25689790]
26. Alqahtani F, Orji R. Usability issues in mental health applications. USA: ACM; 2019 Presented at: 27th Conference on User Modeling, Adaptation and Personalization; June 9 - 12, 2019; Larnaca, Cyprus. [doi: 10.1145/3314183.3323676]
27. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. Evid Based Ment Health 2018 Jun 05:116-119. [doi: 10.1136/eb-2018-102891] [Medline: 29871870]
28. Topham P, Caleb-Solly P, Matthews P, Farmer A, Mash C. Mental health app design: A journey from concept to completion. : ACM Press; 2015 Presented at: 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct - MobileHCI '15; August 24 - 27, 2015; Copenhagen, Denmark p. 582-591. [doi: 10.1145/2786567.2787136]
29. Huguet A, Rao S, McGrath PJ, Wozney L, Wheaton M, Conrod J, et al. A Systematic Review of Cognitive Behavioral Therapy and Behavioral Activation Apps for Depression. PLoS One 2016;11(5):e0154248 [FREE Full text] [doi: 10.1371/journal.pone.0154248] [Medline: 27135410]
30. Zhang R, Nicholas J, Knapp AA, Graham AK, Gray E, Kwasny MJ, et al. Clinically Meaningful Use of Mental Health Apps and its Effects on Depression: Mixed Methods Study. J Med Internet Res 2019 Dec 20;21(12):e15644 [FREE Full text] [doi: 10.2196/15644] [Medline: 31859682]
31. Rathbone AL, Clarry L, Prescott J. Assessing the Efficacy of Mobile Health Apps Using the Basic Principles of Cognitive Behavioral Therapy: Systematic Review. J Med Internet Res 2017 Nov 28;19(11):e399 [FREE Full text] [doi: 10.2196/jmir.8598] [Medline: 29187342]
32. Linardon J, Cuijpers P, Carlbring P, Messer M, Fuller-Tyszkiewicz M. The efficacy of app-supported smartphone interventions for mental health problems: a meta-analysis of randomized controlled trials. World Psychiatry 2019 Oct;18(3):325-336 [FREE Full text] [doi: 10.1002/wps.20673] [Medline: 31496095]
33. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [Medline: 11556941]
34. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
35. Lewis JR. Usability: Lessons Learned … and Yet to Be Learned. International Journal of Human-Computer Interaction 2014;30(9):663-684. [doi: 10.1080/10447318.2014.930311]
36. Nielsen J, Landauer TK. A mathematical model of the finding of usability problems. Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '93 New York, New York, USA: ACM Press; 1993 Presented at: SIGCHI conference on Human factors in computing systems - CHI '93; April 24 - 29, 1993; Amsterdam, Netherlands p. 206-213. [doi: 10.1145/169059.169166]
37. Virzi RA. Refining the Test Phase of Usability Evaluation: How Many Subjects Is Enough? Hum Factors 1992;34(4):457-468. [doi: 10.1177/001872089203400407]
38. Tobii Pro AB. Tobii Pro Lab. Danderyd, Sweden; 2020. URL: https://www.tobiipro.com/ [accessed 2020-01-10]
39. Polson PG, Lewis C, Rieman J, Wharton C. Cognitive walkthroughs: a method for theory-based evaluation of user interfaces. International Journal of Man-Machine Studies 1992 May;36(5):741-773. [doi: 10.1016/0020-7373(92)90039-n]
40. Chin JP, Diehl VA, Norman LK. Development of an instrument measuring user satisfaction of the human-computer interface. Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '88 New York, New York, USA: ACM Press; 1988 Presented at: SIGCHI conference on Human factors in computing systems - CHI '88; May 15 - 19, 1988; Washington D.C p. 213-218. [doi: 10.1145/57167.57203]
41. IBM SPSS Statistics for Windows. Armonk, NY: IBM Corp; 2019. URL: https://www.ibm.com/products/spss-statistics [accessed 2020-10-12]
42. Fessenden T, Nielsen Norman Group. Horizontal Attention Leans Left. 2017. URL: https://www.nngroup.com/articles/horizontal-attention-leans-left/ [accessed 2020-09-16]
43. Fitts PM. The information capacity of the human motor system in controlling the amplitude of movement. J Exp Psychol 1954 Jun;47(6):381-391. [Medline: 13174710]
44. Baumel A, Muench F, Edan S, Kane JM. Objective User Engagement With Mental Health Apps: Systematic Search and Panel-Based Usage Analysis. J Med Internet Res 2019 Sep 25;21(9):e14567 [FREE Full text] [doi: 10.2196/14567] [Medline: 31573916]
45. White S, Chen J, Atchison R. Relationship of preventive health practices and health literacy: a national study. Am J Health Behav 2008;32(3):227-242. [doi: 10.5555/ajhb.2008.32.3.227] [Medline: 18067463]

XSL•FO
RenderX

46. McCleary-Jones V, Scheideman-Miller C, Rev Dorn JA, Johnson B, Overall M, Dwyer K. Health information technology use and health literacy among community-dwelling African Americans. ABNF J 2013;24(1):10-16. [Medline: 23589966]
47. Hur I, Lee R, Schmidt JJ. How healthcare technology shapes health literacy? A systematic review. 2015 Presented at: 2015 Americas' Conference on Information System; August 13-15, 2015; Fajardo, Puerto Rico.
48. Goedhart NS, Zuiderent-Jerak T, Woudstra J, Broerse JEW, Betten AW, Dedding C. Persistent inequitable design and implementation of patient portals for users at the margins. J Am Med Inform Assoc 2021 Feb 15;28(2):276-283. [doi: 10.1093/jamia/ocaa273] [Medline: 33463691]
49. Veinot TC, Mitchell H, Ancker JS. Good intentions are not enough: how informatics interventions can worsen inequality. J Am Med Inform Assoc 2018 Aug 01;25(8):1080-1088. [doi: 10.1093/jamia/ocy052] [Medline: 29788380]
50. Hilty DM, Ferrer DC, Parish MB, Johnston B, Callahan EJ, Yellowlees PM. The effectiveness of telemental health: a 2013 review. Telemed J E Health 2013 Jun;19(6):444-454 [FREE Full text] [doi: 10.1089/tmj.2013.0075] [Medline: 23697504]
51. McCall T, Schwartz TA, Khairat S. The Acceptability of Text Messaging to Help African American Women Manage Anxiety and Depression: Cross-Sectional Survey Study. JMIR Ment Health 2020 Feb 03;7(2):e15801 [FREE Full text] [doi: 10.2196/15801] [Medline: 31909720]
52. American Psychological Association. Demographics of the U.S. psychology workforce: Findings from the 2007-16 American Community Survey. Washington, DC; 2018. URL: https://www.apa.org/workforce/publications/16-demographics/report.pdf [accessed 2020-07-03]

Abbreviations

LHL: limited health literacy
mHealth: mobile health
QUIS: Questionnaire for User Interface Satisfaction
UNC: University of North Carolina

Edited by G Eysenbach; submitted 21.09.20; peer-reviewed by X He, PKH Mo, J op den Buijs; comments to author 07.11.20; revised version received 28.02.21; accepted 15.06.21; published 17.08.21.

Please cite as:
McCall T, Ali MO, Yu F, Fontelo P, Khairat S
Development of a Mobile App to Support Self-management of Anxiety and Depression in African American Women: Usability Study
JMIR Form Res 2021;5(8):e24393
URL: https://formative.jmir.org/2021/8/e24393
doi:10.2196/24393
PMID:34133313

©Terika McCall, Muhammad Osama Ali, Fei Yu, Paul Fontelo, Saif Khairat. Originally published in JMIR Formative Research (https://formative.jmir.org), 17.08.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Utilization of a Directly Supervised Telehealth-Based Exercise Training Program in Patients With Nonalcoholic Steatohepatitis: Feasibility Study

Victoria Motz[1], MD; Alison Faust[1], MD; Jessica Dahmus[1], MD; Benjamin Stern[1], DO; Christopher Soriano[1], MD; Jonathan G Stine[1], MD, MSc

Penn State Milton S Hershey Medical Center, Hershey, PA, United States

Corresponding Author:
Jonathan G Stine, MD, MSc
Penn State Milton S Hershey Medical Center
500 University Dr
Hershey, PA, 17033
United States
Phone: 1 717 531 1017
Email: jstine@pennstatehealth.psu.edu

Abstract

Background: Most patients with nonalcoholic fatty liver disease (NAFLD) are physically inactive despite the well-known benefits of physical activity. Telehealth offers promise as a novel way to deliver an exercise training program and increase physical activity. However, the feasibility, safety, and efficacy of telehealth-based exercise programs is unknown in patients with NAFLD.

Objective: The aim of this study was to determine the feasibility of a directly supervised exercise training program delivered exclusively with telehealth to patients with nonalcoholic steatohepatitis (NASH), the progressive form of NAFLD.

Methods: In response to COVID-19 research restrictions, we adapted an existing clinical trial and delivered 20 weeks of moderate-intensity aerobic training 5 days a week under real-time direct supervision using an audio–visual telehealth platform. Aerobic training was completed by walking outdoors or using a home treadmill. Fitness activity trackers with heart rate monitors ensured exercise was completed at the prescribed intensity with real-time feedback from an exercise physiologist.

Results: Three female patients with biopsy-proven NASH were enrolled with a mean age of 52 (SD 14) years. The mean body mass index was 31.9 (SD 5.1) kg/m2. All patients had metabolic syndrome. All patients completed over 80% of exercise sessions (mean 84% [SD 3%]) and no adverse events occurred. Body weight (mean –5.1% [SD 3.7%]), body fat (mean –4.4% [SD 2.3%]), and waist circumference (mean –1.3 in. [SD 1.6 in.]) all improved with exercise. The mean relative reduction in magnetic resonance imaging-proton density fat fraction (MRI-PDFF) was 35.1% (SD 8.8%). Mean reductions in hemoglobin A1c and Homeostatic Model Assessment for Insulin Resistance were also observed (–0.5% [SD 0.2%] and –4.0 [SD 1.2], respectively). The mean peak oxygen consumption (VO2peak) improved by 9.9 (SD 6.6) mL/kg/min.

Conclusions: This proof-of-concept study found that supervised exercise training delivered via telehealth is feasible and safe in patients with NASH. Telehealth-based exercise training also appears to be highly efficacious in patients with NASH, but this will need to be confirmed by future large-scale trials.

Trial Registration: ClinicalTrials.gov NCT03518294; https://clinicaltrials.gov/ct2/show/NCT03518294

(JMIR Form Res 2021;5(8):e30239) doi:10.2196/30239

KEYWORDS

physical activity; fatty liver; telemedicine; liver; nonalcoholic fatty liver disease; liver disease; fatty liver disease; aerobic training; telehealth; fitness; feasibility; steatohepatitis

Introduction

To date, there is no effective drug therapy nor cure for nonalcoholic fatty liver disease (NAFLD) or its progressive form, nonalcoholic steatohepatitis (NASH). Lifestyle modification, which includes both dietary change and increasing physical activity, remains the most effective treatment for NAFLD and is recommended for all patients. Despite physical

XSL•FO
RenderX

activity's well-known benefits, over 80% of patients with NAFLD are physically inactive [1,2]. Consequently, disease progression is common. There is a clear unmet need to increase physical activity in order to improve patient outcomes, especially in light of the rapidity of weight gain from increased sedentary behavior attributable to the novel COVID-19 pandemic [3]. Out of necessity, telehealth has emerged at the forefront of health care delivery during the pandemic. Telehealth offers additional promise for patients with NAFLD to (1) remove self-identified barriers preventing physical activity [1] and (2) deliver an exercise training program. However, telehealth-delivered exercise training programs remain unexplored in patients with NAFLD and their feasibility, safety, and efficacy are unknown.

Methods

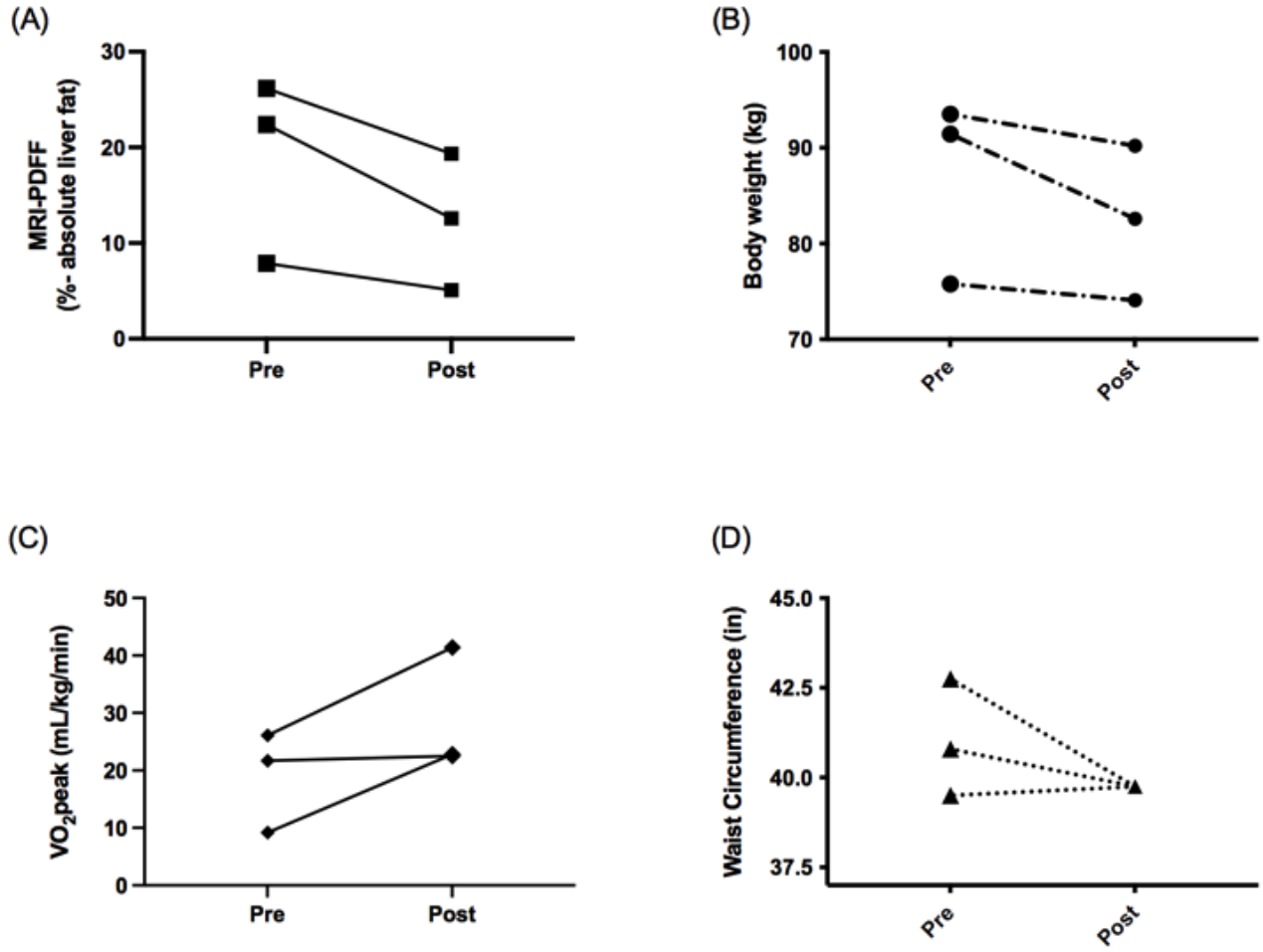
Because institutional restrictions prevented in-person clinical trial activity at the height of the pandemic, we adapted an existing clinical trial that was actively recruiting patients (prior to COVID-19 restrictions, 25 patients were recruited under the original clinical trial protocol) [4] and delivered 20 weeks of moderate-intensity aerobic training 5 days a week under real-time direct supervision by an exercise physiologist using an Institutional Review Board–approved audio–visual (A–V) telehealth platform with 2-way communication. Aerobic training was completed by walking outdoors or using a home treadmill. Fitness activity trackers with heart rate monitors ensured exercise was completed at the prescribed intensity with real-time feedback from an exercise physiologist. Each exercise session lasted 30 minutes, and was preceded by a warm-up and ended with a cool-down in accordance with the original study protocol. Feasibility was defined as completing 80% or more of exercise sessions [5]. Secondary clinical outcomes were captured according to the existing study protocol [4]. Patients also received dietary counseling according to the original study protocol which allowed for telehealth as a way to provide the nutritional feedback.

Results

Three patients with biopsy-proven NASH were recruited and enrolled during the period of COVID-19 research restrictions. All patients were female with a mean age of 52 (SD 14) years. The mean body mass index was 31.9 (SD 5.1) kg/m^2. All patients had metabolic syndrome. Liver histology was as follows: NAFLD activity score of 4 (n=2) and 5 (n=1); and fibrosis stage of 1 (n=2) and 0 (n=1). All patients completed 80% or more of exercise sessions (mean 84% [SD 3%]) and no adverse events occurred. Body weight (mean –5.1% [SD 3.7%]), body fat (mean –4.4% [SD 2.3%]), and waist circumference (mean –1.3 [SD 1.6] in.) all improved with exercise (Figure 1). The mean relative reduction in liver fat measured by magnetic resonance imaging-proton density fat fraction (MRI-PDFF) was 35.1% (SD 8.8%). Mean reductions in hemoglobin A1c, aspartate aminotransferase, alanine aminotransferase, Homeostatic Model Assessment for Insulin Resistance were –0.5% (SD 0.2%), –8.5 (SD 3.2) IU/L, –12.5 (SD 6.7) IU/L, and –4.0 (SD 1.2), respectively. The mean peak oxygen consumption (VO_2peak) improved by 9.9 (SD 6.6) mL/kg/min. Owing to the small sample size, no change in secondary clinical outcomes achieved statistical significance (P>.05).

XSL•FO
RenderX

Figure 1. Observed clinical benefits of a supervised 20-week telehealth-delivered exercise training program. (A) All patients reduced MRI-PDFF-measured liver fat with exercise training. (B) Exercise training reduced body weight in all patients. (C) Cardiorespiratory fitness improved in all patients (mean VO2 peak +9.9 [SD 6.6] mL/kg/min). (D) Waist circumference was reduced by a mean 1.3 (SD 1.6) in. MRI-PDFF: magnetic resonance imaging proton density fat fraction; VO2 peak: peak oxygen consumption.



Discussion

This proof-of-concept study found supervised exercise training delivered via an A–V telehealth platform to be feasible and safe in patients with NASH. All patients met the a priori definition of feasibility and none experienced an adverse event. Importantly, remote exercise training also appears to be efficacious in patients with NASH. The observed reduction in MRI-PDFF–measured liver fat, insulin resistance, body weight, and body fat in parallel with gains in physical fitness even exceed those that are published for supervised in-person exercise training programs of similar length [2]. While we look to future large-scale studies to validate the efficacy of this small study, these data are nonetheless promising and further suggest that A–V telehealth has a role in the routine care of patients with NASH.

While this is the first study to employ remote supervision using real-time A–V telehealth technology to deliver an exercise training program in patients with NASH, other studies have used web-based or mobile health (mHealth) to deliver an unsupervised exercise training program to a more heterogenous population of patients with all types of NAFLD [6,7]. Pfirrmann et al [6] enrolled 44 patients with NAFLD into an 8-week web-based exercise training program, in which patients completed progressive amounts of aerobic and resistance training. While exercise sessions were not directly supervised, weekly feedback was provided to individualize the exercise program and prevent injury. This web-based exercise program was feasible (74% of patients completed ≥80% of the exercise sessions) and safe (no adverse events). Modest gains in physical fitness were seen in parallel with a slight reduction in body weight (<5%) and body fat. Changes in liver fat were not reported, although transient elastography was performed, which demonstrated a small reduction in liver stiffness but no change in liver fibrosis stage.

Lim et al [7] recently explored the efficacy of Nutritionist Buddy, an mHealth app–delivered lifestyle intervention, through which patients with NAFLD were given progressive daily step goals up to 10,000 steps/day, on top of real-time dietitian support services for behavioral change. And while the authors did not provide information regarding changes in daily steps or physical activity as a whole, the mHealth intervention was successful in achieving at least 5% weight loss with corresponding improvement in metabolic parameters in 25% and 40% of patients at 3 and 6 months, respectively. However, liver-specific benefits were not measured beyond a reduction in liver enzymes;

XSL•FO
RenderX

besides, change in liver fat, physical fitness, or insulin resistance were not investigated.

Another important question that remains unanswered is whether or not direct supervision by a fitness professional using telehealth outperforms unsupervised mHealth-based lifestyle modification programs. While it is plausible to imagine regular interaction with a fitness professional over any medium, including A–V telehealth technology, may lead to increased accountability and the potential for greater exercise adherence, no direct head-to-head comparison between unsupervised mHealth-based and supervised telehealth exercise training programs has been made to date. Future study investigating this question would seem of high importance given the rapidity at which telehealth is being incorporated into routine medical care and the expected increase in NAFLD and NASH after the COVID-19 pandemic resolves.

Our study has several limitations worth noting: (1) The sample size of 3 patients limits large-scale conclusions. (2) Patients were highly selected and exclusively female, mobile device literate, and English speaking which limits the generalizability of our preliminary findings. (3) While not powered to determine change in clinical outcomes, no secondary clinical outcome achieved statistical significance and future studies are required to confirm our signal of clinical efficacy.

In conclusion, our findings demonstrate proof of concept that exercise training is feasible, safe, and likely efficacious as well. These findings require validation in a larger randomized controlled trial. If validated, new telehealth-delivered exercise training programs have the possibility to increase exercise adherence and sustainability which we would anticipate to alter the natural history of NAFLD and NASH and improve patient-oriented outcomes.

Acknowledgments

We thank the following individuals for their contributions to the research presented in this manuscript: Chris Sicca and Jeff Vesek at the Penn State Center for NMR Research; Megan Beyer, Breianna Hummer, Zachary Pattison, Gloriany Rivas, and Heather Tressler in the Penn State Division of Gastroenterology and Hepatology; and Kristin Slavoski in the Penn State Physical Medicine & Rehabilitation Research Laboratory. This grant was funded in part by the NIH grant L30 DK118601. This project is also funded, in part, under a grant with the Pennsylvania Department of Health using Tobacco CURE Funds. The Department specifically disclaims responsibility for any analyses, interpretations, or conclusions. The study was supported by NIH/NCATS grants UL1TR000127 and UL1TR002014.

Authors' Contributions

VM contributed to manuscript drafting/editing and approval of final version; AF contributed to study conduct, manuscript drafting/editing, and approval of final version; JD contributed to study conduct, manuscript drafting/editing, and approval of final version; BS contributed to study conduct, manuscript drafting/editing, and approval of final version; CS contributed to study conduct, manuscript drafting/editing, and approval of final version; JS contributed to study design, study conduct, data analysis, manuscript drafting/editing, and approval of final version.

Conflicts of Interest

None declared.

References

1. Stine JG, Soriano C, Schreibman I, Rivas G, Hummer B, Yoo E, et al. Breaking Down Barriers to Physical Activity in Patients with Nonalcoholic Fatty Liver Disease. Dig Dis Sci 2020 Oct 23:e1. [doi: 10.1007/s10620-020-06673-w] [Medline: 33098023]
2. Thorp A, Stine JG. Exercise as Medicine: The Impact of Exercise Training on Nonalcoholic Fatty Liver Disease. Curr Hepatol Rep 2020 Dec 09;19(4):402-411. [doi: 10.1007/s11901-020-00543-9] [Medline: 33767944]
3. Lin AL, Vittinghoff E, Olgin JE, Pletcher MJ, Marcus GM. Body Weight Changes During Pandemic-Related Shelter-in-Place in a Longitudinal Cohort Study. JAMA Netw Open 2021 Mar 01;4(3):e212536 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.2536] [Medline: 33749764]
4. Stine JG, Schreibman I, Navabi S, Kang M, Dahmus J, Soriano C, et al. Nonalcoholic steatohepatitis Fitness Intervention in Thrombosis (NASHFit): Study protocol for a randomized controlled trial of a supervised aerobic exercise program to reduce elevated clotting risk in patients with NASH. Contemp Clin Trials Commun 2020 Jun;18:100560 [FREE Full text] [doi: 10.1016/j.conctc.2020.100560] [Medline: 32309672]
5. Kruger C, McNeely ML, Bailey RJ, Yavari M, Abraldes JG, Carbonneau M, et al. Home Exercise Training Improves Exercise Capacity in Cirrhosis Patients: Role of Exercise Adherence. Sci Rep 2018 Jan 08;8(1):99 [FREE Full text] [doi: 10.1038/s41598-017-18320-y] [Medline: 29311671]
6. Pfirrmann D, Huber Y, Schattenberg JM, Simon P. Web-Based Exercise as an Effective Complementary Treatment for Patients With Nonalcoholic Fatty Liver Disease: Intervention Study. J Med Internet Res 2019 Jan 02;21(1):e11250 [FREE Full text] [doi: 10.2196/11250] [Medline: 30602434]

XSL•FO
RenderX

7. Lim SL, Johal J, Ong KW, Han CY, Chan YH, Lee YM, et al. Lifestyle Intervention Enabled by Mobile Technology on Weight Loss in Patients With Nonalcoholic Fatty Liver Disease: Randomized Controlled Trial. JMIR Mhealth Uhealth 2020 Apr 13;8(4):e14802 [FREE Full text] [doi: 10.2196/14802] [Medline: 32281943]

Abbreviations

A–V: audio–visual
mHealth: mobile health
MRI-PDFF: magnetic resonance imaging proton density fat fraction
NAFLD: nonalcoholic fatty liver disease
NASH: nonalcoholic steatohepatitis

Edited by G Eysenbach; submitted 06.05.21; peer-reviewed by P Simon, C Han; comments to author 31.05.21; revised version received 07.06.21; accepted 06.07.21; published 17.08.21.

Please cite as:
Motz V, Faust A, Dahmus J, Stern B, Soriano C, Stine JG
Utilization of a Directly Supervised Telehealth-Based Exercise Training Program in Patients With Nonalcoholic Steatohepatitis: Feasibility Study
JMIR Form Res 2021;5(8):e30239
URL: https://formative.jmir.org/2021/8/e30239
doi:10.2196/30239
PMID:34402795

©Victoria Motz, Alison Faust, Jessica Dahmus, Benjamin Stern, Christopher Soriano, Jonathan G Stine. Originally published in JMIR Formative Research (https://formative.jmir.org), 17.08.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Satisfaction and Usability of an Information and Communications Technology–Based System by Clinically Healthy Patients With COVID-19 and Medical Professionals: Cross-sectional Survey and Focus Group Interview Study

Ye Seul Bae[1], MD, MPH; Kyung Hwan Kim[2,3], MD, PhD; Sae Won Choi[1], MD; Taehoon Ko[4], PhD; Jun Seo Lim[1], RN; Meihua Piao[1], RN, PhD

[1]Office of Hospital Information, Seoul National University Hospital, Seoul, Republic of Korea
[2]Department of Thoracic & Cardiovascular Surgery, Seoul National University Hospital, Seoul, Republic of Korea
[3]Department of Thoracic & Cardiovascular Surgery, Seoul National University College of Medicine, Seoul, Republic of Korea
[4]Department of Medical Informatics, Catholic University of Korea, Seoul, Republic of Korea

Corresponding Author:
Kyung Hwan Kim, MD, PhD
Department of Thoracic & Cardiovascular Surgery
Seoul National University Hospital
101 Daehak-ro, Jongno-gu
Seoul, 03080
Republic of Korea
Phone: 82 220727546
Email: kkh726@snu.ac.kr

Abstract

Background: Digital health care is an important strategy in the war against COVID-19. South Korea introduced living and treatment support centers (LTSCs) to control regional outbreaks and care for patients with asymptomatic or mild COVID-19. Seoul National University Hospital (SNUH) introduced information and communications technology (ICT)–based solutions to manage clinically healthy patients with COVID-19.

Objective: This study aims to investigate satisfaction and usability by patients and health professionals in the optimal use of a mobile app and wearable device that SNUH introduced to the LTSC for clinically healthy patients with COVID-19.

Methods: Online surveys and focus group interviews were conducted to collect quantitative and qualitative data.

Results: Regarding usability testing of the wearable device, perceived usefulness had the highest mean score of 4.45 (SD 0.57) points out of 5. Regarding usability of the mobile app, perceived usefulness had the highest mean score of 4.62 (SD 0.48) points out of 5. Regarding satisfaction items for the mobile app among medical professionals, the “self-reporting” item had the highest mean score of 4.42 (SD 0.58) points out of 5. In focus group interviews of health care professionals, hospital information system interfacing was the most important functional requirement for ICT-based COVID-19 telemedicine.

Conclusions: Improvement of patient safety and reduction of the burden on medical staff were the expected positive outcomes. Stability and reliability of the device, patient education, accountability, and reimbursement issues should be considered as part of the development of remote patient monitoring. In responding to a novel contagious disease, telemedicine and a wearable device were shown to be useful during a global crisis.

(JMIR Form Res 2021;5(8):e26227) doi:10.2196/26227

KEYWORDS

COVID-19; mobile app; telemedicine; wearable device; vital sign; satisfaction; usability

XSL•FO
RenderX

Introduction

The outbreak of COVID-19 has caused major concerns worldwide. On March 11, 2020, the World Health Organization (WHO) designated COVID-19 as a pandemic and it continues to rapidly spread in almost every country across the globe [1]. As of November 29, 2020, up to 61 million confirmed cases and 1.4 million deaths were reported according to the WHO [2]. In late February 2020, many infections occurred in the Daegu-Gyeongbuk region located in the southeastern part of South Korea due to mass spread at religious facilities. Although every patient should be treated in a negative pressure isolation room in order to minimize spread, there were insufficient medical facilities and medical personnel due to the rapid increase in the number of confirmed patients [3,4].

The Korean government recommended that mild or asymptomatic patients with a positive COVID-19 test be admitted to a living and treatment support center (LTSC) and be managed each center [4]. At the government's request, Seoul National University Hospital (SNUH) operated the third LTSC at the SNUH Human Resource Development Center in Mungyeong, Gyeongsangbuk Province, 180 km southeast of Seoul and 100 km northwest of Daegu, from March 5 to April 9, 2020.

SNUH introduced novel strategies applying information and communications technology (ICT)–based remote patient management systems to a COVID-19 LTSC according to patient clinical pathways [5]. These approaches included cloud-based medical image sharing when a patient was admitted or transferred, communication through mobile apps and wearable monitoring devices for remote consultation, electronic health record templates in hospital information systems (HISs), dashboards for patient monitoring, and an e-prescription system to facilitate management of clinically healthy patients with COVID-19 [5].

Digital health care is an important strategy in the war against COVID-19, as it can minimize the spread of infection and contribute to diagnosis [6-9], treatment [10,11], and management [12,13] after discharge. This study aims to provide insight into the optimal use of the mobile apps and wearable devices that SNUH introduced to the LTSC through surveys and focus group interviews.

Methods

Overview

In this study, online surveys and focus group interviews were conducted to collect quantitative and qualitative data exploring experiences of the ICT-based clinical system for the LTSC operated by SNUH. The study was approved by the institutional review board of SNUH (2004-026-1115).

Participants

Patients

For quantitative data collection, an online survey was conducted. All respondents were adult males and females who agreed to participate in the survey. From the time the LTSC opened on March 5, 2020, until it closed on April 9, 2020, a total of 118 patients had been admitted [5]. Since clinically healthy patients with COVID-19 were admitted to the LTSC, there was no definite indication for use of a wearable device according to severity of symptoms. We allocated a wearable vital-sign data recorder (VDR)—the VDR-1000 (TriBell Labs)—to each of the first 10 rooms, which could accommodate 12 patients each. During the 36 days of LTSC operation, 24 patients admitted to the 10 rooms used the VDR-1000. The survey period was from April 20 to 24, 2020. Mobile text messages were sent to all patients asking them to visit a given web link to access the mobile app survey; 24 patients used the wearable device for the continuous monitoring survey with additional instructions. Finally, 12 respondents completed the mobile app usability survey, and 11 completed surveys on continuous remote monitoring. We explained the study details and obtained informed consent from patients who agreed to participate in the study; participants received ₩10,000 (US $8.60) as compensation.

Medical Staff

All respondents were medical staff (ie, physicians and nurses) who had worked at the LTSC of SNUH. The survey period was from April 20 to April 24, 2020. An SMS message was sent to participants asking them to visit a given web link with additional instructions, and those who agreed to participate were invited to complete the survey. A total of 24 respondents answered the questionnaire. Participants in the study received ₩10,000 (US $8.60) as compensation.

Quantitative Data Collection

Two separate online surveys were designed for patients: one regarding the mobile app used for self-reporting, communication, and notifications, and another regarding the wearable device used for remote monitoring. For both surveys, questions about perceived usefulness, perceived ease of use, and satisfaction were included. Perceived usefulness is a subjective belief that the productivity and efficiency of work will be increased by introducing a new technology or system. Perceived ease of use is the subjective belief that using a new system will not require much mental and physical effort. Medical staff were surveyed regarding their satisfaction with the overall ICT system of the LTSC [14-16].

Respondents rated their level of perceived importance of the device and mobile app using a 5-point Likert scale with the following response options and scores: strongly disagree (1), disagree (2), undecided (3), agree (4), and strongly agree (5) [17]. The online questionnaire also included open-ended questions about the advantages and limitations of the ICT system. The questionnaires were administered using Google Forms, an online survey administration software. Participants could access questionnaires through a URL and were able to complete the survey at any time or place, thereby ensuring privacy and honesty. Only participants who agreed to the instructions were invited to complete the survey. Results were processed using SPSS Statistics, version 22.0 (IBM Corp). Questionnaire items were analyzed using frequencies, percentages, means, and standard deviations.

XSL•FO
RenderX

Focus Group Interviews

Focus group interviews were used for qualitative data collection. One of the distinct features of this method is group dynamics; hence, the type and range of data generated through the social interactions of the group are often deeper and richer than those obtained from one-on-one interviews [18]. The optimum size of focus groups is 6 to 8 participants, excluding researchers, but focus groups can be successful with as few as 3 and as many as 14 participants [19].

Three members of the research team (YSB, MP, and JSL) facilitated the focus groups. The facilitators all have a background in medical informatics and were experienced in conducting focus groups.

In this study, the participants of the focus group interviews were health care professionals (ie, physicians and nurses) who had experience using ICT-based patient management systems in the LTSC of SNUH. The participants were divided into two groups—physicians (n=5) and nurses (n=5)—and attended two sets of interviews in April 2020.

All participating health care professionals provided consent to participate, and they were presented with structured open-ended questions regarding their needs and possible issues when instituting the ICT-based patient management system in the COVID-19 LTSC of SNUH. Each focus group interview was 60 to 90 minutes in length and ended once the conversation no longer yielded new ideas or opinions (ie, saturation of themes). Each focus group interview was recorded in its entirety, with the researchers writing additional memos when necessary. In order to eliminate bias and improve the reliability and validity of the results, two researchers who participated in a course on qualitative research conducted the interviews. All data coders and analysts were trained in qualitative research. All interviews were recorded with the consent of the participants, and the recordings were transcribed as soon as the interviews were completed. In cases where the transcribed data were not comprehensible or interpretable, another follow-up interview was conducted to enhance the reliability of the data and analysis.

The main open-ended questions were as follows:

1. What do you think should be included in ICT-based patient management systems in LTSCs?
2. What are the limitations in introducing ICT-based patient management systems in LTSCs?
3. What is the expected effect of applying ICT-based patient management systems in LTSCs?
4. What is the most important thing (function, role) of ICT-based patient management systems in LTSCs?
5. What are the anticipated administrative issues when using ICT-based patient management systems in LTSCs?
6. What are the expected clinical problems when using ICT-based patient management systems in LTSCs?
7. Do you think ICT-based patient management systems in LTSCs will improve the efficiency of the COVID-19 treatment process? Specifically, what do you think will help?
8. Do you think the experience of introducing ICT-based patient management systems in LTSCs to future long-term care or home care will be helpful? Specifically, what do you think will help?

Qualitative content analysis was as follows. First, we tried to form an overall opinion by reading the interview contents repeatedly. Second, we carefully read each paragraph and formulated the meaning of each statement. Third, we labeled the codes and categorized them according to the subjects' experience. Lastly, codes were categorized according to their relationship and connectivity, and the arranged codes were organized according to their hierarchy of importance.

Results

Overall ICT-Based System Introduced in the LTSC

Mobile App

An Android-based mobile app was developed for the LTSC patients to enable efficient patient management and communication between patients and medical staff. The app consisted of six features: a general guide for patients admitted to the LTSC, a notice board, a symptom questionnaire, vital sign reporting, questions and answers, and push notifications (Multimedia Appendix 1). Patients were instructed to do self-checkups twice a day. Patients received push notifications when they needed to fill out a self-report questionnaire on symptoms and vital signs, when they needed to answer questions, or when medical staff posted new notices on the bulletin board. When the patient filled out a structured questionnaire on the presence or absence of symptoms through the app and input vital signs, the corresponding data were immediately linked to the HIS. Medical staff could also upload general guidance and notices regarding the LTSC or COVID-19.

Wearable Device

Patients were asked to use a wearable device to measure vital signs and allow medical staff to monitor them remotely. The VDR-1000 was used for this purpose; this a wearable, medical, multi-signal measurement device that can concurrently measure a patient's electrocardiogram (ECG), pulse rate, blood pressure (BP), blood oxygen saturation (SpO_2), respiratory waveform, and respiratory rate. The measured data were transmitted to a central monitoring system (CMS) using Wi-Fi and then forwarded to the SNUH HIS. Medical staff in Mungyeong and Seoul used CMS monitors and the HIS to monitor patient vital signs. CMS software can set alarms with different thresholds for each patient. If a value outside the threshold range was measured, an alarm sounded, allowing medical personnel to respond quickly [5].

Patient Survey

Experiences Using the Wearable, Continuous Vital-Sign Monitoring Device

In total, 12 patients completed the questionnaire regarding the wearable device. Of these, 11 (92%) patients provided general information. The mean age was 25 (SD 6.25) years. Of the 11 patients, 6 (55%) were male and 5 (45%) were female (Table S1 in Multimedia Appendix 2). For usability testing of the wearable device, perceived usefulness had the highest score at

XSL•FO
RenderX

4.45 (SD 0.57) points out of 5, followed by perceived ease of use at 4.30 (SD 0.59) points and satisfaction at 3.98 (SD 0.70) points. Of all wearable device measures, SpO_2 had the highest satisfaction at 4.03 (SD 0.76) points, followed by ECG at 3.94 (SD 0.92) points and BP at 3.76 (SD 0.96) points (Table 1). Items from the perceived usefulness, ease of use, and satisfaction surveys are shown in Tables 2, 3, and 4, respectively. Of the perceived usefulness items, "data measured through wearable devices can improve the quality of care" and "wearable devices will be useful for medical staff" had the highest scores (mean 4.55, SD 0.52). Of the perceived ease of use items, "it was easy to use the wearable device by looking at the manual" and "the wearable device's weight is appropriate for use" had the highest scores (mean 4.64, SD 0.50). In contrast, the item "it is convenient to move while wearing the device" had the lowest score (mean 3.55, SD 1.44). Of the satisfaction items, "no discomfort when going to the bathroom" and "no discomfort when moving" had the lowest scores (mean 3.00, SD 1.67, and mean 3.27, SD 1.56, respectively) for the device items as well as for the partial wearable devices.

Table 1. Usability testing as measured by perceived usefulness, perceived ease of use, and satisfaction (n=11).

Survey category	Score[a], mean (SD)
Perceived usefulness	4.45 (0.57)
Perceived ease of use	4.30 (0.59)
Satisfaction	
Wearable device	3.98 (0.70)
Blood pressure measure	4.03 (0.76)
Electrocardiogram measure	3.94 (0.92)
Blood oxygen saturation measure	3.76 (0.96)

[a]Survey items were rated on a 5-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree).

Table 2. Perceived usefulness of the wearable, continuous vital-sign monitoring device (n=11).

Survey item	Score[a], mean (SD)
The wearable device can quickly and easily measure the biomarkers required for medical staff decision making.	4.36 (0.67)
Data measured through the wearable device can improve the quality of care.	4.55 (0.52)
The biomarkers measured with the wearable device can improve the efficiency of treatment.	4.45 (0.69)
The wearable device can improve telemedicine care.	4.36 (0.67)
The wearable device will be useful for medical staff.	4.55 (0.52)

[a]Survey items were rated on a 5-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree).

Table 3. Perceived ease of use of the wearable, continuous vital-sign monitoring device (n=11).

Survey item	Score[a], mean (SD)
It is easy to use the wearable device by referring to the manual.	4.64 (0.50)
The wearable device is designed to be easy to use.	4.18 (1.25)
The wearable device size is appropriate for use.	4.55 (0.52)
The wearable device weight is appropriate for use.	4.64 (0.50)
It is convenient to move while wearing the device.	3.55 (1.44)
The wearable device is convenient to store and manage.	4.27 (0.79)

[a]Survey items were rated on a 5-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree).

XSL•FO
RenderX

Table 4. Satisfaction with the wearable, continuous vital-sign monitoring device and its measures (n=11).

Survey item	Score, mean (SD)			
	All measures of the device	Blood pressure measure	ECG[a] measure	SpO_2[b] measure
Individual item[c]				
The shape of the wearable device is adequate	4.18 (0.60)	4.36 (0.81)	4.18 (0.60)	4.27 (0.65)
The size of the wearable device is adequate	4.36 (0.50)	4.36 (0.81)	4.27 (0.65)	4.27 (0.79)
The weight of the wearable device is adequate	4.36 (0.50)	4.36 (0.67)	4.27 (0.65)	4.18 (0.75)
The location of the part implementing the function is appropriate	4.36 (0.50)	4.36 (0.67)	4.18 (0.60)	4.27 (0.79)
It is convenient to operate	4.36 (0.67)	4.45 (0.69)	4 (1.10)	4.45 (0.69)
It works stably	4.18 (0.60)	4.27 (0.79)	4.27 (0.65)	4.45 (0.69)
No discomfort when eating food	4.00 (1.18)	3.18 (1.66)	3.55 (1.51)	3.55 (1.51)
No discomfort when going to the bathroom	3.00 (1.67)	2.27 (1.85)	3.55 (1.51)	3.27 (1.42)
No discomfort when moving	3.27 (1.56)	2.45 (1.81)	3.55 (1.51)	3.45 (1.51)
No discomfort when sleeping	3.73 (1.19)	3.36 (1.69)	3.64 (1.50)	3.55 (1.21)
No difficulty in connecting the device without assistance	N/A[d]	N/A	4.18 (0.75)	N/A
When attaching or moving the sensor sticker, there was no strain on the skin	N/A	N/A	3.64 (1.50)	N/A
No difficulty connecting the SpO_2 device cup to the finger without assistance	N/A	N/A	N/A	4.45 (0.69)
Willing to use a home-based wearable device with similar performance in the future	3.91 (1.14)	3.91 (1.22)	3.91 (1.22)	4.18 (0.75)
All items[e]				
Expected score for the wearable device before using it	76.36 (22.92)	85.45 (13.68)	76.00 (21.58)	85.27 (11.39)
Evaluation score for the wearable device after actually using it	89.36 (9.67)	90.55 (10.55)	85.00 (20.12)	91.36 (7.78)

[a]ECG: electrocardiogram.

[b]SpO_2: blood oxygen saturation.

[c]Survey items were rated on a 5-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree).

[d]N/A: not applicable; this survey item did not pertain to either the device itself or to the indicated measure.

[e]Participants responded using a scale where the maximum score was 100 points.

Experience Using the Mobile App

In total, 12 patients completed the questionnaire regarding the mobile app. The mean age was 27.75 (SD 10.24) years, and 5 out of 12 (42%) patients were female.

In terms of usability of the mobile app, perceived usefulness had the highest mean score of 4.62 (SD 0.48) points out of 5, followed by satisfaction with a mean score of 4.08 (SD 1.41) points and perceived ease of use with a mean of 3.81 (SD 0.41) points (Table 5). All of the perceived usefulness items for the app had a higher score than the wearable device itself, with the item “the self-reporting mobile app will be useful for medical staff” scoring the highest (mean 4.75, SD 0.62) (Table 6). Of the perceived ease of use items, “checking cumulative BP, pulse, or body temperature history information,” “checking push messages from medical staff,” “checking for responses from medical staff,” and “searching my notification history” had the highest scores (combined mean 3.92, SD 0.50). Of the satisfaction items, “installing the mobile app,” “log-in,” and “entering measurement results such as BP, pulse, or body temperature” showed the highest scores (mean 4.50, SD 1.17). The item “entering measurement results such as BP, pulse, or body temperature” scored relatively highly, not only in terms of ease of use but also in satisfaction. However, for perceived ease of use items, “installing the mobile app,” “searching for notice information,” and “inquiring and submitting questionnaire data” received the lowest scores (mean 3.67, SD 0.65; mean 3.67, SD 0.49; and mean 3.67, SD 0.49, respectively). The item “searching my notification history” received the lowest score (mean 3.58, SD 2.02) in satisfaction (Table 7).

XSL•FO
RenderX

Table 5. Perceived usefulness, perceived ease of use, and satisfaction with the mobile app (n=12).

Category	Score[a], mean (SD)
Perceived usefulness	4.62 (0.48)
Satisfaction	4.08 (1.41)
Perceived ease of use	3.81 (0.41)

[a]Survey items were rated on a 5-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree).

Table 6. Perceived usefulness of the mobile app (n=12).

Survey item	Score[a], mean (SD)
The self-reporting mobile app can quickly and easily measure the biomarkers required for medical staff decision making.	4.50 (0.67)
Data measured through the self-reporting mobile app can improve the quality of care.	4.50 (0.67)
The biomarkers measured through the self-reporting mobile app can improve the efficiency of treatment.	4.67 (0.65)
The self-reporting mobile app can improve telemedicine care.	4.67 (0.49)
The self-reporting mobile app will be useful for medical staff.	4.75 (0.62)

[a]Survey items were rated on a 5-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree).

Table 7. Perceived ease of use of, and satisfaction with, the mobile app (n=12).

Survey item	Ease of use score[a], mean (SD)	Satisfaction score[a], mean (SD)
Installing the mobile app	3.67 (0.65)	4.50 (1.17)
Log-in	3.75 (0.45)	4.50 (1.17)
Getting information about care center guidelines	3.83 (0.58)	3.67 (2.06)
Searching for notice information	3.67 (0.49)	3.92 (1.73)
Inquiring and submitting questionnaire data	3.67 (0.49)	4.42 (1.16)
Entering measurements such as BP[b], pulse, or body temperature	3.83 (0.39)	4.50 (1.17)
Checking cumulative BP, pulse, or body temperature history information	3.92 (0.51)	4.17 (1.75)
Checking push messages from medical staff	3.92 (0.67)	3.75 (2.09)
Checking for responses from medical staff	3.92 (0.67)	3.75 (2.09)
Searching my notification history	3.92 (0.67)	3.58 (2.02)

[a]Survey items were rated on a 5-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree).
[b]BP: blood pressure.

Survey of Medical Staff

Overview

The medical staff who replied to the questionnaire had average of 13.08 (SD 5.33) years of work experience, and their average age was 37.38 (SD 6.27) years. Among them, 83% (20/24) were nurses and approximately 96% (23/24) were female (Table S2 in Multimedia Appendix 2).

Medical Staff Satisfaction With the Mobile App and Web Monitoring System

Among satisfaction items for the mobile app, “self-reporting” had the highest mean score at 4.42 (SD 0.58) points out of 5, followed by “center guidelines” at a mean of 4.29 (SD 0.62) points, “vital sign check” at a mean of 4.21 (SD 0.72) points, “notice information” at a mean of 3.96 (SD 0.62) points, “medical inquiries” at a mean of 3.88 (SD 0.68) points, and “push notifications” at a mean of 3.83 (SD 0.64) points (Table 8).

With regard to satisfaction with using the web monitoring system, only 15 staff out of 24 (63%) had full or partial experience using the system. Among the features, “notice information” showed the highest mean score at 4.20 (SD 0.68) points out of 5, followed by “center guidelines” at a mean of 4.13 (SD 0.72) points, “patient management” at a mean of 4.13 (SD 0.0.64) points, “medical inquires” at a mean of 4.08 (SD 0.67) points, and “message management” at a mean of 3.85 (SD 0.80) points (Table 9).

The total mean perceived usefulness score for the wearable devices when providing medical care was 82.79 (SD 2.77) points out of 100. Of the perceived usefulness items, “the wearable devices can improve telemedicine care” had the highest score (mean 4.33, SD 0.70), while “data measured through the

XSL•FO
RenderX

wearable devices can improve the quality of care" and "I'm willing to use wearable devices for providing medical care" had the lowest score (mean 4.13, SD 0.68) (Table 10).

Table 8. Medical staff satisfaction with the mobile app (N=24).

Survey item	Score[a], mean (SD)
Self-reporting	4.42 (0.58)
Center guidelines	4.29 (0.62)
Vital sign check	4.21 (0.72)
Notice information	3.96 (0.62)
Medical inquiries	3.88 (0.68)
Push notifications	3.83 (0.64)

[a]Survey items were rated on a 5-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree).

Table 9. Medical staff satisfaction with the web monitoring system (n=19).

Survey item	Score[a], mean (SD)
Notice information	4.20 (0.68)
Center guidelines	4.13 (0.72)
Patient management	4.13 (0.64)
Medical inquiries	4.08 (0.67)
Message management	3.85 (0.80)

[a]Survey items were rated on a 5-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree).

Table 10. Perceived usefulness for the wearable continuous vital sign monitoring device among medical staff (n=19).

Survey item	Score, mean (SD)
Individual items[a]	
The wearable device can quickly and easily measure the biomarkers required for medical staff decision making.	4.17 (0.64)
Data measured through the wearable device can improve the quality of care.	4.13 (0.61)
The biomarkers measured through the wearable device can improve the efficiency of treatment.	4.17 (0.70)
The wearable device can improve telemedicine care.	4.38 (0.65)
The wearable device will be useful for medical staff.	4.33 (0.70)
I am willing to use the wearable device for providing medical care.	4.13 (0.68)
Total perceived usefulness for the wearable device when providing medical care[b]	82.79 (12.77)

[a]Survey items were rated on a 5-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree).
[b]Participants responded using a scale where the maximum score was 100 points.

Among the 24 medical staff members who replied to the survey, 23 (96%) answered that the "connection with HIS, such as clinical observation record" was the most crucial function that a wearable device can perform. A total of 20 staff members out of 24 (83%) responded that "accuracy" was important for wearable devices, followed by "ease of use" (16/24, 67%). However, only 5 staff members out of 24 (21%) replied that "variety of measurement data types" was a critical feature. The open-ended responses regarding the requirements for wearable device function are listed below (Textbox 1).

XSL•FO
RenderX

Textbox 1. The most crucial functions that a wearable device must perform according to the medical staff members.

The most crucial functions that a wearable device must perform:

- I think device accuracy is the most important
- It seems to be possible only if the device is accurate and the patient can use it easily
- Device accuracy, HIS (hospital information system) linkage
- Wireless. Patients are too fragmented because of the cable
- Alarm function when it is not attached properly
- HIS linkage is essential for quick response
- Alarm on error
- The accuracy of the device should be high
- HIS linkage for clinical observation recording and verification
- Accuracy, usability, stability, convenience, and privacy
- Accurate measurement and HIS linkage
- Accuracy should be the top priority
- Function for detecting body temperature
- Linkage with all electronic health records that contain patient vital signs
- Function for detecting blood pressure and pulse rate
- Capability of acquiring specific vital signs personalized to each patient
- Vital sign linkage function
- Automatically analyzes the ECG (electrocardiogram) rhythm based on vital signs and alerts medical staff
- HIS linkage and alarm function
- Personal health records
- Function for detecting vital signs

Focus Group Interviews of Health Care Professionals

Overview

Thematic analysis of the focus group interviews with 5 physicians and 4 nurses yielded three themes: (1) major function requirements, (2) expected outcomes, and (3) potential issues. The themes are summarized and described in Table 11.

Table 11. Needs of health care professionals and possible issues with information and communications technology–based management systems for patients with COVID-19.

Themes	Subthemes
Major function requirements	Hospital information system interface
	Features of the wearable device • Additional measurement functions • Alarms
	Features of the mobile app • Presenting a reference range • Messenger
Expected outcomes	Improvement of patient safety
	Contribution to reducing the burden on medical staff
Potential issues	Stability and reliability of the device
	Patient education
	Accountability
	Cost and reimbursement

XSL•FO
RenderX

Major Function Requirements

HIS Interface

The majority of participants mentioned that the HIS interface was the most important major function requirement. The system implemented in the LTSC interfaced with the information acquired from wearable devices and apps with the HIS in real time. Most medical professionals emphasized that information obtained from wearable devices or mobile apps should be immediately linked to the HIS to alert staff to changes in patient status while simultaneously reducing staff workload.

> *It is essential for remote patient monitoring to properly interface and integrate vital sign data generated from a wearable device or patient-generated data collected in the app (eg, symptom self-reporting and vital sign self-measurement data).*
>
> *For remote patient monitoring, when an abnormal signal is detected by the patient, the clinical information of the patient in HIS must be inquired at any time.*
>
> *HIS linkage can reduce human errors that may occur during rewriting and reduce the burden on medical staff.*

In addition, this is likely to be important not only for infectious diseases such as COVID-19 but also for other chronic diseases and mental health issues.

> *If the patient self-reports symptoms such as depression or self-measured blood pressure, blood sugar, etc, and these data are linked to the HIS in real time, it will help medical staff identify changes in the patient's condition promptly.*

Since not all the data generated by various devices or apps could be linked to the HIS, it is necessary to structure systems so that important surrogate markers for each disease can be linked. Moreover, it is crucial to distinguish whether the data interfaced with the HIS are self-reported or measured by medical professionals.

Features of the Wearable Device

Additional Measurement Functions

Although patients admitted to the LTSC were afebrile and had no symptoms or only mild rhinorrhea and cough, 2 patients were later admitted to a nearby hospital due to sudden progression of dyspnea and pneumonia. Therefore, many participants suggested that the ability to continuously and concurrently measure body temperature, oxygen saturation, and respiratory rate of isolated patients is a critical value of the wearable device.

> *In order to treat COVID-19 patients not face-to-face, it is most important to have a device that can measure body temperature and oxygen saturation well.*
>
> *If the patient self-measures the respiratory rate, it is often a high or low value because the patient has difficulty with self-measuring the respiratory rate.*
>
> *It would be nice if there was an auscultation function that remotely hears lung sounds when the patient puts the device to the chest. I think that would help us know what is going on with pneumonia.*

Alarm

When an abnormal value is detected, an alert sound can be used to notify both medical staff and the patient. If the alarm goes off due to wearing the device incorrectly, the patient can reattach the device after checking the manual. If an abnormal value is recorded, medical staff can preemptively respond to the alarm.

> *By adding an alarm function, the patient can recognize whether it is a false signal, and medical staff can check whether it is an error or an actual abnormality.*

Features of the Mobile App

Presenting a Reference Range

Patients measured vital signs by themselves using the symptom questionnaire and vital sign reporting functions, entered them into the app, and self-reported symptoms. Patients wondered if their vital signs were within normal range. In addition, when a value outside the reference range is measured, medical staff should be able to recognize it at a glance, such as using color indicators.

> *In the case of vital signs, if they deviate from the reference range, it would be better to display them in a different color.*

Messenger

Some medical staff emphasized integrating a messenger function into the mobile app. LTSCs were originally public or private facilities that were modified to accommodate and quarantine patients with COVID-19. Therefore, continuous education was required for patients on how to use the facility, how to self-measure vital signs, rules to be observed during quarantine in facilities other than medical institutions, how to dispose of waste, and how to communicate with medical staff when abnormal symptoms occurred.

Expected Outcomes

Improvement in Patient Safety

Since the data measured by the wearable device are directly linked to the HIS, it reduces the potential for human errors that can occur when manually inputting data into the HIS.

Reducing the Burden on Medical Staff

One nurse had to virtually meet about 20 patients at least twice a day. In each consultation, the patient's vital signs, respiratory symptoms potentially related to COVID-19, digestive symptoms potentially related to COVID-19, and mental health concerns, such as depression and anxiety, were checked. Therefore, each virtual consultation took a considerable amount of time, and the burden on the medical staff was substantial. After the introduction of the electronic medical examination system and the patient mobile app, patients could report their symptoms on their own before starting a scheduled consultation and

XSL•FO
RenderX

automatically link them to the HIS so that medical staff could check before starting the virtual consultation.

> *After the patients could input self-monitoring data into the app and transmit it directly to HIS, the time required was reduced from more than 20 minutes per patient to less than 10 minutes if there were no particular problems.*

Potential Issues

Stability and Reliability of the Device

The stability of the wearable device and the reliability of the measured values are very important. The VDR-1000 used in the LTSC could measure ECG, BP, respiratory rate, heart rate, and SpO_2 at the same time. For accurate measurement, patients had to connect to wired ECG, BP, and SpO_2 sensors on the device by themselves while sitting still for 3 to 5 minutes. It is a very sensitive device in which the measured value changes even with small movements.

> *Because the LTSC had active patients with mild disease, they found it difficult to sit still for five minutes and measure vital signs.*

Patient Education

Training on management and education of wearable devices or apps is required, as patient familiarity with information technology (IT) varies. Patients who were unfamiliar with using video calls or mobile apps took a considerable amount of time to get used to non–face-to-face treatment, and the medical staff in charge had to repeatedly educate the patient.

Accountability

Most medical practitioners agreed that the responsibilities of telemedicine, including diagnosis and prescription, should be established. Remote medical treatments provided by LTSCs were temporarily permitted for COVID-19 outbreaks in certain areas, but the scope of responsibility of medical personnel who perform remote medical treatments should be clarified in preparation for the post–COVID-19 era.

Cost and Reimbursement

It is necessary to set an appropriate price for wearable devices that the patient will use. Even if a patient has the opportunity to use a wearable device as part of a non–face-to-face treatment, if it is too expensive, the actual patient's needs will not be met. In addition, if an appropriate fee for telemedicine is not established, the use of various remote medical solutions and wearable devices capable of remote monitoring will be limited.

Discussion

Principal Findings

We administered a questionnaire to clinically healthy patients with COVID-19 and medical staff that included items measuring perceived usefulness, ease of use, and satisfaction with the ICT-based system introduced by SNUH in an LTSC. In addition, focus group interviews were conducted with medical staff to obtain qualitative insights in order to seek future development directions.

Since the COVID-19 pandemic, telemedicine has been spotlighted as a useful way to respond to infectious diseases [20]. Originally, telemedicine was introduced for medically underprivileged areas. With the development of technology over recent decades, various wearable devices, sensors, and platforms have gradually expanded the application of telemedicine to noncommunicable diseases [21,22], infectious diseases, and psychiatric diseases [23]. Clinical consultations through video calls are associated with high patient satisfaction [21,24], and there are no differences in clinical outcomes [22,23,25] compared to face-to-face treatment [26]. However, this is the first time that telemedicine has been performed during a global catastrophe like COVID-19, and research evaluating objective effects is sparse. One study investigated patient satisfaction with telemedicine during the COVID-19 pandemic [27]. However, there have been few studies evaluating the usefulness or satisfaction of both patients and medical professionals. Strengths of our study include evaluation of patient satisfaction as well as evaluation of medical staff satisfaction and conducting of focus group interviews. We found that patients expected that the use of a wearable device would improve quality of care and be helpful in medical treatment. Most respondents reported no major problems with the use of the wearable device, but they complained of discomfort when moving while wearing the device. The wearable device we adopted has multiple lines to accurately measure several vital signs at the same time, which may cause inconvenience to users who are relatively healthy, like those admitted to the LTSC. Therefore, in the future, it is crucial to consider introducing a simple device, such as a wrist monitor, for convenient remote measurement of vital signs in clinically healthy patients with COVID-19. However, the VDR-1000 allowed medical staff to monitor patients' vital signs at a glance through the CMS, even from the Seoul central monitoring center. In addition, it is possible to selectively search and view past data, or to set an alarm that would sound when the device measures a value outside a specific range for each patient, helping with medical treatment. Above all, the measured data were directly linked to the vital sign sheet in the SNUH HIS, which facilitated medical treatment and reduced potential human errors that may occur during the normal process of recording, transcription, and data entry. In the mobile app satisfaction survey, "push notification" had the lowest score among both patients and medical staff. This notification function is the most recently implemented in the mobile app and was developed to provide advanced notice to patients to self-report symptoms and vital signs at the appointed times twice a day. Unfortunately, at the introduction of the mobile app, only the initial function of sending a message or push alarm to all patients in the LTSC was implemented, and it was not possible to give specific alarms to patients assigned to each medical staff member. The focus group interview results showed the need for a messenger function, including SMS and an alarm function, to deliver and communicate patient-specific messages. Development of such functions should be integrated into a future non–face-to-face care platform for management of patients with COVID-19.

The expected outcomes obtained through focus group interviews regarding the ICT-based system introduced in the LTSC of SNUH could contribute to improving patient safety and reducing

XSL•FO
RenderX

the burden on medical staff. As a potential issue, there was an opinion that the device used for remote patient monitoring should be stable and reliable. In addition, because each patient has a different degree of familiarity with IT, different amounts of education will be required. In fact, the United States Health Insurance Portability and Accountability Act (HIPAA) has allowed the use of commonly used video call systems, such as FaceTime, Google Hangouts, and Skype, for video consultation [26]. In addition, a new type of position, such as a technological liaison or coordinator, may be required to overcome the hurdle of patient unfamiliarity with telemedicine [28,29].

Telemedicine in Korea

Korea is one of the few countries where telemedicine is completely banned. The government, the medical community, civic groups, and politicians have different opinions. In particular, there is an opinion that Korea should take a different approach from those in countries where telemedicine is active, due to the characteristics of Korea's medical system, which provides easy medical access with low medical costs. However, one of the best ways in which various digital technologies, such as communication, sensors, the cloud, and information security, can be integrated in the medical field is through the application of telemedicine. In Korea, even if an implantable defibrillator with a remote monitoring function is implanted in patients with arrhythmias, the function is turned off due to policy regulations. Due to these sanctions, there has been little discussion about payment structures related to telemedicine, medical information systems, cost and reimbursement, and security issues. During the recent COVID-19 pandemic, studies have suggested that active use of telemedicine technologies for triage, monitoring, communication, and critical care management [30] may be useful in Korea. It is necessary to selectively allow telemedicine in areas where there is a medical need, and to systematically establish and operate a related technical base. In Korea, at the end of February 2020, phone consultation was temporarily allowed to ensure access to medical care in the COVID-19 situation. By October 25, 2020, up to 950,000 cases of non–face-to-face treatment have been implemented [31]. The Korean government is actively promoting telemedicine. By 2025, 18 smart hospitals using ICT are planned to be built; by 2021, imaging equipment will be provided to 5000 clinic-level medical institutions. The government plans to increase the number of clinics by 500 in 2021, to create a total of 1000 clinics. The budget for this is US $93 million by the end of 2021 [31]. However, considering the complicated medical system of Korea, telemedicine must be addressed carefully. Since there are many stakeholders related to telemedicine, a process for social consensus is necessary. The government should play a leading role in this process, and consensus through in-depth medical, technical, financial, regulatory, and industrial expert discussions considering the complexity and specificity of this issue will be needed. Current evolving IT could be very useful for collecting meaningful data on large cohorts as well as infectious diseases [32]. It can be used not only in relation to COVID-19 but also throughout the patient chain of care, such as during prehospital [33], inpatient [34], and postdischarge stages [35,36]. In addition, selective consideration may be needed to adopt telemedicine in health care. Infectious diseases as well as chronic diseases, mental illness, postoperative patient management, and home care fall into a "gray zone" of existing medical care where telemedicine is particularly useful. In order to collect patient-derived data produced by various devices, sensors, and platforms, a vendor-neutral platform is needed, and two-way communication between medical staff and patients is possible only when data are integrated with a standardized protocol to be linked to the HIS [37,38]. Moreover, it is necessary to evaluate the evidence-based effectiveness of telemedicine and compare it to existing face-to-face treatment: whether patient outcomes of telemedicine are similar or improved, whether the overall quality of medical care is improved, and whether it helps to improve medical productivity and costs. It must be proven to be effective.

Strengths and Limitations

The most powerful strength of this study is evaluation of the perspectives of both patients and medical staff who had participated in newly developed COVID-19–specific non–face-to-face consultation solutions. Our findings will be helpful when setting up telemedicine for contagious diseases as well as noncommunicable diseases. One limitation is that there were insufficient COVID-19–specific investigations. While applying telemedicine to COVID-19, it is important to understand equipment, legislative considerations, coding, logistic concerns, quality of care, cost-effectiveness, and clinical outcomes. In addition, we used a cross-sectional design with which we could not evaluate long-term outcomes due to the short 36-day operation period of the LTSC. In this context, further studies are needed to evaluate telemedicine specific to COVID-19 to improve overall clinical care and health management.

Conclusions

We demonstrated patient and medical professional satisfaction with, and usability of, an ICT-based system for clinically healthy patients with COVID-19. Our findings support the usefulness of telemedicine and wearable devices during a global infectious crisis.

Acknowledgments

This research was supported by a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute, funded by the Ministry of Health and Welfare, South Korea (grant HI19C0790).

Authors' Contributions

KHK had full access to all the data in the study and took responsibility for the integrity of the data and the accuracy of the data analysis. YSB, SWC, and TK conceived and designed the study. YSB, MP, and JSL acquired, analyzed, and interpreted the data.

XSL•FO
RenderX

YSB and MP drafted the manuscript. Critical revision of the manuscript was provided by YSB, MP, and TK. YSB and KHK provided administrative, technical, or material support.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Screenshots of the mobile app in Korean (left) and English (right) showing the six features.
[PNG File , 349 KB - formative_v5i8e26227_app1.png]

Multimedia Appendix 2
Supplementary tables.
[DOCX File , 18 KB - formative_v5i8e26227_app2.docx]

References

1. Coronavirus disease (COVID-19): Weekly epidemiological update and weekly operational update. World Health Organization. 2020. URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/ [accessed 2020-11-29]
2. WHO Coronavirus (COVID-19) Dashboard. World Health Organization. 2020. URL: https://covid19.who.int/ [accessed 2020-11-29]
3. Current Status of Response to COVID-19 and Future Plans. Sejong, Korea: Ministry of Health and Welfare; 2020 Mar 09. URL: https://www.mofa.go.kr/eng/brd/m_22591/down.do?brd_id=20543&seq=10&data_tp=A&file_seq=1 [accessed 2020-03-09]
4. Coronavirus Disease-19, Republic of Korea: Weekly updates for countries with major outbreaks. Ministry of Health and Welfare. URL: http://ncov.mohw.go.kr/en/ [accessed 2020-11-29]
5. Bae YS, Kim KH, Choi SW, Ko T, Jeong CW, Cho B, et al. Information technology-based management of clinically healthy COVID-19 patients: Lessons from a living and treatment support center operated by Seoul National University Hospital. J Med Internet Res 2020 Jun 12;22(6):e19938 [FREE Full text] [doi: 10.2196/19938] [Medline: 32490843]
6. Jin C, Chen W, Cao Y, Xu Z, Tan Z, Zhang X, et al. Development and evaluation of an artificial intelligence system for COVID-19 diagnosis. Nat Commun 2020 Oct 09;11(1):5088 [FREE Full text] [doi: 10.1038/s41467-020-18685-1] [Medline: 33037212]
7. Seshadri DR, Davies EV, Harlow ER, Hsu JJ, Knighton SC, Walker TA, et al. Wearable sensors for COVID-19: A call to action to harness our digital infrastructure for remote patient monitoring and virtual assessments. Front Digit Health 2020 Jun 23;2:1-11. [doi: 10.3389/fdgth.2020.00008]
8. Mei X, Lee H, Diao K, Huang M, Lin B, Liu C, et al. Artificial intelligence-enabled rapid diagnosis of patients with COVID-19. Nat Med 2020 Aug;26(8):1224-1228 [FREE Full text] [doi: 10.1038/s41591-020-0931-3] [Medline: 32427924]
9. Wong CK, Ho DTY, Tam AR, Zhou M, Lau YM, Tang MOY, et al. Artificial intelligence mobile health platform for early detection of COVID-19 in quarantine subjects using a wearable biosensor: Protocol for a randomised controlled trial. BMJ Open 2020 Jul 22;10(7):e038555 [FREE Full text] [doi: 10.1136/bmjopen-2020-038555] [Medline: 32699167]
10. Tsikala Vafea M, Atalla E, Georgakas J, Shehadeh F, Mylona EK, Kalligeros M, et al. Emerging technologies for use in the study, diagnosis, and treatment of patients with COVID-19. Cell Mol Bioeng 2020 Jun 24:1-9 [FREE Full text] [doi: 10.1007/s12195-020-00629-w] [Medline: 32837582]
11. Vaishya R, Javaid M, Khan IH, Haleem A. Artificial intelligence (AI) applications for COVID-19 pandemic. Diabetes Metab Syndr 2020;14(4):337-339 [FREE Full text] [doi: 10.1016/j.dsx.2020.04.012] [Medline: 32305024]
12. Singh RP, Javaid M, Haleem A, Suman R. Internet of things (IoT) applications to fight against COVID-19 pandemic. Diabetes Metab Syndr 2020;14(4):521-524 [FREE Full text] [doi: 10.1016/j.dsx.2020.04.041] [Medline: 32388333]
13. Nasajpour M, Pouriyeh S, Parizi RM, Dorodchi M, Valero M, Arabnia HR. Internet of things for current COVID-19 and future pandemics: An exploratory study. J Healthc Inform Res 2020 Nov 12:1-40 [FREE Full text] [doi: 10.1007/s41666-020-00080-6] [Medline: 33204938]
14. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319. [doi: 10.2307/249008]
15. Rahimi B, Nadri H, Lotfnezhad Afshar H, Timpka T. A systematic review of the technology acceptance model in health informatics. Appl Clin Inform 2018 Jul;9(3):604-634 [FREE Full text] [doi: 10.1055/s-0038-1668091] [Medline: 30112741]
16. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: Toward a unified view. MIS Q 2003;27(3):425. [doi: 10.2307/30036540]
17. McLeod S. Likert scale definition, examples and analysis. Simply Psychology. 2019. URL: https://www.simplypsychology.org/likert-scale.html [accessed 2020-03-12]
18. Kitzinger J. The methodology of focus groups: The importance of interaction between research participants. Sociol Health Illn 1994 Jan;16(1):103-121. [doi: 10.1111/1467-9566.ep11347023]

XSL•FO
RenderX

19. Gill P, Stewart K, Treasure E, Chadwick B. Methods of data collection in qualitative research: Interviews and focus groups. Br Dent J 2008 Mar 22;204(6):291-295. [doi: 10.1038/bdj.2008.192] [Medline: 18356873]
20. Hollander JE, Carr BG. Virtually perfect? Telemedicine for Covid-19. N Engl J Med 2020 Apr 30;382(18):1679-1681. [doi: 10.1056/nejmp2003539]
21. Dorsey ER, Venkataraman V, Grana MJ, Bull MT, George BP, Boyd CM, et al. Randomized controlled clinical trial of "virtual house calls" for Parkinson disease. JAMA Neurol 2013 May;70(5):565-570 [FREE Full text] [doi: 10.1001/jamaneurol.2013.123] [Medline: 23479138]
22. Su D, Zhou J, Kelley MS, Michaud TL, Siahpush M, Kim J, et al. Does telemedicine improve treatment outcomes for diabetes? A meta-analysis of results from 55 randomized controlled trials. Diabetes Res Clin Pract 2016 Jun;116:136-148. [doi: 10.1016/j.diabres.2016.04.019] [Medline: 27321329]
23. Fortney JC, Pyne JM, Kimbrell TA, Hudson TJ, Robinson DE, Schneider R, et al. Telemedicine-based collaborative care for posttraumatic stress disorder: A randomized clinical trial. JAMA Psychiatry 2015 Jan;72(1):58-67. [doi: 10.1001/jamapsychiatry.2014.1575] [Medline: 25409287]
24. Sabesan S, Simcox K, Marr I. Medical oncology clinics through videoconferencing: An acceptable telehealth model for rural patients and health workers. Intern Med J 2012 Jul;42(7):780-785. [doi: 10.1111/j.1445-5994.2011.02537.x] [Medline: 21627743]
25. de Jong MJ, van der Meulen-de Jong AE, Romberg-Camps MJ, Becx MC, Maljaars JP, Cilissen M, et al. Telemedicine for management of inflammatory bowel disease (myIBDcoach): A pragmatic, multicentre, randomised controlled trial. Lancet 2017 Sep;390(10098):959-968. [doi: 10.1016/s0140-6736(17)31327-2]
26. Calton B, Abedini N, Fratkin M. Telemedicine in the time of coronavirus. J Pain Symptom Manage 2020 Jul;60(1):e12-e14 [FREE Full text] [doi: 10.1016/j.jpainsymman.2020.03.019] [Medline: 32240756]
27. Ramaswamy A, Yu M, Drangsholt S, Ng E, Culligan PJ, Schlegel PN, et al. Patient satisfaction with telemedicine during the COVID-19 pandemic: Retrospective cohort study. J Med Internet Res 2020 Sep 09;22(9):e20786 [FREE Full text] [doi: 10.2196/20786] [Medline: 32810841]
28. Vidal-Alaball J, Acosta-Roja R, Pastor Hernández N, Sanchez Luque U, Morrison D, Narejos Pérez S, et al. Telemedicine in the face of the COVID-19 pandemic. Aten Primaria 2020;52(6):418-422 [FREE Full text] [doi: 10.1016/j.aprim.2020.04.003] [Medline: 32402477]
29. Greenhalgh T, Koh GCH, Car J. Covid-19: A remote assessment in primary care. BMJ 2020 Mar 25;368:m1182. [doi: 10.1136/bmj.m1182] [Medline: 32213507]
30. Hong Z, Li N, Li D, Li J, Li B, Xiong W, et al. Telemedicine during the COVID-19 pandemic: Experiences from Western China. J Med Internet Res 2020 May 08;22(5):e19577 [FREE Full text] [doi: 10.2196/19577] [Medline: 32349962]
31. 950,000 temporary phone prescriptions this year. MEDI:GATE NEWS. URL: http://medigatenews.com/news/2384351293 [accessed 2020-12-20]
32. Fagherazzi G, Goetzinger C, Rashid MA, Aguayo GA, Huiart L. Digital health strategies to fight COVID-19 worldwide: Challenges, recommendations, and a call for papers. J Med Internet Res 2020 Jun 16;22(6):e19284 [FREE Full text] [doi: 10.2196/19284] [Medline: 32501804]
33. Chou E, Hsieh Y, Wolfshohl J, Green F, Bhakta T. Onsite telemedicine strategy for coronavirus (COVID-19) screening to limit exposure in ED. Emerg Med J 2020 Jun;37(6):335-337. [doi: 10.1136/emermed-2020-209645] [Medline: 32366616]
34. Zhai Y, Wang Y, Zhang M, Hoffer Gittell J, Jiang S, Chen B, et al. From isolation to coordination: How can telemedicine help combat the COVID-19 outbreak? medRixv. Preprint posted online on February 23, 2020 [FREE Full text] [doi: 10.1101/2020.02.20.20025957]
35. Singh RP, Javaid M, Haleem A, Suman R. Internet of things (IoT) applications to fight against COVID-19 pandemic. Diabetes Metab Syndr 2020;14(4):521-524 [FREE Full text] [doi: 10.1016/j.dsx.2020.04.041] [Medline: 32388333]
36. Nasajpour M, Pouriyeh S, Parizi RM, Dorodchi M, Valero M, Arabnia HR. Internet of things for current COVID-19 and future pandemics: An exploratory study. J Healthc Inform Res 2020 Nov 12:1-40 [FREE Full text] [doi: 10.1007/s41666-020-00080-6] [Medline: 33204938]
37. Jamil F, Ahmad S, Iqbal N, Kim D. Towards a remote monitoring of patient vital signs based on IoT-based blockchain integrity management platforms in smart hospitals. Sensors (Basel) 2020 Apr 13;20(8):2195 [FREE Full text] [doi: 10.3390/s20082195] [Medline: 32294989]
38. Soltan AAS, Kouchaki S, Zhu T, Kiyasseh D, Taylor T, Hussain ZB, et al. Rapid triage for COVID-19 using routine clinical data for patients attending hospital: Development and prospective validation of an artificial intelligence screening test. Lancet Digit Health 2021 Feb;3(2):e78-e87. [doi: 10.1016/S2589-7500(20)30274-0]

Abbreviations

BP: blood pressure
CMS: central monitoring system
ECG: electrocardiogram
HIPAA: Health Insurance Portability and Accountability Act

XSL•FO
RenderX

HIS: hospital information system
ICT: information and communications technology
IT: information technology
LTSC: living and treatment support center
SNUH: Seoul National University Hospital
SpO_2: blood oxygen saturation
VDR: vital-sign data recorder
WHO: World Health Organization

Edited by G Eysenbach; submitted 02.12.20; peer-reviewed by SY Shin, H Mehdizadeh; comments to author 13.12.20; revised version received 22.12.20; accepted 06.07.21; published 26.08.21.

Please cite as:
Bae YS, Kim KH, Choi SW, Ko T, Lim JS, Piao M
Satisfaction and Usability of an Information and Communications Technology–Based System by Clinically Healthy Patients With COVID-19 and Medical Professionals: Cross-sectional Survey and Focus Group Interview Study
JMIR Form Res 2021;5(8):e26227
URL: https://formative.jmir.org/2021/8/e26227
doi:10.2196/26227
PMID:34254946

©Ye Seul Bae, Kyung Hwan Kim, Sae Won Choi, Taehoon Ko, Jun Seo Lim, Meihua Piao. Originally published in JMIR Formative Research (https://formative.jmir.org), 26.08.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Cyberbullying Prevention for Adolescents: Iterative Qualitative Methods for Mobile Intervention Design

Megan L Ranney[1,2], MPH, MD; Sarah K Pittman[2], BA; Isabelle Moseley[1], BA; Kristen E Morgan[2], MPH; Alison Riese[1,2], MPH, MD; Michele Ybarra[3], MPH, PhD; Rebecca Cunningham[4], MD; Rochelle Rosen[1,5], PhD

[1]Center for Digital Health, Brown University, Providence, RI, United States

[2]Rhode Island Hospital, Providence, RI, United States

[3]Center for Innovative Public Health Research, San Clemente, CA, United States

[4]University of Michigan, Ann Arbor, MI, United States

[5]Center for Behavioral and Preventive Medicine, The Miriam Hospital, Providence, RI, United States

Corresponding Author:
Megan L Ranney, MPH, MD
Center for Digital Health
Brown University
139 Point St
Providence, RI, 02903
United States
Phone: 1 4014445159
Email: megan_ranney@brown.edu

Abstract

Background: Cybervictimization among adolescents is associated with multiple negative mental health consequences. Although pediatricians often screen for cyberbullying, validated and acceptable programs to reduce the frequency and impact of adolescent cybervictimization are lacking.

Objective: This study uses *agile* qualitative methods to refine and evaluate the acceptability of a mixed-modality intervention, initiated within the context of usual pediatric care, for adolescents with a history of cyberharassment and cyberbullying victimization.

Methods: Three groups of adolescents were successively recruited from an urban primary care clinic to participate in three consecutive iterations (1, 2, and 3) of the program, which consisted of a brief in-clinic intervention followed by 8 weeks of daily, automated SMS text messaging. After 2 weeks of messaging, iteration 1 (I1) participants completed semistructured interviews regarding intervention experiences. Participant feedback was evaluated via framework matrix analysis to guide changes to the program for iteration 2 (I2). Feedback from 2-week interviews of I2 participants was similarly used to improve the program before initiating iteration 3 (I3). Participants in all 3 iterations completed the interviews after completing the program (8 weeks). Daily response rates assessed participant engagement, and satisfaction questionnaires assessed acceptability.

Results: A total of 19 adolescents (aged 13-17 years) reporting past-year cybervictimization were enrolled: 7 in I1, 4 in I2, and 8 in I3. Demographic variables included the following: a mean age of 15 (SD 1.5) years; 58% (11/19) female, 42% (8/19) male, 63% (12/19) Hispanic, 37% (7/19) non-Hispanic, 79% (15/19) people of color, and 21% (4/19) White. A total of 73% (14/19) self-identified as having a low socioeconomic status, and 37% (7/19) self-identified as lesbian, gay, or bisexual. The average past 12-month cybervictimization score at baseline was 8.2 (SD 6.58; range 2-26). Participant feedback was used to iteratively refine intervention content and design. For example, participants in I1 recommended that the scope of the intervention be expanded to include web-based *conflicts* and *drama*, rather than narrowly focusing on cyberbullying prevention. On the basis of this feedback, the I2 content was shifted toward more general de-escalation skills and bystander empowerment. Overall, 88.34% (940/1064) of the daily queries sent to participants across all 3 iterations received a reply. Participant satisfaction improved considerably with each iteration; 0% (0/7) of I1 participants rated the overall quality of Intervention to Prevent Adolescent Cybervictimization with Text message as *excellent*, compared to 50% (2/4) of I2 participants and 86% (6/7) of I3 participants. Engagement also improved between the first and third iterations, with participants replying to 59.9% (235/392) of messages in I1, compared to 79.9% (358/488) of messages in I3.

XSL•FO
RenderX

Conclusions: This study shows the value of structured participant feedback gathered in an agile intervention refinement methodology for the development of a technology-based intervention targeting adolescents.

(JMIR Form Res 2021;5(8):e25900) doi:10.2196/25900

KEYWORDS

adolescent; mobile health; digital health; cyberbullying; user-centered design; qualitative; mobile phone

Introduction

Background

Cyberbullying, defined as the intentional harm of others through computers, cell phones, and other electronic devices, has become increasingly common [1]. Although rates of cyberbullying vary widely, from 6% to 72% depending on the study, approximately 25% of American adolescents have reported being victims of cyberbullying and web aggression (hereafter referred to as *cybervictimization*) in the past year [2,3]. Cybervictimization is strongly associated with negative consequences, including depressive symptoms and suicidality, posttraumatic stress, alcohol and other drug use, physical violence, and dating violence [4,5]. Although school-based cyberbullying prevention programs may be effective, many youths find it difficult to engage or retain in school-based prevention programs [6].

Almost 80% of adolescents have yearly well-child visits with their pediatrician [7]. Pediatricians play a long-recognized role in behavioral counseling [8]. The American Academy of Pediatrics recommends advising families about cybervictimization [9], but pediatricians lack both time and validated interventions [10]. An easy-to-use, clinically initiated program that enhances users' protective skill sets (eg, emotion regulation and positive social support) against cybervictimization may be helpful [11] to pediatricians. Delivering preventive interventions through technology places fewer demands on staff time than in-person delivery, has inherently higher fidelity, and maybe lower cost [12]. KiVa, a school-based antibullying program that includes web-based components such as games, video clips, and infographics, has been shown to significantly reduce the frequency of cyberbullying and cybervictimization among Finnish youth [13]. Given the success of this technology-based approach, SMS text messaging is a logical medium for delivering interventions aimed at reducing cyberbullying and cybervictimization. Furthermore, our prior work indicates that delivering interventions in the same modality as bullying (ie, electronically) may increase intervention efficacy [14]. SMS text messaging is almost universally used by adolescents from all socioeconomic and racial or ethnic backgrounds [15]. SMS text message–augmented interventions are feasible, acceptable, and may be effective in reducing in-person fights [16-18]. Grounded in previous evidence that electronic interventions are effective for cyberbullying prevention and cybervictimization support and that SMS text messaging is a reliable and efficient way to deliver behavioral change interventions, the following clinical trial applies the technology of SMS text messaging to cyberbullying.

Objective

In the larger project that serves as the use case for this manuscript, our aims are to develop, iteratively refine, and then pilot an SMS text message–based intervention to help adolescents recognize, cope with, and prevent cyberbullying. Our initial intervention prototype was developed from prior SMS text message–based interventions for physical violence and bullying, using existing cyberbullying prevention resources and expert consultation. To refine our prototype, we conducted hour-long qualitative interviews with 23 adolescents with past-year histories of cybervictimization and web-based conflict. Participants shared their own coping strategies for dealing with cyberbullying and offered constructive criticism of intervention content and design. Data from participant feedback were analyzed and used to update the structure and content of the intervention. In this study, we used qualitative methods—which classically occur as a discrete step in the research process—in 3 agile iterations to seek feedback from the intervention audience as the program is further refined. *Agile* methods are a development strategy used in software design, project management, manufacturing, and recently, health behavior research [19,20]. In agile methods, aspects of development occur collaboratively, instead of in isolation (as in traditional development), to incrementally build the product [20]. Iteration, a key component of the agile development process, facilitates usability testing. Review, planning, and testing can all occur iteratively. This paper describes methods used to gather, analyze, and use qualitative and quantitative data in real time for iterative intervention refinement of a cyberbullying intervention.

Methods

Study Setting and Recruitment

This study was conducted from July to November 2017 in a pediatric primary care clinic within an urban teaching hospital in the northeastern United States. Adolescents aged 13-17 years who presented with primary care (well-child and sick visits) were potentially eligible. Adolescents were excluded if they or their parents did not speak English, were in the custody of police or child protective services, had a diagnosis of intellectual or developmental disability, or did not have a parent or legal guardian present. After adolescent verbal assent and parent verbal permission, adolescents completed a screening survey on a tablet computer. Adolescents were eligible for participation if they screened positive for past-year cybervictimization [21], owned a cell phone with text messaging capabilities, and provided written assent and parental permission, per our local institutional review board's request.

Three consecutive iterations of the SMS text messaging program were pilot-tested by adolescents; in each phase, recruitment

XSL•FO
RenderX

continued sequentially until qualitative saturation was reached. First, 8 participants were enrolled for iteration 1 (I1) of the program. Enrolled participants completed a brief in-clinic intervention followed by up to 8 weeks of automated text messaging. After 2 weeks of receiving messages, I1 participants completed semistructured interviews that sought feedback on the program. Participant feedback was then analyzed via a framework matrix and used to guide changes to the content and structure of both the in-clinic intervention and the SMS text-messaging program. A second group consisting of 4 participants was recruited for iteration 2 (I2) to test the impact of these changes. Feedback from 2-week interviews of I2 participants was similarly used to improve the program before initiating iteration 3 (I3), for which a group of 9 participants was recruited. Participants in all 3 iterations were asked to complete an in-person semistructured interview after the receipt of text messages was concluded (8 weeks). This strategy shortened design-to-delivery time as per *generative* and *prototyping* phase of agile intervention development [20]. Participants in all iterations received a US $25 gift card at baseline, US $20 for completing the 8-week interview, US $40 for completing the follow-up survey, and US $10/month for the cost of text messaging or data use. The study materials and procedures were approved by the institutional review board.

Intervention Structure and Content

Intervention to Prevent Adolescent Cybervictimization with Text message (iPACT) has two components: a brief (15-minute) in-clinic PowerPoint-guided session conducted by a trained, bachelors-level research assistant based on motivational interviewing and basic principles of cognitive behavioral therapy, a type of psychotherapy focused on strategies to identify and overcome destructive thought patterns (eg, using thoughts and actions to change feelings) and an 8-week, daily, automated, two-way text messaging curriculum. In addition to daily automated messages, participants could *pull* additional message content by texting keywords. Content was based on in-person and SMS text message–based violence prevention and cyberbullying prevention interventions [6,13,16-18]. Additional details, including the original structure of the intervention before the agile development process, are described elsewhere [14].

Measures

Cyberbullying

The Cyberbullying Scale [21], used to determine study eligibility, is a validated 16-item self-report measure of past-year cybervictimization. Total scores are calculated by summing the participant responses to questions 3-16. Cronbach α was .88. Adolescents were eligible if they reported a frequency of *2=Sometimes* or more on any Cyberbullying Scale item.

Semistructured Interview

For I1 and I2 participants, semistructured qualitative interviews were completed after 2 weeks and 8 weeks of interaction with the intervention, respectively. Participants in I3 underwent 8-week interviews only. Interviews were conducted by a research assistant trained in qualitative data collection and interview facilitation. The interview agenda was designed to generate actionable data for intervention refinement by focusing on how participants perceived and responded to the intervention content. Discussions explored participants' perceptions about, and usability of, key intervention components (in-person and SMS text message content; message tone, frequency, and applicability; daily queries; links to external content; and extra support messages). In the interviews, participants were asked to reflect on message wording and preferences; how, when, and why they engaged with the program; and what they were thinking about when they responded to daily messages. The interviews were digitally recorded. The interviewer completed a written debriefing after each interview.

Acceptability and Feasibility

Feasibility was determined by response rates to daily SMS text message queries, frequency of participant-requested extra support messages, and follow-up rates. To assess acceptability, the Client Satisfaction Questionnaire [22], a self-report measure (range 8-32) with high validity and internal consistency, was administered at the 8-week follow-up [22].

Demographics

Participants self-reported age, school grade, socioeconomic status (SES), race, ethnicity, birth sex, gender identity, and sexual orientation using validated measures [23-26].

Analysis

A framework matrix analysis of qualitative interview data was used to continuously refine the intervention [27,28]. The framework matrix captured participant reactions to intervention components and their answers to the interview questions. Each participant's comments were summarized in a matrix row; matrix columns represented the intervention components discussed in the interview. The interview facilitator listened to the audio recordings, summarizing participant comments in relevant cells. Verbatim quotes were included to provide context for the participants' own words. To ensure credibility, a second analyst reviewed the matrix and audio for each interview, confirming the accuracy and completeness of the summaries. This data reduction technique quickly summarized qualitative information into a single usable source that was shared with coinvestigators and used in weekly study team meetings, allowing us to use the feedback from 2-week interviews of I1 participants to guide intervention refinement before I2 and the 2-week interviews of I2 participants before I3. At these team meetings, conflicting qualitative data were discussed, and a team consensus was reached on how it should be prioritized for intervention refinement. Notes indicating changes made based on the coinvestigator review of participant comments were added to the matrix to track changes.

Descriptive statistics were calculated for all quantitative measures. In accordance with prior work [29,30], acceptability was defined as 80% or greater scores on the Client Satisfaction Questionnaire, and feasibility was defined as 80% response to texts and 80% retention at the 8-week follow-up.

XSL•FO
RenderX

Results

Overview

Of the 121 adolescent patients who completed the study screening, 31 (25.6%) were eligible. Most enrollment refusals were because of *lack of time*, given the clinic setting. No otherwise eligible adolescents who completed the screening survey had to be excluded because of lack of parental permission.

Consistent with qualitative research sampling, we purposefully enrolled a diverse sample to capture potential outlier perspectives. A total of 8 participants enrolled in I1, 4 in I2, and 9 enrolled in I3. The mean age of all enrolled participants was 15 (range 13-17) years; the majority self-identified as people of color (15/19, 79%), Hispanic (12/19, 63%), female (11/19, 58%), and low SES (14/19, 73%). A total of 37% (7/19) of participants were identified as lesbian, gay, or bisexual. The average past 12-month cybervictimization score at baseline was 8.2 (SD 6.58; range 2-26).

Two participants withdrew (1 in I1 and 1 in I3) because they lost or broke their phones before starting the texts. In I1, all 7 participants completed the 2-week interview and 8-week survey; in I2, all 4 participants completed the 2-week interview and 3 completed the 8-week survey; in I3, 5 participants completed the 8-week interview and 7 of 8 completed the 8-week survey.

Intervention Refinement: Qualitative Data

Overview

The key components of the intervention were (1) content, (2) tailoring, and (3) design and delivery. Qualitative data, as summarized in the framework matrix, were used to significantly revise each of these intervention components. The specifications of this process are described below; Table 1 provides additional details of these changes, including an example of the qualitative feedback given and the resulting changes made for each key topic.

XSL•FO
RenderX

Table 1. Exemplars of themes from framework matrix analysis, organized by topic and iteration designated by iteration 1, iteration 2, and iteration 3.

Topics addressed	Exemplar quotes	Examples of changes made
General intervention content		
I1[a]: focus on general web-based drama and conflict rather than only on cyberbullying specifically	• "...school drama and online drama kinda go hand in hand because if you were to have social media, you would of course have, like, most of your friends from school on social media." [ID 29; 13 year; female]	• The word cyberbullying was replaced with online drama on several slides of the PowerPoint used in the in-clinic session. The research assistant conducting the in-clinic intervention began framing the program as prevention for online drama and cyberbullying (rather than as a program about dealing with cyberbullying) and focused more on web-based drama in discussion • The order of the messages was changed so that cyberbullying-specific content was emphasized during the last week of the texting program
I1: teens thought too much of the content was specific to cyberbullying and requested more general self-efficacy content	• "Give something more like-like common sense. Cause nowadays people lack that so much." [ID 2; 17 years; male]	• A message offering possible responses to web-based bullies was replaced with a reminder that jokes or roasting can go too far and that teens should not be afraid to speak up in such situations • Messages were added or changed to enhance self-efficacy; for example, "everyone is affected by drama" was changed to "you can make a difference"
I2[b]: participants liked random messages that were positive and inspirational	• "...have, like, a nice inspirational quote like those, or, like, motivational quotes...I really liked them because it's the..., it's, like, the middle of the school day for me, so it's something that gets me, like, to keep going." [ID 12; 17 years; female]	• Inspirational quotes from other teens such as "If it won't matter in 5 years, don't spend more than 5 minutes thinking about it" were used in the daily messages • Additional random inspirational messages were added
I1, I2, I3[c]: participants thought positive content elevated their mood and also helped them avoid conflict	• "Like, there was one day when I, like, woke up, and I was feeling very mad at this person on social media, and [the study text] was, like, something about...'just think positive,' and...something, 'don't go on there and be a bully.' And I think it helped 'cause I was gonna go on social media and talk to that person, but [the study text] it just made me think, oh, well that's not gonna change anything. What she said is still what she said, so it kinda redirected me in a different path." [ID 1; 15 years; female]	• Messages such as "only have/keep good people on your page" were added to encourage participants to surround themselves with supportive friends
I3: teens suggested adding more content relating to stress and anxiety	• "With anxiety, maybe you're at, like, a big party, and you're feeling really stressed out cause of all the people there." [ID 4; 14 years; male]	• Tips about changing thoughts and feelings were sent earlier in the program (previously scheduled for Weeks 6-7)
Tailoring		

XSL•FO
RenderX

Topics addressed	Exemplar quotes	Examples of changes made
I1: each participant defined cyberbullying differently and had unique experiences with it	• "So cyberbullying is, like, when people, like, try to, like, get you online, right?" [ID 7; 15 years; female] • "Some people be just, like, really, like, sittin' at their computer hurtin' people...It's, like, sad, and, like, people, like...eventually they just gonna snap, and...be like, 'Oh, why should I even be on, like, this earth anymore.'" [ID 31; 15 years; female] • "I feel like once people get, like, really overwhelmed, they just kinda give up on everything, and that's probably when they'll jump to, like, social media..." [ID 12; 17 years; female]	• The research assistant began tailoring the content to each participant based on the participant's age and past-year versus past 2-month Cyberbullying Scale scores. The research assistant also incorporated each participants' language into their intervention session
I2: participants wanted to provide their own strategies for making sense of intervention content rather than relying on confusing graphics in the in-clinic session	• "Sometimes you can have a negative thought but then a smart action." [ID 7; 15 years; female] • "Like, stay positive...Like, don't be a bully...I'm not a bully, so I was like, you know what, lemme just leave it alone." [ID 17; 15 years; female]	• Participants were asked to provide their own example for the research assistant's explanation of the Thoughts-Feelings-Actions Triangle during the in-clinic session • New intervention content was incorporated to allow participants to outline their own positive goals for social media
I2: participants were using different time frames over which to evaluate the daily cyberbullying SMS text message query because some received it in the morning and some in the afternoon	• "Like, I know that there's not gonna be drama at, like, 5:30 in the morning, so I kinda, like, play it back to the day, like, the night, before." [ID 17; 15 years; female]	• The language of the daily SMS text message drama question was changed from "Any drama today, yes or no?" to "Any drama in the past 24 hours, yes or no?" to make sure that every participant's assessment period was the same length
I1 and I2: participants forgot about the on-demand messages and suggested sending a reminder	• "I wish I knew or remembered that I could've texted back. I completely forgot. I didn't know that I could've done that, and if I knew that I could've done that, I definitely would've texted back." [ID 15; 14 years; female]	• Additional reminders were added, for example, "Remember, if you ever want extra advice you can text us any time: ANGRY, SAD, HAPPY, or STRESSED."
I1, I2, I3: participants generally felt the daily messages matched their day; however, they reported that on days with web-based drama, they wanted specific advice about their conflict	• "If someone's really having a bad day, they'll get annoyed, but maybe just, like, one extra question, like, letting it all out in, like, one huge paragraph thing." [ID 5; 13 years; female] • "I feel like it's good for you guys to, like, check up on us and, like, asking us, like, how our day is and how we feeling and thoughts and everything." [ID 10; 14 years; male]	• This suggestion was difficult to incorporate because the study was not designed for staff review of individual participants' experiences; some examples of responses and actions in specific conflicts were added in website links in the texts
Design and delivery of the intervention		

XSL•FO
RenderX

Topics addressed	Exemplar quotes	Examples of changes made
I1: participants preferred that the messages sounded like they were coming from someone older than them, and they did not want them to sound too automated	• "I wouldn't trust it, kind of [if it were coming from someone my age]. 'Cause it's like – it's kinda weird, 'cause, like, they're like me, and they also need help, but, for an adult, they've already been through a situation like this, so they know what to do on it." [ID 1; 15 years; female] • "I liked it because it was – it, like, it kinda felt like somebody, you know, cared to ask me how I was doing and how my day was." [ID 16; 16 years; female]	• A message that began with the rhetorical question, "How do you move on after online drama?" (which sounded like a computer) was rephrased to begin, "Some teens feel hopeless or sad after online drama." • The message "Useful tip (which you might know already): You can block or unfriend people who are bothering you online." Was rephrased as "If someone is bothering you, you can block or unfollow them. We've got some great tips on controlling your social media page" to sound less curt and more sympathetic, like an adult or older sibling
I1, I2, and I3: participants would click on links only if they felt relevant, quick, and interesting	• "[I clicked on a link] when I wasn't having a good day, and I needed it, like, I felt like I needed advice or a video or anything to just, to distract myself from everything else and to help me." [ID 15; 14 years; female] • "I think they're helpful, like, having the link there for you if you don't – if you, like, need more help or something. Like, if I needed more tips, or I felt I needed more tips, I would click on the links, but in some days when, like, I didn't need any more tips, like, that was great enough." [ID 1; 15 years; female)	• The advertising industry principle of clickbait was used in editing several of the messages to entice participants to click on the links more often
I2 and I3: participants requested app-based intervention and messaging instead of in-person interventions and texting	• "It (an app) would be easier because...not a lot of teenagers will answer your text messages...everybody uses, like, Snapchat and everything more than that, so more people are likely to go on to the app than to go onto the text messages." [ID 11; 14 years; female]	• This suggestion was difficult to incorporate because of inherent study design issues
I2 and I3: participants recommended the program to friends	• "I told one of my friends about it...cause, like, she's always involved in drama. I told her, I was like, 'hey, you should try this,' and I showed her one of the websites. She liked it, and she asked me, like, how did I get it on my phone." [ID 16; 16 years; female] • "All my friends who actually, like, would need it or like it, and they were like, 'oh that's cool I wish I got that.'" [ID 15; 14 years; female] • "It's a cool program...a lot of people (my friends) were interested...I wish I could still do it." [ID 12; 17 years; female]	• The program's final message was changed to end with, "Remember all you've learned these past 8 weeks and pass it on to your friends!"

[a]I1: iteration 1.

[b]I2: iteration 2.

[c]I3: iteration 3.

General Intervention Content

In I1 qualitative interviews (Table 1), participants recommended discussing web-based *conflicts* and *drama* rather than narrowly focusing on cyberbullying prevention. On the basis of this feedback, I2 content was changed to discuss *web-based drama* in addition to *cyberbullying* and to provide more general de-escalation skills and bystander empowerment. In-person brief intervention language was also changed (eg, instead of asking "how many other teens do you think have been cyberbullied?" teens were asked to report their own daily experience with *web-based drama*). Cyberbullying-specific statistics were moved to the later weeks of the text program. Content was reframed to enhance the self-efficacy of those

XSL•FO
RenderX

witnessing cyberbullying, as well as those experiencing it. These changes garnered positive feedback from users in I2 and I3.

In I2 qualitative interviews, participants further emphasized the importance of empowerment: teens said they wanted content that would help them take positive actions, boost confidence, and know how to intervene in web-based conflicts. They suggested content focused on using social media for good, sharing examples of supportive content they had found on the web. In response to this feedback, I3 in-clinic sessions incorporated a discussion of how healthy social media habits could help teens reach long-term personal goals, and positive rather than negative valence was emphasized. After these changes, feedback on content in I3 was overall positive; the majority of I3 participants described messages as *educational*, *motivational*, *helpful*, and *inspirational*. Numerous participants described how intervention messages motivated them to change their behavior by targeting their thoughts and feelings or to improve their emotional state by taking action.

Intervention Tailoring

In I1, the participants wanted greater personalization of both portions of the intervention. Accordingly, the in-person intervention was significantly reworked (Table 1). For example, the participants' personal definitions of cyberbullying were incorporated. Specific changes were made to in-person session graphics to allow for greater incorporation of participant feedback. The SMS text message algorithms were also adjusted. For instance, a baseline intervention feature provided an option to text the program in the moment and to receive on-demand support for certain mood states; however, only 2 participants used this feature. In I1 and I2 qualitative interviews, most indicated that they simply forgot the feature existed and suggested sending a reminder. Consequently, multiple reminders of the on-demand feature were added to I3, and nearly all (n=7) participants used it.

Our initial design tailored messages based on each participant's answers to 2 daily assessment questions. Participants were asked to rate their day (on a scale of 1=really bad to 5=great) and to indicate whether they experienced cyberbullying that day (yes or no). On the basis of feedback from I1 users, those questions were revised to ask about *drama or conflict on the web* rather than cyberbullying, and the qualifier *in the past 24 hours* was included because we found the recall period varied when *today* was the prompt. The language change was essential and effective: most interview participants (n=12) said the daily check-in made them feel as if *someone cared*. We were able to see increasing proportions of participants indicating that the daily message matched how they rated their day (3/7, 43% interview participants in I1; 3/4, 75% interview participants in I2; and 4/5, 80% interview participants in I3).

About half of all participants wanted the ability to give more information about their day, and several wanted to be able to get advice on specific instances of web-based drama. Human subject concerns about the collection of large amounts of free-text data from participants, ethical responsibilities to monitor and respond to them, budget, and staff limitations prevented the incorporation of this suggestion. Instead, links to websites and infographics were added after I2 to include specific conflict resolution examples. Although the addition was perceived as helpful, some I3 participants still wanted specific advice.

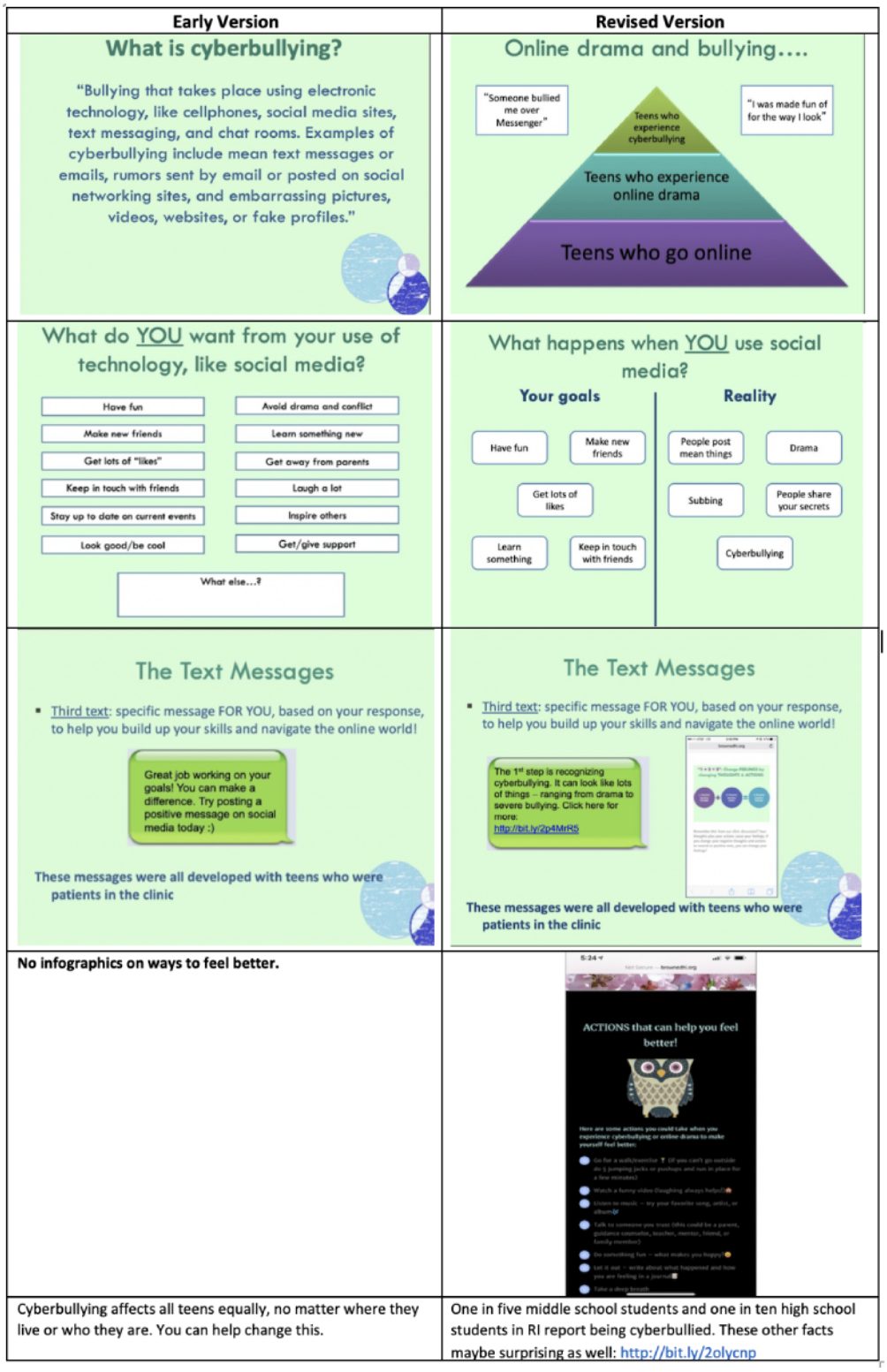
Design and Delivery of Intervention

In I1, the interview participants preferred messages to be written as if a real person was sending them. They also generally preferred a message tone that sounded like an adult or older sibling. The messages were edited in response to this feedback. I1 participants also requested additional reminders about the content of the in-person intervention. The team revised the links and created infographics (Figure 1). Most participants clicked on at least 1 link sent as part of a daily message. Interviewees across all three iterations said they only clicked on links that were personally relevant, visually appealing, and "something that's quick." Participants said they were more likely to click on links when they were having a bad day or were unfamiliar with the topic. In response, we revised the link descriptions to make topics seem more intriguing.

About one-third of all participants said they would prefer to receive the entire intervention through an app rather than an in-person component or text messages. The need for an in-person intervention was identified as a barrier to participation. Many said that their friends used *texting* exclusively through apps because of the lack of reliable cellular services.

XSL•FO
RenderX

Figure 1. Exemplars of changes in design. RI: Rhode Island.



Feasibility and Acceptability: Quantitative Data Across Iterations

Of the 1064 daily queries sent to participants across all three iterations, 940 (88.34%) received a response. Participants responded to both questions (ie, how their day was going and whether cyberdrama had occurred in the past day) on 73.59% (783/1064) days. The rate of response to both daily questions improved from I1, during which, on average, participants replied to both queries on 61% (34/56) of days, to I2 (average of 46/56, 83% days) and I3 (average of 45/56, 81% days). Although we did see a small drop in engagement from I2 to I3, this discrepancy is not statistically significant; both engagement rates were significantly better compared with I1. All participants (n=17) who completed the 8-week follow-up survey rated the overall quality of iPACT as *good* or *excellent*. The ratings improved from I1 (7/7, 100% *good*) to I2 (2/4, 50% *good*; 2/4, 50% *excellent*) to I3 (1/7, 14% *good*; 6/7, 86% *excellent*). All

XSL•FO
RenderX

participants who completed the 8-week survey agreed that they got the *kind of service* they wanted from iPACT, and all said they would come back to iPACT again if they could. Of the participants, 9 reported that they had already recommended the program to a friend or said they would do so.

Discussion

Principal Findings

To our knowledge, iPACT is the first SMS text message–based cyberbullying intervention for adolescents aged 13-17 years, specifically designed for initiation in clinical settings. Our application of an iterative design process through ongoing testing and refinement improved the qualitative and quantitative measures of success. By I3, participant engagement and ratings of acceptability and satisfaction increased: participants found the intervention *motivational* and highly rated both portions of the intervention [31].

Our findings support the importance of ongoing patient involvement in iterative intervention design [29]. Traditional intervention design requires full completion of the 8-week intervention and painstaking interview transcription and analysis before revising content and programming [28]. Our approach decreased the time required to make significant changes to both components of the program. Substantial changes included the addition of infographics, changes in daily SMS text message assessment programming, the addition of random text messages, and restructuring of the order of the SMS text message curriculum. The final iPACT intervention was designed to progressively and acceptably enhance participants' ability to identify emotions, challenge and change negative thoughts and behaviors, engage in prosocial web-based habits, and support peers.

The area in which we were unable to adequately revise the intervention because of both scope and budget was participants' desire for more personalized advice. Participants wanted to be able to solicit guidance from a live person when they became embroiled in a web-based conflict or were having a particularly bad day. This desire for a *just-in-time* adaptive intervention is not unique to our work [32]. Barriers to this real-time response include concerns about the need for 24/7 monitoring of messaging in case of human subjects' concerns, lack of accurate automated identification of moments of conflict, and lack of budget to create this level of tailoring [33]. Research on how best to balance automated and human components in SMS text message–based mental health interventions is warranted [17,32].

Many participants requested delivery of the intervention through a smartphone app. Participants said their friends used apps for communication (ie, Facebook messenger, WhatsApp, and Snapchat) more often than SMS text messaging. Delivering the intervention through an app could integrate additional features such as games and quizzes, provide more nuanced on-demand content, and allow greater intervention personalization (eg, through background images). Apps may be feasible and acceptable in behavioral interventions. In future work, the iterative intervention refinement methods described herein could be relevant for developing app-delivered prevention content. The operating system–specific nature of mobile apps, however, as well as the need for internet data to function, may be important limitations.

Although the intervention topic was unrelated to patients' pediatric visits, the intervention was initiated in a familiar setting where adolescents and parents often expect behavioral counseling to occur [9]. This location may enhance engagement. Future work should also consider the impact on clinical workflow and physicians' and nurses' perceptions of the program, as clinical staff engagement and buy-ins are essential for successful clinical implementation.

Limitations

The findings should be appraised within the context of study limitations. The study was conducted at a single site using a convenience sample of participants. Certain groups of adolescents, including those in police and state agencies (eg, foster care or group home), were not eligible to participate because of institutional review board restrictions. Although this study demonstrates feasibility and acceptability, it may not be generalizable to other populations, given the small sample size of participants drawn from 1 city. The current design cannot speak to the program's efficacy in reducing cybervictimization and its consequences; this would need to be measured in a randomized controlled trial.

Comparison With Previous Work

Contrary to recent reports that participant engagement decreases over time, our data suggest that adolescents continued to engage with over the full 8-week period of I3. With an 88.34% (940/1064) response rate to any daily assessment and 84% (16/19) of participants completing postintervention interviews, iPACT had higher engagement and retention rates than many digital interventions [34-36]. By paying close attention to the participants' input, the study team was able to add information that resonated with users [30,37]. During the follow-up interview, some participants reported using intervention elements to actively avoid web-based drama. However, it is important to note that the I3 arm had a very small sample size, and rates of engagement in a larger sample may not be the same, especially given that this sample was recruited in the context of a research study.

Although cyberbullying is equally common among youth of minority race or ethnicity and of low SES, few interventions have been developed specifically for this population [38]. These groups are less frequently included in research studies, and if and when they are included, they are often insufficient for *meaningful analyses* [39]. Cybervictimization interventions developed with largely White or high-SES populations may need tailoring for minority or low-SES populations because of differences in social norms, prevalence of in-person violence and disenfranchised neighborhoods, and underlying racism and mistrust [40]. This study recruited a sample of predominantly minority and low-SES adolescents. Engaging in the collaborative design and refinement process with this group is paramount.

XSL•FO
RenderX

Conclusions

Agile qualitative methods were used to iteratively develop iPACT, a mixed-modality digital intervention for adolescents affected by cybervictimization. Improvements in participants' satisfaction and level of daily participation suggest the importance of agile development methods and suggest the acceptability of initiating digitally augmented preventive interventions in the pediatric primary care setting. Future work should investigate the efficacy of iPACT in preventing cybervictimization and its consequences.

Acknowledgments

Funding for this study was provided by the National Institute of Child Health and Human Development (grant 1R21HD088739-01). This funding source had no involvement in the study design, data collection, analysis and interpretation, manuscript writing, or the decision to submit this paper for publication.

Conflicts of Interest

None declared.

References

1. Schneider S, O'Donnell L, Smith E. Trends in cyberbullying and school bullying victimization in a regional census of high school students, 2006-2012. J Sch Health 2015 Sep;85(9):611-620. [doi: 10.1111/josh.12290] [Medline: 26201758]
2. Prevent bullying. Centers for Disease Control. 2018. URL: https://www.cdc.gov/features/prevent-bullying/index.html [accessed 2019-06-16]
3. A majority of teens have experienced some form of cyberbullying. Pew Research Center. 2018. URL: http://www.pewinternet.org/2018/09/27/a-majority-of-teens-have-experienced-some-form-of-cyberbullying/ [accessed 2019-06-16]
4. Bottino SM, Bottino CM, Regina CG, Correia AV, Ribeiro WS. Cyberbullying and adolescent mental health: systematic review. Cad Saude Publica 2015 Mar;31(3):463-475 [FREE Full text] [doi: 10.1590/0102-311x00036114] [Medline: 25859714]
5. Kowalski RM, Giumetti GW, Schroeder AN, Lattanner MR. Bullying in the digital age: a critical review and meta-analysis of cyberbullying research among youth. Psychol Bull 2014 Jul;140(4):1073-1137. [doi: 10.1037/a0035618] [Medline: 24512111]
6. Cantone E, Piras AP, Vellante M, Preti A, Daníelsdóttir S, D'Aloja E, et al. Interventions on bullying and cyberbullying in schools: a systematic review. Clin Pract Epidemiol Ment Health 2015;11(Suppl 1 M4):58-76 [FREE Full text] [doi: 10.2174/1745017901511010058] [Medline: 25834628]
7. America's children: Key national indicators of well-being, 2015: well-child and well-adolescent visits. Federal Interagency Forum on Child and Family Statistics. 2015. URL: https://www.childstats.gov/pdf/ac2015/ac_15.pdf [accessed 2019-06-16]
8. Hagan JF, Shaw JS, Duncan PM, editors. Bright Futures Guidelines for Health Supervision of Infants, Children, and Adolescents. 4th Ed. Elk Grove Village, IL: American Academy of Pediatrics; 2017:1-31.
9. O'Keeffe GS, Clarke-Pearson K, Council on Communications Media. The impact of social media on children, adolescents, and families. Pediatrics 2011 Apr;127(4):800-804 [FREE Full text] [doi: 10.1542/peds.2011-0054] [Medline: 21444588]
10. Yarnall KS, Pollak KI, Østbye T, Krause KM, Michener JL. Primary care: is there enough time for prevention? Am J Public Health 2003 Apr;93(4):635-641. [doi: 10.2105/ajph.93.4.635] [Medline: 12660210]
11. Baroncelli A, Ciucci E. Unique effects of different components of trait emotional intelligence in traditional bullying and cyberbullying. J Adolesc 2014 Aug;37(6):807-815. [doi: 10.1016/j.adolescence.2014.05.009] [Medline: 25086457]
12. Portnoy DB, Scott-Sheldon LA, Johnson BT, Carey MP. Computer-delivered interventions for health promotion and behavioral risk reduction: a meta-analysis of 75 randomized controlled trials, 1988-2007. Prev Med 2008 Jul;47(1):3-16 [FREE Full text] [doi: 10.1016/j.ypmed.2008.02.014] [Medline: 18403003]
13. Williford A, Elledge LC, Boulton AJ, DePaolis KJ, Little TD, Salmivalli C. Effects of the KiVa antibullying program on cyberbullying and cybervictimization frequency among Finnish youth. J Clin Child Adolesc Psychol 2013;42(6):820-833 [FREE Full text] [doi: 10.1080/15374416.2013.787623] [Medline: 23659182]
14. Ranney M, Pittman SK, Riese A, Ybarra M, Huang J, Spirito A, et al. Mobile health intervention development principles: lessons from an adolescent cyberbullying intervention. In: Social Media and Healthcare Technology. Manoa: University of Hawai'i; Jan 03, 2018:1-10.
15. Lenhart A. Pew Research Center. 2015. URL: https://www.pewresearch.org/internet/2015/04/09/teens-social-media-technology-2015/ [accessed 2019-06-16]
16. Cunningham RM, Chermack ST, Zimmerman MA, Shope JT, Bingham CR, Blow FC, et al. Brief motivational interviewing intervention for peer violence and alcohol use in teens: one-year follow-up. Pediatrics 2012 Jun;129(6):1083-1090 [FREE Full text] [doi: 10.1542/peds.2011-3419] [Medline: 22614776]
17. Ranney ML, Thorsen M, Patena JV, Cunningham RM, Boyer EW, Walton MA, et al. "You need to get them where they feel it": Conflicting perspectives on how to maximize the structure of text-message psychological interventions for adolescents.

XSL•FO
RenderX

In: Proceedings of the 48th Hawaii International Conference on System Sciences. 2015 Jan Presented at: 48th Hawaii International Conference on System Sciences; Jan. 5-8, 2015; Kauai, HI, USA p. 3247-3255 URL: http://europepmc.org/abstract/MED/26640419 [doi: 10.1109/HICSS.2015.391]
18. Riese A, Gjelsvik A, Ranney ML. Extracurricular activities and bullying perpetration: results from a nationally representative sample. J Sch Health 2015 Aug;85(8):544-551. [doi: 10.1111/josh.12282] [Medline: 26149310]
19. Arain M, Campbell MJ, Cooper CL, Lancaster GA. What is a pilot or feasibility study? A review of current practice and editorial policy. BMC Med Res Methodol 2010;10(1):67 [FREE Full text] [doi: 10.1186/1471-2288-10-67] [Medline: 20637084]
20. Hekler EB, Klasnja P, Riley WT, Buman MP, Huberty J, Rivera DE, et al. Agile science: creating useful products for behavior change in the real world. Transl Behav Med 2016 Jun;6(2):317-328 [FREE Full text] [doi: 10.1007/s13142-016-0395-7] [Medline: 27357001]
21. Stewart RW, Drescher CF, Maack DJ, Ebesutani C, Young J. The development and psychometric investigation of the cyberbullying scale. J Interpers Violence 2014 Aug;29(12):2218-2238. [doi: 10.1177/0886260513517552] [Medline: 24424252]
22. Attkisson CC, Zwick R. The client satisfaction questionnaire. Psychometric properties and correlations with service utilization and psychotherapy outcome. Eval Program Plann 1982;5(3):233-237. [doi: 10.1016/0149-7189(82)90074-x] [Medline: 10259963]
23. Gender Identity in U.S. Surveillance Group. Gender-Related Measures Overview. Los Angeles, CA: Williams Institute, UCLA School of Law; 2013.
24. Harris K, Halpern C, Whitsel E. Study design. The National Longitudinal Study of Adolescent to Adult Health (Add Health). 2009. URL: http://www.cpc.unc.edu/projects/addhealth/design [accessed 2019-06-16]
25. 2011 National Survey of Children's Health. National Center for Health Statistics at the Centers for Disease Control. 2012. URL: http://www.cdc.gov/nchs/data/slaits/2011NSCHQuestionnaire.pdf [accessed 2019-06-16]
26. Badgett MV, Baker KE, Conron KJ, Gates GJ, Gill A, Greytak E, et al. Best practices for asking questions to identify transgender and other gender minority respondents on population-based surveys. The GenIUSS Group, UCLA Williams Institute. 2014. URL: https://williamsinstitute.law.ucla.edu/publications/geniuss-trans-pop-based-survey/ [accessed 2021-08-09]
27. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol 2013 Sep 18;13:117 [FREE Full text] [doi: 10.1186/1471-2288-13-117] [Medline: 24047204]
28. Smith J, Firth J. Qualitative data analysis: the framework approach. Nurse Res 2011;18(2):52-62. [doi: 10.7748/nr2011.01.18.2.52.c8284] [Medline: 21319484]
29. Ranney ML, Choo EK, Cunningham RM, Spirito A, Thorsen M, Mello MJ, et al. Acceptability, language, and structure of text message-based behavioral interventions for high-risk adolescent females: a qualitative study. J Adolesc Health 2014 Jul;55(1):33-40 [FREE Full text] [doi: 10.1016/j.jadohealth.2013.12.017] [Medline: 24559973]
30. Ranney ML, Freeman JR, Connell G, Spirito A, Boyer E, Walton M, et al. A depression prevention intervention for adolescents in the Emergency Department. J Adolesc Health 2016 Oct;59(4):401-410. [doi: 10.1016/j.jadohealth.2016.04.008] [Medline: 27267141]
31. Seko Y, Kidd S, Wiljer D, McKenzie K. Youth mental health interventions via mobile phones: a scoping review. Cyberpsychol Behav Soc Netw 2014 Sep;17(9):591-602. [doi: 10.1089/cyber.2014.0078] [Medline: 25007383]
32. Gonul S, Namli T, Huisman S, Erturkmen GB, Toroslu IH, Cosar A. An expandable approach for design and personalization of digital, just-in-time adaptive interventions. J Am Med Inform Assoc 2019 Mar 01;26(3):198-210 [FREE Full text] [doi: 10.1093/jamia/ocy160] [Medline: 30590757]
33. Yoon S, Parsons F, Sundquist K, Julian J, Schwartz JE, Burg MM, et al. Comparison of different algorithms for sentiment analysis: psychological stress notes. Stud Health Technol Inform 2017;245:1292 [FREE Full text] [Medline: 29295377]
34. Kobak KA, Mundt JC, Kennard B. Integrating technology into cognitive behavior therapy for adolescent depression: a pilot study. Ann Gen Psychiatry 2015;14:37 [FREE Full text] [doi: 10.1186/s12991-015-0077-8] [Medline: 26535048]
35. Waller R, Gilbody S. Barriers to the uptake of computerized cognitive behavioural therapy: a systematic review of the quantitative and qualitative evidence. Psychol Med 2009 May;39(5):705-712. [doi: 10.1017/S0033291708004224] [Medline: 18812006]
36. Whittaker R, Merry S, Stasiak K, McDowell H, Doherty I, Shepherd M, et al. MEMO--a mobile phone depression prevention intervention for adolescents: development process and postprogram findings on acceptability from a randomized controlled trial. J Med Internet Res 2012;14(1):e13 [FREE Full text] [doi: 10.2196/jmir.1857] [Medline: 22278284]
37. Head KJ, Noar SM, Iannarino NT, Grant HN. Efficacy of text messaging-based interventions for health promotion: a meta-analysis. Soc Sci Med 2013 Nov;97:41-48. [doi: 10.1016/j.socscimed.2013.08.003] [Medline: 24161087]
38. Duarte C, Pittman SK, Thorsen MM, Cunningham RM, Ranney ML. Correlation of minority status, cyberbullying, and mental health: a cross-sectional study of 1031 adolescents. J Child Adolesc Trauma 2018 Mar;11(1):39-48 [FREE Full text] [doi: 10.1007/s40653-018-0201-4] [Medline: 32318136]

XSL•FO
RenderX

39. Lau AS, Chang DF, Okazaki S. Methodological challenges in treatment outcome research with ethnic minorities. Cultur Divers Ethnic Minor Psychol 2010 Oct;16(4):573-580 [FREE Full text] [doi: 10.1037/a0021371] [Medline: 21058823]
40. Hong JS. Feasibility of the olweus bullying prevention program in low-income schools. J Sch Violence 2009 Jan 02;8(1):81-97 [FREE Full text] [doi: 10.1080/15388220802067953]

Abbreviations

I1: iteration 1
I2: iteration 2
I3: iteration 3
iPACT: Intervention to Prevent Adolescent Cybervictimization with Text message
SES: socioeconomic status

Edited by G Eysenbach; submitted 19.11.20; peer-reviewed by P Chai, A Radovic-Stakic; comments to author 08.01.21; revised version received 04.06.21; accepted 05.07.21; published 27.08.21.

Please cite as:
Ranney ML, Pittman SK, Moseley I, Morgan KE, Riese A, Ybarra M, Cunningham R, Rosen R
Cyberbullying Prevention for Adolescents: Iterative Qualitative Methods for Mobile Intervention Design
JMIR Form Res 2021;5(8):e25900
URL: https://formative.jmir.org/2021/8/e25900
doi:10.2196/25900
PMID:34448702

©Megan L Ranney, Sarah K Pittman, Isabelle Moseley, Kristen E Morgan, Alison Riese, Michele Ybarra, Rebecca Cunningham, Rochelle Rosen. Originally published in JMIR Formative Research (https://formative.jmir.org), 27.08.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Dietary Intake and Health Status of Elderly Patients With Type 2 Diabetes Mellitus: Cross-sectional Study Using a Mobile App in Primary Care

Joane Diomara Coleone[1], BSc, MSc; Ericles Andrei Bellei[2], BTech, MSc; Mateus Klein Roman[2], BSc; Vanessa Ramos Kirsten[3,4], BSc, MSc, PhD; Ana Carolina Bertoletti De Marchi[1,2], BSc, MSc, PhD

[1]School of Physical Education and Physiotherapy, University of Passo Fundo, Passo Fundo, RS, Brazil
[2]Institute of Exact Sciences and Geosciences, University of Passo Fundo, Passo Fundo, RS, Brazil
[3]Department of Foods and Nutrition, Federal University of Santa Maria, Palmeira das Missões, RS, Brazil
[4]Graduate Program in Gerontology, Federal University of Santa Maria, Santa Maria, RS, Brazil

Corresponding Author:
Ericles Andrei Bellei, BTech, MSc
Institute of Exact Sciences and Geosciences
University of Passo Fundo
B5 Building, São José, Highway BR285, Km 292
Passo Fundo, RS, 99052-900
Brazil
Phone: 55 5433168354
Email: 168729@upf.br

Abstract

Background: Healthy dietary intake reduces the risk of complications of diabetes mellitus. Using assessment methods helps to understand these circumstances, and an electronic application may optimize this practice.

Objective: In this study, we aimed to (1) assess the dietary intake and health status of elderly patients with type 2 diabetes mellitus (T2DM) in primary care, (2) use a mobile app as a tool for data collection and analysis in the context of primary care, and (3) verify the perceptions of multidisciplinary health professionals regarding app use.

Methods: First, we developed a mobile app comprised of the questions of the Food and Nutrition Surveillance System (SISVAN) of Brazil, which includes a food frequency questionnaire of food categories with a recall of the previous 7 days. Thereafter, we used the app to collect data on the health status and dietary intake of 154 participants, aged 60-96 years, diagnosed with T2DM, and under treatment in primary care centers in the northern region of Rio Grande do Sul, Brazil. We also collected participants' demographic, anthropometric, biochemical, and lifestyle variables. The associations between dietary intake and other variables were tested using chi-square tests with a 5% significance level. Regarding the app, we assessed usability and acceptance with 20 health professionals.

Results: Between August 2018 and December 2018, participants had an intake in line with recommended guidelines for raw salads (57.1%), fruits (76.6%), milk products (68.2%), fried foods (72.7%), savory biscuits (60.4%), cookies or sweets (72.1%), and sugary drinks (92.9%) Meanwhile, the consumption of beans (59.7%), pulses and cooked vegetables (73.4%), and processed meat products (59.7%) was not in line with the guidelines. There were statistically significant differences in meeting the recommended guidelines among participants of different genders (P=.006 and P=.035 for the intake of fried foods and sugary drinks, respectively), place of residence (P=.034 for the intake of cookies and sweets), family history of diabetes (P<.001 for the intake of beans), physical activity engagement (P=.003 for the intake fresh fruits), history of smoking (P=.001 for the intake of raw salads), and presence of coronary disease (P=.050 for the intake of pulses and cooked vegetables). The assessment of usability resulted in a mean score of 71.75 points. Similarly, the assessment of the 15 acceptance questions revealed high scores, and the qualitative questions revealed positive perceptions.

Conclusions: We identified that most participants complied with recommended intake guidelines for 7 of 10 categories in the SISVAN guidelines. However, most participants were overweight and had nutritional and clinical disorders, which justifies further investigations in this population. The app was well-rated by health professionals and considered a useful and promising tool for collecting and analyzing data in primary care settings.

XSL•FO
RenderX

(JMIR Form Res 2021;5(8):e27454) doi:10.2196/27454

KEYWORDS

eating; mobile applications; primary health care; aged; type 2 diabetes mellitus

Introduction

An aging population, the prevalence of obesity, sedentary lifestyles, and urbanization processes have been contributing to the increase in type 2 diabetes mellitus (T2DM) worldwide [1]. In public health, especially in developing countries such as Brazil, T2DM results in high economic and social costs for its treatment and care, due to the association with several complications [2]. A healthy dietary intake reduces the risk of complications to maintain acceptable health standards and functional capacity [3]. However, patients with T2DM, particularly the elderly, have difficulty adhering to a healthy diet, as they may perceive diet plans as prohibitive, restrictive, and challenging [4]. In these circumstances, understanding dietary habits requires appropriate assessment tools.

A systematic review [5] revealed that the main method for assessing dietary intake in public health settings for elderly patients with T2DM can be classified as a food frequency questionnaire. However, the application of the questionnaire is paper-based in most cases, which is time-consuming. New technologies offer great opportunities to fully assess the intake of foods and nutrients of large populations at relatively low cost and in real time [6]. These tools have fast access, full-time availability, and potential access through mobile apps, showing practicality for hospitals, clinics, and outpatient clinics [7]. In Brazil's public health system, primary care is the main strategy of surveillance, which also aims to provide comprehensive health care, specialized services, and hospital care, as well as health promotion and disease prevention activities [8]. In these settings, although still not widespread, the use of technology is promising [9,10], in particular to facilitate the work of multidisciplinary health professionals. Therefore, initiatives to introduce digital tools are relevant to test possibilities and improve primary care in Brazil.

In this study, we aimed to (1) assess the dietary intake and health status of elderly patients with T2DM in primary care, (2) use a mobile app as a tool for data collection and analysis in the context of primary care, and (3) document the perceptions of multidisciplinary health professionals regarding app use.

Methods

Overview

We considered the sequences and processes from several studies in the literature involving software applications [11-13] to select, develop, and apply an assessment tool complying with the DIET@NET partnership guidelines [14]. We performed a cross-sectional study between August 2018 and December 2018 in public primary health centers in 4 small towns in the northern region of Brazil's Rio Grande do Sul, namely Estação, Erebango, Getúlio Vargas, and Ipiranga do Sul, which had 27,079 inhabitants in the last official census. This study was conducted according to the guidelines laid down in the Declaration of Helsinki, and all procedures involving research study participants were approved by the ethics committee of the University of Passo Fundo under opinion number 2660304. Written informed consent was obtained from all participants. According to national regulations, participants were not compensated.

Sampling, Subjects, and Recruitment

For the sample calculation, we considered the total population of elderly patients who were diagnosed with T2DM by the public health control from the cities investigated and attended a primary care center (N=257), with an expectation of a dietary intake 50% in line with the Food and Nutrition Surveillance System (SISVAN) of Brazil guidelines, acceptable error of 5%, and a confidence level of 95%, resulting in a target sample size of 154 individuals. The inclusion criteria were 60 years of age or older; T2DM diagnosis, either by self-report or confirmed by examinations in the patient history (fasting glycemia or glycated hemoglobin [HbA_{1c}]); and cognitive ability confirmed by the Mini-Mental State Examination [15]. Through phone calls, we randomly called participants to attend the primary health center for an interview. We invited participants until we reached the required sample. All participants signed an informed consent form.

Development of a Mobile App for Dietary Assessment

Initially, we performed a systematic mapping to investigate the different methods for assessing the dietary intake of adults and the elderly [5]. We observed that the main assessment method was a food frequency questionnaire. However, the questionnaires are usually applied manually. To optimize time and resources, we wanted to use an electronic method for collecting and processing the questionnaires.

In this study, we used the Dietary Intake Form proposed by SISVAN [16], which is an initiative from the Brazilian Ministry of Health. The form is a food frequency questionnaire with a recall period of the previous 7 days and is comprised of 10 categories, which relate to a healthy diet (eg, daily intake of beans, fresh fruits, vegetables, and milk) or to nonrecommended practices (eg, frequent intake of fried foods, processed meat products, cookies, sweets, and soft drinks). SISVAN guidelines recommend an intake of 7 serves per week of raw salad, pulses and cooked vegetables, fresh fruit, and milk or yogurt; 5 or more serves per week of beans; and 0-1 serves per week of fried foods, processed meat products, savory biscuits, cookies or sweets, and sugary drinks.

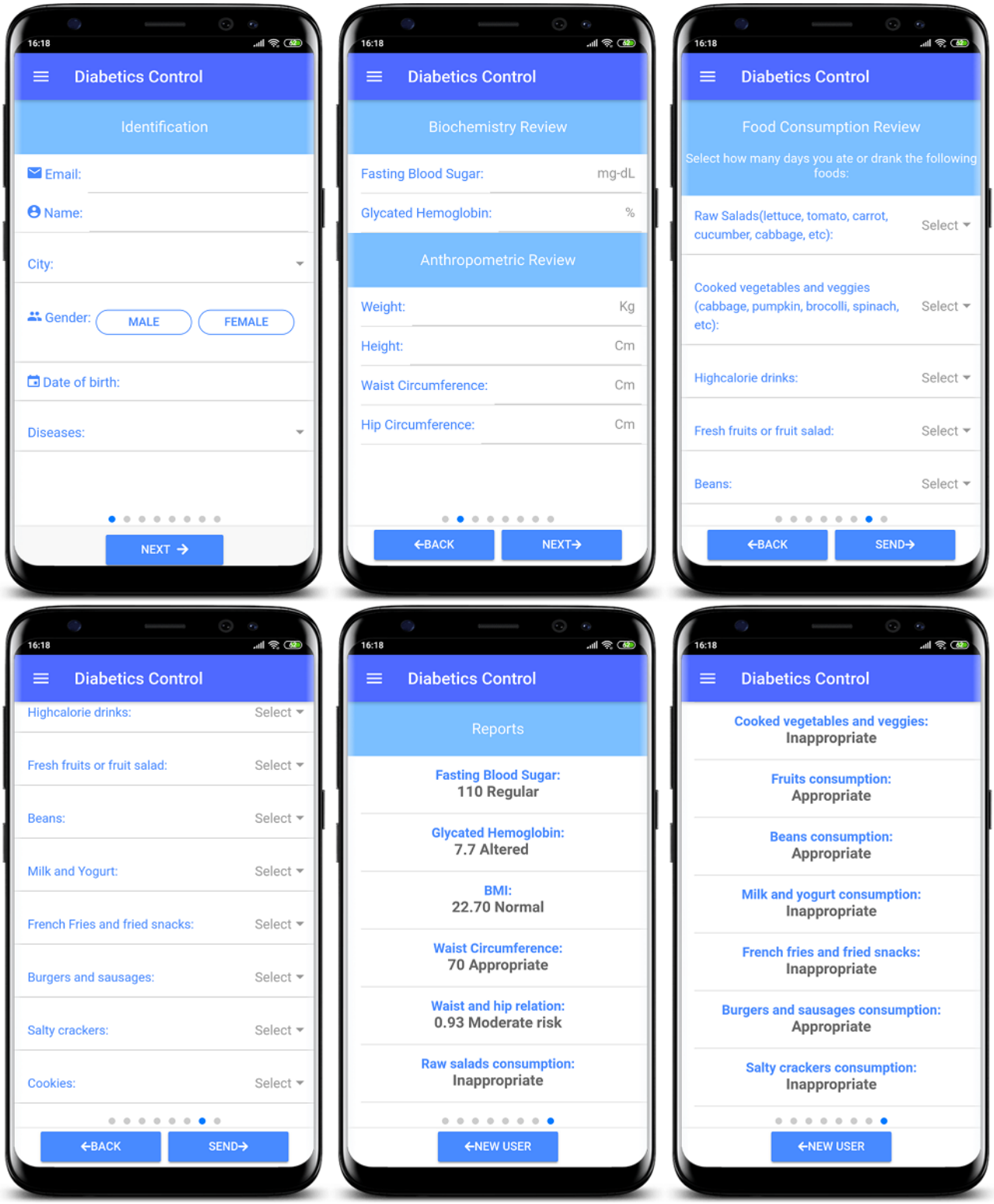
Previously, a mobile app addressing SISVAN had been developed for use with inpatients [7]. Hence, we developed a new version, with improved functionalities adapted to primary care outpatients, called Diabetes Food Control 2. The app now includes questions on demographic and economic characteristics (eg, marital status, self-reported skin color, education, income, and occupation), clinical and nutritional history, time from

XSL•FO
RenderX

diagnosis, family history of diabetes, medication intake, presence of comorbidities, nutritional monitoring, lifestyle, physical activity, smoking, and alcohol consumption. The Brazilian Institute of Industrial Property granted us a software registration certificate under number BR5120190010540.

The app (Figure 1) was designed so that health care professionals could use it in primary health care centers to assess the dietary intake and health status of outpatients with T2DM, addressing the recommendations of Brazil's Ministry of Health [16]. Using the app, we intended to optimize data collection and make the usage of apps a more consistent practice in primary care settings. The app's report provides straightforward information on whether the patient meets the SISVAN criteria related to food intake and anthropometric measurements. It also provides personalized feedback for patients in the form of a report they can receive by email.

Figure 1. Screenshots of the Diabetes Food Control 2 mobile app showing patient data collection forms and feedback on the data collected.



XSL•FO
RenderX

Procedures

We collected all data with the SISVAN questionnaire using the Diabetes Food Control 2 app. In the public primary health centers selected for the study, the same dietitian researcher interviewed each participant to simultaneously fill out the questionnaire on the app. The dietitian also collected demographic and economic data such as marital status, self-reported skin color, level of education, income, and occupation. The participants followed the use of the app during the entire interview and interacted by filling out their demographic information.

To evaluate nutritional status, we analyzed body weight (kg), height (cm), and waist and hip circumferences (cm). For stature measurement, we used a 200-cm stadiometer with a scale of 0.2 centimeters. To assess weight, we used a Welmy digital scale with a capacity of 150 kg. We took all measurements according to the recommendations of the Brazilian Ministry of Health [17]. Using the measured height and weight, the app calculated the BMI and the classification of nutritional status according to the guidelines of the Pan American Health Organization [18].

The dietitian assessed waist circumference with a measuring tape at the smallest curvature between the ribs and the iliac crest, without compressing the skin. For waist circumference, the categories were low risk (men <94 cm; women <80 cm), high risk (men ≥94 cm; women ≥80 cm), and very high risk (men ≥102 cm; women ≥88 cm). We measured hip circumference in centimeters in a horizontal plane at the region of the greatest gluteal protuberance. We calculated the waist-to-hip ratio according to the cutoff points recommended by the World Health Organization [19]. For waist-to-hip ratio, the categories were low risk (men <0.9; women <0.8), moderate risk (men 0.9-0.95; women 0.8-0.85), and high risk (men >0.95; women >0.85).

To assess glycemia, we asked the primary health center in each town to conduct a HbA_{1c} test. All blood tests were collected and evaluated by trusted outsourced laboratories that were contracted by the primary health center. All handling procedures were carried out in accordance with national quality and safety standards and under the supervision of the researchers. A value of 8.0% was the standard reference value, as recommended by the American Diabetes Association for individuals aged over 60 years [20]. We also assessed fasting capillary glycemia, which a nursing professional verified at the time of the questionnaire application using an Accu-Check Active glucometer. In this case, 70-110 mg/dL was the standard reference value [20].

During the interview, the dietitian also investigated the clinical and nutritional history of participants, time from T2DM diagnosis, family history of diabetes, and medication intake. The dietitian registered information on the presence of comorbidities, such as hypertension and dyslipidemia, and whether the participant had received any nutritional guidance, engaged in physical activities routinely, smoked, and consumed alcohol. At the end of the interview, the participant received nutritional instructions for better glycemic control and a report generated by the app showing his or her dietary and health status according to the SISVAN guidelines.

Assessment of the Mobile App With Health Professionals

We also assessed the mobile app with 20 health professionals, including 4 medical doctors, 8 registered nurses, and 8 community health agents (registered nurses and licensed practical nurse) who worked in the primary health centers. Individually, we presented the app and described its functionalities and form of use. We let each professional test and interact with the app and fill out the questionnaire once with dummy data, which we discarded later. After the professionals tested the app, we applied a questionnaire to assess usability. To assess usability, we used an adaptation of the System Usability Scale [21]. To analyze acceptance, we used a questionnaire from the Technology Acceptance Model (TAM) [22]. Last, we asked the health professionals the following 3 descriptive questions: Have you ever used a mobile app in clinical practice? How was your experience with using the Diabetes Food Control 2? What changes do you suggest for this app to make it more useful or applicable to clinical practice? The opinions of professionals from the primary health centers would allow us to understand the potential for implementing the app in clinical practice.

Data Analysis

Participants' dietary intakes were summarized and compared by all the demographic, nutritional, biochemical, and lifestyle variables. Intake was classified according to whether it was in line with SISVAN guidelines [16]. We used SPSS 22.0 statistical software (IBM Corp, Armonk, NY) to analyze the quantitative data. Basic quantitative data are described as mean, SD, and median. Categorical data are described by simple frequencies. The associations between dietary intake and demographic, nutritional, biochemical, and lifestyle variables were tested using chi-square tests, considering a 5% significance level for all analyses. For the qualitative data, we assessed the data in groups, following the model proposed by Minayo [23].

Results

Participants

Table 1 presents the demographic characteristics. Most of the participants were women, married, with self-reported white skin color, residents of an urban area, with <4 years of education, retired, with family income between 4 and 10 minimum wages, and presenting with a family history of diabetes.

XSL•FO
RenderX

Table 1. Demographic characteristics of 154 patients with type 2 diabetes mellitus in primary care in Rio Grande do Sul, Brazil in 2018.

Characteristic	n (%)
Gender	
Female	107 (69.5)
Male	47 (30.5)
Age (years)	
60-69	67 (43.5)
70-79	63 (40.9)
80-89	22 (14.3)
≥90	2 (1.3)
Marital status	
Single	5 (3.2)
Married	101 (65.6)
Divorced	5 (3.2)
Widow(er)	43 (27.9)
Self-reported skin color	
White	130 (84.4)
Brown	15 (9.7)
Black	9 (5.8)
Place of residence	
Urban area	126 (81.8)
Rural area	28 (18.2)
Education level	
Elementary school	137 (89.0)
High school	4 (2.5)
Higher education	3 (1.9)
Illiterate	10 (6.5)
Family income (minimum wages in Brazil)	
Between 10 and 20 minimum wages	12 (7.8)
Between 4 and 10 minimum wages	70 (45.5)
Between 2 and 4 minimum wages	51 (33.1)
Up to 2 minimum wages	21 (13.6)

Lifestyle, Anthropometric, and Glucose Measurements

Table 2 presents the data on clinical characterization and lifestyle. The majority reported using only oral hypoglycemic agents for T2DM treatment, having dyslipidemia and systemic arterial hypertension, and using medications for hypertension. Most respondents also reported having received some nutritional guidance but not performing nutritional monitoring. Most participants reported having never smoked, not consuming alcohol, and not performing physical activity.

XSL•FO
RenderX

Table 2. Clinical characteristics and lifestyle of 154 patients with type 2 diabetes mellitus in primary care in Rio Grande do Sul, Brazil in 2018.

Characteristic	n (%)
Family history of diabetes (yes)	99 (64.3)
Diabetes medication	
Oral hypoglycemic agents	98 (63.6)
Insulin only	8 (5.2)
Insulin and hypoglycemic agents	35 (22.7)
None	13 (8.4)
Self-reported hypertension (yes)	130 (84.4)
Medication use for hypertension (yes)	76 (97.4)
Self-reported dyslipidemia (yes)	88 (57.1)
Has received nutritional guidance (yes)	89 (57.8)
Professional providing guidance	
Physician	37 (24.0)
Dietitian	37 (24.0)
Nurse	4 (2.6)
Nutritional monitoring (yes)	2 (1.3)
Smoking status	
Never smoked	107 (69.5)
Former smoker	38 (24.7)
Smokes currently	9 (5.8)
Frequency of alcohol consumption	
Never	121 (78.6)
Less than 1 dose per month	22 (14.3)
Between 1 and 3 doses per month	7 (4.5)
Between 4 and 7 doses per month	4 (2.6)
Engaged in physical activity (yes)	63 (40.9)

Table 3 presents the characterization of nutritional status. We identified that most participants were obese and had a waist circumference indicating a very high risk of metabolic disorders, a waist-to-hip ratio categorized as high risk, altered capillary glycemia, and altered glycated hemoglobin. As for comorbidities, we identified that 51.9% (80/154) of respondents reported none of the diseases investigated, 31.2% (48/154) with coronary disease, 16.9% (26/154) with diabetic retinopathy, 16.9% (16/154) with depression, 5.8% (9/154) with kidney disease, 1.3% (2/154) with diabetic neuropathy, and 1.3% (2/154) had received a diagnosis of diabetic foot.

XSL•FO
RenderX

Table 3. Characterization of nutritional and biochemical status of 154 patients with type 2 diabetes mellitus in primary care in Rio Grande do Sul, Brazil in 2018.

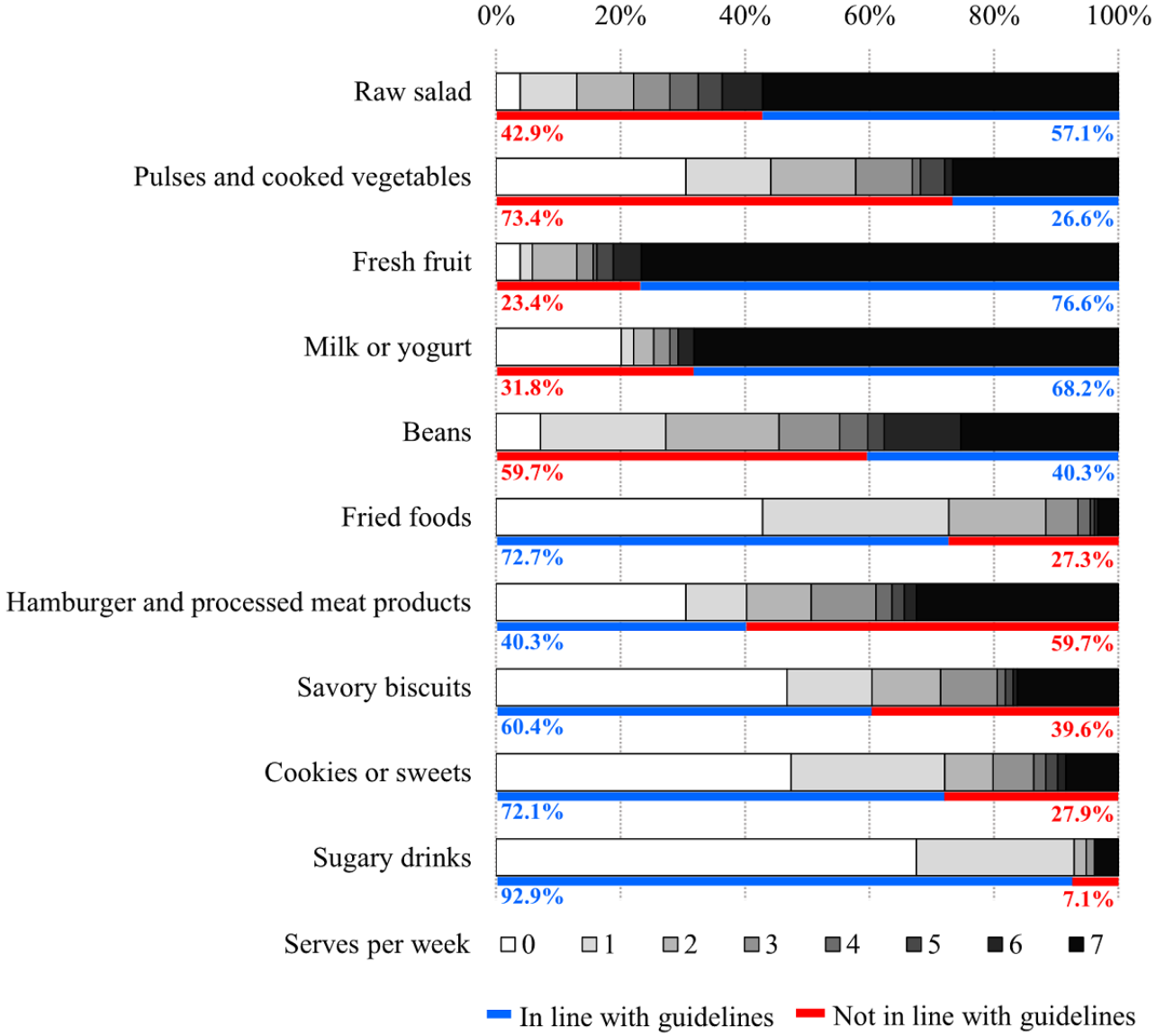
Indicator	Value
Body mass index (kg/m^2), mean (SD)	31.48 (5.95)
Body mass index, n (%)	
Underweight (<23 kg/m^2)	7 (4.5)
Normal weight (≥23 kg/m^2 to <28 kg/m^2)	34 (22.1)
Overweight (≥28 kg/m^2 to <30 kg/m^2)	25 (16.2)
Obesity (≥30 kg/m^2)	88 (57.1)
Waist circumference (cm), mean (SD)	108.2 (13.2)
Waist circumference, n (%)	
Low risk (men <94 cm; women <80 cm)	5 (3.2)
High risk (men ≥94 cm; women ≥80 cm)	12 (7.8)
Very high risk (men ≥102 cm; women ≥88 cm)	137 (89.0)
Waist-to-hip ratio, mean (SD)	0.999 (0.081)
Waist-to-hip ratio, n (%)	
Low risk (men <0.9; women <0.8)	1 (0.6)
Moderate risk (men 0.90:0.95; women 0.80:0.85)	6 (3.9)
High risk (men >0.95; women >0.85)	147 (95.5)
Capillary glycemia (mg/dL), mean (SD)	155 (48)
Capillary glycemia, n (%)	
Normal (70-110 mg/dL)	54 (35.1)
Altered (<70 and >110 mg/dL)	100 (64.9)
HbA_{1c}, mean (SD)	7.4 (1.4)
Normal (≤8.0%)	45 (29.2)
Altered (>8.0%)	50 (32.5)
Not collected	59 (38.3)

Dietary Intake

Among the participants, 82.5% (127/154) ate breakfast, 50% (77/154) ate a morning snack, 100% (154/154) ate lunch, 70.1% (108/154) ate dinner, and only 9.1% (14/154) ate supper. Most of them (139/154, 90.3%) described a usual diet in the week prior to the study, 85.1% (131/154) reported not adding salt in prepared food, and 52.6% (81/154) stated using vegetable oil or butter. Most participants had an intake of raw salads, fruits, milk products, fried foods, savory biscuits, cookies or sweets, and sugary drinks in line with recommended guidelines. Meanwhile, the intake of beans, pulses and cooked vegetables, and processed meat products were not in line with recommended guidelines. Figure 2 summarizes the food frequency data for the number of serves per week and the classification according to the SISVAN guidelines. Absolute values are available in Multimedia Appendix 1.

XSL•FO
RenderX

Figure 2. Food frequency according to the number of servings per week and classification according to the Food and Nutrition Surveillance System of Brazil (SISVAN) guidelines from a sample of 154 patients with type 2 diabetes mellitus in primary care in Rio Grande do Sul, Brazil in 2018.



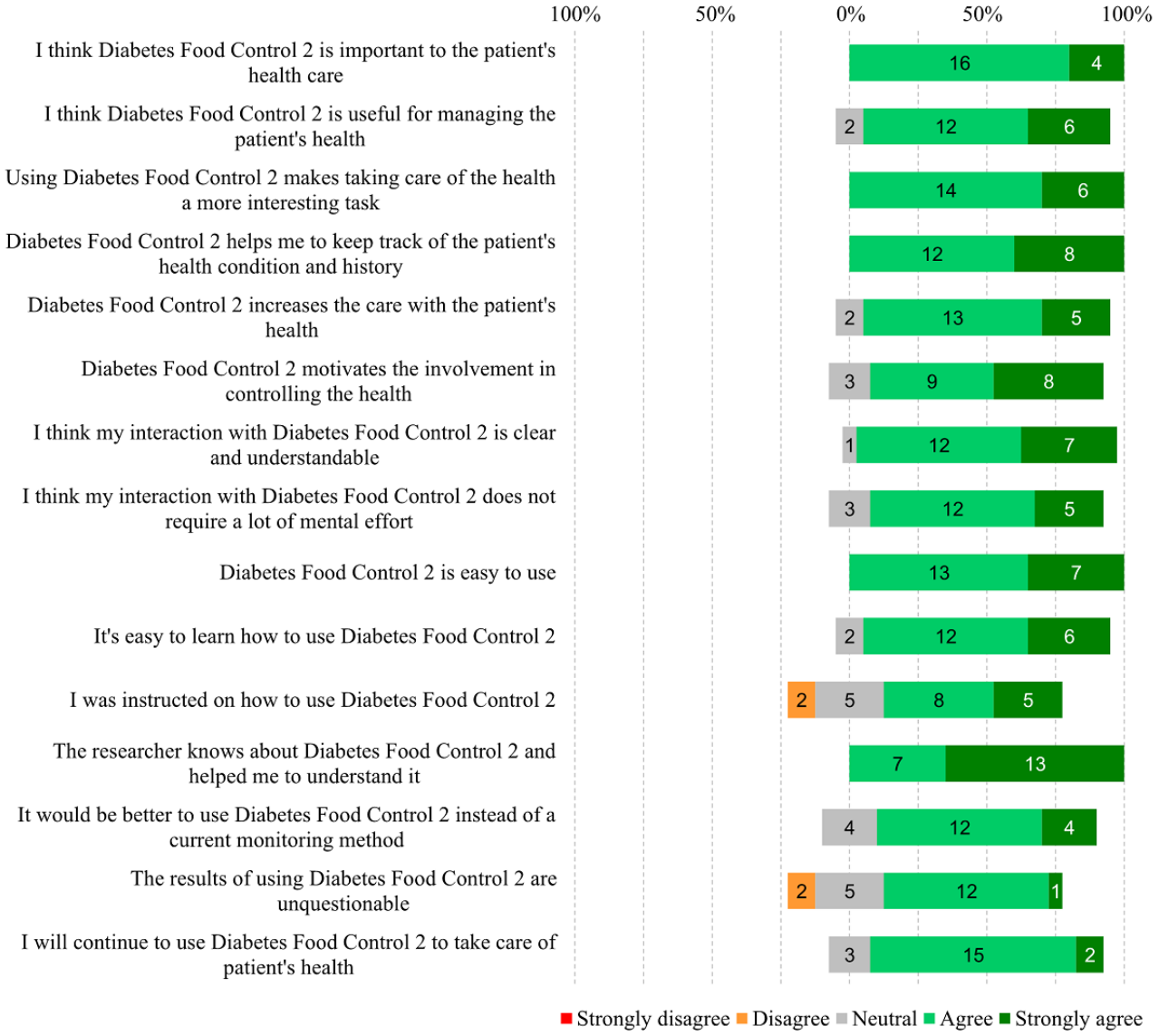
We compared the association between dietary intake with the demographic variables (Multimedia Appendix 2) and with lifestyle, nutritional, and biochemical variables (Multimedia Appendix 3). Statistically, we identified that women had healthier behaviors regarding the intake of fried foods (χ^2_1=7.963, *P*=.006) and sugary drinks (χ^2_1=6.127, *P*=.035). As for the place of residence, respondents living in a rural area had a more appropriate intake of cookies and sweets than residents living in an urban area (χ^2_1=5.035, *P*=.034). Participants without a family history of T2DM had a more appropriate intake of beans (χ^2_1=15.170, *P*<.001). We detected a more appropriate fruit intake in participants who engaged in physical activity (χ^2_1=8.955, *P*=.003). We also found that the lower the rate of smoking, the more adequate the intake of raw salad (χ^2_1=13.034, *P*=.001). Participants with coronary disease reported a more appropriate intake of pulses and cooked vegetables (χ^2_1=4.223, *P*=.050).

Health Professionals' Perceptions About Using the Mobile App

The usability assessment with the System Usability Scale generates a score that should be ≥68 points to determine satisfactory results. The assessment of the Diabetes Food Control 2 by health professionals resulted in a mean of 71.75 points, confirming satisfactory usability. Figure 3 presents the responses from health professionals to the TAM questionnaire [22]. The health professionals also answered 3 descriptive questions, depicted in Textbox 1.

XSL•FO
RenderX

Figure 3. Responses from 20 multidisciplinary, primary care health professionals to the acceptance questionnaire assessing the Diabetes Food control app in Rio Grande do Sul, Brazil in 2018.



Textbox 1. Health professionals' perceptions about using an app in primary care.

1. "Have you ever used a mobile app in clinical practice?" No participant had used apps in professional practice.
2. "How did you experience the use of Diabetes Food Control 2?" Among the responses, the professionals mentioned they would have improved knowledge on dietary monitoring of patients, agility and practicality in data collection, higher information reliability, and a more significant scientific basis to guide patients.
3. "What changes do you suggest for this app to make it more useful or applicable to clinical practice" Most professionals said there was no need to change the app because it was already adequate. Among the suggestions for improvements, some professionals mentioned the possibility of having an option to increase the font size in the app (accessibility option), having instructions on how to measure anthropometric variables (eg, where to measure the waist and where to measure the hip), showing the percentage of questionnaire completion, and showing the total completion time.

Discussion

Principal Findings

This study assessed various aspects of the nutrition of elderly patients diagnosed with T2DM in the context of primary health in line with SISVAN guidelines, which is a national parameter in Brazil. For data collection, we developed and evaluated a new app based on recommendations from medical associations and literature studies. Interesting insights are presented about the use of the app and the health professionals' perceptions about its features and the possibility of using it in the context of primary health.

On nutritional assessments, the profile found was similar to surveys performed in the same circumstances involving adult

XSL•FO
RenderX

[24] and elderly [25] patients with T2DM. The anthropometric assessment identified that 73.3% of the study population was overweight, which is a rate higher than that presented in similar studies. This stresses the need for greater attention to elderly patients with T2DM assisted in basic health care, especially due to the evident metabolic syndrome and the worsened transition from adulthood to the elderly.

In this study, we identified that men had a greater intake of fried foods than women with a similar dietary behavior, as in the study by Linde et al [26], in which men with higher BMIs consumed more fried foods than women. Similar findings were also reported by Cahill et al [27] and Sayon-Orea et al [28]. Regarding the intake of sugary drinks, we identified that women had a more appropriate intake than men, corroborating other studies [29,30]. As for the place of residence, respondents living in a rural area had a more appropriate intake of cookies and sweets than residents in an urban area, which was similar to the findings of Farrell et al [31] regarding foods and beverages with high sugar content.

The literature has already shown that people who engage in physical activity have better dietary habits [32]. In our study, respondents who reported engaging in physical activity had a more appropriate fruit intake. These data corroborate the studies by Tan et al [33] and Alakaam and Lemacks [34], which identified a positive relationship between physical activity and the intake of fruits and vegetables. We also identified an inversely proportional association between salad intake and smoking. Sebastian et al [35] had similar findings, in which salad intake seemed more prevalent among former smokers and people who had never smoked. Besides, the intake of pulses was significantly higher in participants with cardiovascular disease. Other studies have also associated the intake of fruits and vegetables with a lower prevalence of coronary disease and the intake of fiber with a lower prevalence of acute myocardial infarction [36,37]. Interestingly, our findings also suggest that respondents without a family history of diabetes had a more appropriate intake of beans. Some prior studies discuss the relationship between T2DM and the intake of beans. For instance, Ley et al [38] affirmed that a diet rich in vegetables, legumes, and unrefined grains was associated with a lower risk of T2DM, and for diagnosed patients, the intake of such foods improves glycemic levels and lipid control.

The app we developed has several benefits emphasized by health professionals. Other researchers who used mobile apps in dietary intake assessments cited similar advantages. Rangan et al [39] mentioned a reduction in data entry and coding workload for researchers. Meanwhile, Bucher Della Torre et al [40] reaffirmed that electronic assessment helps to prevent some pitfalls of paper-based food records, such as errors in transcription and difficulties in reading what participants have written. There is a growing interest in using smartphones to collect data on dietary assessment [6], considering mobile technology offers a wide range of suitable options to achieve the benefits that technology can deliver [41-43].

Among the 20 respondents, none had used apps in their professional practice. This suggests the need for implementing the use of technology in public health, considering the potential help and benefits provided to professionals. Health professionals also reported that, by using the app, they could thoroughly monitor patients in the long term, considering that the questionnaire is easy to apply and contains all the questions needed for proper monitoring in the context of primary health. Community health agents, who are professionals who visit patients' homes, mentioned that the app could be used by them to counsel the general population regarding nutrition. These professionals receive little prior training on nutrition, and the app would provide a scientific basis so they could provide such care to patients.

Limitations

The physiological aspect of aging was a limitation of the present study. The potential memory loss of the respondents may have affected their answers to retrospective dietary intake. The length of the questionnaire may have also negatively affected the disposition of participants to respond. Another limitation was the impossibility of measuring the HbA_{1c} of all respondents in the laboratory, due to the lack of financial resources. The app was only used one time; therefore, the long-term effectiveness and usefulness should be studied further.

Conclusions

The dietary intake of the assessed participants is partly appropriate. However, this population was substantially overweight and presented with metabolic syndrome and poor results in the anthropometric and biochemical assessments. Moreover, there was a low intake of pulses and cooked vegetables and beans and a high intake of processed meat products. Mostly, the findings of this study are similar to others in the literature, which reinforces a concerning reality of the health conditions of elderly patients with T2DM. These findings can guide dietary interventions and health education in similar settings. In addition, this study adds to the health informatics literature that a new app can be used in the context of public health in Brazil. The mobile app was useful in the study and well-rated by health professionals in primary health care settings. Its features are encouraging tools for use in future studies. In this study, the app was developed to completely focus on T2DM in primary care settings in Brazil. For application in other settings or for other health conditions, the development method could be adapted or used in a similar way, but considering the specificity of other chronic diseases and according to the specific guidelines available. Future studies can use this study as an example of a location and features where technology is satisfactorily accepted by health care professionals in the first introduction initiative.

Acknowledgments

The authors would like to thank the patients and professionals who participated in the assessments. This work was supported in part by the National Council for Scientific and Technological Development (CNPq); State Funding Agency of Rio Grande do

XSL•FO
RenderX

Sul (FAPERGS); and Coordination and Improvement of Higher Level or Education Personnel – Brazil (CAPES) – finance code 001. The authors are solely responsible for designing and performing this study and all its analyses, for drafting and editing the manuscript, and for the final content.

Authors' Contributions

JC and AM conceived and designed the study. JC and MR acquired the data. JC, EB, VK, and AM analyzed and interpreted the data. JC, EB, and MR drafted the manuscript. JC, EB, VK, and AM critically reviewed the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Food frequency in absolute number of servings per week.
[DOCX File , 14 KB - formative_v5i8e27454_app1.docx]

Multimedia Appendix 2
Intake in line with recommended guidelines categorized by demographic and economic variables.
[DOCX File , 22 KB - formative_v5i8e27454_app2.docx]

Multimedia Appendix 3
Intake in line with recommended guidelines categorized by clinical, nutritional, biochemical, and lifestyle variables.
[DOCX File , 54 KB - formative_v5i8e27454_app3.docx]

References

1. Moura EC, Pacheco-Santos LM, Peters LR, Serruya SJ, Guimarães R. Research on chronic noncommunicable diseases in Brazil: meeting the challenges of epidemiologic transition. Rev Panam Salud Publica 2012 Mar;31(3):240-245. [doi: 10.1590/s1020-49892012000300009] [Medline: 22569699]
2. Bahia LR, Araujo DV, Schaan BD, Dib SA, Negrato CA, Leão MPS, et al. The costs of type 2 diabetes mellitus outpatient care in the Brazilian public health system. Value Health 2011;14(5 Suppl 1):S137-S140 [FREE Full text] [doi: 10.1016/j.jval.2011.05.009] [Medline: 21839888]
3. Pinto de Souza Fernandes D, Duarte MSL, Pessoa MC, Franceschini SDCC, Ribeiro AQ. Evaluation of diet quality of the elderly and associated factors. Arch Gerontol Geriatr 2017 Sep;72:174-180. [doi: 10.1016/j.archger.2017.05.006] [Medline: 28688369]
4. Zanetti ML, Arrelias CCA, Franco RC, Santos MAD, Rodrigues FFL, Faria HTG. [Adherence to nutritional recommendations and sociodemographic variables in patients with diabetes mellitus]. Rev Esc Enferm USP 2015 Aug;49(4):619-625 [FREE Full text] [doi: 10.1590/S0080-623420150000400012] [Medline: 26353099]
5. Coleone JD, Bellei EA, De Marchi ACB. Food consumption and glycemic testing of adults and elderly diabetic patients from Public Health: A systematic review of assessment methods. Diabetes Metab Syndr 2019;13(5):3005-3010. [doi: 10.1016/j.dsx.2018.07.013] [Medline: 30057070]
6. Cade JE. Measuring diet in the 21st century: use of new technologies. Proc Nutr Soc 2017 Aug;76(3):276-282. [doi: 10.1017/S0029665116002883] [Medline: 27976605]
7. Zanchim MC, Kirsten VR, Marchi ACBD. Consumption of dietary intake markers by patients with diabetes assessed using a mobile application. Cien Saude Colet 2018 Dec;23(12):4199-4208 [FREE Full text] [doi: 10.1590/1413-812320182312.01412017] [Medline: 30540003]
8. Paim J, Travassos C, Almeida C, Bahia L, Macinko J. The Brazilian health system: history, advances, and challenges. Lancet 2011 May 21;377(9779):1778-1797. [doi: 10.1016/S0140-6736(11)60054-8] [Medline: 21561655]
9. Debon R, Bellei EA, Biduski D, Volpi SS, Alves ALS, Portella MR, et al. Effects of using a mobile health application on the health conditions of patients with arterial hypertension: A pilot trial in the context of Brazil's Family Health Strategy. Sci Rep 2020 Apr 07;10(1):6009 [FREE Full text] [doi: 10.1038/s41598-020-63057-w] [Medline: 32265476]
10. Volpi SS, Biduski D, Bellei EA, Tefili D, McCleary L, Alves ALS, et al. Using a mobile health app to improve patients' adherence to hypertension treatment: a non-randomized clinical trial. PeerJ 2021;9:e11491. [doi: 10.7717/peerj.11491] [Medline: 34123593]
11. Sauceda A, Frederico C, Pellechia K, Starin D. Results of the Academy of Nutrition and Dietetics' Consumer Health Informatics Work Group's 2015 Member App Technology Survey. J Acad Nutr Diet 2016 Aug;116(8):1336-1338. [doi: 10.1016/j.jand.2016.04.009] [Medline: 27236643]

XSL•FO
RenderX

12. Bellei EA, Biduski D, Lisboa HRK, De Marchi ACB. Development and Assessment of a Mobile Health Application for Monitoring the Linkage Among Treatment Factors of Type 1 Diabetes Mellitus. Telemed J E Health 2020 Feb;26(2):205-217. [doi: 10.1089/tmj.2018.0329] [Medline: 30724717]
13. Cechetti NP, Bellei EA, Biduski D, Rodriguez JPM, Roman MK, De Marchi ACB. Developing and implementing a gamification method to improve user engagement: A case study with an m-Health application for hypertension monitoring. Telematics and Informatics 2019 Aug;41:126-138. [doi: 10.1016/j.tele.2019.04.007]
14. Cade JE, Warthon-Medina M, Albar S, Alwan NA, Ness A, Roe M, DIET@NET consortium. DIET@NET: Best Practice Guidelines for dietary assessment in health research. BMC Med 2017 Nov 15;15(1):202 [FREE Full text] [doi: 10.1186/s12916-017-0962-x] [Medline: 29137630]
15. Brucki SMD, Nitrini R, Caramelli P, Bertolucci PHF, Okamoto IH. [Suggestions for utilization of the mini-mental state examination in Brazil]. Arq Neuropsiquiatr 2003 Sep;61(3B):777-781 [FREE Full text] [doi: 10.1590/s0004-282x2003000500014] [Medline: 14595482]
16. Ministério da Saúde. Protocolos do Sistema de Vigilância Alimentar e Nutricional - SISVAN na assistência à saúde. Brasilia: Ministério da Saúde; 2008. URL: https://aps.saude.gov.br/biblioteca/visualizar/MTI4Ng==
17. Ministério da Saúde. Orientações para coleta e análise de dados antropométricos em serviços de saúde: norma técnica do sistema de Vigilância Alimentar e Nutricional - SISVAN. Brasília: Ministério da Saúde; 2011. URL: https://aps.saude.gov.br/biblioteca/visualizar/MTI4MQ==
18. Multicenter survey aging, health and wellbeing in Latin America and the Caribbean (SABE): preliminary report. Pan American Health Organization. 2001. URL: https://iris.paho.org/handle/10665.2/45879 [accessed 2021-07-24]
19. Physical status: the use of and interpretation of anthropometry, report of a WHO expert committee. World Health Organization. 1995. URL: https://apps.who.int/iris/handle/10665/37003 [accessed 2021-07-24]
20. American Diabetes Association. Standards of Medical Care in Diabetes—2019: Abridged for Primary Care Providers. Clin Diabetes 2019 Jan;37(1):11-34 [FREE Full text] [doi: 10.2337/cd18-0105] [Medline: 30705493]
21. Brooke J. In: Jordan PW, Thomas B, Weerdmaster BA, McClelland IL, editors. SUS: a quick and dirty usability scale. London, England: Taylor & Francis; 1996:189-194.
22. Davis FD, Bagozzi RP, Warshaw PR. User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. Management Science 1989 Aug;35(8):982-1003. [doi: 10.1287/mnsc.35.8.982]
23. Minayo MCDS, Gomes SFDR. Pesquisa Social, Teoria, Metodo e Criatividade. Petrópolis: Editora Vozes; 1993.
24. Osonoi Y, Mita T, Osonoi T, Saito M, Tamasawa A, Nakayama S, et al. Relationship between dietary patterns and risk factors for cardiovascular disease in patients with type 2 diabetes mellitus: a cross-sectional study. Nutr J 2016 Feb 04;15:15 [FREE Full text] [doi: 10.1186/s12937-016-0132-6] [Medline: 26847556]
25. Vitoi NC, Fogal AS, Nascimento CDM, Franceschini SDCC, Ribeiro AQ. Prevalence and associated factors of diabetes in the elderly population in Viçosa, Minas Gerais, Brazil. Rev Bras Epidemiol 2015;18(4):953-965 [FREE Full text] [doi: 10.1590/1980-5497201500040022] [Medline: 26982308]
26. Linde JA, Utter J, Jeffery RW, Sherwood NE, Pronk NP, Boyle RG. Specific food intake, fat and fiber intake, and behavioral correlates of BMI among overweight and obese members of a managed care organization. Int J Behav Nutr Phys Act 2006 Nov 26;3:42 [FREE Full text] [doi: 10.1186/1479-5868-3-42] [Medline: 17125525]
27. Cahill LE, Pan A, Chiuve SE, Sun Q, Willett WC, Hu FB, et al. Fried-food consumption and risk of type 2 diabetes and coronary artery disease: a prospective study in 2 cohorts of US women and men. Am J Clin Nutr 2014 Aug;100(2):667-675 [FREE Full text] [doi: 10.3945/ajcn.114.084129] [Medline: 24944061]
28. Sayon-Orea C, Martinez-Gonzalez MA, Gea A, Flores-Gomez E, Basterra-Gortari FJ, Bes-Rastrollo M. Consumption of fried foods and risk of metabolic syndrome: the SUN cohort study. Clin Nutr 2014 Jun;33(3):545-549. [doi: 10.1016/j.clnu.2013.07.014] [Medline: 23954218]
29. Özen AE, Bibiloni MDM, Bouzas C, Pons A, Tur JA. Beverage Consumption among Adults in the Balearic Islands: Association with Total Water and Energy Intake. Nutrients 2018 Aug 23;10(9):1 [FREE Full text] [doi: 10.3390/nu10091149] [Medline: 30142929]
30. Ejtahed HS, Bahadoran Z, Mirmiran P, Azizi F. Sugar-Sweetened Beverage Consumption Is Associated with Metabolic Syndrome in Iranian Adults: Tehran Lipid and Glucose Study. Endocrinol Metab (Seoul) 2015 Sep;30(3):334-342 [FREE Full text] [doi: 10.3803/EnM.2015.30.3.334] [Medline: 26435135]
31. Farrell P, Negin J, Awoke M, Thow AM, Taua M, Faumuina T, et al. Associations between sociodemographic and behaviour factors, and dietary risk factors for overweight and obesity, in Samoan women. Appetite 2019 Mar 01;134:155-161. [doi: 10.1016/j.appet.2018.12.037] [Medline: 30593836]
32. Ferretto LR, Bellei EA, Biduski D, Bin LCP, Moro MM, Cervi CR, et al. A Physical Activity Recommender System for Patients With Arterial Hypertension. IEEE Access 2020;8:61656-61664. [doi: 10.1109/access.2020.2983564]
33. Tan SL, Storm V, Reinwand DA, Wienert J, de Vries H, Lippke S. Understanding the Positive Associations of Sleep, Physical Activity, Fruit and Vegetable Intake as Predictors of Quality of Life and Subjective Health Across Age Groups: A Theory Based, Cross-Sectional Web-Based Study. Front Psychol 2018;9:977 [FREE Full text] [doi: 10.3389/fpsyg.2018.00977] [Medline: 29967588]

XSL•FO
RenderX

34. Alakaam AAH, Lemacks JL. Fruit and Vegetable Consumption, Fat Intake, and Physical Activity Participation in Relation to Socio-demographic Factors Among Medically Underserved Adults. AIMS Public Health 2015;2(3):402-410 [FREE Full text] [doi: 10.3934/publichealth.2015.3.402] [Medline: 29546116]
35. Sebastian RS, Wilkinson Enns C, Goldman JD, Hoy MK, Moshfegh AJ. Findings from What We Eat in America, National Health and Nutrition Examination Survey 2011-2014 support salad consumption as an effective strategy for improving adherence to dietary recommendations. Public Health Nutr 2019 Apr;22(6):976-987. [doi: 10.1017/S1368980018003695] [Medline: 30767843]
36. Rimm EB, Ascherio A, Giovannucci E, Spiegelman D, Stampfer MJ, Willett WC. Vegetable, fruit, and cereal fiber intake and risk of coronary heart disease among men. JAMA 1996 Feb 14;275(6):447-451. [doi: 10.1001/jama.1996.03530300031036] [Medline: 8627965]
37. Miedema MD, Petrone A, Shikany JM, Greenland P, Lewis CE, Pletcher MJ, et al. Association of Fruit and Vegetable Consumption During Early Adulthood With the Prevalence of Coronary Artery Calcium After 20 Years of Follow-Up: The Coronary Artery Risk Development in Young Adults (CARDIA) Study. Circulation 2015 Nov 24;132(21):1990-1998 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.114.012562] [Medline: 26503880]
38. Ley SH, Hamdy O, Mohan V, Hu FB. Prevention and management of type 2 diabetes: dietary components and nutritional strategies. The Lancet 2014 Jun;383(9933):1999-2007. [doi: 10.1016/S0140-6736(14)60613-9]
39. Rangan AM, Tieleman L, Louie JCY, Tang LM, Hebden L, Roy R, et al. Electronic Dietary Intake Assessment (e-DIA): relative validity of a mobile phone application to measure intake of food groups. Br J Nutr 2016 Jun;115(12):2219-2226. [doi: 10.1017/S0007114516001525] [Medline: 27121045]
40. Bucher Della Torre S, Carrard I, Farina E, Danuser B, Kruseman M. Development and Evaluation of e-CA, an Electronic Mobile-Based Food Record. Nutrients 2017 Jan 18;9(1):1 [FREE Full text] [doi: 10.3390/nu9010076] [Medline: 28106767]
41. Maunder K, Walton K, Williams P, Ferguson M, Beck E. eHealth readiness of dietitians. J Hum Nutr Diet 2018 Aug;31(4):573-583. [doi: 10.1111/jhn.12542] [Medline: 29473238]
42. De Marchi ACB, Alves ALSA, Gonçalves CBC, Cervi CR, Biduski D, Bellei EA, et al. An Electronic Health Platform for Monitoring Health Conditions of Patients With Hypertension in the Brazilian Public Health System: Protocol for a Nonrandomized Controlled Trial. JMIR Res Protoc 2020 Jan 20;9(1):e15299 [FREE Full text] [doi: 10.2196/15299] [Medline: 31958068]
43. Roman MK, Bellei EA, Biduski D, Pasqualotti A, De Araujo CDSR, De Marchi ACB. "Hey assistant, how can I become a donor?" The case of a conversational agent designed to engage people in blood donation. J Biomed Inform 2020 Jul;107:103461 [FREE Full text] [doi: 10.1016/j.jbi.2020.103461] [Medline: 32504669]

Abbreviations

CAPES: Coordination and Improvement of Higher Level or Education Personnel
CNPq: National Council for Scientific and Technological Development
FAPERGS: State Funding Agency of Rio Grande do Sul
HbA_{1c}: glycated hemoglobin
SISVAN: Food and Nutrition Surveillance System of Brazil
T2DM: type 2 diabetes mellitus
TAM: Technology Acceptance Model

Edited by G Eysenbach; submitted 30.01.21; peer-reviewed by M Mauriello, Z Zhang, M Friehs; comments to author 31.03.21; revised version received 23.05.21; accepted 05.07.21; published 27.08.21.

Please cite as:
Coleone JD, Bellei EA, Roman MK, Kirsten VR, De Marchi ACB
Dietary Intake and Health Status of Elderly Patients With Type 2 Diabetes Mellitus: Cross-sectional Study Using a Mobile App in Primary Care
JMIR Form Res 2021;5(8):e27454
URL: https://formative.jmir.org/2021/8/e27454
doi:10.2196/27454
PMID:34448711

©Joane Diomara Coleone, Ericles Andrei Bellei, Mateus Klein Roman, Vanessa Ramos Kirsten, Ana Carolina Bertoletti De Marchi. Originally published in JMIR Formative Research (https://formative.jmir.org), 27.08.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR

XSL•FO
RenderX

Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Evaluation and Refinement of a Bank of SMS Text Messages to Promote Behavior Change Adherence Following a Diabetes Prevention Program: Survey Study

Megan MacPherson[1], MSc; Kaela Cranston[1], MSc; Cara Johnston[1]; Sean Locke[2], PhD; Mary E Jung[1], PhD

[1]School of Health and Exercise Sciences, Faculty of Health and Social Development, University of British Columbia, Kelowna, BC, Canada
[2]Department of Kinesiology, Brock University, St Catharines, ON, Canada

Corresponding Author:
Mary E Jung, PhD
School of Health and Exercise Sciences
Faculty of Health and Social Development
University of British Columbia
1147 Research Road
Kelowna, BC, V1V 1V7
Canada
Phone: 1 250 807 9670
Fax: 1 250 807 8085
Email: mary.jung@ubc.ca

Abstract

Background: SMS text messaging is a low-cost and far-reaching modality that can be used to augment existing diabetes prevention programs and improve long-term diet and exercise behavior change adherence. To date, little research has been published regarding the process of SMS text message content development. Understanding how interventions are developed is necessary to evaluate their evidence base and to guide the implementation of effective and scalable mobile health interventions in public health initiatives and in future research.

Objective: This study aims to describe the development and refinement of a bank of SMS text messages targeting diet and exercise behavior change to be implemented following a diabetes prevention program.

Methods: A bank of 124 theory-based SMS text messages was developed using the Behaviour Change Wheel and linked to active intervention components (behavior change techniques [BCTs]). The Behaviour Change Wheel is a theory-based framework that provides structure to intervention development and can guide the use of evidence-based practices in behavior change interventions. Once the messages were written, 18 individuals who either participated in a diabetes prevention program or were a diabetes prevention coach evaluated the messages on their clarity, utility, and relevance via survey using a 5-point Likert scale. Messages were refined according to participant feedback and recoded to obtain an accurate representation of BCTs in the final bank.

Results: 76/124 (61.3%) messages were edited, 4/124 (3.2%) were added, and 8/124 (6.5%) were removed based on participant scores and feedback. Of the edited messages, 43/76 (57%) received minor word choice and grammar alterations while retaining their original BCT code; the remaining 43% (33/76, plus the 4 newly written messages) were recoded by a reviewer trained in BCT identification.

Conclusions: This study outlines the process used to develop and refine a bank of SMS text messages to be implemented following a diabetes prevention program. This resulted in a bank of 120 theory-based, user-informed SMS text messages that were overall deemed clear, useful, and relevant by both individuals who will be receiving and delivering them. This formative development process can be used as a blueprint in future SMS text messaging development to ensure that message content is representative of the evidence base and is also grounded in theory and evaluated by key knowledge users.

(JMIR Form Res 2021;5(8):e28163) doi:10.2196/28163

XSL•FO
RenderX

KEYWORDS

text messaging; prediabetic state; telemedicine; telecommunications; exercise; diet; preventive medicine; mHealth; intervention development; behavior change; mobile phone

Introduction

Background

The global prevalence of prediabetes is 375 million [1], which increases the likelihood of developing type 2 diabetes (T2D) by up to 70% [2]. Dietary and physical activity behavior change programs have been shown to prevent or delay the progression of prediabetes to T2D [3-5]. A recent meta-analysis that assessed the effects of mobile health (mHealth) diabetes prevention programs (including videos, web-based resources, videoconferencing, phone calls, and SMS text messaging) found that the use of technology-based diabetes prevention strategies led to clinically significant weight loss in individuals at risk of developing T2D [6].

With more than 5 billion mobile phone users worldwide [7] and more than 18 billion SMS text messages sent daily [8], SMS text messaging is one of the widest reaching mHealth interventions [7,9] and has been shown to improve dietary and physical activity behavior change [10-13]. Within diabetes care, text messaging has been shown to improve participants' self-awareness, knowledge, and control over their condition and has been deemed a suitable technology for diabetes management; however, more research is needed to explore which message design features may be most effective in changing behaviors [14].

It is widely accepted that behavior change interventions should be developed using theory, past evidence, and formative research [15,16] and that they report sufficient detail regarding the intervention components [17,18]; however, many mHealth interventions lack rigor, theory, and the inclusion of experts in content development [6]. More specifically, few SMS text messaging interventions detail formative development and evaluation and typically lack information pertaining to content development [19-22], resulting in text messaging interventions with insufficient details to allow for replication or evaluation.

Prior SMS Text Message Development Research

Within the diabetes prevention literature, many SMS text messaging interventions do not provide information on how messages are developed. When formative development is described, messages are generally based on previous literature or health authority recommendations and are developed by a team of experts including clinicians, researchers, and individuals at risk of developing T2D [23-31]. Following message development by content experts, some studies take an additional step and have end users evaluate the messages to refine content based on key stakeholder preferences [23,25,26,29,32-34]. For example, O'Reilly and Laws [34] had participants rank messages as useful, unsure, or not acceptable and then conducted focus groups to improve messages rated as not acceptable. The use of theory within diabetes prevention messaging content is sparse, and only one study to date has linked their messages to behavior change techniques (BCTs) [23].

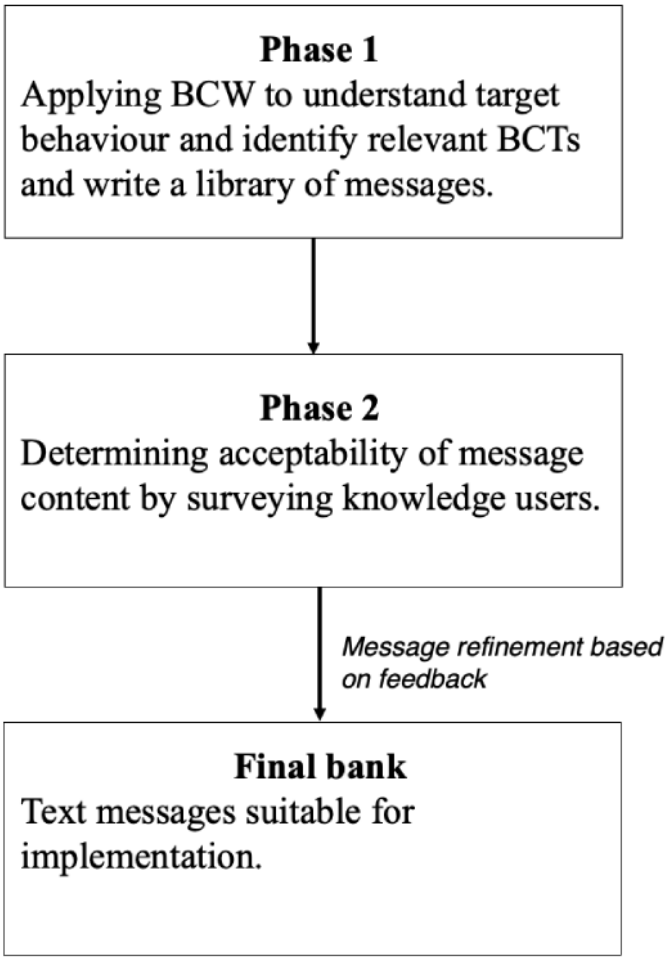
Although not in the field of diabetes prevention, Chai et al [35] used the Theoretical Domains Framework (TDF) and Behaviour Change Wheel (BCW) to integrate theory more effectively into SMS text messaging. The TDF was developed as a synthesis of 33 theories and 128 behavior change constructs combined into 14 distinct domains [36]. Following the identification of relevant domains within the TDF, Chai et al [35] used the BCW to identify relevant intervention functions and had a team of experts develop messages to map onto the three domains and four intervention functions they felt were relevant to their bank of messages. In phase 2 of their development, messages were reviewed by end users for message clarity, usefulness, and relevance. Once evaluated, the messages were refined based on participant responses before implementation and effectiveness testing.

This Work

This study aims to outline the evaluation and refinement of a bank of SMS text messages designed for individuals at risk for T2D who have taken part in the Small Steps for Big Changes (SSBC) diabetes prevention program [37,38]. The iterative design used in the current text messaging content development included two broad phases: (1) application of the BCW framework to develop a bank of messages and (2) message evaluation and refinement by those representing SMS text message recipients and senders (Figure 1). Phase 1 has already been published (MacPherson et al [39]) but will be briefly described.

XSL•FO
RenderX

Figure 1. Text messaging content development and evaluation process. BCT: behavior change technique; BCW: Behaviour Change Wheel.



The BCW was developed using expert consensus and validation and has been used to design a myriad of behavior change interventions [40-44]. The BCW allows for the systematic and structured development of complex behavioral interventions and can improve subsequent implementation and evaluation through the selection of BCTs to target specific barriers and facilitators to engage in a behavior [45]. BCTs are the building blocks or *active components* within an intervention designed to modify a behavior. When developing the taxonomy of BCTs, Michie et al [45] identified 93 distinct BCTs, which can be grouped together in 16 different BCT categories.

During the message development phase, the research team identified the target behaviors of dietary and physical activity behavior change adherence. Barriers and facilitators to engage in these behaviors were identified through previous qualitative research conducted with individuals following their participation in the SSBC program [46]. Relevant BCTs were identified through systematic reviews addressing BCTs within diet or physical activity interventions for populations with or at risk of developing T2D [47-54] and BCTs currently being used within the SSBC diabetes prevention program [55]. All identified BCTs were assessed for inclusion by 2 members of the research team using the acceptability, practicability, effectiveness and cost-effectiveness, affordability, safety and side-effects, equity (APEASE) criteria [56]. Subsequently, messages were written pursuant to the included BCTs. This study identified 43 BCTs and 124 messages targeting diet and physical activity behavior change in individuals at risk of developing T2D. To improve intervention affordability and equity, messages were developed as one way messages, were limited to 160 characters (the maximum number of characters for a single SMS text message for most carriers—characters include but are not limited to numbers, letters, and spaces), and were at a less than an eighth-grade reading level.

Methods

Recruitment

A total of 25 potential knowledge users—5/25 (20%) diabetes prevention coaches and 20/25 (80%) past SSBC clients who consented to be contacted for future research studies—were contacted via email to participate in this study. Consenting individuals were asked to complete a web-based survey, which took approximately 60 minutes for diabetes prevention coaches (who were asked to review the entire message bank) and 20 minutes for past SSBC clients (who were asked to evaluate a

XSL•FO
RenderX

random subset of approximately one-third of messages). Participants were compensated for their time with an electronic gift card in the amount of Can $40 (US $32.14) and Can $20 (US $16.07) for diabetes prevention coaches and SSBC clients, respectively.

Procedures

Surveys were administered through a web-based survey platform (Qualtrics) in which study participants were asked to review messages on (1) readability or clarity, (2) usefulness, and (3) relevance by rating each message on a 5-point Likert scale (1=strongly disagree and 5=strongly agree). This process of message evaluation was used by Chai et al [35] and adapted from Redfern [57]. These questions were chosen as they target message acceptability, a key criterion within the APEASE criteria in assessing intervention content and delivery. This was built on the previous research stage, which assessed APEASE from the research team's perspective, not end users' perspective. For each message, following the quantitative questions, participants were provided with an open text box to provide additional feedback or suggestions regarding the message content.

At the conclusion of the survey, diabetes prevention coaches were asked how many months they would be willing to send messages (similar to the ones they reviewed) to their participants and how many per week they would be willing to send if they were required to do so without any automation. SSBC clients were similarly asked about the number of months they would like to receive messages and how many days per week they would find messages helpful.

Feedback regarding the prototype bank of messages was summarized, and messages were modified once Likert scale responses were addressed. A total sum score was created for each message by combining scores for all three questions (clarity, usefulness, and relevance), resulting in total sum scores per message ranging from 3 (if participants strongly disagreed for all questions) to 15 (participants strongly agreed for all questions). Before analysis, it was decided that any message with an average total sum score above 14 across all participants was to be retained with the option for revisions based on participant feedback. Total sum scores above 14 out of 15 would indicate that most end users rated the messages with a perfect score. Any message with a total score of ≤14 was to be refined to ensure that all messages to be used in the SSBC diabetes prevention program were highly rated by those sending and receiving the messages. This cutoff was chosen a priori to limit the impact that potential social desirability bias had on which messages were refined. Refinement of messages began with addressing any participant feedback; if a message with a score of ≤14 did not have any feedback, they were refined based on quantitative scores (eg, if a message had a lower readability score, it was altered to improve the clarity in which it was written). To ensure that all messages were accurately linked with BCTs, any revised messages were recoded by an independent reviewer for BCTs.

Results

Overview

A total of 18 adults—5/18 (28%) diabetes prevention coaches and 13/18 (72%) past SSBC clients—participated in the evaluation (7 females, 9 males, and 2 missing; mean age 47 years, SD 15), and an additional 7 past SSBC clients were contacted but did not participate in the survey. Participants were asked to provide demographic information including sex, age, ethnicity, highest level of education, occupation, and marital status (Table 1).

XSL•FO
RenderX

Table 1. Descriptive statistics for individuals that participated in the intervention.

Characteristics	All (N=18)	Diabetes prevention coaches (n=5)	Past Small Steps for Big Changes clients (n=13)
Age (years), mean (SD)	47.16 (15.21)	29.60 (4.72)	59.58 (6.32)
Sex, n (%)			
Male	9 (50)	2 (40)	7 (54)
Female	7 (39)	3 (60)	4 (31)
Did not answer	2 (11)	0 (0)	2 (15)
Ethnic origin, n (%)			
White	13 (72)	3 (60)	10 (77)
Latin American	1 (6)	1 (20)	0 (0)
Asian	2 (11)	1 (20)	1 (8)
Indigenous	2 (11)	0 (0)	2 (15)
Annual income (Can $ [US $]), n (%)			
0-24,999 (0-19,987.58)	1 (6)	1 (20)	0 (0)
25,000-49,999 (19,988.38-39975.95)	1 (6)	1 (20)	0 (0)
50,000-74,999 (39,976.75-59,964.33)	4 (2)	1 (20)	3 (23)
75,000-99,999 (59,965.13-79,952.70)	4 (2)	1 (20)	3 (23)
≥100,000 (≥79,953.50)	7 (39)	1 (20)	6 (46)
Missing	1 (6)	0 (0)	1 (8)
Education, n (%)			
High school	2 (17)	0 (0)	2 (15)
College diploma	6 (33)	0 (0)	6 (46)
University degree	4 (22)	2 (40)	2 (15)
Postgraduate degree	6 (33)	3 (60)	3 (23)
Marital status, n (%)			
Single	4 (22)	3 (60)	1 (8)
Married	8 (44)	0 (0)	8 (62)
Common law	5 (28)	2 (40)	3 (23)
Divorced	1 (6)	0 (0)	1 (8)

Message Evaluation and Refinement

A total of 124 messages were included in the initial message bank. Each message was evaluated between 7 and 12 times, and messages received a total score of 13.77 (SD 0.76), with diabetes prevention coaches scoring messages higher than past SSBC clients (Table 2 provides scores on readability, relevance, and usefulness). On average, each message received approximately two comments (ranging from 0 to 5 comments per message).

Table 2. Mean scores for the message evaluations for diabetes prevention coaches, past Small Steps for Big Changes participants, and all participants combined.

Evaluation	All (N=18), mean (SD)	Diabetes prevention coaches (n=5), mean (SD)	Past Small Steps for Big Changes clients (n=13), mean (SD)
Total score (out of 15)	13.77 (0.76)	14.11 (0.97)	13.30 (0.92)
Readability (out of 5)	4.60 (0.30)	4.68 (0.39)	4.48 (0.30)
Relevance (out of 5)	4.60 (0.25)	4.73 (0.33)	4.40 (0.34)
Usefulness (out of 5)	4.58 (0.27)	4.70 (0.33)	4.41 (0.36)

A total of 65 messages received a score of ≥14 (Multimedia Appendix 1 [39,45] provides more information on individual message scores and BCTs). These highly rated messages primarily fall within the following BCT categories: *goals and*

XSL•FO
RenderX

planning (16/65, 25%), *self-belief* (11/65, 17%), *repetition and substitution* (10/65, 15%), and *feedback and monitoring* (7/65, 11%). The remaining 59 messages were scored <14 and fell within the BCT categories: *self-belief* (10/59, 17%), *reward and threat* (9/59, 15%), *goals and planning* (8/59, 13%), and *identity* (7/59, 12%).

A total of 76 messages were edited (76/124, 61% of the initial message bank): 59 received a total score of ≤14 and were refined based on participant feedback and individual quantitative scores; 20 messages received a total score of >14 but had suggestions for improvement and were refined based on participant feedback. An additional 4 messages were added to the library based on participant suggestions, and 8 messages were removed as they received consistent comments regarding their applicability to the program by both diabetes prevention coaches and past SSBC clients. Of the 76 messages edited, 43/76 (57%) were altered based on minor word choice and grammatical errors and therefore retained their original BCT codes; the remaining 43% (33/76; plus the additional 4 messages that were written) were coded by a reviewer trained in BCT identification (Multimedia Appendix 2 [39,45] provides final message bank and corresponding BCTs).

The edited bank of messages (Multimedia Appendix 2) includes 120 messages based on 41 distinct BCTs. The BCT categories most used included *goals and planning* (26/120, 21.7%), *self-belief* (22/120, 18.3%), *repetition and substitution* (16/120, 13.3%), *social support* (11/120, 9.2%), and *natural consequences* (11/120, 9.2%). Table 3 provides information on all BCT categories within both the initial and edited message banks.

Table 3. Number of messages (%) within the initial and edited message bank within each of behavior change technique category.[a]

	Initial message bank scored ≥14 (n=65), n (%)	Initial message bank scored <14 (n=59), n (%)	Total initial message bank (n=124), n (%)	Edited message bank (n=120), n (%)
Goals and planning	16 (24.6)	8 (13.6)	24 (19.3)	26 (21.7)
Feedback and monitoring	7 (10.8)	1 (1.7)	8 (6.4)	6 (5)
Social support	4 (6.2)	6 (10.2)	10 (8.1)	11 (9.1)
Shaping knowledge	4 (6.2)	1 (1.7)	5 (4)	5 (4.1)
Natural consequences	6 (9.2)	3 (5.1)	9 (7.3)	11 (9.2)
Comparison of behavior	5 (7.7)	2 (3.4)	7 (5.7)	8 (6.7)
Associations	2 (3.1)	1 (1.7)	3 (2.4)	3 (2.5)
Repetition and substitution	10 (15.4)	6 (10.2)	16 (13)	15 (13)
Comparison of outcomes	4 (6.2)	5 (8.5)	9 (7.3)	10 (8.3)
Reward and threat	3 (4.6)	9 (15.3)	12 (9.7)	6 (5)
Regulation	2 (3.1)	0 (0)	2 (1.6)	3 (2.5)
Antecedents	4 (6.2)	3 (5.1)	7 (5.7)	5 (4.2)
Identity	1 (1.5)	7 (11.9)	8 (6.5)	8 (6.7)
Scheduled consequences	2 (3.1)	6 (10.2)	8 (6.5)	5 (4.2)
Self-belief	11 (16.9)	10 (16.9)	21 (16.9)	22 (18.3)
Covert learning	0 (0)	1 (1.7)	1 (0.8)	1 (0.8)

[a]Note that some messages had multiple behavior change techniques and thus may fall within more than one behavior change technique category.

Diabetes prevention coaches noted that if they were required to send similar messages to their own participants, they would be willing to send them for an average of 5 months (mode 1, median 3, range 1-12) and for an average of 3 messages per week (mode 3, median 3, range 1-5). When SSBC clients were asked about their preference for receiving messages, they indicated, on average, that they would prefer to receive messages for 7 months following program completion (mode 12, median 6, range 2-12) with 3 messages per week (mode 2, median 2, range 1-5).

Discussion

Principal Findings

The primary objective of this study was to outline the development, evaluation, and refinement process used to develop a bank of SMS text messages targeting health behaviors among individuals at risk of developing T2D. Message content development and refinement involved a range of diabetes prevention coaches, researchers, and past SSBC clients who reviewed content and provided suggestions to improve message content and clarity. Overall, diabetes prevention coaches and individuals who participated in the SSBC diabetes prevention program identified that the bank of messages was clear, relevant, and useful for individuals at risk of developing T2D. A total of 52.4% (65/124) of messages received a total sum score of ≥14

XSL•FO
RenderX

of 15. Of these 65 highly rated messages, 20/65 (31%) were physical activity messages, 19/65 (29%) were dietary messages, and 26/65 (40%) were general messages. Of those 59 messages scoring <14 of 15, 21/59 (36%) were physical activity messages, 22/59 (37%) were dietary, and 16/59 (27%) were general. Overall, the entirety of the message bank was rated 13.77 of 15, indicating general acceptability of most messages; however, the highest proportion of message scores of ≥14 contained general content, not targeting only diet or physical activity. Future research should examine whether general messages (eg, "Your first plan will not work 100% of the time. Continue to change your goals until you find what works best for you!") versus targeted behavioral messages (eg, "Think about where, when and how you'll get your exercise in today!") are more effective in behavior change adherence following the SSBC program.

BCT Composition of the Message Bank

On the basis of total sum scores and participant feedback, 4 messages were added, 8 messages were removed, and 76 messages were edited based on knowledge user feedback, resulting in 120 theory-based, user-informed SMS text messages to be used following the SSBC diabetes prevention program. All 8 messages that were removed consisted of BCTs within the category *reward and threat*, for example, the lowest scoring message reads: "Did you put effort into meeting your exercise goals this week? Good for you! Don't forget to reward yourself." Messages similar to this one received consistent comments that this language and the use of external rewards were inconsistent with the program. The counseling portion of SSBC is informed by motivational interviewing, a collaborative communication style that draws on an individual's own motivations and commitment to change. To stay consistent with the aims of SSBC to facilitate an individual's autonomous motivation, messages including a focus on external motivators and rewards were removed. The final bank of messages includes either highly rated messages or messages that were tailored based on participant suggestions and relate to all 16 BCT categories. More than 70.8% (85/120) of this final message bank falls within five BCT categories: *goals and planning*, *self-belief*, *repetition and substitution*, *social support*, and *natural consequences*. Future research using this bank of messages can manipulate the BCT categories that are being used to provide additional experimental evidence for the categories that may be more or less effective in behavior change for individuals at risk of developing T2D.

Although BCTs within the categories of *goals and planning, feedback, and monitoring* have been cited as the most commonly used within physical activity behavior change interventions [58], there is not a large evidence base for which specific BCT categories or combinations of BCT categories may optimally influence long-term behavior change adherence because of lack of experimental evidence for different BCT categories. For example, Howlett et al [58] recently conducted a systematic review assessing BCTs within physical activity randomized controlled trials and found that studies including *biofeedback, demonstration of behavior, graded tasks, action planning, instruction on how to perform the behavior, prompts/cues, behavior practice*, and *self-reward* showed larger effect sizes compared with studies that did not. However, no BCTs within categories 11-16 (*regulation, antecedents, identity, scheduled consequences, self-belief*, or *overt learning*) were included in these analyses. The authors identified a total of 240 BCTs within the 26 included studies, with <3% corresponding to BCTs within the categories 11-16, thereby limiting researchers' knowledge of whether or how these techniques may influence physical activity behavior change. Furthermore, researchers have shown that BCTs that are most frequently used in digital interventions may not adequately address user needs.

Asbjørnsen et al [59,60] provided a comprehensive analysis of both BCTs being consistently used, and the BCTs identified as relevant and meaningful to stakeholders to promote weight loss maintenance within digital interventions. The authors started by conducting a review that identified *feedback and monitoring, goals and planning, social support, shaping knowledge, associations*, and *repetition and substitution* as consistently applied BCT categories within digital interventions targeting weight loss motivation, adherence, and maintenance [60]. Following this, Asbjørnsen et al [59] conducted individual interviews and focus groups with end users to identify values and needs, as they relate to digital interventions targeting weight loss maintenance. On the basis of key values identified by end users, BCT categories identified as potentially relevant for digital interventions to support long-term behavior change maintenance included *goals and planning, feedback and monitoring, social support, self-belief, natural consequences*, and *identity*. This work highlights that although there is some overlap, BCTs that researchers are using do not necessarily overlap with what end users want or need to help them maintain long-term behavior change.

Although the primary BCT categories within the current message bank do not correspond to those that are commonly used in physical activity interventions [58], the systematic development and inclusion of key stakeholders may have resulted in more tailored BCT categories emerging as the most used. The inclusion of key stakeholders has been highlighted as an imperative step in designing digital interventions, as it allows the resulting intervention to reflect the values and support the goals identified by the end users [61,62].

Limitations and Future Research

This paper outlines the iterative development of the content of SMS text messages; however, there is still an insufficient amount of research examining when or how frequently these messages should be sent and which specific messages or techniques are more effective than others. Although the inclusion of BCT mapping and knowledge users in the development and refinement of the bank of messages is a strength of this study, it is not yet known whether this bank of messages is effective in improving adherence to diet and physical activity recommendations following the SSBC diabetes prevention program. Future research should be conducted to identify further adaptations and refinements of the current bank of messages to facilitate long-term implementation. Such work could provide an in-depth understanding of *how* participants engage with the messages over time, which BCTs and specific content are most

XSL•FO
RenderX

or least helpful, and how SMS text messages can influence behavior change over time.

Another limitation of this work is that content was tailored for deployment in SMS text messages, but study recruitment and assessment of the message bank was conducted via email and a web-based survey. Furthermore, as smartphones and associated apps for communication (eg, WhatsApp and Facebook Messenger) are becoming increasingly popular and traditional text messaging may become less accepted over time, possibly limiting the reach and usability of the current research program. That being said, the risk of T2D increases with age [63], and older adults are increasingly using text messaging [64] but are still less likely to own a smartphone compared with their younger counterparts [65], making text messaging a viable option for a diabetes prevention program. Furthermore, the current bank of messages can be easily adapted in future iterations for delivery across a myriad of platforms to improve accessibility and reach for varying populations.

It is important to note that on average, past SSBC clients rated messages lower than diabetes prevention coaches (Table 2). The precise rationale behind these ratings is not known. This could be because the responses of past SSBC clients are based on the knowledge of lived experience with the program; they may have a more accurate idea of how future SSBC clients will interact with the message bank and which content is most or least helpful from a participant standpoint. In addition, because of the training that diabetes prevention coaches receive and their overall passion for the subject matter, it is notable that differences in knowledge may have contributed to the overall augmented ratings of the bank of messages. Future studies could benefit from the contributions of past SSBC clients during the text messaging content development process. Despite these considerations, the message bank was rated favorably by both past SSBC clients and diabetes prevention coaches and as a result is believed to be suitable for implementation in the SSBC program.

Finally, given the large spread in the number of months and days per week, coaches were willing to send messages and participants wanted to receive them (ranging from 1 to 12 months of messages for 1 to 5 days per week), more empirical evidence is needed to identify the duration and frequency of messages that optimally influence behaviors following the SSBC program. In addition, although our group intends to implement and test this message bank following completion of the SSBC program, the final message bank may also be suitable for prospective or current participants, and future research can assess the time point at which these messages are most effective.

Comparison With Previous Work

Many mobile phone text messaging interventions provide little to no detail regarding the processes used to develop their message content [19]. In many interventions, it is unclear whether content was developed ad hoc, if it was informed by behavior change theoretical frameworks, if content was reviewed by health experts, or whether feedback from stakeholders was sought to inform content. Although robust development of mobile behavior change interventions can be time consuming [66,67], this step is necessary to ensure that sufficient attention is placed on theoretical underpinnings and active intervention components, resulting in an mHealth intervention with a strong evidence base.

The process outlined in this paper includes the initial development of a bank of messages based on the BCW to link each message to a theoretical mechanism of action, followed by evaluation and refinement based on the opinions of researchers as well as those of past SSBC clients and diabetes prevention coaches. Broadly, the BCW has been used in the development of text messaging interventions [68], and the evaluation and refinement process has been successfully used in previous studies targeting parents and individuals with cardiovascular disease [35,57,69]. More specifically, the content of diabetes prevention SMS text messages typically includes the provision of information relating to diet and physical activity, reinforcement of key concepts following lessons, and reminders about goals and self-monitoring behavior; however, the authors cautioned the use of such mHealth interventions, as the marketplace is filled with products whose development lacks rigor, theory, and the inclusion of experts in content development [6].

Furthermore, to improve the quality and utility of mHealth interventions, it has been recommended that researchers specify not only theoretical mechanisms and provide explicit details on active intervention components [17,70] but that they also include knowledge users throughout the development process [16,71-74]. In this research program, message content was based on target barriers and facilitators identified by individuals who participated in the SSBC program [46], and content was further refined based on the preferences of those who represented those who would be receiving and sending the messages. Such collaborative approaches have been shown to improve the uptake of research into practice [75,76] and have been described as an integral part of program development to improve intervention relevance, impact, and efficiency [77].

Conclusions

Understanding how to optimally intervene via SMS text messages requires rigorous development and consistent stakeholder engagement, which is often underreported [19-22]. This study reports on the iterative development and refinement of a bank of SMS text messages that are suitable for use or further effectiveness testing following the SSBC diabetes prevention program. Furthermore, the resulting bank of SMS text messages from this research has been written in a nonprogram-specific manner and can be implemented or tailored to other programs wishing to implement a text messaging intervention to improve adherence to behavior change. The authors request that any researchers planning to use this bank of messages acknowledge this paper in addition to the development paper [39]. The two-phase method used in this study to develop useful, clear, and relevant SMS text messages will provide a practical framework that can be used not only for future diabetes prevention research but also for other behavior change interventions. Overall, this study resulted in a bank of messages deemed acceptable by those who delivered and potentially received them. These messages are based on evidence and behavior change theory and have been refined

XSL•FO
RenderX

based on feedback from those with lived experiences as diabetes prevention coaches and individuals who have participated in a diabetes prevention program.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Average and total sum scores per message.
[DOCX File , 34 KB - formative_v5i8e28163_app1.docx]

Multimedia Appendix 2
Final message bank and associated behavior change technique.
[DOCX File , 28 KB - formative_v5i8e28163_app2.docx]

References

1. Saeedi P, Petersohn I, Salpea P, Malanda B, Karuranga S, Unwin N, IDF Diabetes Atlas Committee. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9 edition. Diabetes Res Clin Pract 2019 Nov;157:107843. [doi: 10.1016/j.diabres.2019.107843] [Medline: 31518657]
2. Nathan DM, Davidson MB, DeFronzo RA, Heine RJ, Henry RR, Pratley R, American Diabetes Association. Impaired fasting glucose and impaired glucose tolerance: implications for care. Diabetes Care 2007 Mar;30(3):753-759. [doi: 10.2337/dc07-9920] [Medline: 17327355]
3. Diabetes Prevention Program Research Group, Knowler WC, Fowler SE, Hamman RF, Christophi CA, Hoffman HJ, et al. 10-year follow-up of diabetes incidence and weight loss in the Diabetes Prevention Program Outcomes Study. Lancet 2009 Nov 14;374(9702):1677-1686 [FREE Full text] [doi: 10.1016/S0140-6736(09)61457-4] [Medline: 19878986]
4. Li G, Zhang P, Wang J, Gregg EW, Yang W, Gong Q, et al. The long-term effect of lifestyle interventions to prevent diabetes in the China Da Qing Diabetes Prevention Study: a 20-year follow-up study. Lancet 2008 May 24;371(9626):1783-1789. [doi: 10.1016/S0140-6736(08)60766-7] [Medline: 18502303]
5. Lindström J, Louheranta A, Mannelin M, Rastas M, Salminen V, Eriksson J, et al. The Finnish Diabetes Prevention Study (DPS): Lifestyle intervention and 3-year results on diet and physical activity. Diabetes Care 2003 Dec;26(12):3230-3236. [Medline: 14633807]
6. Bian RR, Piatt GA, Sen A, Plegue MA, De Michele ML, Hafez D, et al. The Effect of Technology-Mediated Diabetes Prevention Interventions on Weight: A Meta-Analysis. J Med Internet Res 2017 Mar 27;19(3):e76 [FREE Full text] [doi: 10.2196/jmir.4709] [Medline: 28347972]
7. 7 TME2. The Mobile Economy 2019. The Mobile Economy 2019. URL: https://www.gsma.com/r/mobileeconomy/ [accessed 2020-08-23]
8. How Many Texts Do People Send Every Day?. URL: http://www.trtest.xyz/blog/how-many-texts-people-send-per-day/ [accessed 2020-08-23]
9. Cole-Lewis H, Kershaw T. Text messaging as a tool for behavior change in disease prevention and management. Epidemiol Rev 2010;32:56-69 [FREE Full text] [doi: 10.1093/epirev/mxq004] [Medline: 20354039]
10. Armanasco AA, Miller YD, Fjeldsoe BS, Marshall AL. Preventive Health Behavior Change Text Message Interventions: A Meta-analysis. Am J Prev Med 2017 Mar;52(3):391-402. [doi: 10.1016/j.amepre.2016.10.042] [Medline: 28073656]
11. De Leon E, Fuentes LW, Cohen JE. Characterizing periodic messaging interventions across health behaviors and media: systematic review. J Med Internet Res 2014;16(3):e93 [FREE Full text] [doi: 10.2196/jmir.2837] [Medline: 24667840]
12. Hall AK, Cole-Lewis H, Bernhardt JM. Mobile text messaging for health: a systematic review of reviews. Annu Rev Public Health 2015 Mar 18;36:393-415 [FREE Full text] [doi: 10.1146/annurev-publhealth-031914-122855] [Medline: 25785892]
13. Head KJ, Noar SM, Iannarino NT, Grant HN. Efficacy of text messaging-based interventions for health promotion: a meta-analysis. Soc Sci Med 2013 Nov;97:41-48. [doi: 10.1016/j.socscimed.2013.08.003] [Medline: 24161087]
14. Sahin C, Courtney KL, Naylor PJ, E Rhodes R. Tailored mobile text messaging interventions targeting type 2 diabetes self-management: A systematic review and a meta-analysis. Digit Health 2019;5:2055207619845279 [FREE Full text] [doi: 10.1177/2055207619845279] [Medline: 31041110]
15. Collins L. Optimization of Behavioral, Biobehavioral, and Biomedical Interventions: The Multiphase Optimization Strategy (MOST) (Statistics for Social and Behavioral Sciences). Gewerbestrasse 11, 6330 Cham, Switzerland: Springer; 2018.
16. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. Int J Nurs Stud 2013 May;50(5):587-592. [doi: 10.1016/j.ijnurstu.2012.09.010] [Medline: 23159157]

XSL•FO
RenderX

17. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014;348:g1687 [FREE Full text] [Medline: 24609605]
18. Michie S, Fixsen D, Grimshaw JM, Eccles MP. Specifying and reporting complex behaviour change interventions: the need for a scientific method. Implement Sci 2009;4:40 [FREE Full text] [doi: 10.1186/1748-5908-4-40] [Medline: 19607700]
19. Maar MA, Yeates K, Toth Z, Barron M, Boesch L, Hua-Stewart D, et al. Unpacking the Black Box: A Formative Research Approach to the Development of Theory-Driven, Evidence-Based, and Culturally Safe Text Messages in Mobile Health Interventions. JMIR Mhealth Uhealth 2016 Jan 22;4(1):e10 [FREE Full text] [doi: 10.2196/mhealth.4994] [Medline: 26800712]
20. Ricci-Cabello I, Bobrow K, Islam SMS, Chow CK, Maddison R, Whittaker R, et al. Examining Development Processes for Text Messaging Interventions to Prevent Cardiovascular Disease: Systematic Literature Review. JMIR Mhealth Uhealth 2019 Mar 29;7(3):e12191 [FREE Full text] [doi: 10.2196/12191] [Medline: 30924790]
21. Whittaker R, Merry S, Dorey E, Maddison R. A development and evaluation process for mHealth interventions: examples from New Zealand. J Health Commun 2012;17 Suppl 1:11-21. [doi: 10.1080/10810730.2011.649103] [Medline: 22548594]
22. Fitts WJ, Furberg R. Underdeveloped or underreported? Coverage of pretesting practices and recommendations for design of text message-based health behavior change interventions. J Health Commun 2015 Apr;20(4):472-478. [doi: 10.1080/10810730.2014.977468] [Medline: 25749250]
23. Morton K, Sutton S, Hardeman W, Troughton J, Yates T, Griffin S, et al. A Text-Messaging and Pedometer Program to Promote Physical Activity in People at High Risk of Type 2 Diabetes: The Development of the PROPELS Follow-On Support Program. JMIR Mhealth Uhealth 2015 Dec 15;3(4):e105 [FREE Full text] [doi: 10.2196/mhealth.5026] [Medline: 26678750]
24. MacPherson MM, Merry KJ, Locke SR, Jung ME. Effects of Mobile Health Prompts on Self-Monitoring and Exercise Behaviors Following a Diabetes Prevention Program: Secondary Analysis From a Randomized Controlled Trial. JMIR Mhealth Uhealth 2019 Sep 05;7(9):e12956 [FREE Full text] [doi: 10.2196/12956] [Medline: 31489842]
25. Pfammatter A, Spring B, Saligram N, Davé R, Gowda A, Blais L, et al. mHealth Intervention to Improve Diabetes Risk Behaviors in India: A Prospective, Parallel Group Cohort Study. J Med Internet Res 2016 Aug 05;18(8):e207 [FREE Full text] [doi: 10.2196/jmir.5712] [Medline: 27496271]
26. Abebe NA, Capozza KL, Des Jardins TR, Kulick DA, Rein AL, Schachter AA, et al. Considerations for community-based mHealth initiatives: insights from three Beacon Communities. J Med Internet Res 2013;15(10):e221 [FREE Full text] [doi: 10.2196/jmir.2803] [Medline: 24128406]
27. Nanditha A, Thomson H, Susairaj P, Srivanichakorn W, Oliver N, Godsland IF, et al. A pragmatic and scalable strategy using mobile technology to promote sustained lifestyle changes to prevent type 2 diabetes in India and the UK: a randomised controlled trial. Diabetologia 2020 Mar;63(3):486-496 [FREE Full text] [doi: 10.1007/s00125-019-05061-y] [Medline: 31919539]
28. Alzeidan R, Shata Z, Hassounah MM, Baghdadi LR, Hersi A, Fayed A, et al. Effectiveness of digital health using the transtheoretical model to prevent or delay type 2 diabetes in impaired glucose tolerance patients: protocol for a randomized control trial. BMC Public Health 2019 Nov 21;19(1):1550 [FREE Full text] [doi: 10.1186/s12889-019-7921-8] [Medline: 31752774]
29. Cheung NW, Blumenthal C, Smith BJ, Hogan R, Thiagalingam A, Redfern J, et al. A Pilot Randomised Controlled Trial of a Text Messaging Intervention with Customisation Using Linked Data from Wireless Wearable Activity Monitors to Improve Risk Factors Following Gestational Diabetes. Nutrients 2019 Mar 11;11(3):590 [FREE Full text] [doi: 10.3390/nu11030590] [Medline: 30862052]
30. Wong CKH, Fung CSC, Siu SC, Lo YYC, Wong KW, Fong DYT, et al. A short message service (SMS) intervention to prevent diabetes in Chinese professional drivers with pre-diabetes: a pilot single-blinded randomized controlled trial. Diabetes Res Clin Pract 2013 Dec;102(3):158-166. [doi: 10.1016/j.diabres.2013.10.002] [Medline: 24466598]
31. Catley D, Puoane T, Tsolekile L, Resnicow K, Fleming K, Hurley EA, et al. Adapting the Diabetes Prevention Program for low and middle-income countries: protocol for a cluster randomised trial to evaluate 'Lifestyle Africa'. BMJ Open 2019 Nov 11;9(11):e031400 [FREE Full text] [doi: 10.1136/bmjopen-2019-031400] [Medline: 31719084]
32. Ritchie ND, Gutiérrez-Raghunath S, Durfee MJ, Fischer H. Supplemental Text Message Support With the National Diabetes Prevention Program: Pragmatic Comparative Effectiveness Trial. JMIR Mhealth Uhealth 2020 Jun 18;8(6):e15478 [FREE Full text] [doi: 10.2196/15478] [Medline: 32554385]
33. Gupta Y, Kapoor D, Josyula LK, Praveen D, Naheed A, Desai AK, et al. A lifestyle intervention programme for the prevention of Type 2 diabetes mellitus among South Asian women with gestational diabetes mellitus [LIVING study]: protocol for a randomized trial. Diabet Med 2019 Feb;36(2):243-251. [doi: 10.1111/dme.13850] [Medline: 30368898]
34. O'Reilly SL, Laws R. Health-e mums: Evaluating a smartphone app design for diabetes prevention in women with previous gestational diabetes. Nutr Diet 2019 Nov;76(5):507-514. [doi: 10.1111/1747-0080.12461] [Medline: 30109762]
35. Chai LK, May C, Collins CE, Burrows TL. Development of text messages targeting healthy eating for children in the context of parenting partnerships. Nutr Diet 2019 Nov;76(5):515-520. [doi: 10.1111/1747-0080.12498] [Medline: 30426627]

XSL•FO
RenderX

36. Cane J, O'Connor D, Michie S. Validation of the theoretical domains framework for use in behaviour change and implementation research. Implement Sci 2012;7:37 [FREE Full text] [doi: 10.1186/1748-5908-7-37] [Medline: 22530986]
37. Bean C, Dineen T, Locke SR, Bouvier B, Jung ME. An Evaluation of the Reach and Effectiveness of a Diabetes Prevention Behaviour Change Program Situated in a Community Site. Can J Diabetes 2021 Jun;45(4):360-368. [doi: 10.1016/j.jcjd.2020.10.006] [Medline: 33323314]
38. Jung ME, Locke SR, Bourne JE, Beauchamp MR, Lee T, Singer J, et al. Cardiorespiratory fitness and accelerometer-determined physical activity following one year of free-living high-intensity interval training and moderate-intensity continuous training: a randomized trial. Int J Behav Nutr Phys Act 2020 Feb 26;17(1):25 [FREE Full text] [doi: 10.1186/s12966-020-00933-8] [Medline: 32102667]
39. MacPherson MM, Cranston KD, Locke SR, Bourne JE, Jung ME. Using the behavior change wheel to develop text messages to promote diet and physical activity adherence following a diabetes prevention program. Transl Behav Med 2021 May 19:ibab058. [doi: 10.1093/tbm/ibab058] [Medline: 34008852]
40. Fahim C, Acai A, McConnell MM, Wright FC, Sonnadara RR, Simunovic M. Use of the theoretical domains framework and behaviour change wheel to develop a novel intervention to improve the quality of multidisciplinary cancer conference decision-making. BMC Health Serv Res 2020 Jun 24;20(1):578 [FREE Full text] [doi: 10.1186/s12913-020-05255-w] [Medline: 32580767]
41. Fulton EA, Brown KE, Kwah KL, Wild S. StopApp: Using the Behaviour Change Wheel to Develop an App to Increase Uptake and Attendance at NHS Stop Smoking Services. Healthcare (Basel) 2016 Jun 08;4(2):31 [FREE Full text] [doi: 10.3390/healthcare4020031] [Medline: 27417619]
42. Gould GS, Bar-Zeev Y, Bovill M, Atkins L, Gruppetta M, Clarke MJ, et al. Designing an implementation intervention with the Behaviour Change Wheel for health provider smoking cessation care for Australian Indigenous pregnant women. Implement Sci 2017 Sep 15;12(1):114 [FREE Full text] [doi: 10.1186/s13012-017-0645-1] [Medline: 28915815]
43. Munir F, Biddle SJH, Davies MJ, Dunstan D, Esliger D, Gray LJ, et al. Stand More AT Work (SMArT Work): using the behaviour change wheel to develop an intervention to reduce sitting time in the workplace. BMC Public Health 2018 Mar 06;18(1):319 [FREE Full text] [doi: 10.1186/s12889-018-5187-1] [Medline: 29510715]
44. Robinson H, Hill E, Peel J, Direito A, Jones AW. Use of the Behaviour Change Wheel to develop an intervention to promote physical activity following pulmonary rehabilitation in patients with COPD. European Respiratory Journal 2019 Sep;54(Suppl 63):PA653. [doi: 10.1183/13993003.congress-2019.pa653]
45. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
46. Bean C, Dineen T, Jung ME. "It's a Life Thing, Not a Few Months Thing": Profiling Patterns of the Physical Activity Change Process and Associated Strategies of Women With Prediabetes Over 1 Year. Can J Diabetes 2020 Dec;44(8):701-710. [doi: 10.1016/j.jcjd.2020.09.001] [Medline: 33189581]
47. Cradock KA, ÓLaighin G, Finucane FM, Gainforth HL, Quinlan LR, Ginis KAM. Behaviour change techniques targeting both diet and physical activity in type 2 diabetes: A systematic review and meta-analysis. Int J Behav Nutr Phys Act 2017 Feb 08;14(1):18 [FREE Full text] [doi: 10.1186/s12966-016-0436-0] [Medline: 28178985]
48. Greaves CJ, Sheppard KE, Abraham C, Hardeman W, Roden M, Evans PH, et al. Systematic review of reviews of intervention components associated with increased effectiveness in dietary and physical activity interventions. BMC Public Health 2011;11:119 [FREE Full text] [doi: 10.1186/1471-2458-11-119] [Medline: 21333011]
49. Job JR, Fjeldsoe BS, Eakin EG, Reeves MM. Effectiveness of extended contact interventions for weight management delivered via text messaging: a systematic review and meta-analysis. Obes Rev 2018 Apr;19(4):538-549. [doi: 10.1111/obr.12648] [Medline: 29243354]
50. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. Health Psychol 2009 Nov;28(6):690-701. [doi: 10.1037/a0016136] [Medline: 19916637]
51. Olander EK, Fletcher H, Williams S, Atkinson L, Turner A, French DP. What are the most effective techniques in changing obese individuals' physical activity self-efficacy and behaviour: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 2013;10:29 [FREE Full text] [doi: 10.1186/1479-5868-10-29] [Medline: 23452345]
52. Samdal GB, Eide GE, Barth T, Williams G, Meland E. Effective behaviour change techniques for physical activity and healthy eating in overweight and obese adults; systematic review and meta-regression analyses. Int J Behav Nutr Phys Act 2017 Dec 28;14(1):42 [FREE Full text] [doi: 10.1186/s12966-017-0494-y] [Medline: 28351367]
53. Van Rhoon L, Byrne M, Morrissey E, Murphy J, McSharry J. A systematic review of the behaviour change techniques and digital features in technology-driven type 2 diabetes prevention interventions. Digit Health 2020 Dec;6:2055207620914427 [FREE Full text] [doi: 10.1177/2055207620914427] [Medline: 32269830]
54. Williams SL, French DP. What are the most effective intervention techniques for changing physical activity self-efficacy and physical activity behaviour--and are they the same? Health Educ Res 2011 Apr;26(2):308-322 [FREE Full text] [doi: 10.1093/her/cyr005] [Medline: 21321008]

XSL•FO
RenderX

55. MacPherson MM, Dineen TE, Cranston KD, Jung ME. Identifying Behaviour Change Techniques and Motivational Interviewing Techniques in Small Steps for Big Changes: A Community-Based Program for Adults at Risk for Type 2 Diabetes. Can J Diabetes 2020 Dec;44(8):719-726. [doi: 10.1016/j.jcjd.2020.09.011] [Medline: 33279096]
56. Michie S, Atkins L, West R. The Behaviour Change Wheel: A Guide to Designing Interventions. London: Silverback Publishing; 2014.
57. Redfern J, Thiagalingam A, Jan S, Whittaker R, Hackett ML, Mooney J, et al. Development of a set of mobile phone text messages designed for prevention of recurrent cardiovascular events. Eur J Prev Cardiol 2014 Apr;21(4):492-499. [doi: 10.1177/2047487312449416] [Medline: 22605787]
58. Howlett N, Trivedi D, Troop NA, Chater AM. Are physical activity interventions for healthy inactive adults effective in promoting behavior change and maintenance, and which behavior change techniques are effective? A systematic review and meta-analysis. Transl Behav Med 2018 Feb 28;9(1):147-157. [doi: 10.1093/tbm/iby010] [Medline: 29506209]
59. Asbjørnsen RA, Wentzel J, Smedsrød ML, Hjelmesæth J, Clark MM, Solberg Nes L, et al. Identifying Persuasive Design Principles and Behavior Change Techniques Supporting End User Values and Needs in eHealth Interventions for Long-Term Weight Loss Maintenance: Qualitative Study. J Med Internet Res 2020 Nov 30;22(11):e22598 [FREE Full text] [doi: 10.2196/22598] [Medline: 33252347]
60. Asbjørnsen RA, Smedsrød ML, Solberg Nes L, Wentzel J, Varsi C, Hjelmesæth J, et al. Persuasive System Design Principles and Behavior Change Techniques to Stimulate Motivation and Adherence in Electronic Health Interventions to Support Weight Loss Maintenance: Scoping Review. J Med Internet Res 2019;21(6):e14265 [FREE Full text]
61. van Gemert-Pijnen L, Kelders S, Kip H, Sanderman R. eHealth research, theory and development: a multi-disciplinary approach. New York: Routledge; 2018.
62. van Gemert-Pijnen JE, Nijland N, van LM, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]
63. Kirkman MS, Briscoe VJ, Clark N, Florez H, Haas LB, Halter JB, et al. Diabetes in older adults. Diabetes Care 2012 Dec;35(12):2650-2664 [FREE Full text] [doi: 10.2337/dc12-1801] [Medline: 23100048]
64. Gerber BS, Stolley MR, Thompson AL, Sharp LK, Fitzgibbon ML. Mobile phone text messaging to promote healthy behaviors and weight loss maintenance: a feasibility study. Health Informatics J 2009 Mar;15(1):17-25 [FREE Full text] [doi: 10.1177/1460458208099865] [Medline: 19218309]
65. Smartphone use and smartphone habits by gender and age group 2019. Government of Canada. 2019. URL: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=2210011501 [accessed 2020-12-07]
66. Fjeldsoe BS, Miller YD, O'Brien JL, Marshall AL. Iterative development of MobileMums: a physical activity intervention for women with young children. Int J Behav Nutr Phys Act 2012;9:151 [FREE Full text] [doi: 10.1186/1479-5868-9-151] [Medline: 23256730]
67. Pfaeffli L, Maddison R, Whittaker R, Stewart R, Kerr A, Jiang Y, et al. A mHealth cardiac rehabilitation exercise intervention: findings from content development studies. BMC Cardiovasc Disord 2012;12:36 [FREE Full text] [doi: 10.1186/1471-2261-12-36] [Medline: 22646848]
68. Nelligan RK, Hinman RS, Atkins L, Bennell KL. A Short Message Service Intervention to Support Adherence to Home-Based Strengthening Exercise for People With Knee Osteoarthritis: Intervention Design Applying the Behavior Change Wheel. JMIR Mhealth Uhealth 2019 Oct 18;7(10):e14619 [FREE Full text] [doi: 10.2196/14619] [Medline: 31628786]
69. Redfern J, Thiagalingam A, Jan S, Whittaker R, Hackett ML, Mooney J, et al. Development of a set of mobile phone text messages designed for prevention of recurrent cardiovascular events. Eur J Prev Cardiol 2014 Apr;21(4):492-499. [doi: 10.1177/2047487312449416] [Medline: 22605787]
70. Baranowski T, Lin LS, Wetter DW, Resnicow K, Hearn MD. Theory as mediating variables: Why aren't community interventions working as desired? Annals of Epidemiology 1997 Oct;7(7):S89-S95. [doi: 10.1016/s1047-2797(97)80011-7]
71. Blandford A, Gibbs J, Newhouse N, Perski O, Singh A, Murray E. Seven lessons for interdisciplinary research on interactive digital health interventions. Digit Health 2018;4:2055207618770325 [FREE Full text] [doi: 10.1177/2055207618770325] [Medline: 29942629]
72. Campbell NC, Murray E, Darbyshire J, Emery J, Farmer A, Griffiths F, et al. Designing and evaluating complex interventions to improve health care. BMJ 2007 Mar 3;334(7591):455-459 [FREE Full text] [doi: 10.1136/bmj.39108.379965.BE] [Medline: 17332585]
73. Pagliari C. Design and evaluation in eHealth: challenges and implications for an interdisciplinary field. J Med Internet Res 2007;9(2):e15 [FREE Full text] [doi: 10.2196/jmir.9.2.e15] [Medline: 17537718]
74. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. J Med Internet Res 2015;17(1):e30 [FREE Full text] [doi: 10.2196/jmir.4055] [Medline: 25639757]
75. Innvaer S, Vist G, Trommald M, Oxman A. Health policy-makers' perceptions of their use of evidence: a systematic review. J Health Serv Res Policy 2002 Oct;7(4):239-244. [doi: 10.1258/135581902320432778] [Medline: 12425783]

XSL•FO
RenderX

76. Ouimet M, Landry R, Amara N, Belkhodja O. What factors induce health care decision-makers to use clinical guidelines? Evidence from provincial health ministries, regional health authorities and hospitals in Canada. Soc Sci Med 2006 Feb;62(4):964-976. [doi: 10.1016/j.socscimed.2005.06.040] [Medline: 16314015]
77. Banner D, Bains M, Carroll S, Kandola DK, Rolfe DE, Wong C, et al. Patient and Public Engagement in Integrated Knowledge Translation Research: Are we there yet? Res Involv Engagem 2019;5:8 [FREE Full text]

Abbreviations

APEASE: acceptability, practicability, effectiveness and cost-effectiveness, affordability, safety and side-effects, equity
BCT: behavior change technique
BCW: Behaviour Change Wheel
mHealth: mobile health
SSBC: Small Steps for Big Changes
T2D: type 2 diabetes
TDF: Theoretical Domains Framework

Edited by G Eysenbach; submitted 23.02.21; peer-reviewed by A Beatty, KL Mauco; comments to author 20.05.21; revised version received 04.06.21; accepted 05.07.21; published 27.08.21.

Please cite as:
MacPherson M, Cranston K, Johnston C, Locke S, Jung ME
Evaluation and Refinement of a Bank of SMS Text Messages to Promote Behavior Change Adherence Following a Diabetes Prevention Program: Survey Study
JMIR Form Res 2021;5(8):e28163
URL: https://formative.jmir.org/2021/8/e28163
doi:10.2196/28163
PMID:34448713

©Megan MacPherson, Kaela Cranston, Cara Johnston, Sean Locke, Mary E Jung. Originally published in JMIR Formative Research (https://formative.jmir.org), 27.08.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Scoping the Need for a Tailored mHealth App to Improve Health and Well-being Behavioral Transformation in the Police: Exploring the Views of UK Police Workers via Web-Based Surveys and Client Meetings

Emma Swanston[1], BSc (Hons); Andy Pulman[1], MA, PhD; Huseyin Dogan[1], BSc, MSc, PG Cert, EngD; Jane Murphy[2], BSc (Hons), PhD; Fiona Bitters[3], BSc (Hons), MSt

[1]Faculty of Science and Technology, Bournemouth University, Poole, United Kingdom

[2]Faculty of Health and Social Sciences, Bournemouth University, Bournemouth, United Kingdom

[3]Hampshire Constabulary, Southampton, United Kingdom

Corresponding Author:
Andy Pulman, MA, PhD
Faculty of Science and Technology
Bournemouth University
Poole House, Talbot Campus,
Fern Barrow
Poole, BH12 5BB
United Kingdom
Phone: 44 1202 961047
Email: apulman@bournemouth.ac.uk

Abstract

Background: Police officers often work long, unsocial hours in a highly pressurized environment and may experience difficulties in managing their health and well-being. Their jobs can be highly stressful and feature unusual working hours and multiple shift patterns. When considering the policing environment of 2021, many roles that were previously the domain of warranted officers are now being carried out by nonwarranted police staff equivalents. These police staff roles are relatively new to policing but put staff under some of the same stresses as police officers. A UK police force requested help to investigate technologies that could be used to improve health and well-being and research how these technologies could be used to measure and track health behavior change.

Objective: Historical research studies need to be appraised in light of this new policing environment, and new research also needs to include this shift in dynamics when considering aspects of policing, including their health and well-being. This study explores police officer and staff attitudes toward and their use of existing health-related technology, highlights existing practices, and gathers views about how technology could be used more effectively.

Methods: A web-based survey was completed by police officers and staff (N=213) during the initial period of the UK lockdown in 2020. The survey was designed to find the solutions that participants used outside of those supplied by their employer, identify issues or problems, and find what they would like a hypothetical app to focus on. Additional requirements data were captured through client meetings, including discussions concerning previously attempted solutions and those currently in place. Thematic analysis was undertaken to identify the key themes.

Results: Attitudes toward and uses of existing health-related technology were captured, and existing practices were highlighted. Participants identified a need for an app to consider that a user was on shift—an important point, as many issues and problems with elements of their health and well-being involved shift work. Data also highlighted that a multifunctional tool would be more beneficial to participants than focusing on just 1 element. The key features and four domains were identified for app coverage. The prioritized order of importance of the four domains was activity, food and diet, sleep, and fluid intake.

Conclusions: For police officers and staff, research data suggest that there is a previously unidentified requirement for a mobile app that could provide an easily accessible platform for them to use, regardless of the current location; one that could provide guidelines on diet, lifestyle habits, and health behavior to help the user make informed decisions to assist in personalized behavior change. Notably, one which is multifunctional and which also aligns effectively with the irregular shift patterns of its users.

XSL•FO
RenderX

(JMIR Form Res 2021;5(8):e28075) doi:10.2196/28075

KEYWORDS

nutrition; food; behavior change; mobile health; police; lifestyle management; well-being; mobile phone

Introduction

Policing and Health and Well-being

Police officers often work long, unsocial hours in a highly pressurized environment, and as a result, they can experience difficulties in managing their health and well-being [1,2]. Their jobs can be highly stressful and may feature unusual working hours and irregularly changing shift patterns [3]. Police officers are more likely to experience stress because of exposure to dangerous situations and traumatic events [4] and other stress-related issues, such as working in understaffed environments [5]. Stress can have multiple knock-on effects, including insomnia, fatigue, and lapsed concentration—all of which make performing the job harder than it already is, thereby adding to the original problem [6]. When considering the policing environment of 2021, many roles previously the domain of warranted officers are now carried out by nonwarranted police staff equivalents. These roles are relatively new to policing but put staff under some of the same stresses as police officers. However, stress is not the only mental health condition that police officers and staff are more likely to experience. A 2012 study found that 46.7% of US police officers had required help for mental health conditions during their lifetime [7]. In the United Kingdom, a 2018 survey of serving police officers and staff suggested that symptoms of posttraumatic stress disorder (PTSD) and complex PTSD (CPTSD) might be present in as many as 1 in 5 workers [8]. A companion study noted that CPTSD was more common than PTSD in police officers, and the data supported a cumulative burden model of CPTSD [9].

Police officers also have a greater risk of being overweight or obese and experiencing long-term chronic conditions, such as cardiovascular disease [10] and cancer [11]. Risk factors are heightened by working in a high-stress environment [4], which increases the likelihood or severity of these issues [12,13]. A 1996 study of police officers found that prolonged exposure to body armor and mobile patrols could increase the risk of back pain [14]. Comparing later data from a December 2020 internal commissioning client well-being report confirms this as still being an issue—with muscular skeletal reasons being the fourth highest cause of sickness absence. This is an issue that could have a major effect over a longer period, as low back pain is a health problem leading to the greatest disability in the United Kingdom [15].

Although some risk factors relating to health and an individual's risk of certain health issues are uncontrollable (age, gender, and genetic family history of disease), others are modifiable risks and can be managed by individuals to reduce risk. These include managing diet, alcohol intake, and the amount of physical activity undertaken [16] to help improve health and well-being.

Policing and Health and Well-being (During a Pandemic)

In January 2020, the World Health Organization Emergency Committee declared a global health emergency based on the growing case notification rates of COVID-19 [17]. Since the start of 2020, a number of different lockdown laws have been enacted globally and within the United Kingdom—fluctuating periods of restrictions, lockdowns, and rule relaxations made law [18]. Global and UK policies for policing during COVID-19 have similarly altered during this time [19]. Emergent research during the last year suggests that stress levels of policing during COVID-19 will further exacerbate health and well-being issues [20,21]. COVID-19 policing was hypothesized by Stogner et al [20] as a significant stressor for officers and compounding the general and organizational stresses associated with the occupation. A web-based survey of 2567 police officers across 5 European countries showed that law enforcement agencies and police officers were unprepared for the potential mental stress of policing during COVID-19 [21]. The research paper by De Camargo [22], using data from 18 UK interviews during summer 2020, suggested that unsupportive line managers and colleagues—demonstrated by examples of ridicule and the downplaying of the virus's seriousness—could only elevate officers' stress and anxiety, warranted or not.

Rationale and Aims

During 2020, a UK police force approached our team and requested help in investigating new technologies that could be used to help them keep track and manage various aspects of health. The initial research aims were as follows:

1. Exploring current police officer and staff attitudes toward and their use of existing health-related technology.
2. Highlighting existing practices and gathering views about how technology can be used more effectively in this area.

Theoretical and Practical Contributions

This paper discusses the initial scoping phase of our research, which took place during the first lockdown period in the United Kingdom in 2020. It describes the methods used to undertake these aims, collecting data from 213 participants from the UK police force via web-based surveys and from face-to-face client meetings. We then discuss the results in three thematic sections—exploring attitudes toward and their use of existing health-related technology, highlighting existing practices, and gathering views about how technology could be used more effectively in this area. The results are discussed in light of the current research.

As already noted, both previous studies and future research in health and well-being for this user group need to be considered in the context of a policing environment where many roles previously the domain of warranted officers are now carried out by nonwarranted police staff equivalents. The policing environment of 2021 is not the same as it was earlier. Our new

XSL•FO
RenderX

and unique contribution to the research pool is conducting initial scoping in the midst of the pandemic and being mindful of the needs to include both this impact and the shift in work dynamics and roles when considering aspects of health and well-being for all police workers and the unique demands of their jobs.

Methods

Study Design

Methods chosen were considered with the participants in mind—those that could be completed in 1 session, where time could be set aside when they were free. We used a human-centered approach to optimize understanding and accommodate the perspectives of potential users [23].

Participants, Recruitment, and Consent

Data were gathered in the southern region of the United Kingdom where the commissioning organization was based. Participants were recruited through a purposive sampling strategy that targeted personnel within the organization. Survey participants were recruited via a gatekeeper from the commissioning organization. The gatekeeper role was to initiate communication between the researcher and police officers and staff who wished to participate—via distribution of study literature—without compromising anonymity or affecting the veracity of web-based responses. Participants were voluntary and remained anonymous to both the gatekeeper and the organization.

Although other methods were initially considered to capture the original requirements, because of complications arising from COVID-19 in terms of logistics, lockdown restrictions, and operational pressures on the commissioning force, requirements were also captured from client meetings.

Procedures and Measures

A web-based survey was chosen as the method of data collection. This was owing to the initial UK COVID-19 lockdown period restricting in-person meetings, the number of people we were able to gather data on, and the shortened time to complete surveying (in comparison with undertaking interviews). The participants provided consent before the survey was completed. Using a web-based survey as the main method of data collection allowed initial ideas and requirements to be documented in depth and helped tailor prototype ideas to accurately reflect participant needs. Content validity (also referred to as face validity) indicates whether a questionnaire appears logical to a group of experts [24]. A pilot version was sent to 5 participants working in the field, who completed the questionnaire and provided feedback during an interview. To ensure suitable content validity, questions were asked around their ability to understand questions, whether categories allowed users to give the answer they wanted, and general ease of understanding. This feedback was used to make changes and improvements before it was sent back to the client for additional feedback. Improvements were then implemented into a revised version, which was sent to the client for further feedback. All participants completed the survey in less than 30 minutes. This was important as it was required to not take up too much participant time because of operational availability. Once the final version was confirmed, a second test run was completed by the same participants (n=5) who had taken the draft survey, but this time on Google Forms.

The survey included a *mixture* of open-ended, qualitative and closed questions. Qualitative questions were used to collect information relating to the features and functions that participants used most on apps and their benefits and limitations, also reflecting on issues and problems which a new app could potentially help them with. Closed questions were used to collect demographic statistics and details on currently used technology, such as accessible devices, operating systems used, and device preferences. Owing to the nature of the questions being asked, such as providing opinions on health and well-being apps and listing popular and unpopular features, we did not use a Likert scale within the closed questions included.

Client meetings were intended as a combination of discussion on clients' expectations, their background with the project or similar projects, and answering preprepared questions. The initial meeting took place with a client representative and a representative from Bournemouth University and covered the background to the project and what direction it might be expected to go in. Subsequent meetings took place with a researcher and the client representative and covered potential existing solutions, solutions the client organization had already tried or had in place, data gathering, and expectations for evaluating developing prototypes.

Sample Size

The recommended sample size was calculated to ensure that the number of participants for the survey was acceptable. The sample size was calculated with a 95% confidence level and a margin of error of 5%. As the intended audience for the final product was a UK police force, the workforce value supplied by the client organization was rounded to the nearest 1000 and used to calculate an ideal sample size. The value used for the workforce was 5000. Owing to the high workforce number, it was expected that a confidence level of 95% and margin of error of 5% might not be possible because of restrictions with accessing participants. A margin of error of 5% was calculated to be used as the goal value to reach (95% confidence level+5% margin of error; n=375 participants), but a margin of error of 7.5% was accepted as the lowest value (95% confidence level+7.5% margin of error; n=165 participants) acceptable for this study [25].

Data Analysis

Survey data were analyzed and summarized using descriptive statistics. Data were tested for normal distribution and then presented as means, SDs, ranges, and percentages. For age, as under 21 years and over 60 years did not have confirmed lower and upper values, respectively, 18 years and 65 years were used to calculate the mean, SD, and range (18 years being the lowest age of UK employment and 65 years being the current age of UK retirement).

Open questions were thematically analyzed using a deductive approach that focused on the areas covered in the questions (such as design, functionality, and content). A generic qualitative approach to thematic analysis was used [26] with interresearcher

XSL•FO
RenderX

interpretation. Following familiarization with the qualitative data, a member of the team charted the initial themes. A second researcher subsequently familiarized themselves with the data and initial themes. They developed a coding scheme using an analytical framework that combined a priori issues from the original surveys and emerging themes [27]. We attended to the principles of sampling, saturation, negative cases, confirmation, and logical progression throughout so that bias and any other errors identified were eradicated during inquiry.

Results

Overview

The number of participants was 213, which was above the lower threshold for an acceptable number of participants for the sample size [23]. With this number of participants and an estimated population of 5000, there is a 95% confidence level with a final margin of error of 6.5%.

Participant Characteristics

We recruited 37.1% (79/213) male and 62.9% (134/213) female participants. The "other" or "prefer not to say" option was also included within the gender question—no responses were received. The mean age of the participants was 41.9 (SD 10.0) years (Table 1).

Table 1. Participant demographic data (n=213).

Characteristics	Participants
Age (years)[a]	
Value, mean (SD; range)	41.9 (10.0; 21-65)
Age group, n (%)	
Under 21	0 (0)
21-30	26 (12.2)
31-40	75 (35.2)
41-50	67 (31.5)
51-60	38 (17.8)
Over 60	7 (3.3)
Gender , n (%)	
Male	79 (37.1)
Female	134 (62.9)
Other or prefer not to say	0 (0)

[a]As under 21 years and over 60 years did not have confirmed lower and upper values, respectively, 18 years and 65 years were used to calculate the mean, SD, and range (with 18 years being the lowest age of UK employment and 65 years being the current age of UK retirement).

Attitudes Toward and Use of Existing Health-Related Technology

Technology Use

Most (206/213, 96.7%) of the participants noted they had access to a smartphone, compared with 78.4% (167/213) for a PC or laptop and 58.2% (124/213) for a tablet. Alongside these data, 78.9% (168/213) of participants reported that out of the options available to them, they used a smartphone the most. Regarding operating systems, 56.3% (120/213) of participants said their smartphone ran on Android as opposed to 43.2% (92/213) running Apple iOS.

Out of all the participants, 43.2% (92/213) did not use a smart watch or device of any kind; therefore, at this point in time, a large number of potential users would not be able to use or take advantage of features targeted specifically at those devices. Reasons given for not using these ranged from lack of awareness of the options available, the technological skills required to make use of them, and cost restrictions. Where some participants had previously used these devices, there were also reasons given for nonuse currently, such as reliability: "Used to use Fitbit but broke it and haven't got a new one" [P145]. Issues with interoperability, features, and functionality: "I used to have an Apple Watch but found constant messages coming through to it was distracting when trying to work or relax and therefore no long wear it" [P188]. There were also a range of other explanations:

> *The electronic wrist device gave me an upset stomach and a small irritated patch of skin. I don't feel all the tech is good for my body.* [P40]

Apps in Use

Participants noted a number of different mobile health (mHealth) apps that they currently used to help with aspects of their health and well-being (Table 2).

XSL•FO
RenderX

Table 2. Mobile health apps mentioned in relation to usage (n=213).

Categories	Participants, n (%)
General health	
Fitbit[a]	21 (9.8)
Garmin[b]	16 (7.5)
Apple Health[c]	9 (4.2)
Samsung Health	4 (1.8)
Huawei Health	1 (0.4)
Google Fitness	1 (0.4)
ProFit[d]	1 (0.4)
Virgin Health–Max Buzz	1 (0.4)
Activity	
MyFitnessPal	20 (9.3)
Strava	18 (8.4)
MapMy (Walk, Run, or Ride)	7 (3.2)
Runkeeper	2 (0.9)
Peloton	2 (0.9)
Fitnotes	1 (0.4)
Interval Timer	1 (0.4)
Fitbit Coach	1 (0.4)
Pacer	1 (0.4)
Polar Beat	1 (0.4)
Food or diet	
Weight Watchers	2 (0.9)
NHS[e] Well Man	1 (0.4)
NHS BMI Tracker	1 (0.4)
NHS Calorie Checker	1 (0.4)
Noom	1 (0.4)
Nutracheck	1 (0.4)
Slimming World	1 (0.4)
Rest, relaxation, or mental health	
Headspace	7 (3.2)
Mindfulness UK	2 (0.9)
Breathe2Rrelax	1 (0.4)
Calm	1 (0.4)
Waking Up	1 (0.4)
Fluid	
Drink Free Days	1 (0.4)
Other	
COVID-19 Symptom Tracker	1 (0.4)
GP24	1 (0.4)
Tinnitus Therapy	1 (0.4)

[a]Including the Fitbit app.

XSL•FO
RenderX

[b]Including connect.

[c]Including workouts and activity.

[d]Including VeryFitPro.

[e]NHS: National Health Service.

Barriers to Use and Success

Barriers to engagement and adherence highlighted included the risk of organizational infrastructure causing difficulties in improving aspects of health and well-being. These could be physical building limitations:

> *Not enough different options in the [Police Investigation Centre] PIC canteens.* [P112]
>
> *Our police station has no adequate, clean, hygienic area to prepare food, so the default is to snack or binge on treats. Our police station has no canteen, so no healthy options are available even though the organisation want and stress that officers should eat well and not rely on fast food.* [P162]

Or also managerial issues affecting the ability to change:

> *I have suggested that the app includes screen breaks and breaks from work. The problem with this is that whilst senior management seem to support this, day to day management [sgts and inspectors] do not, so if you are seen getting up and taking breaks you are considered not to be working. For example just two days ago I was told by a sgt that I was not entitled to a lunch break even though we are being encouraged to eat lunch away from our screens! The app sounds great, but without support from line managers it's not going to work.* [P188]
>
> *I do not believe that technological solutions are the biggest factor in solving staff health and wellbeing. A flexible approach within management for the health and wellbeing of staff would be a more beneficial solution.* [P129]

Data also suggested that there could be an in-built resistance for some who were not in the right mindset to want to engage with a technological solution offered to them:

> *Pointless applications, how have we, the human race, survived this long without being told that we need to drink water by some artificial device. I know if I haven't slept enough, I'm tired. These are issues that do not need managing, they are issues invented by the manufacturers of needless devices.* [P98]
>
> *Apps will only make you think about your wellbeing. It's down to whatever inspires the individual to get them off the sofa and start moving, to give them the availability [time management] to home cook and to afford to eat well/healthy. Recipes are good and prompts are just that, prompts.* [P54]

In addition, some participants also offered a variety of other reasons, which might have caused them to avoid engaging with a new solution offered on a smartphone. This again highlighted the technology experience concerns:

> *Money, Privacy, Guilt induction by failure to see progress.* [P45]
>
> *I don't know how to set it up I'm not very technically minded.* [P203]

Some participants noted an element of technology fatigue (where they were having to use technology most of the time): "Fed up with everything being online, using my time up to log in/passwords/checking etc" [P63]. Another barrier to success in the area of emotional and mental well-being was the continual demands of the job:

> *But this is a job like no other [few other occupations including other emergency services get so emotionally, verbally and physically abused] and get micro examined, as the Police. That will take a change in [the small number of] societies attitude and behaviour that would make the greatest help to mental and emotional wellbeing, not to mention time and effort.* [P54]

Highlighting Existing Practice—Employer Initiatives

The client organization had investigated other solutions to help improve employee health and well-being. They had participated in the Virgin Pulse Corporate Global Challenge [28], but participation in this had now finished. As this had run for a limited period, the opportunities of longer-term effects from this had in effect stopped at that point in time, as noted by the following responses:

> *Time limited - health challenge.* [P110]
>
> *Global challenge finished so could no longer log into the application.* [P151]

In client meetings, another of the existing solutions listed was investing in retreats focusing on mental health and well-being. However, this particular aspect was currently limited to the numbers who could take advantage against the total number of staff working for the force. No mention was made of these events in the survey responses.

Gathering Views About How Technology Could Be Used More Effectively

App Coverage

The areas of health and well-being that participants said they felt they could benefit from having support from the most in an app were prioritized by response. The four highest prioritized domains were food and diet (76/213, 35.6%), activity (68/213, 31.9%), sleep (27/213, 12.6%), and fluid intake (27/213, 12.6%; Textbox 1). Other lower scoring, suggested areas for support included relaxation techniques and rest breaks (2/213, 0.9%), general health and fitness goals (1/213, 0.4%), differing aspects of mental health (1/213, 0.4%), and specific gender-related health issues (1/213, 0.4%). No preference (10/213, 4.6%) accounted for the other responses.

XSL•FO
RenderX

Textbox 1. Domains to support health-related behavior change identified for app coverage.

Food and Diet

- "Shift work making organising meals more difficult. My own lack of self-discipline of making better choices. The convenience and availability of fast food and takeaways." [P166]
- "Keeping it simple and using realistic meals that are quick and easy. Other apps have very overly complicated meals." [P145]
- "Stress Eating." [P120]

Activity

- "My own motivation - I used to get my exercise from my previous job as I was on my feet all day, but now I have a sedentary job, I have gained weight as a result, but my diet hasn't changed an awful lot." [P59]
- "Time constraints." [P104]
- "Feeling bad when I don't exercise!" [P24]

Fluid Intake

- "I know that I don't usually drink as much water as I should and drink a lot of coffee. I also don't track how much water I drink." [P32]
- "Because we are in an office all day you forget to go to the kitchen to get water." [P188]
- "Lack of prompting to drink." [P102]

Sleep

- "Medication and insomnia at times." [P184]
- "Not being able to switch off my mind/relax." [P117]
- "Shift patterns and concerns about work." [P182]

What Is Missing Currently

Notably, when analyzing responses, participants wanted an app to take into consideration the fact that a user was on shifts—an important point, as many issues and problems participants had within each of the 4 identified domains involved shift work:

> *I like eating and work shifts, meals aren't always practical as suggested on apps.* [P37]
>
> *I'm a shift worker and find that none of these apps really know how to advise those who work shifts. They are aimed at those who work mon-fri 9-5.* [P61]

Web-based data collected also highlighted that a multifunctional tool would be more beneficial to participants than focusing on just one element:

> *It would be great to have a one stop shop for health and wellbeing.* [P46]
>
> *Having an app that combines all aspects is what is required. There are apps for counting calories, but don't give enough inspiration for exercise or keeping active daily.* [P185]

Considering a wider focus for an app, rather than just focusing on one or two specific subsections, was suggested within the data because of the nature of the research problem and the environment and job requirements of participants:

> *Other apps are all about exercise, but don't help with the food aspect.* [P185]
>
> *Most apps are aimed at very active very fit people, I am inactive, have arthritis and need an app to set realistic targets that build over time with strength, repetition and confidence. I usually can never achieve the targets which is off putting.* [P14]

Key Features

Dashboard and Visualization

Participants wanted the ability to visualize how their behavior might be changing and how aspects they used could be personalized, such as those relating to changes in individual shift patterns:

> *When you have beaten a personal best you get a notification or you have a dashboard of your best efforts to see easily.* [P144]
>
> *There are lots of people that work shifts so could benefit from helpful advice explaining how to shift sleep pattern during night shifts and what to eat, when to eat, etc.* [P61]

Goal-Setting and Gamification

The option of setting regular goals within each of the four domains identified (eg, fluid intake, snacks avoided or consumed over the course of a day, or planned physical activities achieved) was also described as a useful feature within an app. Nearly half of the participants (106/213, 49.7%) said that they did not enjoy gamification features in an app. Of the 28.6% (61/213) of participants who said they enjoyed gamification features, the most popular types of gamification were goals (44/61, 72%), daily challenges (42/61, 69%), and leaderboards (40/61, 65%) with competition-style gamification, such as teams (23/61, 38%) and experience points (22/61, 36%) being the least popular:

XSL•FO
RenderX

> *I would like to see goals / medals / points achieved for weight loss and same for good nutrition, which also shows the improvement scale over a period of time.* [P108]

> *I really dislike group participation. I am very self-disciplined, and I like to set my own targets and goals and strive to beat my own PB's. I enjoy exercise, but enforced group participation would demotivate me to a point of non-engagement.* [P5]

Diary Options

Participants thought that the ability to enter diary-based information into the app would be useful, such as notes associated with differing shift times. One barrier to using the existing apps highlighted was their adherence to keeping daily routines the same and not allowing for users who worked nonstandard shift patterns:

> *Most applications of this nature cater for either athletes or people that work regular hours during the week, they do not cater for people who work weekends, overnight or irregular/shifts. Most of the research is not catered for those who have no choice so they get regular sleep, eat at odd times.* [P6]

Customizable Notifications

Of the 113 participants who used a health or well-being app, 47 (41.6%) said that they checked an app multiple times during a day and that they thought that notifications would help them to remember to do certain tasks. Shift-related customized reminders were also suggested to be helpful:

> *It couldn't change to my shift patterns, for instance on nights I don't need a 07:00hrs reminder to drink water.* [P7]

Discussion

Principal Findings

Attitudes Toward and Their Use of Existing Health-Related Technology

This research has highlighted that for participants, a mobile app rather than other mediums would provide the best platform for user access. Smartphone apps are increasingly considered useful health promotion tools owing to their accessibility and scalability [29] and with a view to integrating health within a work-based framework [30]. Penetration of smart devices (including watches) lags behind phones [31], with more associated barriers involved in their wider use at this point—highlighted by participants in comparison with more widely adopted and user-friendly phone technology—although this might change as the market matures.

An increasing number of apps and technologies for managing aspects of health and well-being have been launched over the past 10 years [32]. There are currently a number of health and fitness solutions available on multiple platforms in personal use by police officers and staff, as evidenced by our data (Table 2). The fact that most apps focus on one particular element of health and well-being is notable. They have not been designed to address the specific issues that police officers and staff face, nor are they configurable for the types of routines and working patterns that they regularly encounter [33]. Recent literature reviews, such as a 2020 review of research on gamification and mHealth apps for emergency service personnel and police officers across 6 major databases, have highlighted a lack of literature in this area for these groups [34].

Spanning both police officers and staff, our research has highlighted concerns that the working environment makes it harder for those affected to make healthy choices, such as physical building limitations and managerial challenges. Therefore, the problem includes not only thinking of a solution to help manage personalized risk issues but also ensuring that the solution would not be intrusive for the user during and outside of work, to counteract the in-built resistance some participants noted when engaging with this type of media. As became clear from the data, solving this issue will not be easy because of the fact that one solution would not work effectively for all police officers and staff. Barriers to usage and workplace limitations for those who work in the police provide a challenge to solving this issue. Considering activity, perceived pressure of work and organizational culture appears to be a sturdy barrier to reducing sedentary time [35]. Police officers in previous studies have expressed a need for more opportunities to take breaks and more encouragement from managers or supervisors [36], as some of our participants also noted. In terms of pre- and postpandemic mental health, particularly challenging are the potential difficulties of highlighting and raising this topic within a culture where broaching mental health difficulties have sometimes been viewed as a sign of weakness or undesirable topic to discuss [37].

Highlighting Existing Practice

As a part of our project, we performed 2 reviews of research concerning different domains of police officers and staff health and well-being—separate but often interlinked domains thematically identified in our research to date. This also included a review of current mHealth technology in use by police forces to help address these domains. Article length precludes a more detailed discussion, but we are able to summarize some notable points here in relation to our data analysis at this stage.

The client organization had already investigated solutions to help improve employee health and well-being. They had participated in the Virgin Pulse Corporate Global Challenge [28], with each employee being offered a smart tracker. A desktop or mobile app is used to track the progress of teams throughout the challenge and a self-health assessment used initially—to help users understand the goals and improvements they should aim to make [38]. Notably, this was limited to the period of the actual challenge rather than a long-term intervention with the issues this poses. For example, Rossomanno et al [39] found that only 51% of officers reported meeting recommendations for physical activity at their study end and theorized that it was possible that participants may have lacked motivation, time, and incentive to continue the training necessary past end date. Recent UK research in this area involving 2 other UK police forces [36,40] has included a physical activity wearables study, which used a combination of

XSL•FO
RenderX

a Fitbit activity monitor and *Bupa Boost* smartphone app to promote physical activity and reduce sedentary behavior in police officers.

In client meetings, another of the existing solutions listed was investing in retreats focusing on mental health and well-being. However, this particular aspect is currently limited to the numbers who can take advantage as opposed to the number of staff working for the force overall. The College of Policing recently conducted a randomized controlled trial, giving 5 other police forces access to either Headspace or Mindfit Cop. This study found that after 6 months, the average well-being, life satisfaction, resilience, and performance of individuals in the intervention groups improved compared with the control group [41]. Whereas a UK-based mobile app (Backup Buddy) allowing police officers to informally view static audio and video information and signposted support options on common mental health issues is used by 14 other UK forces [42]. No research data are currently available on its usage and impact. A research study is also currently in progress to better understand police officer experiences of the 87% well-being app, which supports employee well-being strategies [43].

In the United Kingdom, the National Police Wellbeing Service [44] covers 8 core elements to deliver tailored support and guidance for policing. Although offering a broad approach to support the health and well-being of policing personnel, it does not presently incorporate or recommend the use of digital solutions [34]. From the literature reviewed, each UK force has a certain degree of autonomy in promoting health and well-being and working with the research community and external providers. This can lead to differences in approaches and solutions used, as pointed out in the analysis of the Blue Light Wellbeing Framework [45]. Initiatives within the United Kingdom have seemingly worked in silos, with subsets of forces piloting interventions before a wider roll out—note the number of differing ongoing mental health initiatives listed earlier. This approach can reduce opportunities for collaboration and knowledge exchange and increase the chances of multiple initiatives overlapping—targeting the same areas at the same time with different technology and devices.

Gathering Views About How Technology Could Be Used More Effectively

The 4 highest prioritized domains were food and diet, activity, fluid intake, and sleep (Textbox 1). Behavior change is likely to play a large part in making an app focusing on health and well-being successful, with the suggestion that increased implementation of behavioral change techniques (BCTs) could improve interventions and achieve higher levels of user engagement [46]. Evidence also suggests that interventions incorporating multiple BCTs are more effective in meeting the challenge of long-term, sustainable change than those using a few or single BCT [47]. Antezana et al [46] suggested that project teams should embrace different BCTs and establish synergies by linking different features. This seems to be borne out in the limited number of BCTs found in many apps [32], suggesting an opportunity for improvement in new designs. It is also suggested to avoid the pitfalls of badly designed and unengaging interventions that project teams gather in-depth feedback about users' views of all elements of interventions so they and developers can understand what users required [23,48]—part of the reason for our human-centered approach to this study [23].

As a 24/7 service, police officers typically work a 40-hour week, in shifts, including weekends and bank holidays [49]. Shift patterns might have 12 hours between shifts and have a minimum of either a continuous 24-hour rest period each week, or an uninterrupted rest period of 48 hours in any reference period of 14 days. They might also need to support the delivery of operational demand. A common pattern for policing is 2 day shifts, 2 late shifts and 2 night shifts, followed by 4 days off, referred to as a *six on four off*. Currently, few apps allow for customization concerning these patterns to help meet the needs and requirements of shift workers [33] and specifically police officers [1-5]. This correlates with participants' responses and literature demonstrating a dearth of relevant research being published in this area [34], which our project helps to redress.

Mirroring the thoughts of Antezana et al [46] on multiple BCTs, participants noted that a multifunctional tool would be more beneficial to them than focusing on a particular individual element. Many current apps focus on one specific area of health and well-being, meaning that to focus on general health and well-being, the user needs to download and use different apps rather than personalizing just one. An individual could theoretically customize a suite of apps to use for managing aspects of health they were interested in or adopt a generic multifunctional tool such as Apple or Samsung Health as a very small number of participants had (Textbox 1). However, a much larger proportion of participants felt that this did not provide a suitable solution when coupled with the particular lifestyles and working environments of their jobs [1-5].

In terms of key feature requirements, data showed that participants wanted the ability to visualize how their behavior might be changing and how aspects used could be personalized, such as those relating to changes in individual shift patterns. The option of personalized customization is recognized as one of the advantages of mHealth interventions [38].

Half of the participants stated that they did not enjoy gamification within an app, and a large emphasis on gamification in association with goal-setting features would be unwise. As people get older, they are less likely to have experience of using digital games and the nuances of gamification elements (having not grown up within this culture). They might have different assumptions [50], making it essential to account for their experience with games to conceptualize a successful and suitable gamified intervention. Moreover, goals and priorities differ between younger and older people [51], which might have an impact on what motivates or affects their enjoyment of gamification. It is also possible that elements of the culture of working within the police might view gaming as a frivolous activity [37], as evidenced in some participant responses. However, it could be possible to implement less intrusive gamification elements, such as goals for the user, without a focus on teams or competition. For example, previous research looking at a number of physical activity apps noted that combinations of BCTs embedded within gamified apps

XSL•FO
RenderX

were reported as self-monitoring and goal setting, with the addition of either a focus on past success or nonspecific rewards and incentives [52].

Participants also viewed the ability to enter diary-based information, such as notes associated with different shift times, into the app as being very useful. This area deserves further investigation, especially in relation to mental health.

In terms of customizable notifications, participants thought they would help them to remember to do certain tasks, whereas shift-related customized reminders were also reported as likely to be very helpful. As this particular area seems polarizing based on individual work-life practices highlighted, a good middle ground would allow users to customize the notifications received. This way, it would not be off-putting for users who either liked or disliked checking an app regularly. This would also allow users to be notified about areas of health and well-being that they designated as being of interest—a popular technique for incorporation into digital health apps [52-54].

Limitations

A limitation of this study is that for the age descriptive calculation (Table 1), police officers usually retire at a much younger age in the United Kingdom (usually between 50 and 55 years) than the current UK age of 65 years used, which is when police staff might retire. As we did not formerly differentiate between police officers and staff during data collection, this might mean that the upper limit used for this calculation is not entirely representative. Although the authors believe that the results of this study can be generalized for other police forces across the country, we acknowledge that there is a limitation of only accessing data from a subsection of one regional UK police force, which might have inherent organizational biases toward health and well-being. However, this approach can be expanded upon to cover more regions in due course. Concerning bias with purposeful sampling—where the belief is that qualitative research should describe the medium or the norm—the point to underline is that new phenomena are being described, so we needed to purposively select the best examples of what we were interested in. This gave us the clearest cases with the least *noise* or extraneous errors and allowed for the identification of characteristics and boundaries [55]. A further limitation of this study is the greater number of female responses (134/213, 62.9%) than male responses when compared with the police workforce in England and Wales numbers as of 2016, which noted that only 28.6% of all officers were female [56]. We will attempt to improve upon this gender imbalance in future stages of the study.

Conclusions

For police officers and staff, research data suggest that there is a previously unidentified requirement for a mobile app that provides an easily accessible platform for them to use regardless of the current location. One that could provide guidelines on diet, lifestyle habits, and health behavior to help the user make informed decisions to assist in personalized behavior change. A multifunctional app, which aligns with the irregular shift patterns of its users, can be customized effectively for their specific needs. More detailed requirements gathering is therefore expected to take place while funding is sought for the creation and prototyping of a solution. This work will include drafting initial requirements and preferences for devising a new targeted solution in this area as a starting point for future prototyping, then creating and prototyping a solution. It will continue to ensure the participation of police officers and staff, working toward a human-centered design methodology at each stage of the development cycle.

Acknowledgments

The authors would like to thank all the participants who participated in the web-based survey. Ethical approval was obtained from Bournemouth University (ref. 29976). This research project was cofunded by Bournemouth University and the commissioning client.

Authors' Contributions

FB, JM, and HD conceived and designed the study. ES managed the creation of the survey and data collection. ES and AP conducted data analysis. AP led the preparation of this paper. All authors critically reviewed and revised the manuscript and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Violanti JM, Charles LE, McCanlies E, Hartley TA, Baughman P, Andrew ME, et al. Police stressors and health: a state-of-the-art review. Int J Police Strat Manag 2017 Nov 20;40(4):247-261. [doi: 10.1108/pijpsm-06-2016-0097]
2. Demou E, Hale H, Hunt K. Understanding the mental health and wellbeing needs of police officers and staff in Scotland. Police Pract Res 2020 Jun 09;21(6):702-716 [FREE Full text] [doi: 10.1080/15614263.2020.1772782] [Medline: 33408592]
3. Ma CC, Andrew ME, Fekedulegn D, Gu JK, Hartley TA, Charles LE, et al. Shift work and occupational stress in police officers. Saf Health Work 2015 Mar;6(1):25-29 [FREE Full text] [doi: 10.1016/j.shaw.2014.10.001] [Medline: 25830066]

XSL•FO
RenderX

4. Galbraith N, Boyda D, McFeeters D, Galbraith V. Patterns of occupational stress in police contact and dispatch personnel: implications for physical and psychological health. Int Arch Occup Environ Health 2021 Feb 12;94(2):231-241. [doi: 10.1007/s00420-020-01562-1] [Medline: 33044570]
5. Kula S. Occupational stress, supervisor support, job satisfaction, and work-related burnout: perceptions of Turkish National Police (TNP) members. Police Pract Res 2016 Nov 07;18(2):146-159. [doi: 10.1080/15614263.2016.1250630]
6. Mayhew C. Occupational health and safety risks faced by police officers. In: Trends & Issues in Crime and Criminal Justice. Canberra: Australian Institute of Criminology; 2001.
7. Fox J, Desai MM, Britten K, Lucas G, Luneau R, Rosenthal MS. Mental-health conditions, barriers to care, and productivity loss among officers in an urban police department. Conn Med 2012 Oct;76(9):525-531 [FREE Full text] [Medline: 23155671]
8. Miller J. Police Care UK. 2018. URL: https://www.policingtrauma.sociology.cam.ac.uk/survey [accessed 2021-02-02]
9. Brewin CR, Miller JK, Soffia M, Peart A, Burchell B. Posttraumatic stress disorder and complex posttraumatic stress disorder in UK police officers. Psychol Med 2020 Sep 07:1-9. [doi: 10.1017/s0033291720003025]
10. Zimmerman F. Cardiovascular disease and risk factors in law enforcement personnel: a comprehensive review. Cardiol Rev 2012;20(4):159-166. [doi: 10.1097/CRD.0b013e318248d631] [Medline: 22314143]
11. Wirth M, Vena J, Burch J. Risk of cancer incidence and cancer mortality among police officers. In: Violanti JM, editor. Dying for the Job: Police Work Exposure and Health. Springfield, IL: Charles C Thomas Publisher, Ltd; Jan 01, 2014:57-70.
12. Ramey S. Relationship between cardiovascular disease morbidity, risk factors, and stress in a law enforcement cohort. Retrospective Thesis and Dissertations, Iowa State University. 2002. URL: https://lib.dr.iastate.edu/cgi/viewcontent.cgi?article=1967&context=rtd [accessed 2021-08-04]
13. Ramey SL. Cardiovascular disease risk factors and the perception of general health among male law enforcement officers: encouraging behavioral change. AAOHN J 2003 May;51(5):219-226. [Medline: 12769168]
14. Burton AK, Tillotson KM, Symonds TL, Burke C, Mathewson T. Occupational risk factors for the first-onset and subsequent course of low back trouble. A study of serving police officers. Spine (Phila Pa 1976) 1996 Nov 15;21(22):2612-2620. [doi: 10.1097/00007632-199611150-00011] [Medline: 8961449]
15. United Kingdom. Institute for Health Metrics and Evaluation. Seattle, WA: University of Washington URL: http://www.healthdata.org/united-kingdom [accessed 2021-08-04]
16. Roger VL, Go AS, Lloyd-Jones DM, Adams RJ, Berry JD, Brown TM, American Heart Association Statistics CommitteeStroke Statistics Subcommittee. Heart disease and stroke statistics--2011 update: a report from the American Heart Association. Circulation 2011 Feb 01;123(4):e18-e209 [FREE Full text] [doi: 10.1161/CIR.0b013e3182009701] [Medline: 21160056]
17. Velavan TP, Meyer CG. The COVID-19 epidemic. Trop Med Int Health 2020 Mar 16;25(3):278-280 [FREE Full text] [doi: 10.1111/tmi.13383] [Medline: 32052514]
18. Baker C, Brown JS. Research Briefing - Coronavirus: a history of English lockdown laws. House of Commons Library. URL: https://commonslibrary.parliament.uk/research-briefings/cbp-9068/ [accessed 2021-08-04]
19. COVID-19 restrictions. College of Policing. URL: https://www.college.police.uk/guidance/covid-19-restrictions [accessed 2021-03-22]
20. Stogner J, Miller BL, McLean K. Police stress, mental health, and resiliency during the COVID-19 pandemic. Am J Crim Justice 2020 Jun 26;45(4):1-13 [FREE Full text] [doi: 10.1007/s12103-020-09548-y] [Medline: 32837167]
21. Frenkel MO, Giessing L, Egger-Lampl S, Hutter V, Oudejans RR, Kleygrewe L, et al. The impact of the COVID-19 pandemic on European police officers: Stress, demands, and coping resources. J Crim Justice 2021 Jan;72:101756 [FREE Full text] [doi: 10.1016/j.jcrimjus.2020.101756] [Medline: 33100418]
22. De Camargo C. 'It's tough shit, basically, that you're all gonna get it': UK virus testing and police officer anxieties of contracting COVID-19. Policing Soc 2021 Feb 21:1-17. [doi: 10.1080/10439463.2021.1883609]
23. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. J Med Internet Res 2015 Jan;17(1):e30 [FREE Full text] [doi: 10.2196/jmir.4055] [Medline: 25639757]
24. Chung KC, Pillsbury MS, Walters MR, Hayward RA. Reliability and validity testing of the Michigan Hand Outcomes Questionnaire. J Hand Surg Am 1998 Jul;23(4):575-587. [doi: 10.1016/S0363-5023(98)80042-7] [Medline: 9708370]
25. Fugard AJ, Potts HW. Supporting thinking on sample sizes for thematic analyses: a quantitative tool. Int J Soc Res Methodol 2015 Feb 10;18(6):669-684. [doi: 10.1080/13645579.2015.1005453]
26. Caelli K, Ray L, Mill J. 'Clear as Mud': Toward greater clarity in generic qualitative research. Int J Qual Methods 2016 Nov 29;2(2):1-13. [doi: 10.1177/160940690300200201]
27. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
28. The Global Challenge. Virgin Pulse. URL: https://www.virginpulse.com/en-gb/global-challenge/ [accessed 2021-02-02]
29. Sullivan AN, Lachman ME. Behavior change with fitness technology in sedentary adults: a review of the evidence for increasing physical activity. Front Public Health 2016;4:289 [FREE Full text] [doi: 10.3389/fpubh.2016.00289] [Medline: 28123997]

XSL•FO
RenderX

30. Jimenez P, Bregenzer A. Integration of eHealth tools in the process of workplace health promotion: proposal for design and implementation. J Med Internet Res 2018 Feb 23;20(2):e65 [FREE Full text] [doi: 10.2196/jmir.8769] [Medline: 29475828]
31. Chuah SH, Rauschnabel PA, Krey N, Nguyen B, Ramayah T, Lade S. Wearable technologies: The role of usefulness and visibility in smartwatch adoption. Comput Hum Behav 2016 Dec;65:276-284. [doi: 10.1016/j.chb.2016.07.047]
32. McKay FH, Wright A, Shill J, Stephens H, Uccellini M. Using health and well-being apps for behavior change: a systematic search and rating of apps. JMIR Mhealth Uhealth 2019 Jul 04;7(7):e11926 [FREE Full text] [doi: 10.2196/11926] [Medline: 31274112]
33. Nunes F, Ribeiro J, Braga C, Lopes P. Supporting the self-care practices of shift workers. In: Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia. 2018 Nov 25 Presented at: MUM 2018: 17th International Conference on Mobile and Ubiquitous Multimedia; November 25 - 28, 2018; Cairo, Egypt p. 71-81. [doi: 10.1145/3282894.3282914]
34. Marston HR, Hadley R, Pike G, Hesketh I. Police J 2020 Jul 06:1-16. [doi: 10.1177/0032258x20937327]
35. Cole JA, Tully MA, Cupples ME. "They should stay at their desk until the work's done": a qualitative study examining perceptions of sedentary behaviour in a desk-based occupational setting. BMC Res Notes 2015 Nov 17;8(1):683 [FREE Full text] [doi: 10.1186/s13104-015-1670-2] [Medline: 26576730]
36. Buckingham SA, Morrissey K, Williams AJ, Price L, Harrison J. The Physical Activity Wearables in the Police Force (PAW-Force) study: acceptability and impact. BMC Public Health 2020 Nov 03;20(1):1645 [FREE Full text] [doi: 10.1186/s12889-020-09776-1] [Medline: 33143665]
37. Velazquez E, Hernandez M. Effects of police officer exposure to traumatic experiences and recognizing the stigma associated with police officer mental health. Policing Int J 2019 Aug 12;42(4):711-724. [doi: 10.1108/pijpsm-09-2018-0147]
38. Huang Y, Benford S, Blake H. Digital interventions to reduce sedentary behaviors of office workers: scoping review. J Med Internet Res 2019 Feb 07;21(2):e11079 [FREE Full text] [doi: 10.2196/11079] [Medline: 30730294]
39. Rossomanno C, Herrick J, Kirk S. A 6-month supervised employer-based minimal exercise program for police officers improves fitness. J Streng Condit Res 2012 Sep 01;26(9):2338-2344. [doi: 10.1519/jsc.0b013e31823f2b64]
40. Buckingham S, Morrissey K, Williams A, Price L, Harrison J. OP67 The physical activity wearables in the police force (PAW-force) trial: feasibility, acceptability and impact. J Epidemiol Community Health 2019;73(1):32-33. [doi: 10.1136/jech-2019-SSMabstracts.68]
41. Fitzhugh H. Does mindfulness work in policing? GOV.UK Blog: What Works. 2020 Mar 02. URL: https://whatworks.blog.gov.uk/2020/03/02/does-mindfulness-work-in-policing/ [accessed 2021-02-17]
42. Blog: Backup Buddy: Police mental health app. National Police Wellbeing Service. 2020 Aug 26. URL: https://oscarkilo.org.uk/blog-backup-buddy-police-mental-health-app/ [accessed 2021-08-04]
43. 87% app pilot study participant information. Police Federation. URL: https://www.polfed.org/87percentpilotstudyterms [accessed 2021-02-17]
44. National Police Wellbeing Service Introduction. National Police Wellbeing Service. URL: https://oscarkilo.org.uk/national-police-wellbeing-service-2/ [accessed 2021-02-02]
45. Wellbeing in policing: Blue light wellbeing framework. University of Central Lancashire and Lancashire Constabulary. 2018. URL: https://oscarkilo.org.uk/app/uploads/2019/04/BLWF-Interim-Report-1.pdf [accessed 2021-02-25]
46. Antezana G, Venning A, Blake V, Smith D, Winsall M, Orlowski S, et al. An evaluation of behaviour change techniques in health and lifestyle mobile applications. Health Informatics J 2020 Mar 30;26(1):104-113 [FREE Full text] [doi: 10.1177/1460458218813726] [Medline: 30501364]
47. Pagoto S, Bennett GG. How behavioral science can advance digital health. Transl Behav Med 2013 Sep;3(3):271-276 [FREE Full text] [doi: 10.1007/s13142-013-0234-z] [Medline: 24073178]
48. Yardley L, Morrison L, Muller I, Bradbury K. Maximizing user engagement with behavior change interventions. In: Hagger M, Cameron K, Hamilton L, Hankonen N, Lintunen T, editors. The Handbook of Behavior Change (Cambridge Handbooks in Psychology). Cambridge: Cambridge University Press; Jul 23, 2020:361-371.
49. Force Policy - Shift working and working time for police staff. West Yorkshire Police. URL: https://www.westyorkshire.police.uk/sites/default/files/2017-12/shift_working_and_working_time_for_police_staff_0.pdf [accessed 2021-08-04]
50. Brox E, Konstantinidis ST, Evertsen G. User-centered design of serious games for older adults following 3 years of experience with exergames for seniors: a study design. JMIR Serious Games 2017 Jan 11;5(1):e2 [FREE Full text] [doi: 10.2196/games.6254] [Medline: 28077348]
51. Altmeyer M, Lessel P, Krüger A. Investigating gamification for seniors aged 75+. In: Proceedings of the 2018 Designing Interactive Systems Conference. 2018 Jun 9 Presented at: DIS '18: Designing Interactive Systems Conference 2018; June 9 - 13, 2018; Hong Kong China p. 453-458. [doi: 10.1145/3196709.3196799]
52. Pickering K, Pringle A. Gamification for physical activity behaviour change. Perspect Public Health 2018 Nov;138(6):309-310. [doi: 10.1177/1757913918801447] [Medline: 30412020]

XSL•FO
RenderX

53. Thomas S, Pulman A, Dogan H, Jiang N, Thomas P, Passmore D, et al. Creating a digital toolkit to reduce fatigue and promote quality of life in multiple sclerosis: mixed methods. JMIR Preprints. Preprint posted online April 9, 2020 [FREE Full text] [doi: 10.2196/preprints.19230]
54. Attwood S, Parke H, Larsen J, Morton KL. BMC Public Health 2017 Dec 17;17(1):394 [FREE Full text] [doi: 10.1186/s12889-017-4358-9] [Medline: 28511698]
55. Morse J. Qualitative Health Research: Creating a New Discipline. Walnut Creek, CA: Left Coast Press; Jun 15, 2012:1-176.
56. Police Workforce, England and Wales, 31 March 2016. Home Office Statistical Bulletin 05/16. 2016 Jul 21. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/544849/hosb0516-police-workforce.pdf [accessed 2021-08-04]

Abbreviations

BCT: behavioral change technique
CPTSD: complex posttraumatic stress disorder
mHealth: mobile health
PTSD: posttraumatic stress disorder

Edited by G Eysenbach; submitted 19.02.21; peer-reviewed by A Akinbi, T Saheb; comments to author 10.03.21; revised version received 01.06.21; accepted 06.07.21; published 31.08.21.

Please cite as:
Swanston E, Pulman A, Dogan H, Murphy J, Bitters F
Scoping the Need for a Tailored mHealth App to Improve Health and Well-being Behavioral Transformation in the Police: Exploring the Views of UK Police Workers via Web-Based Surveys and Client Meetings
JMIR Form Res 2021;5(8):e28075
URL: https://formative.jmir.org/2021/8/e28075
doi:10.2196/28075
PMID:

©Emma Swanston, Andy Pulman, Huseyin Dogan, Jane Murphy, Fiona Bitters. Originally published in JMIR Formative Research (https://formative.jmir.org), 31.08.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Mobile Health App (WYZ) for Engagement in Care and Antiretroviral Therapy Adherence Among Youth and Young Adults Living With HIV: Single-Arm Pilot Intervention Study

Parya Saberi[1], PharmD, MAS; Nadra E Lisha[2], PhD; Xavier A Erguera[1]; Estie Sid Hudes[1], PhD; Mallory O Johnson[1], PhD; Theodore Ruel[3], MD; Torsten B Neilands[1], PhD

[1]Department of Medicine, University of California, San Francisco, San Francisco, CA, United States

[2]Center for Tobacco Control Research and Education, University of California, San Francisco, San Francisco, CA, United States

[3]Department of Pediatrics, University of California, San Francisco, San Francisco, CA, United States

Corresponding Author:
Parya Saberi, PharmD, MAS
Department of Medicine
University of California, San Francisco
550 16th street
San Francisco, CA, 94143
United States
Phone: 1 415 502 1000 ext 17171
Email: Parya.Saberi@ucsf.edu

Abstract

Background: Youth are globally recognized as being vulnerable to HIV. Younger age has been correlated with worse health outcomes. Mobile health (mHealth) interventions have the potential to interact with youth where they are, using a device they already access.

Objective: Using predefined benchmarks, we sought to evaluate the feasibility and acceptability of WYZ, an mHealth app, for improved engagement in care and antiretroviral therapy (ART) adherence among youth and young adults living with HIV. WYZ was designed and developed with input from youth and young adults living with HIV using a human-centered design approach and was based on the information, motivation, and behavioral skills framework to address common barriers to care and ART adherence among youth and young adults living with HIV.

Methods: We recruited youth and young adults living with HIV (18-29 years old) from the San Francisco Bay Area to take part in a 6-month pilot trial. Their participation included completing baseline and exit surveys, and participating in seven phone check-ins about their use of WYZ.

Results: Youth and young adults living with HIV (N=79) reported high levels of feasibility and acceptability with WYZ use. We met predefined benchmarks for recruitment (79/84, 94%), mean logins per week (5.3), tracking ART adherence (5442/9393, 57.9%), posting chat topics per week (4.8), and app crashes reported per week (0.24). The ease of app download, install, and setup, and comfort with security, privacy, and anonymity were highly rated (all over 91%). Additionally, participants reported high satisfaction for a research project that was remotely conducted. Participants used the app for shorter timeframes compared to the predefined benchmark.

Conclusions: We noted high feasibility and acceptability with WYZ. Further research to examine the efficacy of WYZ will enable youth and young adults living with HIV and their providers to make informed decisions when using, recommending, and prescribing the app for improved engagement in HIV care and ART adherence.

Trial Registration: ClinicalTrials.gov NCT03587857; https://clinicaltrials.gov/ct2/show/NCT03587857

(JMIR Form Res 2021;5(8):e26861) doi:10.2196/26861

KEYWORDS

youth living with HIV; mobile health; mobile app; engagement in care; antiretroviral therapy adherence; pilot

XSL•FO
RenderX

Introduction

In the United States, youth and young adults carry a significant burden of HIV. Youth and young adults living with HIV experience disparities at all steps of the HIV care continuum, including higher HIV incidence, lower linkage and retention in care, suboptimal antiretroviral therapy (ART) adherence, and lower virologic suppression [1-6]. The consequences of continued disparities include poor health outcomes, development and transmission of drug-resistant viruses, a future generation of adults who are more susceptible to developing AIDS, and further widening of these health disparities. Youth and young adults living with HIV experience many individual, structural (eg, transition to adult health care, inexperience with medical systems, and lack of insurance), social (eg, poverty, unstable housing, food insecurity, social isolation, and stigma), and biological (cognitive developmental stages) challenges that impact their abilities to access and adhere to oral ART [2,3,7-9]. However, there are few effective and tailored interventions that address ART adherence and engagement in HIV care for youth and young adults living with HIV.

In the United States, over 96% of youth and young adults living with HIV own smartphones [10], over two-thirds have downloaded mobile health (mHealth) apps [11], and over 90% are social media users [12]. The nearly ubiquitous access to and use of smartphones represents a powerful platform for the delivery of mHealth interventions to this population. Additionally, given the reduction in transportation costs, time constraints, potential stigma associated with participation in in-person HIV research [13], and missing data, mHealth technology can surmount common barriers, increasing the reach and generalizability of findings. Several mHealth apps are in various stages of development for people living with HIV [14-17], as we have previously summarized [18]. However, despite technology-based behavioral interventions showing promise in older adults living with HIV [19], few interventions have shown efficacy in addressing the unique aspects of youth developmental phases, youth culture, and gravitation of youth toward the use of technology [20]. In this study, we pilot tested an mHealth app to address barriers to engagement in care among youth and young adults living with HIV.

Methods

Study Design and Sample

From July 2019 to May 2020, we conducted a 6-month single-arm pilot study to evaluate the feasibility and acceptability of an mHealth app, named WYZ (pronounced “wise”), to address barriers to engagement in HIV care among individuals aged 18 to 29 years living with HIV in the San Francisco Bay Area [18]. WYZ was designed and developed using a human-centered design (HCD) approach [21-24]; formative research with youth and young adults living with HIV [18,20,25-27]; the information, motivation, and behavioral skills (IMB) [28-30] framework; and mHealth designers and developers from the University of California, San Francisco (UCSF) School of Medicine Technology team (SOM Tech). HCD focuses on creating approaches and delivering solutions to problems based on efforts to understand the specific needs and perspectives of the users. Therefore, HCD seeks to gain insights into the needs of the beneficiaries of an innovation, and creates approaches and delivers solutions to meet their needs.

Details of WYZ design and development, as well as the pilot study protocol, have previously been published [18]. In short, WYZ contains three main features, My Health, My Team, and My Community. My Health allows users to keep track of their ART medication information, visualize their adherence and laboratory data, and understand their health; My Team provides community resources and facilitates communication with health care team members; and My Community allows for social support from peers through anonymous and moderated discussion forums and allows users to stay up-to-date on health-related news. These features were developed with guidance from youth and young adults living with HIV and further refined through focus groups with youth and young adults living with HIV and iterative field testing with our Youth Advisory Panel (YAP), and were chosen to address specific barriers to ART adherence and engagement in HIV care (eg, social isolation and lack of community support).

WYZ design, development, and technological support were provided by UCSF’s SOM Tech. To ensure Health Insurance Portability and Accountability Act (HIPAA) compliance, we used Salesforce as the backend service and for storing sensitive data in a secure cloud-based database. Data about app usage were collected using Flurry (a mobile analytics tool) and Salesforce analytics. To enhance the security and privacy of WYZ, we used a two-step authentication process for downloading, password protection (with each log in), aliases, deletion of all communications over 30 days old, and remote revocation of app access in case of theft, loss, or misuse.

Participants were recruited using various strategies, including flyers at clinics and community-based organizations, emails to clinicians at clinics serving youth and young adults living with HIV, peer referral, and contacting prior study participants who had consented to being notified of future research. Information about the study was also disseminated through the YAP.

Individuals aged 18 to 29 years living with HIV, who lived or received medical care in the San Francisco Bay Area, spoke English, and had access to an Android or iOS smartphone, were included. Those with any evidence of severe cognitive impairment or active psychosis that impeded their ability to provide informed consent were excluded. To confirm an individual’s age and HIV serostatus, the potential participant text messaged a photo identification showing their date of birth and either a clinician’s letter of HIV diagnosis, a copy of laboratory test results (for HIV antibody or HIV viral load), or their ART medication vial. These photos were sent via text message to an encrypted and secure study phone for verification by study staff.

All study activities, including recruitment, screening, enrollment, study assessments, provision of incentives, and exit interviews, were conducted remotely using text message, telephone, email, and videoconference. Participants received a check-in at weeks 1, 2, and 4, followed by monthly check-ins, and up to US $215 for completion of all study activities. All procedures were

XSL•FO
RenderX

reviewed and approved by the UCSF Institutional Review Board with a requirement for electronic consent. At baseline and 6 months, participants completed study assessments using a Qualtrics survey.

Measures

Demographics

Demographic data, including date of birth, sex at birth, sexual identity, race/ethnicity, perceived financial security, and work status (full time, part time, or not working), were collected.

Feasibility Metrics

Feasibility metrics were collected using Flurry and Salesforce analytics. Metrics were based on predefined thresholds [18], including how many people were recruited for the study, mean logins to the app, mean minutes in the app, and use of specific features in the app.

Acceptability Metrics

Acceptability metrics were collected using a Qualtrics survey administered during the last study visit at 6 months. The survey included questions related to satisfaction with WYZ, ease of WYZ use, and satisfaction with the study. Additionally, we asked participants about WYZ acceptability using the System Usability Scale (SUS), with scores ranging from 0 to 100 and scores >68 being considered above average [31,32]. A threshold of 70% or greater satisfaction on all questions was used to determine acceptability.

HIV and Psychosocial Outcomes

HIV and psychosocial outcomes were measured at baseline and 6 months. These included self-reported HIV viral load (detectable or undetectable) [33], self-reported ART adherence [34], depression (Patient Health Questionnaire-9 [PHQ-9]) [35], resilience [36], social provisions [37], social isolation (Patient-Reported Outcomes Measurement Information System [PROMIS]) [38], health care empowerment [39], and unmet subsistence needs and instrumental support [40].

Data Analysis

Descriptive statistics of the baseline demographics of WYZ study participants were calculated. Next, we examined descriptive statistics for feasibility metrics and compared them to predefined benchmarks. We then calculated frequencies for all of our acceptability metrics measured at the exit survey. Lastly, frequencies for HIV and psychosocial outcomes were calculated at baseline and the exit survey (6 months). For these data evaluated at both baseline and 6 months, we compared data from those who were retained in the study until 6 months and the entire group to examine divergent results. Given that this was a pilot study with limited statistical power, based on guidance from the National Institutes of Health and literature regarding wide confidence intervals and instability of effect sizes from pilot studies, tests of statistical significance and the efficacy of the intervention to compare HIV clinical outcomes preintervention and postintervention were not evaluated [41-44]. All analyses were completed using SAS 9.2 (SAS Institute).

Results

Demographics

Demographics of the study participants are presented in Table 1. At baseline, there were 79 participants (mean age 26.9 years, SD 2.9 years). Among the 79 participants, 69 (87%) identified as male, 48 (61%) identified as gay, 33 (43%) identified as Latino, and 16 (21%) identified as Black. Although nearly 57% (45/79) of participants were working, financial insecurity was relatively common, as 62% (49/79) of participants noted “barely getting by” or “not getting by” on the money they have.

Feasibility metrics are presented in Table 2. Of the 92 individuals who were screened, 84 were eligible. Of these 84 individuals, 79 (94%) consented to participate in the study, and 69 (87%) of those who enrolled completed the exit survey at 6 months. All predefined benchmarks were met (Table 2), except for mean minutes in the app per week (benchmark=15 min/week, actual=8.7 min/week). The mean number of logins per week was 5.3 (SD 5.6). In My Health, ART adherence tracking was conducted in 57.9% (5442/9393) of the inquiries. Moreover, the mean number of postings of chat topics on the My Community chat per person per week was 4.8 (range, 1-42), and the number of reported app crashes was less than once per week (0.24).

XSL•FO
RenderX

Table 1. Baseline demographics of the study participants.

Demographic	Value[a] (N=79)
Age (years), mean (SD)	26.9 (2.9)
Sex at birth, n (%)	
Male	69 (87)
Female	9 (11)
Decline to answer	1 (1)
Sexual identity, n (%)	
Gay	48 (61)
Heterosexual	8 (10)
Other	23 (16)
Race/ethnicity, n (%)	
Asian	3 (4)
Black	16 (21)
Latino	33 (43)
Other	13 (17)
White	11 (15)
Financial security, n (%)	
I have enough money to live comfortably	24 (30)
I can barely get by on the money I have	40 (51)
I cannot get by on the money I have	9 (11)
Decline to answer	6 (8)
Work status, n (%)	
Full time	28 (35)
Part time	17 (22)
Not working	28 (36)
Other	4 (5)
Decline to answer	2 (3)

[a]Missing n ranged from 0 to 6 for each item.

XSL•FO
RenderX

Table 2. Feasibility metrics, prespecified threshold for each metric, and actual outcome.

Metric and threshold	Actual value
General	
Screened, N	92
Eligible (N=92), n (%)	84 (91)
Recruited (N=84) (threshold: ≥70%, target N=80 [ie, ≥56]), n (%)	79 (94)
Logins (threshold: 1 login/week), mean (SD)	5.3 (5.6)
Minutes in the app (threshold: 15 min/week), mean (SD)	8.7 (5.0)
My Health	
ART[a] adherence tracking (threshold: ≥3 times/week [ie, ≥43%]), mean/week	58%
My Community	
Post chat topic (threshold: 1 chat topic/person/week), mean number/person/week (range)	4.8 (1-42)
Administrative	
Reported app crashes (threshold: <1 app crash/week), mean number/week	0.24

[a]ART: antiretroviral therapy.

Acceptability metrics are presented in Table 3. Among the 69 participants who completed the study, 77% (n=53) rated their overall experience with the app as excellent to very good, 91% (n= 63) reported the app to be extremely to somewhat easy to download and install, and 96% (n=66) reported that WYZ setup was extremely to somewhat easy. All participants reported being extremely to somewhat comfortable with the security, privacy, and anonymity of WYZ. Moreover, approximately 83% (n=57) stated that they would be extremely to somewhat likely to continue to use WYZ and 94% (n=64) were extremely to somewhat likely to participate in a similar study in the future. Furthermore, 86% (n=59) of participants rated their overall experience with participation in the WYZ study as excellent to very good and 90% (n=62) reported excellent to very good experience with participating in a completely remotely conducted study. The mean SUS score was 75.6, which is considered to be well above average.

HIV and psychosocial metrics are presented in Table 4. At baseline and 6 months, 9% (7/79) and 4% (3/69) of participants, respectively, reported a detectable HIV viral load. During this time, self-reported ART adherence was unchanged. From baseline to 6 months, participants reporting mild depressive symptoms decreased by 9% (30/79, 38% to 19/66, 29%). Moreover, the mean social isolation score decreased by 12.1 points. Overall, we did not note divergent patterns with regard to the HIV and psychosocial metrics between baseline data from the entire sample (N=79) and those who were retained until 6 months (N=69).

XSL•FO
RenderX

Table 3. Acceptability metrics exit survey findings.

Metrics	Value[a] (N=69), n (%)
How would you rate your overall experience with the WYZ app?	
Excellent to very good	53 (77)
Good	0 (0)
Fair	2 (3)
Poor to very poor	14 (20)
How easy or difficult was it to download and install WYZ on your phone?	
Extremely to somewhat easy	63 (91)
Extremely to somewhat difficult	6 (9)
How easy or difficult was it to setup WYZ (ie, create a personal passcode, set med reminders, enter your pharmacy information, etc)?	
Extremely to somewhat easy	66 (96)
Extremely to somewhat difficult	3 (4)
How satisfied were you with the visual design of WYZ?	
Extremely to somewhat satisfied	56 (81)
Extremely to somewhat unsatisfied	12 (19)
Overall, how helpful was the My Health module in supporting your ability to take your medications as prescribed?	
Extremely to very helpful	54 (81)
Moderately helpful	7 (10)
A little to not at all helpful	6 (9)
How satisfied were you with medication and refill reminders?	
Extremely to somewhat satisfied	50 (78)
Extremely to somewhat unsatisfied	14 (22)
How satisfied were you with the adherence calendar?	
Extremely to somewhat satisfied	57 (88)
Extremely to somewhat unsatisfied	8 (12)
How satisfied were you with the lab visualization tools (ie, CD4 graph and viral load graphs)?	
Extremely to somewhat satisfied	41 (82)
Extremely to somewhat unsatisfied	9 (18)
Overall, how helpful was the My Team module in supporting your ability to seek and connect to support and services that you need?	
Extremely to very helpful	40 (62)
Moderately helpful	18 (28)
A little to not at all helpful	7 (11)
How satisfied were you with the listing of community resources (ie, list of local resources and organizations in your area)?	
Extremely to somewhat satisfied	64 (100)
Extremely to somewhat unsatisfied	0 (0)
How satisfied were you with the secure messaging feature with your health care team (ie, being able to send a message to your health care team through WYZ)?	
Extremely to somewhat satisfied	45 (82)
Extremely to somewhat unsatisfied	10 (18)
How would you rate your comfort level with asking your health care providers to support your participation in the study by agreeing to be added to My Team?	
Extremely to somewhat comfortable	51 (88)
Extremely to somewhat uncomfortable	7 (12)

XSL•FO
RenderX

Metrics	Value[a] (N=69), n (%)
Overall, how helpful was the My Community module in supporting your ability to feel connected to other youth living with HIV?	
Extremely to very helpful	34 (52)
Moderately helpful	20 (30)
A little to not at all helpful	12 (15)
How satisfied were you with the news feature within WYZ (ie, being able to see the latest HIV and health news in WYZ)?	
Extremely to somewhat satisfied	52 (84)
Extremely to somewhat unsatisfied	10 (16)
How satisfied were you with the private calendar (ie, being able to add appointments and community events within WYZ)?	
Extremely to somewhat satisfied	44 (70)
Extremely to somewhat unsatisfied	19 (30)
Overall, how helpful was the private calendar in supporting your ability to keep your health appointments?	
Extremely to very helpful	35 (56)
Moderately helpful	14 (20)
A little to not at all helpful	13 (23)
What is your comfort level with the security, privacy, and anonymity provided by WYZ?	
Extremely to somewhat comfortable	67 (100)
Extremely to somewhat uncomfortable	0 (0)
How likely are you to continue to use WYZ after your participation in the study ends?	
Extremely to somewhat likely	57 (83)
Extremely to somewhat unlikely	12(17)
In the future, how likely would you be to participate in a similar study where you are asked to use a mobile health app like WYZ on a regular basis?	
Extremely to somewhat likely	64 (94)
Extremely to somewhat unlikely	4 (6)
How would you rate your overall experience with participation in the WYZ study?	
Excellent to very good	59 (86)
Good	10 (15)
Fair	0 (0)
Poor to very poor	0 (0)
How would you rate your experience participating in a study where everything was conducted remotely (ie, you did not have to come into a clinic or office to complete study activities)?	
Excellent to very good	62 (91)
Good	5 (7)
Fair	1 (1)
Poor to very poor	0 (0)
How would you rate your experience with having to schedule and complete regular check-ins over the phone with a study coordinator?	
Excellent to very good	56 (82)
Good	10 (15)
Fair	2 (3)
Poor to very poor	0 (0)
How would you rate your experience with receiving compensation for your study participation using a reloadable debit card (ie, ClinCard)?	
Excellent to very good	60 (87)
Good	6 (9)

XSL•FO
RenderX

Metrics	Value[a] (N=69), n (%)
Fair	3 (4)
Poor to very poor	0 (0)
How easy or difficult was it to remember to use WYZ regularly for the 6 months of your study participation?	
Extremely to somewhat easy	53 (77)
Extremely to somewhat difficult	16 (23)
How helpful was the communication with study staff?	
Extremely to very helpful	66 (96)
Moderately helpful	2 (3)
A little to not at all helpful	1 (1)

[a]Missing n ranged from 0 to 14 for each item.

Table 4. HIV and psychosocial outcomes at baseline and 6 months.

Outcome	Baseline (N=79)	Baseline (N=69)[a]	6 months (N=69)
HIV viral load, n (%)			
Detectable	7 (9)	6 (9)	3 (4)
Undetectable	70 (89)	59 (88)	60 (87)
Do not know	2 (3)	2 (3)	2 (3)
Decline to answer	0 (0)	0 (0)	4 (6)
ART[b] adherence, mean (SD)	85.0 (16.6)	86.5 (15.3)	85.0 (16.9)
Depression (PHQ-9[c]), n (%)			
No to minimal depression (0-4)	22 (28)	17 (25)	26 (39)
Mild depression (5-9)	30 (38)	26 (38)	19 (29)
Moderate depression (10-14)	12 (15)	10 (15)	12 (18)
Severe to moderately severe depression (15-27)	15 (19)	14 (21)	9 (14)
Resilience, mean (SD)	3.6 (0.7)	3.5 (0.8)	3.6 (0.7)
Social support, mean (SD)	37.9 (6.0)	23.1 (5.3)	37.7 (7.6)
Social isolation, mean (SD)	50.0 (10.0)	35.4 (12.3)	37.9 (7.0)
Health care empowerment, mean (SD)	17.4 (3.1)	18.6 (3.3)	17.4 (3.3)
Relationship with health care provider, mean (SD)	1.4 (0.6)	1.4 (0.6)	1.40 (0.6)
Unmet subsistence needs and instrumental support score, mean (SD)	0.61 (1.2)	0.65 (1.2)	0.67 (1.2)
Unmet subsistence needs and instrumental support, n (%)			
0	54 (70)	46 (70)	37 (65)
1	11 (14)	9 (14)	11 (19)
2	6 (8)	4 (6)	3 (5)
≥3	7 (8)	7 (11)	6 (11)

[a]Baseline data for participants who were retained until the end of the study (6 months).

[b]ART: antiretroviral therapy.

[c]PHQ-9: Patient Health Questionnaire-9.

Discussion

Principal Findings

The use of WYZ was highly feasible and acceptable among youth and young adults living with HIV in the San Francisco Bay Area. We met predefined benchmarks for recruitment, mean logins per week, tracking ART adherence, posting chat topics, and app crashes reported. The ease of app download, installation, and setup, and the overall comfort with security, privacy, and anonymity were highly rated. Additionally, participants reported high satisfaction for a research project that was remotely

XSL•FO
RenderX

conducted. These findings demonstrate high potential for uptake and app functionality, indicating a promising role for WYZ as an intervention for engagement in HIV care and ART adherence among youth and young adults living with HIV.

Participants used the app for shorter timeframes than were predefined; however, our predefined benchmark may have been an overestimate. Additionally, due to the ability to log ART adherence using out-of-app notifications, some of the interactions with WYZ were not captured in the analytical tools used. In the next phase of this study, we will ask participants to further elaborate about app use during exit qualitative interviews.

Small changes in self-reported HIV and psychosocial metrics from baseline to 6 months highlight the limitations of pilot studies, in which examination of the intervention's "preliminary impact" is not meaningful due to wide confidence intervals [41-44]. However, we noted improvements in the social isolation score, which, along with the high level of activity in the My Community Chat section, underscore the importance of this feature and deserve further evaluation in future research.

There are currently few mHealth apps in the early stages of development and pilot testing for enhanced engagement in HIV care, ART adherence, and communication with health care teams for people living with HIV [14-17]. The limitations of some of these mHealth apps include lack of specification of a theoretical framework, limited feasibility and acceptability metrics with no predefined benchmarks, small sample size (N<30), wide age range (≥18 years), and availability for either iOS or Android (not both). We have previously summarized these studies [18]. In developing and pilot testing WYZ, we have addressed these limitations.

In this pilot study, we were able to recruit a diverse group of participants with regard to race/ethnicity; however, participants were mainly gay cis-gender men. The other limitations of our study include a single-arm design (ie, no control group) and a relatively small convenience sample of participants from the San Francisco Bay Area who had access to a smartphone and most of who had an undetectable HIV viral load; therefore, study findings may not be generalizable to other populations. The loss to follow-up was approximately 13%, which is lower than estimates among youth and young adults living with HIV in the HIV Research Network (20%) [45] and in other studies in this population (up to 55%) [46]. We believe that the relatively low loss to follow-up may have been due to the fact that this research was conducted completely remotely, which allowed for flexibility for participation. Since the completion of this pilot study, we have resolved all minor bugs and smartphone compatibility challenges. Additionally, we are updating My Health for those who may use long-acting injectables in the near future and the My Team resources section based on user geolocation.

Conclusion

Youth and young adults living with HIV represent a population that is disproportionately impacted by HIV and requires tailored youth-friendly interventions. There is a dearth of technology-based interventions that address the changing needs of youth and young adults living with HIV. In future research, we will examine the efficacy and effectiveness of WYZ in improving engagement in HIV care and ART adherence among a larger sample of youth and young adults living with HIV taking into account findings from this study. Given the speed of technological advancement and the need for evidence-based solutions for improved HIV health outcomes among youth and young adults living with HIV, we believe that more funding should be allocated to technology-based interventions to move the National Institutes of Health's Behavioral and Social Sciences Research agenda forward.

Acknowledgments

The authors' work was supported by the National Institute of Mental Health (award number R34MH114604). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Authors' Contributions

PS conceptualized the study and received grant funding. PS and XAE collected the data. PS, ESH, and NEL ran the analyses. PS and NEL wrote the first draft of the study. TBN guided the design study and analysis plan. TR and MOJ guided grant funding and study design. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Bekker L, Johnson L, Wallace M, Hosek S. Building our youth for the future. J Int AIDS Soc 2015;18(2 Suppl 1):20027 [FREE Full text] [doi: 10.7448/IAS.18.2.20027] [Medline: 25724512]
2. Kahana SY, Jenkins RA, Bruce D, Fernandez MI, Hightow-Weidman LB, Bauermeister JA, Adolescent Medicine Trials Network for HIV/AIDS Interventions. Structural Determinants of Antiretroviral Therapy Use, HIV Care Attendance, and Viral Suppression among Adolescents and Young Adults Living with HIV. PLoS One 2016 Apr 1;11(4):e0151106 [FREE Full text] [doi: 10.1371/journal.pone.0151106] [Medline: 27035905]

XSL•FO
RenderX

3. Zanoni BC, Mayer KH. The adolescent and young adult HIV cascade of care in the United States: exaggerated health disparities. AIDS Patient Care STDS 2014 Mar;28(3):128-135 [FREE Full text] [doi: 10.1089/apc.2013.0345] [Medline: 24601734]
4. HIV and Youth. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/hiv/group/age/youth/index.html [accessed 2021-08-16]
5. Lally M, van den Berg J, Westfall A, Rudy B, Hosek S, Fortenberry J, Adolescent Medicine Trials Network for HIV/AIDS Interventions (ATN). HIV Continuum of Care for Youth in the United States. J Acquir Immune Defic Syndr 2018 Jan 01;77(1):110-117 [FREE Full text] [doi: 10.1097/QAI.0000000000001563] [Medline: 28991884]
6. Wood SM, Lowenthal E, Lee S, Ratcliffe SJ, Dowshen N. Longitudinal Viral Suppression Among a Cohort of Adolescents and Young Adults with Behaviorally Acquired Human Immunodeficiency Virus. AIDS Patient Care STDS 2017 Sep;31(9):377-383 [FREE Full text] [doi: 10.1089/apc.2017.0078] [Medline: 28891717]
7. Dobroszycki J, Lee P, Romo DL, Rosenberg MG, Wiznia A, Abadi J. Antiretroviral therapy in HIV-infected adolescents: clinical and pharmacologic challenges. Expert Rev Clin Pharmacol 2017 May 14;10(5):509-516. [doi: 10.1080/17512433.2017.1301205] [Medline: 28288535]
8. Rotheram-Borus MJ, Lee S, Swendeman D. Getting to Zero HIV Among Youth: Moving Beyond Medical Sites. JAMA Pediatr 2018 Dec 01;172(12):1117-1118 [FREE Full text] [doi: 10.1001/jamapediatrics.2018.3672] [Medline: 30326015]
9. Reeder C, Neilands TB, Palar K, Saberi P. Food Insecurity and Unmet Needs Among Youth and Young Adults Living With HIV in the San Francisco Bay Area. J Adolesc Health 2019 Aug;65(2):262-266 [FREE Full text] [doi: 10.1016/j.jadohealth.2019.02.023] [Medline: 31196781]
10. Mobile Fact Sheet. Pew Research Center. 2021. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2020-02-25]
11. Rideout V, Fox S. Digital Health Practices, Social Media Use, and Mental Well-Being Among Teens and Young Adults in the U.S. Hopelab and Well Being Trust. 2018. URL: https://assets.hopelab.org/wp-content/uploads/2020/08/a-national-survey-by-hopelab-and-well-being-trust-2018.pdf [accessed 2020-12-07]
12. Social Media Fact Sheet. Pew Research Center. 2021. URL: https://www.pewresearch.org/internet/fact-sheet/social-media/ [accessed 2020-03-24]
13. DiClemente R, Ruiz M, Sales J. Barriers to adolescents' participation in HIV biomedical prevention research. J Acquir Immune Defic Syndr 2010 Jul;54 Suppl 1:S12-S17 [FREE Full text] [doi: 10.1097/QAI.0b013e3181e1e2c0] [Medline: 20571418]
14. Escobar-Viera C, Zhou Z, Morano JP, Lucero R, Lieb S, McIntosh S, et al. The Florida Mobile Health Adherence Project for People Living With HIV (FL-mAPP): Longitudinal Assessment of Feasibility, Acceptability, and Clinical Outcomes. JMIR Mhealth Uhealth 2020 Jan 08;8(1):e14557 [FREE Full text] [doi: 10.2196/14557] [Medline: 31913127]
15. Flickinger TE, Ingersoll K, Swoger S, Grabowski M, Dillingham R. Secure Messaging Through PositiveLinks: Examination of Electronic Communication in a Clinic-Affiliated Smartphone App for Patients Living with HIV. Telemed J E Health 2020 Mar 01;26(3):359-364 [FREE Full text] [doi: 10.1089/tmj.2018.0261] [Medline: 30900961]
16. Whiteley L, Brown LK, Mena L, Craker L, Arnold T. Enhancing health among youth living with HIV using an iPhone game. AIDS Care 2018 Jan 09;30(sup4):21-33 [FREE Full text] [doi: 10.1080/09540121.2018.1503224] [Medline: 30626196]
17. Hightow-Weidman L, Muessig K, Knudtson K, Srivatsa M, Lawrence E, LeGrand S, et al. A Gamified Smartphone App to Support Engagement in Care and Medication Adherence for HIV-Positive Young Men Who Have Sex With Men (AllyQuest): Development and Pilot Study. JMIR Public Health Surveill 2018 Apr 30;4(2):e34 [FREE Full text] [doi: 10.2196/publichealth.8923] [Medline: 29712626]
18. Erguera XA, Johnson MO, Neilands TB, Ruel T, Berrean B, Thomas S, et al. WYZ: a pilot study protocol for designing and developing a mobile health application for engagement in HIV care and medication adherence in youth and young adults living with HIV. BMJ Open 2019 May 05;9(5):e030473 [FREE Full text] [doi: 10.1136/bmjopen-2019-030473] [Medline: 31061063]
19. Kempf M, Huang C, Savage R, Safren SA. Technology-Delivered Mental Health Interventions for People Living with HIV/AIDS (PLWHA): a Review of Recent Advances. Curr HIV/AIDS Rep 2015 Dec 9;12(4):472-480 [FREE Full text] [doi: 10.1007/s11904-015-0292-6] [Medline: 26452644]
20. Saberi P, Ming K, Dawson-Rose C. What does it mean to be youth-friendly? Results from qualitative interviews with health care providers and clinic staff serving youth and young adults living with HIV. Adolesc Health Med Ther 2018;9:65-75 [FREE Full text] [doi: 10.2147/AHMT.S158759] [Medline: 29731672]
21. Norman D, Draper S. User Centered System Design: New Perspectives on Human-computer Interaction. Boca Raton, FL: CRC Press; 1986.
22. LeRouge C, Wickramasinghe N. A review of user-centered design for diabetes-related consumer health informatics technologies. J Diabetes Sci Technol 2013 Jul 01;7(4):1039-1056 [FREE Full text] [doi: 10.1177/193229681300700429] [Medline: 23911188]
23. Matheson GO, Pacione C, Shultz RK, Klügl M. Leveraging human-centered design in chronic disease prevention. Am J Prev Med 2015 Apr;48(4):472-479. [doi: 10.1016/j.amepre.2014.10.014] [Medline: 25700655]

XSL•FO
RenderX

24. Kachirskaia I, Mate K, Neuwirth E. Human-Centered Design and Performance Improvement: Better Together. NEJM Catalyst. 2018. URL: https://catalyst.nejm.org/doi/full/10.1056/CAT.18.0144 [accessed 2019-03-14]
25. Saberi P, Yuan P, John M, Sheon N, Johnson MO. A pilot study to engage and counsel HIV-positive African American youth via telehealth technology. AIDS Patient Care STDS 2013 Sep;27(9):529-532 [FREE Full text] [doi: 10.1089/apc.2013.0185] [Medline: 23991691]
26. Saberi P, Mayer K, Vittinghoff E, Naar-King S, Adolescent Medicine Trials Network for HIV/AIDS Interventions. Correlation between use of antiretroviral adherence devices by HIV-infected youth and plasma HIV RNA and self-reported adherence. AIDS Behav 2015 Jan 31;19(1):93-103 [FREE Full text] [doi: 10.1007/s10461-014-0806-z] [Medline: 24879628]
27. 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]
28. Fisher JD, Fisher WA, Misovich SJ, Kimble DL, Malloy TE. Changing AIDS risk behavior: effects of an intervention emphasizing AIDS risk reduction information, motivation, and behavioral skills in a college student population. Health Psychol 1996 Mar;15(2):114-123. [doi: 10.1037//0278-6133.15.2.114] [Medline: 8681919]
29. Amico KR, Toro-Alfonso J, Fisher JD. An empirical test of the information, motivation and behavioral skills model of antiretroviral therapy adherence. AIDS Care 2005 Aug 20;17(6):661-673. [doi: 10.1080/09540120500038058] [Medline: 16036253]
30. Fisher JD, Amico KR, Fisher WA, Harman JJ. The information-motivation-behavioral skills model of antiretroviral adherence and its applications. Curr HIV/AIDS Rep 2008 Nov 25;5(4):193-203. [doi: 10.1007/s11904-008-0028-y] [Medline: 18838059]
31. Lewis JR. The System Usability Scale: Past, Present, and Future. International Journal of Human–Computer Interaction 2018 Mar 30;34(7):577-590. [doi: 10.1080/10447318.2018.1455307]
32. System Usability Scale (SUS). Usability.gov. URL: http://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html [accessed 2021-03-31]
33. Kalichman SC, Rompa D, Cage M. Reliability and validity of self-reported CD4 lymphocyte count and viral load test results in people living with HIV/AIDS. Int J STD AIDS 2000 Sep 25;11(9):579-585. [doi: 10.1258/0956462001916551] [Medline: 10997499]
34. Wilson IB, Lee Y, Michaud J, Fowler FJ, Rogers WH. Validation of a New Three-Item Self-Report Measure for Medication Adherence. AIDS Behav 2016 Nov 20;20(11):2700-2708 [FREE Full text] [doi: 10.1007/s10461-016-1406-x] [Medline: 27098408]
35. Fine TH, Contractor AA, Tamburrino M, Elhai JD, Prescott MR, Cohen GH, et al. Validation of the telephone-administered PHQ-9 against the in-person administered SCID-I major depression module. J Affect Disord 2013 Sep 25;150(3):1001-1007. [doi: 10.1016/j.jad.2013.05.029] [Medline: 23747208]
36. Smith BW, Dalen J, Wiggins K, Tooley E, Christopher P, Bernard J. The brief resilience scale: assessing the ability to bounce back. Int J Behav Med 2008 Sep;15(3):194-200. [doi: 10.1080/10705500802222972] [Medline: 18696313]
37. Mancini JA, Blieszner R. Social provisions in adulthood: concept and measurement in close relationships. J Gerontol 1992 Jan 01;47(1):P14-P20. [doi: 10.1093/geronj/47.1.p14] [Medline: 1730853]
38. Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, PROMIS Cooperative Group. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. J Clin Epidemiol 2010 Nov;63(11):1179-1194 [FREE Full text] [doi: 10.1016/j.jclinepi.2010.04.011] [Medline: 20685078]
39. Johnson MO, Rose CD, Dilworth SE, Neilands TB. Advances in the conceptualization and measurement of Health Care Empowerment: development and validation of the Health Care Empowerment inventory. PLoS One 2012 Sep 19;7(9):e45692 [FREE Full text] [doi: 10.1371/journal.pone.0045692] [Medline: 23029184]
40. Chen WT, Wantland D, Reid P, Corless IB, Eller LS, Iipinge S, et al. Engagement with Health Care Providers Affects Self-Efficacy, Self-Esteem, Medication Adherence and Quality of Life in People Living with HIV. J AIDS Clin Res 2013 Nov 01;4(11):256 [FREE Full text] [doi: 10.4172/2155-6113.1000256] [Medline: 24575329]
41. Leon AC, Davis LL, Kraemer HC. The role and interpretation of pilot studies in clinical research. J Psychiatr Res 2011 May;45(5):626-629 [FREE Full text] [doi: 10.1016/j.jpsychires.2010.10.008] [Medline: 21035130]
42. Whitehead AL, Sully BG, Campbell MJ. Pilot and feasibility studies: is there a difference from each other and from a randomised controlled trial? Contemp Clin Trials 2014 May;38(1):130-133. [doi: 10.1016/j.cct.2014.04.001] [Medline: 24735841]
43. Kraemer HC, Mintz J, Noda A, Tinklenberg J, Yesavage JA. Caution regarding the use of pilot studies to guide power calculations for study proposals. Arch Gen Psychiatry 2006 May 01;63(5):484-489. [doi: 10.1001/archpsyc.63.5.484] [Medline: 16651505]
44. Eldridge SM, Lancaster GA, Campbell MJ, Thabane L, Hopewell S, Coleman CL, et al. Defining Feasibility and Pilot Studies in Preparation for Randomised Controlled Trials: Development of a Conceptual Framework. PLoS One 2016 Mar 15;11(3):e0150205 [FREE Full text] [doi: 10.1371/journal.pone.0150205] [Medline: 26978655]

XSL•FO
RenderX

45. Agwu AL, Lee L, Fleishman JA, Voss C, Yehia BR, Althoff KN, et al. Aging and loss to follow-up among youth living with human immunodeficiency virus in the HIV Research Network. J Adolesc Health 2015 Mar;56(3):345-351 [FREE Full text] [doi: 10.1016/j.jadohealth.2014.11.009] [Medline: 25703322]
46. Farmer C, Yehia BR, Fleishman JA, Rutstein R, Mathews WC, Nijhawan A, HIV Research Network. Factors Associated With Retention Among Non-Perinatally HIV-Infected Youth in the HIV Research Network. J Pediatric Infect Dis Soc 2016 Mar 19;5(1):39-46 [FREE Full text] [doi: 10.1093/jpids/piu102] [Medline: 26908490]

Abbreviations

ART: antiretroviral therapy
HCD: human-centered design
mHealth: mobile health
SUS: System Usability Scale
UCSF: University of California, San Francisco
YAP: Youth Advisory Panel

Edited by G Eysenbach; submitted 31.12.20; peer-reviewed by JM Simoni, M Medich, P Athilingam; comments to author 02.03.21; revised version received 07.03.21; accepted 26.07.21; published 31.08.21.

Please cite as:
Saberi P, Lisha NE, Erguera XA, Hudes ES, Johnson MO, Ruel T, Neilands TB
A Mobile Health App (WYZ) for Engagement in Care and Antiretroviral Therapy Adherence Among Youth and Young Adults Living With HIV: Single-Arm Pilot Intervention Study
JMIR Form Res 2021;5(8):e26861
URL: https://formative.jmir.org/2021/8/e26861
doi:10.2196/26861
PMID:

©Parya Saberi, Nadra E Lisha, Xavier A Erguera, Estie Sid Hudes, Mallory O Johnson, Theodore Ruel, Torsten B Neilands. Originally published in JMIR Formative Research (https://formative.jmir.org), 31.08.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Informed Decision-making for Health Insurance Enrollment: Survey Study

Coralys M Colón-Morales[1*], BSc, MSc; Wayne C W Giang[1*], BASc, MASc, PhD; Michelle Alvarado[1*], BSc, ME, PhD

Department of Industrial and Systems Engineering, University of Florida, Gainesville, FL, United States

[*]all authors contributed equally

Corresponding Author:
Coralys M Colón-Morales, BSc, MSc
Department of Industrial and Systems Engineering
University of Florida
303 Weil Hall
Gainesville, FL, 32603
United States
Phone: 1 (352) 392 1464
Email: ccolonmorales@ufl.edu

Abstract

Background: Health insurance enrollment is a difficult financial decision with large health impacts. Challenges such as low health insurance literacy and lack of knowledge about choosing a plan further complicate this decision-making process. Therefore, to support consumers in their choice of a health insurance plan, it is essential to understand how individuals go about making this decision.

Objective: This study aims to understand the sources of information used by individuals to support their employer-provided health insurance enrollment decisions. It seeks to describe how individual descriptive factors lead to choosing a particular type of information source.

Methods: An introduction was presented on health insurance plan selection and the sources of information used to support these decisions from the 1980s to the present. Subsequently, an electronic survey of 151 full-time faculty and staff members was conducted. The survey consisted of four sections: *demographics*, *sources of information*, *health insurance literacy*, and *technology acceptance*. Descriptive statistics were used to show the demographic characteristics of the 126 eligible respondents and to study the response behaviors in the remaining survey sections. Proportion data analysis was performed using the Cochran-Armitage trend test to understand the strength of the association between our variables and the types of sources used by the respondents.

Results: In terms of demographics, most of the respondents were women (103/126, 81.7%), represented a small household (1-2 persons; 87/126, 69%), and used their insurance 3-12 times a year (52/126, 41.3%). They assessed themselves as having moderate to high health insurance literacy and high acceptance of technology. The most selected and top-ranked sources were *Official employer or state websites* and *Official Human Resources Virtual Benefits Counselor Alex*. From our data analysis, we found that the use of official primary sources was constant across age groups and health insurance use groups. Meanwhile, the use of friends or family as a primary source slightly decreased as age and use increased.

Conclusions: In this exploratory study, we identified the main sources of health insurance information among full-time employees from a large state university and found that most of the respondents needed 2-3 sources to gather all the information that they desired. We also studied and identified the relationships between individual factors (such as age, gender, and literacy) and 2 dependent variables on the types of primary sources of information. We encountered several limitations, which will be addressed in future studies.

(JMIR Form Res 2021;5(8):e27477) doi:10.2196/27477

KEYWORDS

health insurance; information; sources; survey; literacy

XSL•FO
RenderX

Introduction

Background

Enrolling in a health insurance plan is a complex decision that can have large health and financial impacts [1]. The decision is based on many factors (eg, premium costs, current health status, and the number of people covered) and directly affects future health care decisions such as choosing a provider or treatment. Therefore, informed decision-making is key for effective health insurance enrollment decisions. However, there are barriers to effective and informed enrollment decisions. Poor understanding of basic health insurance terminology (eg, deductible, premium, and copay) is the main barrier [2] because only 4% of the US population accurately understands such terms [3]. Lack of knowledge about how to choose a plan is another barrier to efficient enrollment decisions. Thus, consumers can benefit from specific information regarding health insurance literacy, available plans, and guidance on how to choose a plan based on the needs of individuals. Informed decision-making requires access to, and understanding of, useful information that can support the decision-making process. Over the last few decades, the proliferation of digital resources has changed how health information is accessed, but little research has examined how these changes have influenced health insurance decision-making. Therefore, the first aim of this paper is to understand the history of health insurance enrollment decision research and how decision-making has changed with the development of new digital technologies.

Currently, there are many sources of health insurance information used by individuals; however, these sources can sometimes lack accessibility, accuracy, or completeness of information, which can lead individuals to use more than 1 source. For our study, we considered the following types of sources: meetings with benefits representatives, printed material from health insurance providers, friends and family, websites, and virtual benefits counselors (VBCs). A VBC is a type of digital decision aid that mimics one-on-one conversations with a human resources (HR) benefits counselor using a conversational digital interface. VBCs provide personalized guidance in the decision process by contrasting different health insurance plans while providing increased access to health insurance literacy. Many employers have created customized VBCs to provide decision support to their employees. The VBC in this study refers to the official HR VBC Alex (Jellyvision Labs Inc) [4].

As individuals choose the source of information that they wish to use, their selection can vary depending on their goals, preferences, and knowledge about health insurance [5]. Similarly, a person's attitude toward technology could also have an impact on the source of information that they choose to use as support for their enrollment decision [6,7]. Therefore, the second aim of this paper is to explore where current consumers are searching for information to support their employer-provided health insurance decisions. To achieve this aim, we created and used a survey to gain insight into the current sources of information used by employees of a large state university. This paper presents the pilot study for this project; therefore, the work shown here is exploratory in nature.

To bridge the gap in our understanding of how health insurance consumers achieve informed enrollment decisions, it is essential to understand how individuals go about making these decisions and what information they use to support their choice. During the 1980s and the 1990s, with the rise of health maintenance organizations in the health insurance market, extensive research was conducted on the information factors that influence health plan selection and on how individuals use this information to decide from the options available to them. For example, Scanlon et al [5] compiled and analyzed numerous studies. They identified and categorized a set of variables to rationalize the variation in health plan choices. The primary variables referred to health plan characteristics, which included costs, provider restrictions, different types of coverage and benefits, quality, and convenience. The secondary variables referred to consumer demographics and other characteristics such as health and economic status. Although the research showed that the primary variables, particularly price, influenced the plan choice, many studies also showed that, depending on the age, gender, or health status, groups can have distinct patterns of enrolling in specific health plans [8-10]. Therefore, the distinction between the primary and secondary variables and the interactions between them are important for understanding plan selection bias and the probability of a consumer enrolling in a particular plan. Prominent empirical studies during the 1980s and 1990s used (1) modeling consumer choice under uncertainty [9,11], (2) conditional choice models that supported the estimation of the trade-off between price and other health plan characteristics included in the model [12], or (3) probit or logit analysis, which allowed modeling the probability of enrolling in a plan as a function of price and other plan features [13].

Studies also sought to understand the ideal types of information that consumers preferred when making health insurance enrollment decisions, and Edgman-Levitan and Cleary [14] concluded that consumers need comparative data on the various plans; trade-off evaluations among access, cost, and quality; and methods to determine their out-of-pocket costs based on their health status. In contrast, Isaacs [15] pointed out the need for information on the quality of primary care physicians and specialists, the range of services covered, pre-existing condition exclusions, and consumer ratings of hospitals and physicians. Similarly, Tumlinson et al [16] showed that cost, price, benefits, availability, and quality of providers are essential when comparing plans and making an informed decision. However, across these studies, there was an acknowledgment of the challenges faced when creating ways to provide and present this information to support consumer understanding. This shows that the process of choosing a health insurance plan is complex and typically leads individuals to make poor decisions [17].

Various types and sources of health insurance information were examined during the 1980s and 1990s; for example, digital methods of presenting information to consumers that present different *layers* of information depending on individual interests. As another example, consumer report cards were a popular mechanism for sharing health plan performance measures through consumer ratings and insights into the available health

XSL•FO
RenderX

plans. Although these methods have the potential to provide quality information, it is important to have a thorough comprehension of this information [18]. However, Gibbs et al [19] discussed that consumer satisfaction measures need to be carefully chosen and presented to communicate meaningful information and appeal to specific consumer needs. Compared with report cards, information sources that are impartial to health insurance companies were preferred by the participants in this study; therefore, they valued the input from friends and family. The findings of the study by Isaacs [15] revealed similar preferences, along with trust in health plan information received from the primary physician.

With the diffusion of the internet during the late 1990s and the early 2000s, websites became a source of health and medical information [20], especially among younger consumers [21,22], consumers with chronic conditions, and those who face barriers when accessing health care [23]. Therefore, many scholars analyzed the internet's impact on consumers' search for health information. Studies showed that more people were searching the web for medical information before talking to their doctors [24], and federal websites were widely considered trustworthy sources of health information [25]. In fact, the results of a study by Rains [26] showed that lack of trust in the traditional sources of information (eg, mass media and one's health care provider) was associated with the increased use of websites as a primary information source. However, when it was difficult to find or understand health information on the web, there was less trust in it [27]. When comparing the primary sources for health information among individuals, Dutta-Bergman [28] reported that the internet, newspapers or magazines, and family or friends were the primary sources of health-conscious individuals. These studies focused on the effect of the internet on the search for health information in general and provided a baseline for the main categories of primary sources of information considered in this paper: official sources such as federal websites, internet-based sources, and trusted individuals such as friends and family.

However, few studies during the period between the late 1990s and the early 2000s focused on sources of health insurance information. A study by Mark and Coffey [29] evaluated the available sources of health plan information for consumers. They considered consumer report cards and report filings from state insurance departments, health plan websites, the US Census, independent organizations, and a national survey on health care use. Although each source had its strengths, the authors concluded that all sources would benefit from more practical and cost-effective methods for providing information. It is also important to look at the decisions of employees regarding employer-provided insurance because there are other factors that affect their decisions [30]. A later study by Oetjen et al [31] focused on government employees' access to and use of three health plan information sources: printed information from the state, printed information from the health plans, and web-based information. In this 2003 study, the most accessed and used source was printed information from the health plans, followed by printed information from the state and, finally, web-based information, with 34% of the participants using all 3 sources.

With the passing of the Affordable Care Act in 2010, more focus was given to understanding the complexity of choosing a health insurance plan because the act sought to expand coverage, and new decision support tools began to be developed. This became critical to ensuring enrollment success [32,33]. Studies found that plan choice complexity emerges from the wide variation in health costs and other plan features such as network size, service coverage, and reputation on processing claims [2], as well as the level of engagement from consumers [34]. Therefore, to make an informed decision, consumers must project their health expenses and subsequently evaluate the trade-off among the plan's health costs. However, this is still a challenge for most people because only 4% of the US population understands basic health insurance terminology [3]. In fact, Hero et al [35] showed that among people with low health insurance literacy, at least 54% had difficulty finding the best or most affordable plan, and at least 48% had fair or poor overall experience when choosing their plan. These findings suggest that disparities in the ability to access and understand health insurance information may be a reason why different demographics may have differing plan selections [36].

Thus, to better support health insurance decision-making, different digital decision support tools such as cost estimators, quality ratings, provider lookup, drug lookup, and pop-up definitions started to become more available to consumers. These digital tools also allowed for the sorting and filtering of information by individual characteristics, a task that was much more difficult with printed media. However, in the beginning, these aids were missing from most federal and some state-based websites [37]. In fact, Vardell [38] points out the need for consumers to seek various sources of information (2.8 sources, on average) to fully assess their options. As studies began to highlight the importance of decision aids [39-41], these tools became the standard for informed consumer choice when the information was accurate and understood by consumers. Later studies have also shown a greater adoption of some decision support tools by websites for private and public health insurance marketplaces [42,43], a trend that is likely to continue.

The increased availability of decision aids led to the development of new tools that combine multiple methods for supporting enrollment decisions (eg, cost estimation, provider lookup, and definition of health insurance terms). One example is the Show Me My Health Plans (SMHP) tool, built to provide information about the Affordable Care Act marketplace in Missouri [44-46]. The SMHP tool contains simplified health insurance information (it uses plain language and graphics) for educational purposes, a cost-estimating component, an assessment of plan feature preferences, and plan recommendations. The SMHP decision aid improved the health insurance selection quality by improving decision self-efficacy, health insurance literacy, and confidence in plan choice compared with the health care government website [45]. More recently, companies have also developed aids that walk individuals through the enrollment decision-making process. VBCs are an implementation of these aids that have been used by HR departments to assist with employee plan selection. The continued development of these decision aids is likely to be of benefit to health insurance consumers, but it is still not known

XSL•FO
RenderX

if consumers are adopting and preferring these tools in comparison with other sources of health insurance information.

Objective

As shown in this study, health insurance enrollment choices are complex, multifactorial decisions that require access to different types of information. Although some studies have examined the sources of information that individuals access, the recent availability of new tools such as the SMHP tool and VBCs may have changed the landscape of information that users seek. These issues are of particular interest because employer HR departments are now recommending the use of these tools to their employees as a means of helping to support informed decision-making, especially because 56% of Americans receive their coverage from their employer [47]. Thus, a web-based survey of employees at a large state university was used to understand the sources of information used to support their health insurance enrollment decisions and to study the factors that led them to the sources that they use.

Methods

Health Insurance Information Sources Survey

The Health Insurance Information Sources Survey (HIISS) is a web-based survey containing 27 questions in total and takes approximately 5 minutes to complete on Qualtrics (Qualtrics XM). A sample of the HIISS questions is shown in Table 1, along with their respective response types and options. The full survey questions are available in Multimedia Appendix 1. The HIISS survey consists of four sections:

1. *Demographics and employment status*: The questions in this section are mostly categorical (either nominal or ordinal) and were all author-created. It is important to note that the age category of 55-66 years has a cutoff of 66 years because this is the social security full-benefits retirement age.
2. *Sources of health insurance information used*: This section contains only the 2 questions presented in Table 1. Both questions were author-created.
3. *Health insurance literacy*: Four questions were selected from the Health Insurance Literacy Measure (HILM) developed by Paez et al [48,49] (Table 1, section 3). This measure assesses 2 categories of health insurance literacy (selecting health insurance and using health insurance), each addressing 2 dimensions (confidence choosing and comparing plans in the first category and confidence using and being proactive in the second category). The 4 questions selected were each taken from a different category and dimension. In these scales, 1 represented not confident or not likely at all, whereas 7 represented extremely confident or likely.
4. *Technology acceptance and experience with virtual chatbots and agents*: In this section, the first 3 questions about technology acceptance were obtained from Reimer et al [50-53] and have scaled responses from 1 to 10, with 1 being very inexperienced or distrustful and 10 being very experienced or trustful. The other questions related to experience with virtual chatbots and agents were author created.

The HIISS was designed to gain insights into where employees are looking for health insurance information as well as the respondents' demographic information, health insurance literacy, and interactions with technology. The questionnaire focused on six main sources of health insurance inspired by the background: official employer or state information, an official HR VBC system, friends or family, other nonofficial websites or resources (with space to write in), official HR in-person benefits counselors, and other in-person resources (with space to write in). It is important to note that the HR department at the employer had recommended the use of a VBC for enrollment decisions starting in 2017, which may have contributed to this greater awareness of this new type of technology. All data collected were confidential, and the study was approved by the local institutional review board (IRB #201900070).

XSL•FO
RenderX

Table 1. Health Insurance Information Sources Survey questions.

Questions	Response options
Section 1: demographics	
For your last enrollment period, were you responsible, or did you share responsibility, in making health insurance decisions within your household? (nominal response)	• Yes, I was primarily responsible • Yes, I shared responsibility • No, I did not share responsibility
What is your age? (ordinal response)	• 18-24 years old • 25-34 years old • 35-44 years old • 45-54 years old • 55-66 years old • 67 plus years old • Prefer not to say
Which of these best describes your gender? (nominal response)	• Male • Female • Other • Prefer not to say
Which of these best describes your marital status? (nominal response)	• Single • Married or domestic partnership • Prefer not to say
Do you have dependents beside a spouse or partner? (nominal response)	• Yes • No • Prefer not to say
How large is your household? (ordinal response)	• 1 to 2 • 3 to 5 • 6 or more
How often do you or your dependents use your health insurance? (ordinal response)	• Never • Less than 3 times a year • 3 to 12 times a year • 13 to 24 times a year • More than 24 times a year • Prefer not to say
Section 2: sources of information	
Where do you find information about health insurance plans? (Select all that apply)	• Human Resources' Alex, an online, virtual benefits counselor • Other official or state website • Other online websites or resources (eg, Google, government health care website; please indicate the website or resource) • Human Resources' in-person benefits counselor • Friends or family • Other in-person resources (please indicate the resources)
Please rank the sources of information that you use for health insurance enrollment decisions from most important (1) to least important. (Please select one ranking per source)	• Each source from the question above is listed with 6 possible rankings to the right of each
Section 3: health insurance literacy (ordinal responses)	
(a) How confident would you feel that you understand health insurance terms (eg, copay, deductible, co-insurance, premium)?	• 7-point Likert scale, with 1 being *not confident at all* and 7 being *extremely confident*
(b) When comparing health plans, how confident are you in understanding what needs to be paid for emergency department visits?	• 7-point Likert scale, with 1 being *not confident at all* and 7 being *extremely confident*
(c) How confident are you in knowing what is and is not covered before you receive a health care service?	• 7-point Likert scale, with 1 being *not confident at all* and 7 being *extremely confident*

XSL•FO
RenderX

Questions	Response options
(d) When using your health insurance plan, how likely are you to find out if a doctor is in-network before you see him/her?	• 7-point Likert scale, with 1 being *not likely at all* and 7 being *extremely likely*
Section 4: technology acceptance	
(a) How would you rate your level of experience with technology (eg, cell phones, automatic teller machines, digital cameras, computers, etc)? (ordinal response)	• A scale of 1-10, with 1 being *very inexperienced* and 10 being *very experienced*
(b) Some people prefer to avoid new technologies as long as possible while others like to try them out as soon as they become available. In general, how would you rate yourself as being an avoider or an early adopter of new technology? (ordinal response)	• A scale of 1-10, with 1 being *avoid as long as possible* and 10 being *try as soon as possible*
(c) How would you rate your overall level of trust in technology? (ordinal response)	• A scale of 1-10, with 1 being *very distrustful* and 10 being *very trustful*
Have you ever interacted with a "virtual agent," "chatbot," or "virtual rep" when interacting with a website or web service? Virtual agents provide automated customer service using a conversational interface. (nominal response)	• Yes, multiple times • Yes, I have tried them • No, I have not tried them • Not sure
How would you rate the usefulness of the virtual agents you've interacted with in the past? (ordinal response)	• A scale of 1-10, with 1 being *not useful at all* and 10 being *extremely useful*
How would you rate the ease of using the virtual agents you've interacted with in the past? (ordinal response)	• A scale of 1-10, with 1 being *not easy to use at all* and 10 being *extremely easy to use*

Recruitment

The HIISS was distributed to employees at a large state university through emails and flyers through academic and service departments, as well as through an HR newsletter sent to more than 31,000 employees. Given that this is an exploratory study, we allowed for snowball sampling to occur. Only qualifying respondents were included in the final data set. The inclusion criteria were as follows: (1) full-time employment (measured in full-time equivalents) at the university that qualified for health benefits and (2) being the primary decision-maker or sharing decision-making responsibility in the household for health care. We estimate that approximately 15,500 members are full-time staff or faculty members satisfying inclusion criterion 1.

Statistical Analysis

We present descriptive statistics on the demographic variables, HILM scores, technology acceptance scores, the sources of information selected, and the source ranks assigned by the respondents. This allowed for an understanding of the sources of health insurance information used by employees to enroll in health plans.

As the respondents ranked their sources of information in order of importance, their highest-ranked source was considered their primary source of information. The response or dependent variables in our analysis centered on whether the primary source of information was an official source as opposed to a nonofficial source, as well as whether it consisted of friends or family. These binary factors were identified as important characteristics of the information sources in our background [25,28] and are therefore the main factors of interest in this study. Therefore, the 6 source options provided in the survey were classified as either official or nonofficial sources of information and as friends or family or not friends or family. Under official sources, we classified official employer or state websites, the official HR VBC system, and official HR in-person benefits counselors. Under nonofficial sources, we placed friends or family, other nonofficial websites or resources, and other in-person resources. Although there exist external websites (eg, government health care website) that contain validated general health insurance information and other helpful decision aids, they do not provide information on these specific employer-provided health insurance plans. Therefore, they were considered nonofficial sources of information for the scope of this study. Similarly, under not friends or family, we considered all the listed sources except friends or family.

This categorization allowed for the creation of two binary variables: (1) official (true) and nonofficial (false) primary sources and (2) friends or family (true) and not friends or family (false) as primary sources. Given that all our collected data are categorical variables (mostly ordinal) and that these primary source variables are binary, we required an association test for these specific types of variables. Ultimately, we decided upon the Cochran-Armitage trend test. The goal of these tests was to understand whether a relationship or a trend exists between our binary primary source—dependent variables—and our demographic variables, HILM scores, and technology acceptance scores. With these tests, we sought to identify whether there are respondents' characteristics that drive their primary source decision. The Cochran-Armitage test is used to test whether there is a linear trend when the explanatory variable is ordinal and the response is binary. It assesses whether there is a monotonically increasing or decreasing trend in the proportions of the response Y over the ordered categories X. This test is applied to data in the form of an $r \times 2$ contingency

XSL•FO
RenderX

table, with $r>2$. The null hypothesis is that no trend exists, which means that the proportions for all levels of the explanatory variable X are the same. On considering a 95% CI values, P<.05 indicates that an association or trend exists for the binary response Y over the categorical variable X. In the case of the demographic factor gender, which is a nominal variable, the association with our 2 binary responses was measured using the ϕ coefficient of correlation.

Results

Demographic Factors

Of the 151 individuals who responded to the HIISS, 126 (83.4%) were included in the final data set. Of the 25 participants who were excluded, 4 (16%) did not complete the survey, 4 (16%) did not share responsibility in making health insurance decisions within their household, and 17 (68%) had a current employment status of less than 0.75 full-time equivalent. Table 2 summarizes the respondents' demographic information. Our sample comprised mostly women (103/126, 81.7%), with greater participation from the age groups 25-34 years and 55-66 years. In addition, most of the respondents (79/126, 62.7%) were married or in a domestic relationship and lived in households of 1-2 persons (87/126, 69%). Table 2 also shows how often the respondents' households used their health insurance over the course of a year. Most commonly, the respondents used their insurance 3-12 times a year, as indicated by 41.3% (52/126) of the respondents.

Using Spearman correlation tests [54], Table 3 presents an upper-triangular correlation matrix (ie, there are only values above the diagonal) among the ordinal demographic factor variables and the health insurance use variable. As expected, the number of dependents and household size were positively correlated (ρ=0.81). Furthermore, household size was correlated (ρ=0.32) with higher health insurance use. The results found no evidence that age was associated with an increase in health insurance use.

Table 2. Respondents' demographics (N=126).

Demographics	Health insurance use, n (%)					Total (N=126), n (%)
	Never (n=12)	Fewer than 3 times a year (n=19)	3-12 times a year (n=52)	13-24 times a year (n=26)	More than 24 times a year (n=17)	
Age (years)						
18-24	0 (0)	2 (10.5)	3 (5.8)	1 (3.8)	0 (0)	6 (4.7)
25-34	7 (58.3)	3 (15.8)	19 (36.5)	6 (23.1)	3 (17.6)	38 (30.2)
35-44	1 (8.3)	4 (21.1)	11 (21.1)	6 (23.1)	2 (11.8)	24 (19)
45-54	1 (8.3)	3 (15.8)	9 (17.3)	7 (26.9)	4 (23.5)	24 (19)
55-66	3 (25)	6 (31.6)	10 (19.2)	6 (23.1)	8 (47.1)	33 (26.2)
≥67	0 (0)	1 (5.3)	0 (0)	0 (0)	0 (0)	1 (0.8)
Gender						
Female	9 (75)	14 (73.7)	46 (88.5)	21 (80.1)	13 (76.5)	103 (81.8)
Male	3 (25)	4 (21.1)	5 (9.61)	5 (19.2)	4 (23.5)	21 (16.7)
Other	0 (0)	0 (0)	1 (1.9)	0 (0)	0 (0)	1 (0.8)
Prefer not to say	0 (0)	1 (5.3)	0 (0)	0 (0)	0 (0)	1 (0.8)
Marital status						
Single	10 (83.3)	7 (36.8)	19 (36.5)	8 (30.1)	1 (5.9)	45 (35.7)
Married or domestic partnership	2 (16.7)	10 (52.6)	33 (63.5)	18 (69.2)	16 (94.1)	79 (62.7)
Prefer not to say	0 (0)	2 (10.5)	0 (0)	0 (0)	0 (0)	2 (1.6)
Household size (persons)						
1-2	11 (91.7)	16 (84.2)	38	15 (57.7)	7 (41.2)	87 (69)
3-5	1 (8.3)	3 (15.8)	14	11 (42.3)	10 (58.8)	39 (30.9)

XSL•FO
RenderX

Table 3. Demographic correlation matrix.

Correlation matrix	Age (years)	Household size	Health insurance use
Age (years)	1	0.12	0.14
Household size	—[a]	1	0.32[b]
Health insurance use	—	—	1

[a]Not applicable.

[b]Significant correlations at α=.05.

HILM and Technology Acceptance

The score distributions for the HILM and technology acceptance are shown in Figures 1-4, with higher scores representing higher self-reported confidence or ability in that dimension of health literacy. A score of 5 was the mode among the questions on understanding terms (Figure 1), understanding emergency payments (Figure 2), and understanding coverage (Figure 3). For the question on understanding terms, there was a tendency toward a higher score (≥5), whereas for the questions on understanding emergency payments and understanding coverage, the score frequencies were balanced throughout the scale. However, the question on finding out whether a doctor is in-network (Figure 4) had a different response behavior, with 57.9% (73/126) of the respondents responding that they were extremely likely to find out whether a doctor is in-network before they see them.

Figure 1. Scores for Health Insurance Literacy Measure in understanding terms. HILM: Health Insurance Literacy Measure.

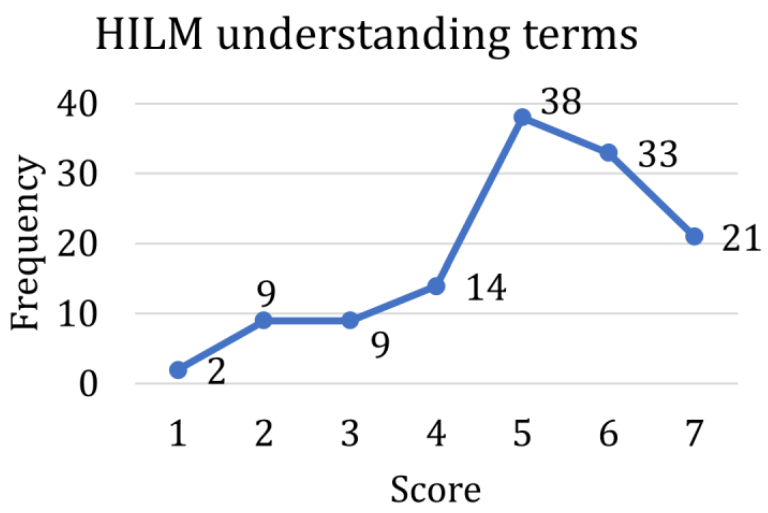
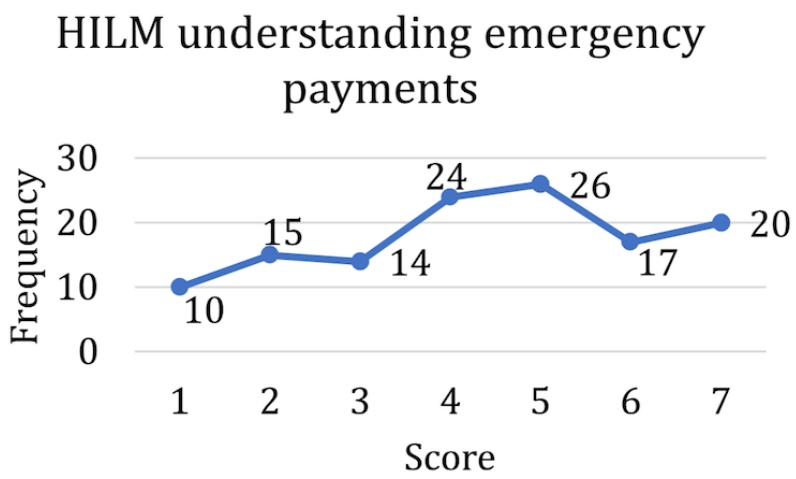


Figure 2. Scores for Health Insurance Literacy Measure in understanding emergency payments. HILM: Health Insurance Literacy Measure.



XSL•FO
RenderX

Figure 3. Scores for Health Insurance Literacy Measure in understanding coverage. HILM: Health Insurance Literacy Measure.

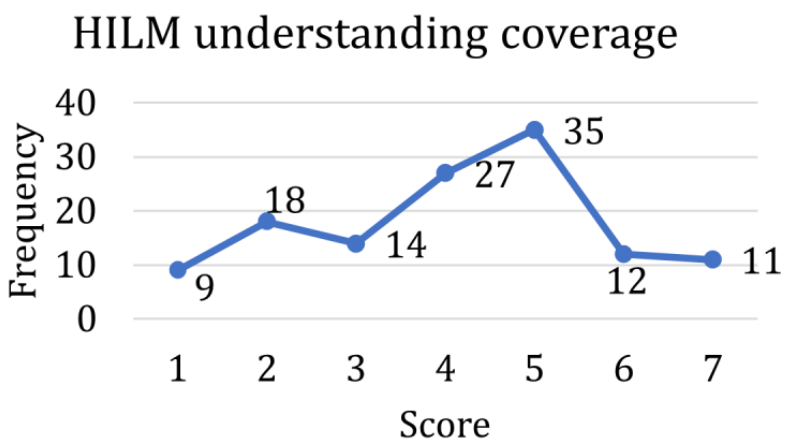
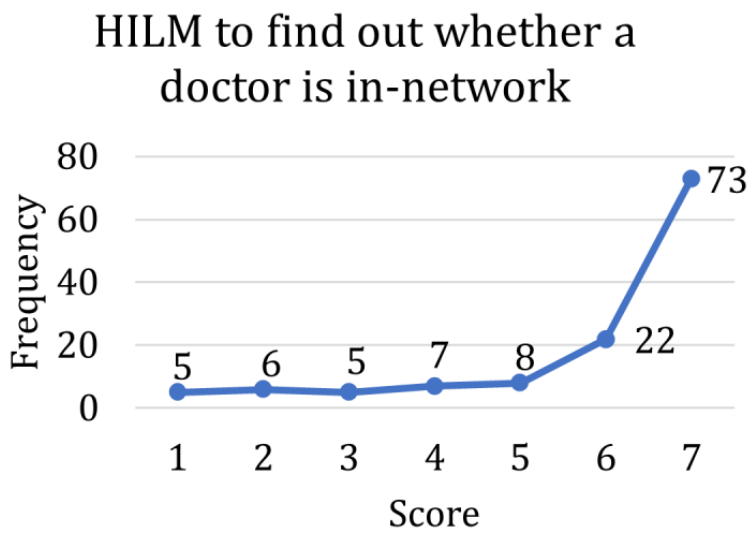


Figure 4. Scores for Health Insurance Literacy Measure to find out whether a doctor is in-network. HILM: Health Insurance Literacy Measure.



The respondents' technology acceptance and experience with virtual agents were also assessed (Table 1, section 4). All respondents had high levels of experience with technology; in fact, not a single response scored lower than 4, and most of the respondents rated their experience with technology as either 9 or 10 (Figure 5). For the most part, both trust in technology (Figure 6) and adoption of new technologies (Figure 7) were also high across the sample. Overall, our respondents—full-time employees at a large state university—had very high technology acceptance.

The respondents in this sample were also largely familiar with virtual agents, with only 15.9% (20/126) stating that they had never used (or were not sure if they had used) a virtual agent. The respondents were also asked to rate the usefulness of virtual agents on a 10-point scale. Of the 126 respondents, 44 (34.9%) had tried virtual agents, and the average reported usefulness was 4.72 (SD 1.92; median 5), and among the 65 (51.6%) respondents who had used them multiple times, the average usefulness was 6.51 (SD 1.99; median 7). Of note, the perceived usefulness among people who had used these agents multiple times was higher than among those who have only tried them.

XSL•FO
RenderX

Figure 5. Scores for the self-reported technology experience question.

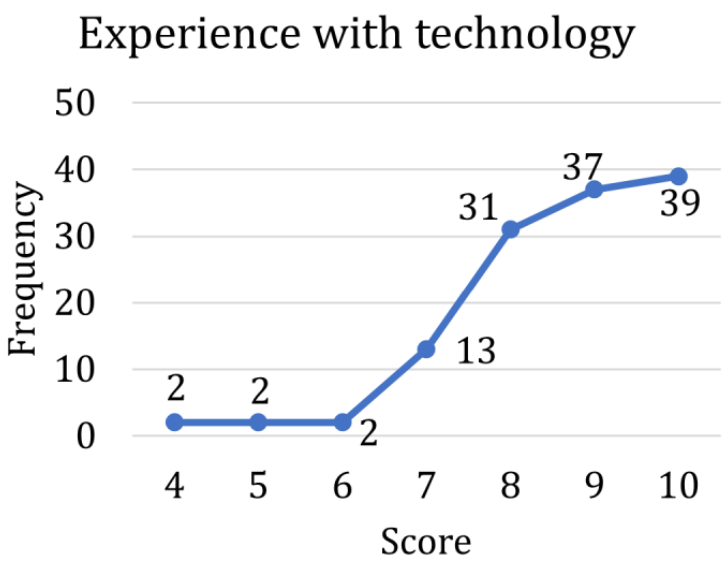
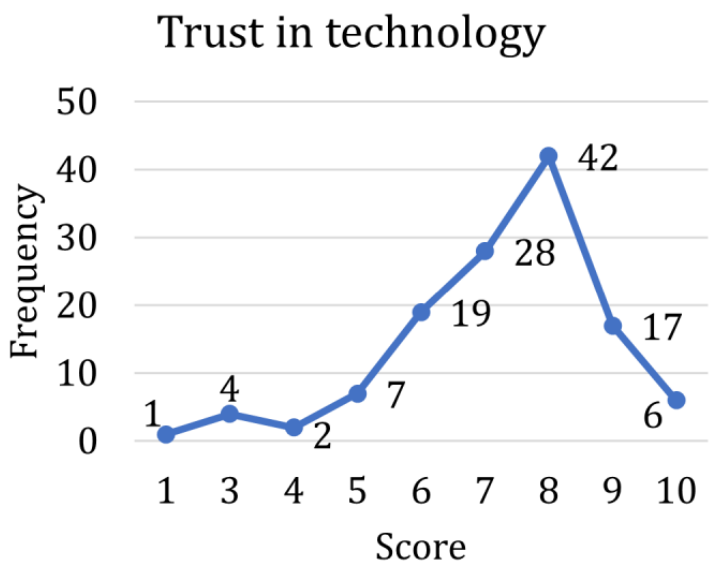
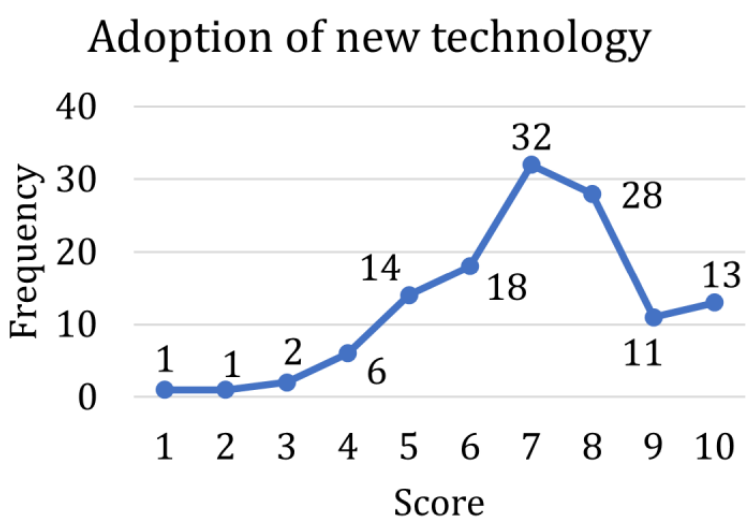


Figure 6. Scores for the self-reported technology trust question.



XSL•FO
RenderX

Figure 7. Scores for the self-reported adoption of new technology question.



Sources of Information

The respondents were also asked where they found information regarding their health insurance plans. They were presented with the 6 options shown in Table 4 and asked to select all sources of information that they had used. On average, the respondents selected 2-3 sources of information, none selected all 6 options, and only 2 selected 5 sources of information (Table 5). The official employer or state websites were the most selected source (96/126, 76.2%), followed by official HR VBC Alex (76/126, 60.3%), both being digital and official sources of information. The third most selected source was friends or family with (58/126, 46%) a nonofficial and nondigital source. These results showed that most respondents made use of more than 1 source of information.

Table 4. Number of respondents who selected each source as an information source for health insurance plans (N=126).

Source of information	Respondents, n (%)
Official employer or state websites	96 (76.2)
Official HR[a] VBC[b] Alex	75 (59.5)
Friends or family	58 (46)
Other websites or resources (eg, Google and health care government website)	28 (22.2)
Official HR in-person benefits counselors	23 (18.2)
Other in-person resources	12 (9.5)

[a]HR: human resources.
[b]VBC: virtual benefits counselor.

Table 5. Number of sources respondents use for health insurance plan information (N=126).

Number of sources	Frequency, n (%)
1	25 (19.8)
2	50 (39.7)
3	39 (30.9)
4	10 (7.9)
5	2 (1.6)
6	0 (0)

The respondents were also asked to rank their selected sources in order of importance (Table 6; see the graph in the Multimedia Appendix 2 for a better visual understanding of Table 6). The official employer or state websites were ranked first by 49.2% (54/126) of the respondents, whereas the official HR VBC Alex and friends or family were ranked first by 21.4% (27/126) of the respondents. These results suggest that the official employer or state websites were the most used and preferred source of information. In contrast, the VBC decision aid seemed to be a supplementary resource because it had more second-place votes

XSL•FO
RenderX

than first-place votes. Finally, although fewer respondents made use of friends or family, the most common rank that this source received was 1 (27/58, 47%). This suggests that there is a strong preference for family or friends among the persons who make use of this source.

Table 6. Ranking occurrences per source.

Source	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6
UF[a] human resources' VBC[b] Alex	27	32	10	5	0	0
Other official UF or state of Florida website	54	25	11	5	0	0
Other websites or web-based resources (eg, Google and health care government website)	9	10	9	0	0	0
UF human resources' in-person benefits counselors	5	8	8	1	1	0
Friends or family	27	18	11	1	1	0
Other in-person resources	4	7	1	0	0	0

[a]UF: University of Florida.

[b]VBC: virtual benefits counselor.

Influence of Decision-Maker Factors on Preferred Health Information Source

We first present the results for the response variable of official versus nonofficial primary sources of information. The explanatory variables considered were age, health insurance use, and technology acceptance. However, the technology acceptance data did not show significant trends or results and are therefore not presented. The Cochran-Armitage results are presented in Table 7. For this response, all *P* values are less than .05, indicating that a trend exists between the explanatory variables and the use of an official primary source.

Table 7. Cochran-Armitage test results for official versus nonofficial primary sources of information response.

Explanatory variable	*z* value	Two-sided *P* value
Age (years)	5.657	<.001
Health insurance use	4.677	<.001
HILM[a] understand terms	3.431	.001
HILM understand emergency payments	2.436	.01
HILM understand coverage	4.438	<.001
HILM finding out whether doctor is in-network	−3.503	.001

[a]HILM: Health Insurance Literacy Measure.

These trends can be increasing linear trends, constant trends over time, or decreasing linear trends. The specific behavior for each explanatory variable is presented in the plots in Figures 8-10. In these plots, the proportions on the Y-axis refer to the proportions of participants who selected an official source as their primary source (true value in the contingency table). For the explanatory variable age, Figure 8 shows a significant increasing trend. The effect of age on this response slightly increases ordinally, which means that as age increases, the proportion of people who use official sources as primary sources increases as well. However, the middle-aged groups remained consistent at approximately 0.70. For the health insurance use explanatory variable in Figure 9, we see that the proportions are somewhat constant as the use of health insurance increases.

As indicated in the *Methods* section, for the association between gender and the dependent variables, we used the ϕ coefficient of correlation. For this response of the official primary source, the coefficient was 0.03, which is extremely low and can be considered a negligible relationship.

The 4 plots for the HILMs are presented in Figure 10. The Cochran-Armitage test results showed that the proportion of people who use official primary sources is dependent on the scores of all 4 measures. The effect of the HILMs on the response variable changed ordinally. They all show increasing trends, except for the HILM finding out whether the doctor is in-network, for which the proportion of people who use official sources as primary sources decreases as the score increases.

XSL•FO
RenderX

Figure 8. Contingency table proportions of official primary source per age group.

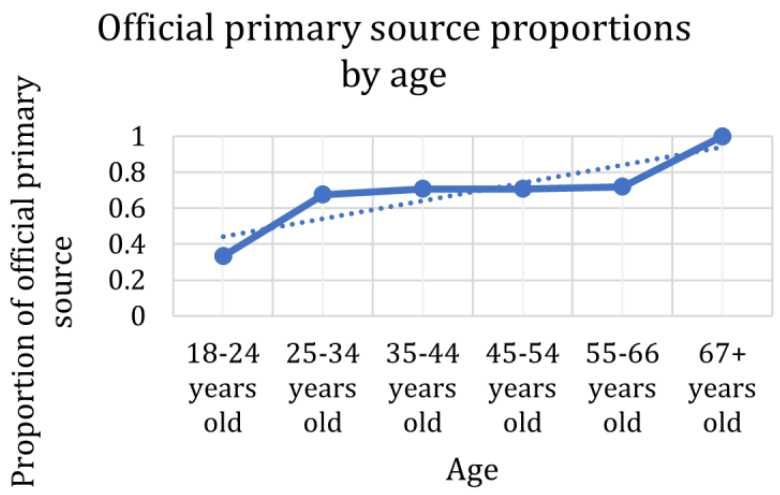
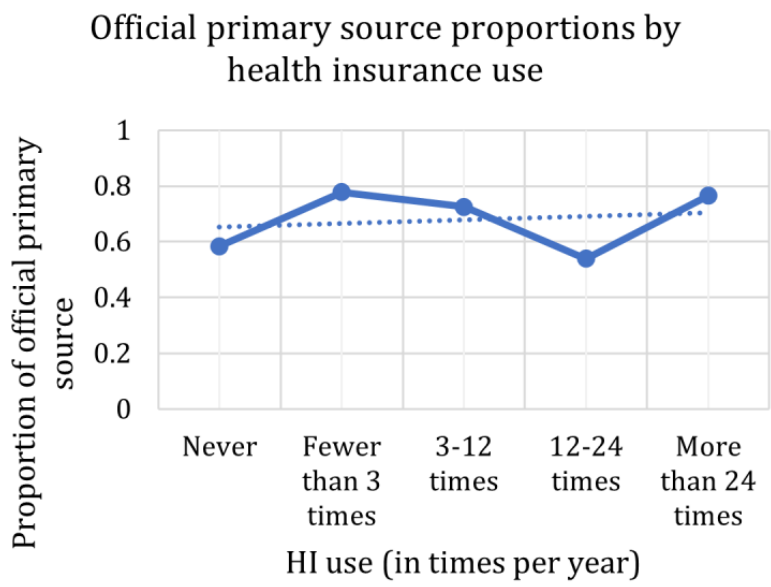
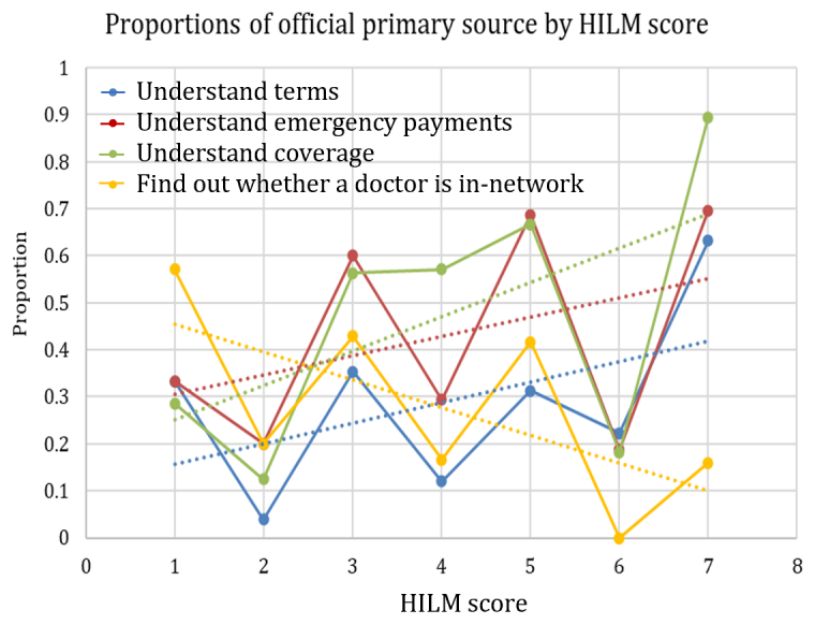


Figure 9. Contingency table proportions of official primary source per health insurance use group. HI: health insurance.



XSL•FO
RenderX

Figure 10. Contingency table proportions of official primary source per each Health Insurance Literacy Measure score. HILM: Health Insurance Literacy Measure.



For the response variable of friends or family versus not friends or family as a primary source of information, the Cochran-Armitage results are presented in Table 8 and the trend plots in Figures 8-10. In this case, the explanatory variables of age, health insurance use, HILM understand coverage, and HILM finding out whether the doctor is in-network had a *P*=.001, indicating that an association exists among these variables and the use of friends or family as primary sources. However, HILM understanding terms and HILM understand emergency payments do not indicate a significant relationship with the response. The specific behavior for the significant explanatory variables is presented in the plots in Figures 11-13. In these plots, the proportions on the Y-axis represent the proportions of participants who selected friends or family as their primary source of information (true value in the contingency table).

Table 8. Cochran-Armitage test results for friend or family vs non–friend or family responses.

Explanatory variable	*z* value	Two-sided *P* value
Age (years)	5.159	<.001
Health insurance use	3.451	.001
HILM[a] understand terms	1.334	.18
HILM understand emergency payments	1.489	.14
HILM understand coverage	3.596	<.001
HILM finding out whether doctor is in-network	−3.225	.001

[a]HILM: Health Insurance Literacy Measure.

For the explanatory variable age, we now see in Figure 11 a slightly decreasing effect size for the friends or family primary source response. As age increases, the proportion of participants who make use of friends or family as a primary source decreases. Similarly, for health insurance use in Figure 12, we also see a slightly decreasing effect size as the use variable increases. Similar to the previous dependent variable, for this response of friends or family as a primary source, the ϕ coefficient for gender was 0.03. Once again, this is extremely low and can be considered a negligible relationship.

In terms of the HILMs, we only present the 2 measures that resulted in a significant association with this response. As mentioned previously, the 2 measures with significant relationships with friends and family as a primary source are HILM understand coverage (represented by the green lines in Figure 13) and HILM finding out whether the doctor is in-network (represented by the yellow lines in Figure 13). For HILM understanding coverage, we see an increasing trend, whereas for HILM finding out whether the doctor is in-network, we once again see a decreasing trend of proportions as the values of the scores increase.

XSL•FO
RenderX

Figure 11. Contingency table proportions of friends or family as the primary source per age group.

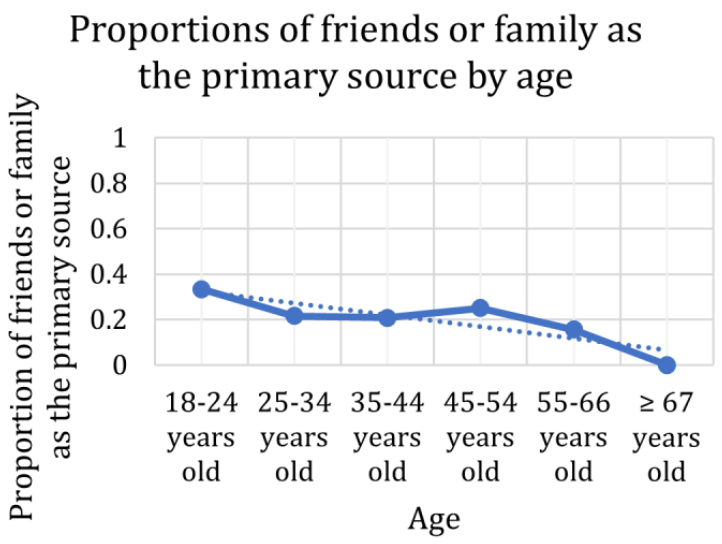
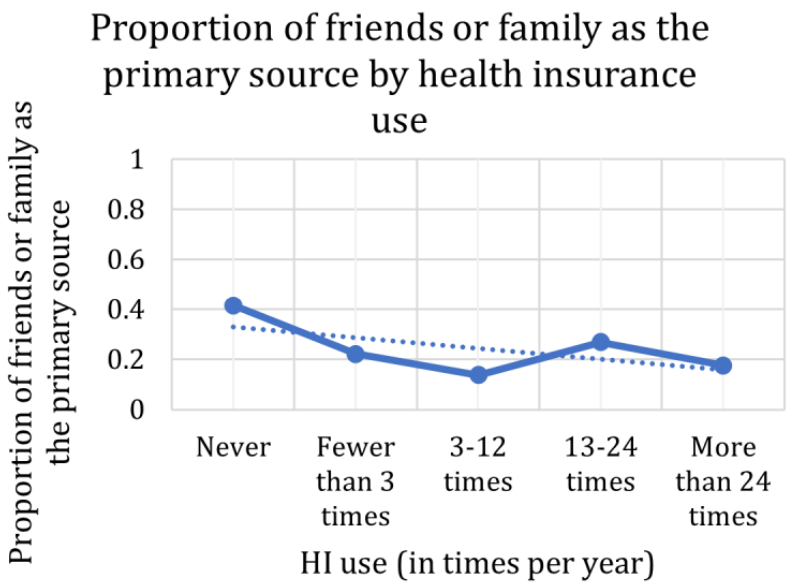
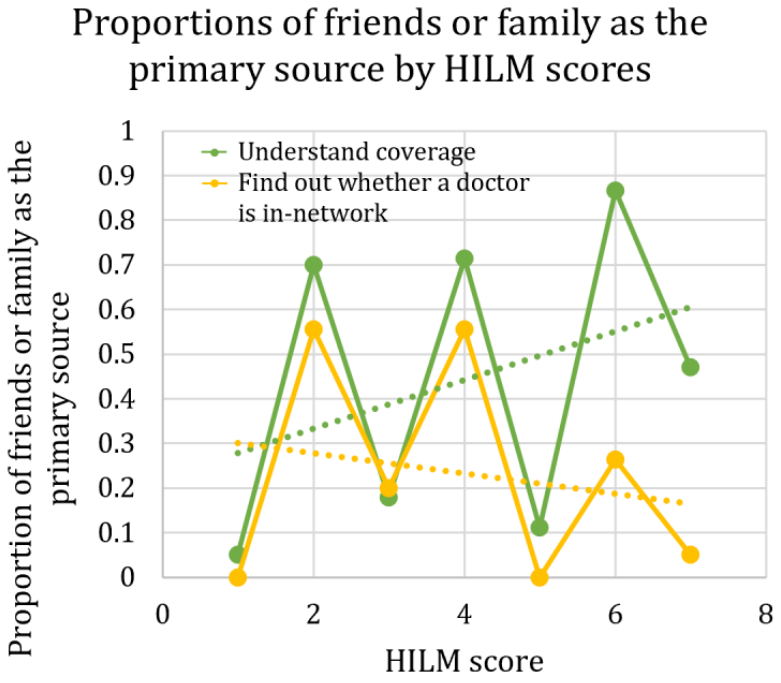


Figure 12. Contingency table proportions of friends or family as the primary source per health insurance use group. HI: health insurance.



XSL•FO
RenderX

Figure 13. Contingency table proportions of friends or family as the primary source for 2 Health Insurance Literacy Measure scores. HILM: Health Insurance Literacy Measure.



Discussion

Principal Findings

Access to useful and accurate information is key to informed decision-making, especially for difficult decisions such as choosing a health insurance plan. Through the years, the health insurance environment has undergone many changes, and individuals have had to adapt their enrollment decision-making process accordingly. Therefore, we present a chronological background that evaluates the health plan information–seeking processes of individuals and the evolution of available sources and tools for health insurance information as an introduction to our work. In general, we learned that as the internet became more accessible, individuals mostly used digital and official sources of health information, as well as friends or family [24-28]. In more recent years, there have been greater efforts to include newer and varying tools within the sources of information. Given that little research has been conducted on the effect of these tools, it is important to study and understand the sources of information that users are currently relying on and the characteristics of users and health plans that contribute to the use of different sources, tools, and information types. With this survey, we take the initial steps toward answering these questions.

In terms of our sample, we had small groups in the youngest and older age categories, which was to be expected given our targeted population. Most of our participants used their insurance often (3-12 times per year), although the other use groups had balanced sample sizes. Overall, we saw medium to high health insurance literacy within our sample and high technology acceptance, which can also be a reflection of the targeted population. It was observed that the perceived usefulness of virtual agents among people who had used them multiple times was higher than among those who had only tried them. This could be because of selection bias (only those who liked using them or found them useful continued to do so) or because repeated interactions with virtual agents allowed people to find them more useful over time.

Our sample of employees from a large state university mainly used official employer or state websites as well as the official HR VBC Alex as sources of health insurance information, which are both digital and official sources. The previous results highlighted a dimension that seemed to differentiate the preferred source of health insurance information: whether the source of information was an official employer-provided source. The respondents who chose friends or family had a strong preference for this source, suggesting that there may be individual differences that led them to choose this unofficial source as their primary resource for health insurance information. The most prominent nonofficial source consisted of friends or family. Almost half of our sample relied on friends or family as part of their information-seeking strategy, and, interestingly, people who relied on friends or family tended to reach out to them first before considering other sources (27/58, 47% of the respondents who selected friends or family). In addition, our respondents mostly selected 2-3 sources of information, which means that within our sample, most people did not find all the information that they desired within a single source. This was to be expected and has been the trend since the early 2000s for general health information and health insurance information [24-28,38], especially observed in the study by Oetjen et al [31].

Our data analysis showed the following interesting trends. Looking at our first response, official versus nonofficial primary

XSL•FO
RenderX

source, the proportion remains consistent at approximately 0.70 for the 3 middle-aged groups. For health insurance use, we observed a somewhat constant trend. These results were to be expected, given that official sources are suggested and encouraged by the employer's HR department and contain the most specific plan information. For the HILM, we see an increasing trend (although not perfectly linear) across 3 of the 4 measures. Higher literacy is associated with increased use of official sources as the primary source. Given that higher literacy means that people understand and know more about health insurance in general, it makes sense that these people are more likely to better understand the details of the specific plans offered to them. Therefore, this type of specific information will be found in official sources. The fourth measure, which did not show an increasing trend, was the HILM finding out whether the doctor is in-network, which had a very right-skewed distribution, which might have caused such different behavior. Looking at our second response, the age variable showed a decreasing trend, indicating that younger people consider their friends or family as their primary source more than older people. We can imagine that younger people with less experience in plan enrollment might seek trusted advice when it comes to this important decision. For health insurance use, there was a slightly decreasing trend as well. This makes sense because more experienced insurance users are more familiar with how a plan works and, hence, would be less likely to seek trusted advice from friends or family. Finally, for the HILM, only 2 measures had significant associations, and both had behaviors similar to the official primary source response.

Limitations

As suggested in the previous sections, our study included several limitations, which was to be expected given that this is an exploratory study. Allowing for snowball sampling to occur took power away from our data. Our skewed responses also contributed to this. The survey response rate was heavily skewed toward female respondents (103/126, 81.7%), resulting in small sample sizes for the other genders, which limited our ability to perform other analyses. This could be attributed to the fact that women are often the primary decision-makers for health concerns within a family [55] or because women are more likely to take the time to respond to health-related survey questions solicited through email and flyers. Future work should explore whether this finding continues to be robust health insurance decision-making, while also recruiting a larger sample to understand the health information–seeking behaviors of other genders. Our sample also had high technology acceptance and use, which also limits the generalizability of our findings.

In contrast, when studying the survey responses, we made note of possible areas of improvement in the way our questions were written, which will be considered when updating the survey for the next distribution. A limitation of this exploratory study concerned the factors explored in our questionnaire. The factors explored in this study (eg, age, household size, health insurance use, health insurance literacy, and technology adoption) were selected because they are most directly related to health insurance use and the methods used for accessing information; thus, they were factors that we hypothesized to have very strong connections to how individuals may seek sources of information. However, there are many demographic factors that can affect how individuals are able to access and understand different information sources that were not within the scope of this study, including race, numeracy, and health status. These factors should be explored in future work to understand both the systematic reasons for why certain types of health insurance sources are used and the underlying cognitive processes that help individuals understand and seek health insurance information. Another limitation and the most defining change to be made is the list of options for sources of health insurance information provided to the respondents. Our survey did not include sources such as extended networks, nor did it include the types of individual tools and decision aids. VBC Alex and the official websites considered in this study contain tools to support the decision-making process. For example, VBC Alex contains a cost estimator and comparison charts. These were not considered separately; therefore, we could not assess whether these aids are the drivers for the most selected sources.

In our next steps, we plan to extend our research to a more general population. We will create a new survey and work with the university's Institute of Food and Agricultural Sciences (IFAS) extension office. The IFAS extension office is a partnership among state, federal, and county governments to provide scientific knowledge and expertise to the public. The IFAS has locations across the state and in every county. By working with the IFAS extension office, we will have access to community health workers who can help us reach a broader sample of the general population, regardless of employment status, for the new survey. This collaboration will result in access to a larger, more diverse sample with respect to race, income, and employment status. With a broader sample, we will have sufficient power to perform inferential statistics. In this new survey, the list of provided sources of information will be expanded and include more specific tools and decision aids to more effectively answer our research questions and address some of the limitations identified in this study. We expect some of our findings to extend to a more general population. Zhao et al [46], who used a more general population, provided the closest comparison to our study; however, the authors focused on the evaluation of a specific web-based decision tool—the SMHP tool—whereas the aim of our study is to understand the sources of information for health insurance plan selection. Nevertheless, there are some common overlaps. First, Zhao et al [46] also reported a high number of female respondents or participants. Second, they found that high-income populations were more likely to seek web-based information to support their health insurance plan decision-making process. As our inclusion criteria required the participants to be employed full-time, we assumed that our population has a higher income than the general population, and thus the responses in our study are likely biased toward official web-based sources of health insurance information. Finally, we expect that among the general population, participants with high HILM scores will still prefer official primary sources of health insurance information. In addition, in our next steps, we will collect data on the outcome of the enrollment decision and examine relationships between the sources of information (and the number of sources) used and the final enrollment decision.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Health Insurance Information Sources Survey.
[DOCX File , 31 KB - formative_v5i8e27477_app1.docx]

Multimedia Appendix 2
Graph of the ranking occurrences per source. UF: University of Florida, VBC: virtual benefits counselor.
[PNG File , 80 KB - formative_v5i8e27477_app2.png]

References

1. Taylor E, Carman K, Lopez A, Muchow A, Roshan P, Eibner C. Consumer Decisionmaking in the Health Care Marketplace. California, United States: RAND Corporation; 2016:1-98.
2. Bhargava S, Loewenstein G. Choosing a health insurance plan: complexity and consequences. J Am Med Assoc 2015 Dec 15;314(23):2505-2506. [doi: 10.1001/jama.2015.15176] [Medline: 26670967]
3. 4 basic health insurance terms 96% of Americans don't understand. Policygenious. 2018. URL: https://www.policygenius.com/health-insurance/health-insurance-literacy-survey/ [accessed 2020-09-15]
4. Jellyvision - Interactive Employee Communication Software. URL: https://www.jellyvision.com/ [accessed 2020-09-15]
5. Scanlon DP, Chernew M, Lave JR. Consumer health plan choice: current knowledge and future directions. Annu Rev Public Health 1997;18:507-528. [doi: 10.1146/annurev.publhealth.18.1.507] [Medline: 9143729]
6. Gücin N, Berk Ö. Technology acceptance in health care: an integrative review of predictive factors and intervention programs. Procedia Soc Behav Sci 2015;195:1698-1704. [doi: 10.1016/j.sbspro.2015.06.263]
7. Yamin CK, Emani S, Williams DH, Lipsitz SR, Karson AS, Wald JS, et al. The digital divide in adoption and use of a personal health record. Arch Intern Med 2011 Mar 28;171(6):568-574. [doi: 10.1001/archinternmed.2011.34] [Medline: 21444847]
8. Barringer MW, Mitchell OS. Workers' preferences among company-provided health insurance plans. ILR Rev 2016 Jun 25;48(1):141-152. [doi: 10.1177/001979399404800110]
9. Feldman R, Finch M, Dowd B, Cassou S. The demand for employment-based health insurance plans. J Hum Res 1989;24(1):115. [doi: 10.2307/145935]
10. Marquis MS, Long SH. Worker demand for health insurance in the non-group market. J Health Econ 1995 May;14(1):47-63. [doi: 10.1016/0167-6296(94)00035-3] [Medline: 10143489]
11. McFadden D. The measurement of urban travel demand. J Public Econ 1974 Nov;3(4):303-328. [doi: 10.1016/0047-2727(74)90003-6]
12. Garnick DW, Lichtenberg E, Phibbs CS, Luft HS, Peltzman DJ, McPhee SJ. The sensitivity of conditional choice models for hospital care to estimation technique. J Health Econ 1989;8(4):377-397. [doi: 10.1016/0167-6296(90)90022-u] [Medline: 10296934]
13. Chakraborty G, Ettenson R, Gaeth G. How consumers choose health insurance. J Health Care Mark 1994;14(1):21-33. [Medline: 10134041]
14. Edgman-Levitan S, Cleary PD. What information do consumers want and need? Health Aff (Millwood) 1996;15(4):42-56. [doi: 10.1377/hlthaff.15.4.42] [Medline: 8991253]
15. Isaacs SL. Consumer's information needs: results of a national survey. Health Aff (Millwood) 1996;15(4):31-41. [doi: 10.1377/hlthaff.15.4.31] [Medline: 8991252]
16. Tumlinson A, Bottigheimer H, Mahoney P, Stone EM, Hendricks A. Choosing a health plan: what information will consumers use? Health Aff (Millwood) 1997;16(3):229-238. [doi: 10.1377/hlthaff.16.3.229] [Medline: 9141340]
17. Lubalin JS, Harris-Kojetin LD. What do consumers want and need to know in making health care choices? Med Care Res Rev 2016 Dec 15;56(1_suppl):67-102. [doi: 10.1177/1077558799056001s04]
18. Hibbard JH, Jewett JJ. Will quality report cards help consumers? Health Aff (Millwood) 1997;16(3):218-228. [doi: 10.1377/hlthaff.16.3.218] [Medline: 9141339]
19. Gibbs DA, Sangl JA, Burrus B. Consumer perspectives on information needs for health plan choice. Health Care Financ Rev 1996;18(1):55-73 [FREE Full text] [Medline: 10165037]
20. Brodie M, Flournoy RE, Altman DE, Blendon RJ, Benson JM, Rosenbaum MD. Health information, the internet, and the digital divide. Health Aff (Millwood) 2000;19(6):255-265. [doi: 10.1377/hlthaff.19.6.255] [Medline: 11192412]
21. Skinner H, Biscope S, Poland B, Goldberg E. How adolescents use technology for health information: implications for health professionals from focus group studies. J Med Internet Res 2003 Dec 18;5(4):e32 [FREE Full text] [doi: 10.2196/jmir.5.4.e32] [Medline: 14713660]

XSL•FO
RenderX

22. Gray NJ, Klein JD, Noyce PR, Sesselberg TS, Cantrill JA. Health information-seeking behaviour in adolescence: the place of the internet. Soc Sci Med 2005 Apr;60(7):1467-1478. [doi: 10.1016/j.socscimed.2004.08.010] [Medline: 15652680]
23. Bundorf MK, Wagner TH, Singer SJ, Baker LC. Who searches the internet for health information? Health Serv Res 2006 Jun;41(3 Pt 1):819-836 [FREE Full text] [doi: 10.1111/j.1475-6773.2006.00510.x] [Medline: 16704514]
24. Hesse BW, Nelson DE, Kreps GL, Croyle RT, Arora NK, Rimer BK, et al. Trust and sources of health information: the impact of the internet and its implications for health care providers: findings from the first Health Information National Trends Survey. Arch Intern Med 2005 Dec;165(22):2618-2624. [doi: 10.1001/archinte.165.22.2618] [Medline: 16344419]
25. Dutta-Bergman M. Trusted online sources of health information: differences in demographics, health beliefs, and health-information orientation. J Med Internet Res 2003;5(3):e21 [FREE Full text] [doi: 10.2196/jmir.5.3.e21] [Medline: 14517112]
26. Rains SA. Perceptions of traditional information sources and use of the world wide web to seek health information: findings from the health information national trends survey. J Health Commun 2007;12(7):667-680. [doi: 10.1080/10810730701619992] [Medline: 17934943]
27. Ye Y. Correlates of consumer trust in online health information: findings from the health information national trends survey. J Health Commun 2011 Jan;16(1):34-49. [doi: 10.1080/10810730.2010.529491] [Medline: 21086209]
28. Dutta-Bergman MJ. Primary sources of health information: comparisons in the domain of health attitudes, health cognitions, and health behaviors. Health Commun 2004 Jul;16(3):273-288. [doi: 10.1207/s15327027hc1603_1]
29. Mark TL, Coffey RM. Studying the effects of health plan competition: are available data resources up to the task? Health Serv Res 2001 Apr;36(1 Pt 2):253-275 [FREE Full text] [Medline: 11327176]
30. Gruber J, McKnight R. Why did employee health insurance contributions rise? J Health Econ 2003 Dec;22(6):1085-1104. [doi: 10.1016/j.jhealeco.2003.06.001] [Medline: 14604562]
31. Oetjen D, Fottler MD, Unruh L. Determinants of the use of health plan information in formulating health plan selection decisions. Health Mark Q 2003;20(3):55-79. [doi: 10.1300/j026v20n03_05] [Medline: 15018002]
32. Barnes AJ, Hanoch Y, Rice T. Can plan recommendations improve the coverage decisions of vulnerable populations in health insurance marketplaces? PLoS One 2016;11(3):e0151095 [FREE Full text] [doi: 10.1371/journal.pone.0151095] [Medline: 27028008]
33. Barnes AJ, Hanoch Y, Rice T. Determinants of coverage decisions in health insurance marketplaces: consumers' decision-making abilities and the amount of information in their choice environment. Health Serv Res 2015 Mar;50(1):58-80 [FREE Full text] [doi: 10.1111/1475-6773.12181] [Medline: 24779769]
34. Hibbard JH, Cunningham PJ. How engaged are consumers in their health and health care, and why does it matter? Res Brief 2008 Oct(8):1-9. [Medline: 18946947]
35. Hero JO, Sinaiko AD, Kingsdale J, Gruver RS, Galbraith AA. Decision-making experiences of consumers choosing individual-market health insurance plans. Health Aff (Millwood) 2019 Mar;38(3):464-472 [FREE Full text] [doi: 10.1377/hlthaff.2018.05036] [Medline: 30830810]
36. Barnes A, Hanoch Y. Knowledge and understanding of health insurance: challenges and remedies. Isr J Health Policy Res 2017 Jul 13;6(1):40 [FREE Full text] [doi: 10.1186/s13584-017-0163-2] [Medline: 28705191]
37. Wong CA, Nirenburg G, Polsky D, Town R, Baker T. Insurance plan presentation and decision support on HealthCare.gov and state-based web sites created for the affordable care act. Ann Intern Med 2015 Aug 18;163(4):327. [doi: 10.7326/l15-5127]
38. Vardell E. Health Iinsurance Literacy: How People Understand and Make Health Insurance Purchase Decisions. In: ProQuest Diss Theses. Chapel Hill: The University of North Carolina; 2017:1-236.
39. Barnes AJ, Hanoch Y, Rice T, Long SK. Moving beyond blind men and elephants: providing total estimated annual costs improves health insurance decision making. Med Care Res Rev 2017 Oct;74(5):625-635. [doi: 10.1177/1077558716669210] [Medline: 27624636]
40. Nelson L. The good, the bad, and the ugly. Critical Constructivism and Postphenomenology: Ethics, Politics, and the Empirical 2020;24(1-2):195-217. [doi: 10.5840/techne2020514124]
41. Wong CA, Kulhari S, McGeoch EJ, Jones AT, Weiner J, Polsky D, et al. Shopping on the public and private health insurance marketplaces: consumer decision aids and plan presentation. J Gen Intern Med 2018 Aug;33(8):1400-1410 [FREE Full text] [doi: 10.1007/s11606-018-4483-x] [Medline: 29845467]
42. Hempstead K. Consumer decision support on the individual market will be more important than ever. Health Affairs Blog 2017:9 [FREE Full text] [doi: 10.1377/hblog20170131.058536]
43. Wong CA, Polsky DE, Jones AT, Weiner J, Town RJ, Baker T. For third enrollment period, marketplaces expand decision support tools to assist consumers. Health Aff (Millwood) 2016 May;35(4):680-687. [doi: 10.1377/hlthaff.2015.1637] [Medline: 27044969]
44. Politi MC, Barker AR, Kaphingst KA, McBride T, Shacham E, Kebodeaux CS. Show Me My Health Plans: a study protocol of a randomized trial testing a decision support tool for the federal health insurance marketplace in Missouri. BMC Health Serv Res 2016 Mar 16;16:55 [FREE Full text] [doi: 10.1186/s12913-016-1314-9] [Medline: 26880251]
45. Politi MC, Kuzemchak MD, Liu J, Barker AR, Peters E, Ubel PA, et al. Show Me My Health Plans: Using a decision aid to improve decisions in the federal health insurance marketplace. MDM Policy Pract 2016 Jul 16;1(1):238146831667999 [FREE Full text] [doi: 10.1177/2381468316679998] [Medline: 28804780]

XSL•FO
RenderX

46. Zhao J, Mir N, Ackermann N, Kaphingst KA, Politi MC. Dissemination of a web-based tool for supporting health insurance plan decisions (Show Me Health Plans): cross-sectional observational study. J Med Internet Res 2018 Jun 20;20(6):e209 [FREE Full text] [doi: 10.2196/jmir.9829] [Medline: 29925498]
47. Berchick E, Barnett J, Upton R. Health Insurance Coverage in the United States: 2018, Washington, DC: U.S. Census Bureau. Current Population Reports - Report No.: P60–267(RV). 2019. URL: https://www.census.gov/content/dam/Census/library/publications/2019/demo/p60-267.pdf [accessed 2021-07-09]
48. Paez KA, Mallery CJ, Noel H, Pugliese C, McSorley VE, Lucado JL, et al. Development of the Health Insurance Literacy Measure (HILM): conceptualizing and measuring consumer ability to choose and use private health insurance. J Health Commun 2014;19 Suppl 2:225-239 [FREE Full text] [doi: 10.1080/10810730.2014.936568] [Medline: 25315595]
49. Bartholomae S, Russell MB, Braun B, McCoy T. Building Health Insurance Literacy: Evidence from the Smart Choice Health Insurance™ Program. J Fam Econ Iss 2016 Feb 20;37(2):140-155. [doi: 10.1007/s10834-016-9482-7]
50. Reimer B, Mehler B. The effects of a production level "Voice-Command" interface on driver behavior: summary findings on reported workload, physiology, visual attention, and driving performance. Assessing the Demands of Voice Based In-Vehicle Interfaces. 2013 Nov. URL: https://bit.ly/3xCQC9n [accessed 2021-07-09]
51. Reimer B, Mehler B, Dobres J, Mcanulty H, Mehler A, Munger D, et al. The effects of a production level "Voice-Command" interface on driver behavior: summary findings on reported workload, physiology, visual attention, and driving performance. Assessing the Demands of Voice Based In-Vehicle Interfaces. 2014 Jul. URL: https://www.researchgate.net/publication/265594707 [accessed 2021-07-09]
52. Sokol N. The effects of noise-robustness of in-car voice-controlled systems on user perceptions and driving behaviour. Master's Theses, University of Toronto. 2017. URL: https://tspace.library.utoronto.ca/handle/1807/79327 [accessed 2021-07-09]
53. Chakraborty J. An exploration of predictors of adoption and use of in-vehicle voice control systems. Master's Theses, University of Toronto. 2010. URL: https://tspace.library.utoronto.ca/handle/1807/103403 [accessed 2021-07-09]
54. cor: Correlation, Variance and Covariance (Matrices). RDocumentation - cor Function. URL: https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/cor [accessed 2021-01-11]
55. Matoff-Stepp S, Applebaum B, Pooler J, Kavanagh E. Women as health care decision-makers: implications for health care coverage in the United States. J Health Care Poor Underserved 2014 Dec;25(4):1507-1513. [doi: 10.1353/hpu.2014.0154] [Medline: 25418222]

Abbreviations

HIISS: Health Insurance Information Sources Survey
HILM: Health Insurance Literacy Measure
HR: human resources
IFAS: Institute of Food and Agricultural Sciences
SMHP: Show Me My Health Plans
VBC: virtual benefits counselor

Edited by G Eysenbach; submitted 31.01.21; peer-reviewed by D Huang, N Ackermann, Y Hanoch, Z Yan; comments to author 17.03.21; revised version received 29.05.21; accepted 14.06.21; published 12.08.21.

Please cite as:
Colón-Morales CM, Giang WCW, Alvarado M
Informed Decision-making for Health Insurance Enrollment: Survey Study
JMIR Form Res 2021;5(8):e27477
URL: https://formative.jmir.org/2021/8/e27477
doi:10.2196/27477
PMID:34387555

©Coralys M Colón-Morales, Wayne C W Giang, Michelle Alvarado. Originally published in JMIR Formative Research (https://formative.jmir.org), 12.08.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Attitudes of Patients and Health Professionals Regarding Screening Algorithms: Qualitative Study

Christina Oxholm[1*], MPhil; Anne-Marie Soendergaard Christensen[1*], MA, DPhil; Regina Christiansen[1], MPhil; Uffe Kock Wiil[2], DPhil; Anette Søgaard Nielsen[3], MA, DPhil

[1]Department for the Study of Culture, University of Southern Denmark, Odense, Denmark

[2]The Maersk Mc-Kinney Moller Institute, University of Southern Denmark, Odense, Denmark

[3]Research Unit of Clinical Alcohol Research, Department of Clinical Research, University of Southern Denmark, Odense, Denmark

[*]these authors contributed equally

Corresponding Author:
Anne-Marie Soendergaard Christensen, MA, DPhil
Department for the Study of Culture
University of Southern Denmark
Campusvej 55
Odense, 5230
Denmark
Phone: 45 65503630
Email: amsc@sdu.dk

Abstract

Background: As a preamble to an attempt to develop a tool that can aid health professionals at hospitals in identifying whether the patient may have an alcohol abuse problem, this study investigates opinions and attitudes among both health professionals and patients about using patient data from electronic health records (EHRs) in an algorithm screening for alcohol problems.

Objective: The aim of this study was to investigate the attitudes and opinions of patients and health professionals at hospitals regarding the use of previously collected data in developing and implementing an algorithmic helping tool in EHR for screening inexpedient alcohol habits; in addition, the study aims to analyze how patients would feel about *asking* and *being asked* about alcohol by staff, based on a notification in the EHR from such a tool.

Methods: Using semistructured interviews, we interviewed 9 health professionals and 5 patients to explore their opinions and attitudes about an algorithm-based helping tool and about asking and being asked about alcohol usage when being given a reminder from this type of tool. The data were analyzed using an ad hoc method consistent with a close reading and meaning condensing.

Results: The health professionals were both positive and negative about a helping tool grounded in algorithms. They were optimistic about the potential of such a tool to save some time by providing a quick overview if it was easy to use but, on the negative side, noted that this type of helping tool might take away the professionals' instinct. The patients were overall positive about the helping tool, stating that they would find this tool beneficial for preventive care. Some of the patients expressed concerns that the information provided by the tool could be misused.

Conclusions: When developing and implementing an algorithmic helping tool, the following aspects should be considered: (1) making the helping tool as transparent in its recommendations as possible, avoiding black boxing, and ensuring room for professional discretion in clinical decision making; and (2) including and taking into account the attitudes and opinions of patients and health professionals in the design and development process of such an algorithmic helping tool.

(JMIR Form Res 2021;5(8):e17971) doi:10.2196/17971

KEYWORDS

screening; algorithms; alcohol; qualitative study; attitudes; opinions; patients; health professionals

XSL•FO
RenderX

Introduction

Background

The more specialized hospital treatments and hospital departments become, the more challenging it will be to maintain an overview of all the information collected in the hospitals' electronic health records (EHRs). It has been suggested that a software algorithm may be a reliable strategy for automatic screening of EHRs [1]. Software using data mining, machine learning, and natural language processing may not only prove to be useful for overseeing the vast amount of data in the EHRs but using software to generate automatic messages to health professionals for decision making may also remove barriers when it comes to talking to their patients about sensitive topics such as alcohol.

Hospitalized Interventions Aimed at Reducing Harmful Alcohol Intake

There is an overwhelming body of evidence proving the negative influence of the substantial use of alcohol in the areas of both public health and health economics. Alcohol is a significant risk factor for bad health and premature death, and alcohol use disorder is responsible for considerable physical morbidity and injuries [2]. Alcohol is leading to, or complicating, at least 60 diagnoses, and excessive drinking in Denmark alone is related to 28,000 hospital admissions, 10,000 emergency room visits, and an additional annual cost of health services of DKK 947 million (US $151.05 million) [3].

In studies on recovery from alcohol problems, health problems and hospital admissions are among the most cited predictors of recovery [4]. Health problems and hospital admissions may open a window for changing alcohol consumption if this opportunity is exploited. In other words, general hospitals may be in a good position to identify individuals with alcohol issues. Hence, the screening, brief intervention, and referral to specialized treatment (SBIRT) approach for alcohol use disorder during hospitalization is considered suitable to address and lower the alcohol intake of patients [5].

The SBIRT approach aims to identify patients with high alcohol intake (the screening component). When identified, the next step is to increase the patients' awareness of their alcohol intake and increase their motivation to lower it (the brief intervention component), which is most often based on the principles of motivational interviewing [6]. The final part of the intervention sends the patients who need treatment for alcohol use disorder to specialized treatment (the referral to treatment component). A series of projects have tried to implement the SBIRT approach, but it has proved indeed very challenging [7,8].

In particular, it seems that health professionals are reluctant to both screen for risky alcohol use among patients due to the time spent on the procedures, and address drinking issues when the patients screen positive [9,10]. Health professionals express, however, that they are willing to talk about alcohol with their patients if they have a reason for doing so [10,11]. One difficult barrier seems to be that the excessive alcohol intake, even daily, is hard to observe and detect because large alcohol intake is often invisible. Therefore, health professionals may be afraid of insulting nondrinkers if they screen for excessive alcohol use systematically. Hence, health professionals' discomfort and avoidance of the topic have led to alcohol problems being ignored [12,13].

Helping Tools for Health Professionals Based on Data From EHRs

Predictive models based on EHR data may be a way to help health professionals at hospitals to identify patients with alcohol use disorder. Data mining and machine learning techniques have been used extensively for predictive models and clinical decision support. Indeed, Escobar et al [14,15] aimed to predict the occurrence of an adverse event to avoid hospitalized patients being transferred to the intensive care unit. Hackmann et al [16] developed a clinical warning system that simultaneously reduces the risk of false positives while ensuring that the right patients are administered into the monitoring program. Mishra et al [17] analyzed discharge summaries to identify diabetes, protocol compliance, and high-risk factors; this was also done using simple concept extraction methods. So far, an algorithm-based helping tool that screens EHR data to inform hospital staff that the patient might have a complicating use of alcohol that should be addressed in order to improve the prognosis of the patient has not been developed [18].

However, using machine learning techniques, it seems possible to develop an algorithm that screens data already stored in patient case notes and that can be the backbone of a clinical decision tool for identifying possible harmful alcohol use [19]. If successful, the clinical decision tool would give a message to health professionals if the patients screen positive for harmful use. The tentative models for developing reliable algorithms central to such tool are promising [19]. However, before developing predictive models and algorithms that are ready to be implemented in clinical practice, we need to know if a clinical decision support tool, which can scan data already stored in patient case notes in EHRs and informing health professionals about indications of harmful use of alcohol, is considered acceptable by patients and health professionals. The risk is that such a tool will be perceived as unethical and as "big brother is watching you."

There are a several reports mapping the ethical issues involved in the development and use of clinical decision support tools on theoretical levels [20], and various studies on how to achieve user acceptance of such tools [21]. However, we lack knowledge on the attitudes and ethical considerations of patients and health care professionals in relation to such tools. Thus, this study will investigate the attitudes and ethical considerations among both health professionals at hospitals and patients toward using patient data already stored in the EHR, to develop an algorithm and a subsequent helping tool to inform staff that harmful alcohol use may be a complicating factor for the patient in question.

Methods

Recruitment Process/Strategy

The participants were primarily recruited at Odense University Hospital, Denmark, in the Department of Neurology, the

XSL•FO
RenderX

Department of Orthopedics and Traumatology, and the Department of Gastroenterology. Our overall goal was to interview 2-3 health professionals and 2-3 patients in each department. Because of the short and varying number of admissions and an unpredictable work environment, we were unable to ensure a specific number and variation of participants in advance.

Unfortunately, there were not enough patients on the days of the interviews who were fit enough or willing to be interviewed. Because we did not get enough diversity of opinions with the admitted patients, we decided to expand our study to include nonadmitted potential patients to achieve thematic saturation. The rationale for including nonadmitted potential patients was that everyone is a *potential* patient and has been seen by a doctor at some point in their lives and can therefore relate to and have opinions about being asked about their alcohol use by a health professional. Furthermore, potential patients often have a greater surplus of mental resources to consider the questions in the interview than ill, hospitalized patients. Therefore, 2 participants in the patient group were recruited through an open call via the University of Southern Denmark's network.

Ethics, Consent, and Permission

Because the project was solely based on voluntary interviews and did not involve any biological material, the Regional Committees on Health Research Ethics for Southern Denmark ruled that there was no need to apply for permission to conduct the interviews.

The heads of each medical department were contacted via email and given thorough information about the project beforehand. Together, we scheduled a date for conducting the interviews. Three separate dates were arranged with each of the medical departments.

To ensure that we did not interview patients who were either not fit enough or unable to give informed consent, we consulted with the health professionals. The health professionals selected which patients would be physically and mentally capable of participating. This gave us ethical assurance that the patients who were interviewed would not suffer in terms of well-being and could, in fact, provide informed consent. The patients were first approached by one of the health professionals who gave them a basic outline of the project and asked them if they would like to learn more and, potentially, participate. This was to ensure that the patients did not feel uncomfortable when being approached by a stranger. Besides, making the health professional establish first contact made our project more trustworthy and credible. If the patient wanted to hear more about the project, the interviewer (CO) presented herself and the project in detail.

Before starting the interviews, we obtained written consent from all the participants. The consent ensured the participants their confidentiality and anonymity. The participants were also informed that they would be able to withdraw their consent at any time in the process.

Data Collection

This study makes use of a hermeneutic-phenomenological approach. Because of the study's clear aim and limited time available for both admitted patients and health professionals, we decided that semistructured interviews would be the best way to approach our research question. We developed the interview guide in a relatively open manner to avoid influencing the participants' attitudes with specific words or phrases. The interview guide was adjusted by conducting a series of test interviews to ensure that the questions were understandable and that the interview was structured in the right way.

Because of the study's explorative focus on patient and health professional's attitudes toward, and opinions on, use of technology in health care rather than technology compliance, we did not base this study on a technology acceptance framework.

In total, 14 participants were interviewed (9 health professionals and 5 patients). The health professionals were selected by the head of each department, depending on their availability. Table 1 provides an overview of the group of health professionals interviewed in the study. The interviews of the health professionals were conducted in a small, quiet meeting room in each department.

A total of 5 patients were interviewed for this study, 3 of whom were patients admitted at the Odense University Hospital. The remaining 2 patients were recruited through the open call. The interviews with the admitted patients took place in their respective rooms in the hospital. The nonadmitted patient interviews were held at the University of Southern Denmark in a private office. Table 2 presents the gender of the patients interviewed.

Table 1. Health professionals: an overview.

Gender	Doctors, n	Nurses, n	Social and health service assistant, n
Female	2	4	2
Male	1	—	—

Table 2. Gender of the patients.

Gender	Patients, n
Female	3
Male	2

XSL•FO
RenderX

All interviews had a maximum duration of 30 minutes and were audio recorded. The interviews focused on patients' and health professionals' views on addressing the patients' alcohol use during hospitalization and the patients' and health professionals' views on developing automatic screening tools based on algorithms that can screen EHRs for signs of problematic alcohol use. All interviews were conducted by the same person (CO) who used a semistructured interview guide. CO does not have a health-related educational background. Thus, she did not have any specific ideas about good and bad practices in health care and was not familiar with the health profession's norms and rules beforehand. This allowed her to be open toward participants and perform the interviews without any prejudice.

Data Analysis

The audio files were transcribed in MSWord and analyzed ad hoc. The analysis primarily involved a process of close reading of the printed transcriptions and meaning condensing and color coding by hand. This process did not include any computer-assisted qualitative software. The analysis was primarily carried out by CO and sent back and forth to A-MC, RC, and AN, who then evaluated and gave their feedback on the analysis independent of each other. A-MC, RC, and AN also had the full transcriptions and, therefore, could make qualified comments on something that might have been missed or that needed further exploration. The data were analyzed for both the patients' and health professionals' views on addressing alcohol use during hospitalization, and for the patients' and health professionals' views on developing an algorithm and screening existing EHRs. The patients' and health professionals' views on addressing alcohol abuse are presented elsewhere [11]. This paper focuses solely on the attitude and ethics regarding automatic screening of EHRs for signs of alcohol abuse.

Results

Overview

To investigate the health professionals' and patients' opinions on and attitudes toward an algorithm-based helping tool for screening alcohol habits, the questions in the interview contained 2 different perspectives. First, the interview questions asked about the respondent's general views on an algorithm-based helping tool and their opinions about *asking* and *being asked*, based on a reminder from the same helping tool.

Health Professionals

Interview Outcomes

To unfold health professionals' opinions on an algorithm-based helping tool, they were asked about the general use of information technology (IT) in health care (mainly EHR and other IT systems regarding registration and documentation). This was done to test whether a possible negative opinion about the general use of IT in health care was affecting their opinion regarding a specific helping tool. Interrogating the health professionals' opinions and attitudes about *asking* about alcohol, based on a reminder from the helping tool, was a way to make them consider the helping tool from a more practical perspective. This would uncover any inconsistencies in their answers.

Opinions on the General Use of IT in Health Care

Regarding the question about their general opinions on the use of IT in health care, 8 out of the 9 respondents had both positive and negative opinions, with the remaining 1 being entirely positive about the general use of IT in health care. On the positive side were, for example, that the EHRs' availability saves valuable time compared with the old paper journals, and that EHRs provide a quick and useful overview of the patients. Even though none of the health professionals explicitly expressed this claim, the fact that all 9 of them mentioned time saving as a positive aspect of IT in health care serves as a strong indicator that time may be limited, and therefore, a valuable resource in health care.

On the negative side, a recurring theme among the health professionals was that the documentation requirements following EHRs were considered time-consuming. Some added that it sometimes was experienced as meaningless, for example, because it doubled documentation and took valuable time away from the patients.

One health professional (HP-2) expressed this last point by saying that having to withdraw from the patients to spend a large amount of time in front of a computer was not why she became a health professional. At the same time, she also expressed a view on IT as having many advantages; for example, it can save a lot of time when reading up on patients. Another health professional (HP-7) who also had mixed opinions about IT pointed to some specific challenges: (1) that the restrictions on access to information are a limitation because this can be a hindrance for some health professionals when accessing valuable information; (2) that some of the IT systems do not work together, which results in professionals having to document the same thing in 2 different systems; and (3) that the high documentation requirements produce a vast amount of data, which can be difficult to navigate through when trying to find specific information. As HP-2, HP-7 also expressed positive views on the use of IT in health care when compared with old paper journals. These 2 examples sum up most of the health professionals' ambivalence toward the general use of IT in health care: the paradox that IT both saves time and requires a lot of time. Although most of the health professionals shared this ambivalent opinion, no one was exclusively negative about the general use of IT in health care.

Opinions on an Algorithm-Based Helping Tool

Six out of the 9 health professionals (HP-1, HP-5, HP-6, HP-7, HP-8, and HP-9) were predominantly positive about the idea of an automatic and algorithm-based helping tool, with 1 (HP-2) being exclusively positive about it. The remaining 2 health professionals (HP-3 and HP-4) were ambivalent. A common denominator among all 9 health professionals in their positive opinions about this type of helping tool was how these opinions almost exactly mirrored their positive opinions about IT in general: the possibility of *saving time* thanks to the automatic sorting of data, getting a *quick overview* of the relevant issues concerning the individual patient, and making sure that they do not forget important tasks by being reminded of them (especially for new health professionals).

XSL•FO
RenderX

Although 6 of the health professionals were predominantly positive and expressed that they would consider the tool helpful in their work, they also expressed that their positive attitudes were conditioned by the following requirements: (1) that the helping tool would be adapted to each department's field of expertise, (2) that it would not be time-consuming, (3) that there would be sufficient time and resources spent on its implementation, (4) that the helping tool would actually work and be useful, (5) that the tool would be designed in a user-friendly way, and (6) that the system would not be implemented in a top–down manner. The conditions were distributed across the answers of 6 of the respondents. Even though the respondents all had some reservations, they were categorized as being predominantly positive, because they did not object to anything related to the helping tool itself, but rather to some conditions surrounding it.

HP-1, who expressed that she would want the helping tool to be adapted to each department's field of expertise, elaborated that she would find it very frustrating if such a tool was implemented in a top–down manner without any consideration of how the tool would affect or benefit the ones using it. In her opinion, not all initiatives would be equally relevant for every field of expertise, and if the tool provides some users with irrelevant information, it would be more of a disturbance than a benefit. If the tool would be adapted according to relevance, she would be positively disposed toward this kind of technology. HP-1's reservation could point to a more general frustration that health professionals experience: not being included in the development and implementation of systems or guidelines that have a substantial influence on how they work.

HP-9, who expressed the condition of a user-friendly design, also mentioned a possible pitfall if the data used for the screenings were inadequate or incomplete. In other words, if the health professionals—for one reason or another—did not document items correctly, sufficiently, or in the right place, then the screenings could potentially produce errors (eg, in the form of faulty recommendations).

Although none of the health professionals were exclusively negative about the idea, 2 (HP-3 and HP-4) were clearly ambivalent in their opinions on the matter and were neither predominantly positive nor predominantly negative. Their reservations about the helping tool were of a more fundamental character than the ones expressed by those with predominantly positive opinions because they touched upon how algorithm-based helping tools could affect the nature of being a health professional.

Even though HP-3 was very positive about this type of helping tool, stating that she would consider such a tool helpful in her work, she also expressed a profound concern that the tool would take away the health professionals' instinct: "I think that it is fine to make them [the helping tools], but you have to be careful that it does not become a false safety for the health professionals, that they are going to use it blindly without using...the instinct". She was concerned that the helping tools would make health professionals "lazy" by following their instructions blindly without exercising professional discretion. She had mixed feelings on the matter, because while having this concern, she could also see great potential benefits from having this type of tool.

HP-4 was also very ambivalent about helping tools. On the one hand, she was very positive about algorithms assisting her with medication interactions, but very negative about algorithms assisting her in the diagnostic process. In her opinion, algorithms cannot make complex assessments regarding a patient's health status. She gave an example of a triage algorithm that evaluates the status of acute patients according to 5 different categories. In her experience, triage algorithms are more of a constraint than help, because they either triage some irrelevant parameter very high or do not catch very serious conditions. According to her, triage algorithms are generally wrong because they cannot evaluate complex medical problems such as a patient's health status. Therefore, she relies more on her expertise and professional discretion than on these algorithms. This triage algorithm, she said, sometimes makes it difficult for her to be a good clinician because it interferes with her professional judgments: "(...) it can be a problem being...to be an understanding...a good clinician when there is too much that becomes algorithm-based (...)". But this does not mean that HP-4 was negative about algorithm-based helping tools per se, but rather that she was negative or skeptical about their *usefulness* and *competence* to do certain tasks, for example, assessing a patient's health status. If a helping tool could help in more "black-and-white" matters (eg, medication interactions), she was very positive about such initiatives.

Most of the health professionals we interviewed had a predominantly positive attitude regarding helping tools based on algorithms, despite most of them being rather ambivalent regarding the *general* use of IT in health care. Nothing suggests that the health professionals' mixed opinions about IT in general affected their opinions and attitudes regarding the idea of an algorithm-based helping tool.

Asking on the Basis of a Reminder

A total of 6 of the 9 respondents were positive about having to ask their patients about alcohol based on a reminder from an algorithm-based helping tool. Most of them stated that they would consider it a help in their work, and some even said that they thought the helping tool would make it *easier* for them to ask patients, because it would give them a sense of having a valid reason or excuse to ask the patients about the patients' alcohol habits (HP-1, HP-3, and HP-7). These answers indicate that asking about alcohol can be a difficult task for some health professionals. Another reason behind this positive attitude was that the helping tool would ensure that they would remember to ask the patients in the first place.

One of the participants (HP-6) expressed mixed opinions about asking patients based on being reminded by an algorithm. She stated that she would be critical about the reminder and first examine if she agreed that alcohol would be a relevant thing to ask about a specific case. If she were to agree with the algorithm, she would be positive about being reminded. But if this was not the case and she did not find it relevant, she stated that she would ignore the reminder and not ask the patient. This skeptical attitude was also present in her answer regarding whether she would consider the helping tool a help or nuisance in her work.

XSL•FO
RenderX

Here, she stated that she would consider the system a help in her work only when she would otherwise have overlooked the relevance of alcohol.

HP-4, who was very negative about algorithm-based helping tools making clinical assessments about patients' health status, would ignore a reminder asking her to talk to the patients about their alcohol habits. Her reasons behind ignoring the algorithm were similar to the ones she gave when asked about the helping tool in general: She did not think that algorithms can provide clinical assessments, because this is too complex a task for a computer system, which is why she would ignore the reminder and rely on her professional intuition instead of an algorithm.

HP-8 stands out here by not having any problem with asking about alcohol based on a reminder, while also pointing out that the alcohol habits of patients are not of any interest to her. This means that HP-8 would not be likely to ask at all—reminded or not—because of the conviction that this issue is of little relevance in her work. The problem for HP-8 is thus not the reminder of the helping tool per se, but alcohol as a subject. HP-8 might be positive about asking about other cases based on a reminder.

Patients

Interview Outcomes

To investigate their opinions about an algorithm-based helping tool, the patients were also asked about their opinions on the general use of IT in health care (mainly EHRs and other IT systems). This question was asked to uncover if there was a potential negative attitude about the helping tool arising out of a negative or skeptical attitude about the use of IT in health care in general.

To discover the patients' opinions of the practical use of the helping tool, they were asked several—but related—questions on this matter. These questions included their views on the following: (1) using a helping tool to screen for increased risk or disposition for specific diseases (eg, Alzheimer, cancer, strokes), (2) using a helping tool to screen for signs of alcohol and lifestyle diseases, (3) having their personal EHRs screened for alcohol habits, and, lastly (4) being *asked* about alcohol habits by a health professional, who was reminded to do so by an algorithm-based helping tool. Questions 1 and 2 were asked to test whether it was the *helping tool* or the *use* of the helping tool that they might have an issue with. Questions 3 and 4 were used to uncover whether their general opinion about the helping tool also applied to cases concerning themselves.

Opinions on the General Use of IT in Health Care

All patients had a positive attitude about the general use of IT in health care. One of the patients was exclusively positive about the general use of IT, with the remaining 4 being predominantly positive but at the same time expressing some reservations about the potential negative side effects connected to using IT in health care. The concerns raised by the 4 respondents were as follows: (1) that IT could take away valuable time from the patients, (2) that EHRs are more susceptible to abuse than the old paper journals, and (3) that the patients' access to their own EHRs may be a cause of unnecessary worry. Again, this listing of concerns is not to be perceived as if all the patients raised each of these concerns.

The last concern was raised by P-5, who was interviewed as a patient, but worked as a health professional. Her point was not that patient access to their EHRs is a negative thing, but that it does have the potential to make patients feel anxious if they read the doctors' notes online and do not understand the medical terms used. This is possibly because the doctors' notes and test results are often available in the online and patient-accessible EHRs before the patients' appointment with their doctor. The concern about patient access to EHRs online was raised because she had experienced this sort of dilemma in her personal life and saw what kind of fear it could spark in a patient—sometimes for no reason.

On the positive side regarding the general use of IT in health care, the patients mentioned the following: (1) that using IT is an inevitability and the right step forward in health care, (2) the accessibility of EHRs is an advantage, and (3) the possibility of giving health professionals a heads-up for any sensitivity or allergies to medicine is an important and positive aspect of the use of IT in health care.

Opinions on an Algorithm-Based Screening of EHRs

Out of the 5 respondents, 2 (P-1 and P-3) were exclusively positive about screening for both an increased risk or disposition for disease and inexpedient alcohol habits. One of the 5 (P-2) was predominantly positive but raised some concerns, while the remaining 2 (P-4 and P-5) were ambivalent, with no clear preference given to neither the positive nor negative aspects they mentioned.

The 2 patients who were exclusively positive about screenings said that they thought the screenings would be a beneficial tool for preventive care and, ultimately, a help for the patients. As for the respondent (P-2) who was predominantly positive, but had raised some concerns, these concerns were primarily linked to the screening of inexpedient alcohol habits; he was concerned that the knowledge following this kind of screening could be the subject of abuse, but did not elaborate on what kind of abuse. Regarding the screening for increased risk for disease, he did not express the same type of concerns.

The 2 patients who were ambivalent said they were positive about the screenings as a useful tool for preventive care. Yet, at the same time, they were somewhat skeptical about the potential of misusing both the information about increased risk of certain diseases and the information about inexpedient alcohol habits by, for example, insurance companies. Another major concern expressed by P-4 was the risk of false positives and how this could affect a patient's life. Having experienced the consequences of a false positive herself, she was naturally concerned about this aspect of screenings. P-5 expressed concerns about how the results of screenings can have negative effects, such as stigmatization of the patient, which could result in inferior treatment. Although both P-4 and P-5 could see the benefits of these screenings, it was not entirely clear whether the positive aspects could outweigh the negative ones. In other words, in their answers to these questions, it is not clear whether they were predominantly positive or negative about screenings.

XSL•FO
RenderX

Therefore, they were categorized as ambivalent. However, as shown in the next section, they might lean a little toward the more positive side.

When asked about whether they would personally agree to let their EHRs be screened for inexpedient alcohol habits, 3 of the patients (P-1, P-2, and P-3) said yes, unconditionally, while the remaining 2 (P-4 and P-5) were predominantly positive but had some reservations on the matter. Two of those who were exclusively positive about letting their personal EHRs be screened for inexpedient alcohol habits were also positive about screening in general. P-2, who expressed concerns about screening in general, did not express the same concerns about having his personal EHRs screened.

P-4 and P-5, who were ambivalent about screening in general, expressed a more positive attitude about letting their personal EHR screen for inexpedient alcohol habits. P-4 stated that her positive attitude was probably conditioned by the fact that she knows that alcohol is not a problem for her. If it was, she would likely feel less comfortable with the screening but would still allow it because she would like to know the result. P-5's reservations were not about the screening itself but rather about the potential misuse of the information that the screening can produce. It was very important for her that the results or information from the screening would only be used in a way that helps and benefits the patient and not in any negative way (eg, condemnation and stigmatization). Although both could see personal advantages and, therefore, would like to know the results of a screening of their personal EHRs, they also expressed reservations about the matter. That they did not altogether reject the idea of having their personal EHRs screened for inexpedient alcohol habits is an indication that they were leaning toward a more positive attitude about screening in general.

Being Asked Based on a Reminder

The last aspect of the overall question about patients' attitudes and opinions about an algorithm-based helping tool was how they would feel about being *asked* about this subject based on a screening. Three out of the 5 patients (P-1, P-2, and P-3) had no reservations about being asked based on a screening and said that they would be comfortable with this. The remaining 2 respondents (P-4 and P-5) had some reservations about the matter; these reservations did not concern the screening per se, but rather about *how* the health professionals would handle the asking, based on being reminded. For them to be comfortable about being asked would depend on *the way* the health professional asked them. However, the importance placed on how they are asked about alcohol habits was important for them, *regardless* of the reasons behind them being asked. In other words, being comfortable about being asked about alcohol habits depended on *how* they were asked. Further, they mentioned that they would like to be informed about the reasons behind being asked—whether the reason is a screening, routine, or suspicion. This means that the reservation they had about being asked was not linked to the screening per se, but to being asked in general.

Discussion

Summary

As a preparatory step before the development of an algorithm and a subsequent helping tool that would present notifications in the EHR by means of algorithms screening the data if the patient was considered to have a possible harmful use of alcohol, we conducted interviews to investigate the attitudes and ethical considerations among patients and health professionals at hospitals toward such a helping tool. The health professionals had both positive and negative opinions about such an algorithm-based helping tool. On the positive side, the health professionals noted that the helping tool would save them some much needed time by providing a quick overview of the information in the EHR and ensuring important tasks such as addressing harmful use of alcohol would not be forgotten. However, this positive attitude was conditioned by a number of requirements: That the tool would not be time-consuming and would be adapted according to relevance and usefulness, that it would work and be user-friendly, and that sufficient resources would be spent on implementation. On the negative side, the health professionals noted concerns that this type of helping tool would take away the health professionals' instinct, because they might follow recommendations blindly without exercising professional discretion, and they also have concerns of a more fundamental nature, questioning the algorithm's capability of making and giving clinical decisions and recommendations. The patients were overall very positive about the idea of an algorithm-based screening tool, saying that they would consider such a tool beneficial for preventive care, thereby ultimately helping patients in need for advice about alcohol habits. However, some expressed concerns that such a tool would provide information that could be misused, that the screenings could result in stigmatization and inferior treatment, and that false positives could impact patients' lives.

Comment on the Practical Nature of the Health Professionals' Answers

Something very characteristic about the health professionals' positive and negative opinions about both the general use of IT in health care and the algorithm-based helping tool was that—with a few exceptions—they were very practical. In other words, their justifications for being both positive and negative were related to how IT and the helping tool would affect them in a practical way. Only 2 health professionals expressed opinions about how the use of IT and, specifically, an algorithm-based helping tool might affect the nature of health care in a more general way. Even though we did not ask them specifically about this more general and ethical perspective, it is interesting that so few brought it up, and that they only gave practical justifications. There can be a number of reasons for this: (1) we did not ask them specifically about a more general perspective on the use of IT in health care, (2) they did not have any reflections on the more general and ethical perspective, or (3) the time frame for the interview was short, and they were at work. However, a focus on the practical perspectives could also be a useful and an important insight for anyone developing and implementing these systems. Indeed, if you want health

XSL•FO
RenderX

professionals to use and comply with a new IT system, these professionals must be included in the development and implementation processes, that is, in how the systems will affect them in a practical way.

Ethical Issues

The introduction and use of algorithms can essentially change the health care system and how medicine is practiced today. Therefore, it is important to take seriously how the use of algorithms can ethically challenge the fundamental aspects of medical practice. Some of the ethical challenges that might arise out of using algorithm-based helping tools are (1) the patient's privacy, (2) the autonomy of health professionals, and (3) the relationship between the patient and health professional.

Using algorithm-based helping tools to screen previously collected health data has obvious gains regarding patient beneficence. At the same time, however, this screening can be a possible breach of privacy because of the flow of data from one context to the other, as data collected in one context may be used for an algorithm-based screening in another context. These screenings can be a useful tool in the preventive treatment of, for example, inexpedient alcohol habits. However, the ethical tension between beneficence and privacy arises because the screening may give health professionals access to information about the patient that he/she has chosen not to disclose or found irrelevant or inappropriate in that specific health care context. Therefore, it is important to weigh the concern for patient beneficence and patient privacy when developing and implementing this type of algorithm-based helping tool.

A central aspect of professional autonomy is the exercise and cultivation of *professional practical wisdom* [22]. Such practical knowledge entails that the professional considers a broad range of possible issues, decisions, and actions when contemplating clinical decisions. Another way of putting this is to say that professional practical wisdom makes good, professional discretion possible. In the context of professional autonomy, it is useful to distinguish between algorithm-based helping tools that give clinical decision *support* via recommendations, and tools that *make* clinical decisions because they have different impacts on professional autonomy, with the latter being an authoritative helping tool. If the tool is used to offer *support*, this can be very constructive for the exercise and cultivation of professional practical wisdom because it can draw attention to important considerations that the professional might not otherwise have given thought to. If the tool is used to *make* clinical decisions, it can become a threat to professional autonomy because it restricts the practitioner's ability to exercise professional discretion. However, constitutional constraints, such as time pressure and the design of the helping tool, can influence whether the clinical support tool is used and perceived as an authoritative tool or not, if, for example, because of time pressure or because the recommendation is phrased in an authoritative way, the professional follows the recommendations blindly without exercising professional discretion. This would, de facto, mean that a clinical decision *supportive* tool would become a clinical decision-*making* tool, thereby posing a threat to professional autonomy, because a clinical decision-making tool leaves no room for exercising professional practical wisdom.

The patient-centered relationship is currently the most widely accepted ideal for the doctor–patient relationship in the Western world [23]. This type of doctor–patient relationship is characterized by the following 5 aspects: (1) the biopsychosocial perspective, (2) the patient-as-person, (3) shared power and responsibility, (4) the therapeutic alliance, and (5) the practitioner-as-person [24]. Introducing and using algorithm-based helping tools can have an impact on 4 out of the 5 central aspects of the patient-centered relationship, leaving only the fifth aspect, the practitioner-as-person, untouched. One of the aspects that could be affected is the shared power and responsibility between the doctor and patient; this aspect entails that the doctor and patient are equal in their autonomy and authority because they both possess expert knowledge—the patient about personal needs and preferences, while the doctor has the required medical knowledge—which is essential to the shared decision-making process. The patient's autonomy in the patient-centered relationship is based on being *heard* and *receiving* expert medical knowledge, making an informed decision possible. By contrast, the doctor's autonomy, in this aspect, is based on *having* medical knowledge. Algorithm-based helping tools can interfere in this central exchange of knowledge if the doctor does not understand the decisions or suggestions of the helping tool because of, for example, black boxing, thereby restricting both the patient's and doctor's autonomy by not being able to respectfully receiving and giving expert knowledge. Introducing and using algorithm-based helping tools can intervene in the doctor–patient relationship in 4 out of these 5 central aspects of the patient-centered relationship and, ultimately, change this relationship into a more paternalistic relationship, where the autonomy and authority are centered around algorithms, not the doctor and patient.

Limitations

This study has 3 main limitations: (1) the small sample size, (2) the patient's sometimes restricted ability to participate in the interviews, and (3) the physical frameworks of the interviews. Even though the correlation among respondent answers was good, this is a limitation. An obvious way of furthering this study would be to make a quantitative investigation based on the same research questions. This would ensure a larger population, thereby strengthening the study. The second limitation relates to the admitted patients who did not necessarily have sufficient energy and strength to participate in an in-depth interview. Even though we ensured that the health professionals approved of the patients' participation, this is not necessarily a guarantee. This limitation was obvious when we interviewed the nonadmitted patients, who had significantly more mental surplus and, therefore, gave more nuanced and lengthy answers. The third limitation concerns the physical frameworks of the interview. Some of the admitted patients were admitted in multibed wards and were, therefore, interviewed with other patients present. This may have influenced their level of comfort in being interviewed and, ultimately, their answers. The interviews with health professionals were conducted in a small office away from patients. Even though the interviews were held at a distance from their respective departments, they

XSL•FO
RenderX

nevertheless were at work and, therefore, had limited time available. Hence, the situation they were in right before the interview—which might have been a difficult one—could have influenced their concentration and mental presence.

Recommendations for Development and Implementation

This study's results highlight the importance of designing algorithmic helping tools to be as transparent as possible. Another way of putting this is to say that the helping tool must not be designed to only provide yes/no answers. This would, de facto, "black box" important information and make the health professionals unable to use professional discretion to evaluate the recommendation made by the helping tool. Therefore, an algorithmic helping tool should provide some insight into why a recommendation was given. This way, the health professional will be allowed to judge whether to agree or disagree with the recommendation based on the more detailed information.

Another key result highlights the importance of including health professionals in the development and design process of algorithmic helping tools; indeed, they hold important and valuable knowledge about what key factors will increase use of such a system. This is important to ensure that the helping tools being implemented are, indeed, helping health professionals and not creating frustration in an already busy work environment.

Conflicts of Interest

None declared.

References

1. Hornbrook MC, Whitlock EP, Berg CJ, Callaghan WM, Bachman DJ, Gold R, et al. Development of an algorithm to identify pregnancy episodes in an integrated health care delivery system. Health Serv Res 2007 Apr;42(2):908-927 [FREE Full text] [doi: 10.1111/j.1475-6773.2006.00635.x] [Medline: 17362224]
2. World Health Organization. Global Status Report on Alcohol and Health. Geneva, Switzerland: World Health Organization; 2014. URL: https://www.who.int/substance_abuse/publications/global_alcohol_report/msb_gsr_2014_1.pdf [accessed 2021-06-20]
3. Juel K, Sørensen J, Brønnum-hansen H. Risikofaktorer og Folkesundhed i Danmark. Copenhagen, Denmark: National Institute of Public Health, University of Southern Denmark URL: https://www.sst.dk/~/media/762BA0CB12714748810B93946E95730F.ashx [accessed 2021-06-20]
4. Sobell LC, Ellingstad T, Sobell M. Natural recovery from alcohol and drug problems: methodological review of the research with suggestions for future directions. Addiction 2000 May;95(5):749-764. [doi: 10.1046/j.1360-0443.2000.95574911.x] [Medline: 10885050]
5. Babor TF, McRee BG, Kassebaum PA, Grimaldi PL, Ahmed K, Bray J. Screening, Brief Intervention, and Referral to Treatment (SBIRT). Substance Abuse 2007 Nov 21;28(3):7-30. [doi: 10.1300/j465v28n03_03]
6. Miller WR, Rollnick S. Motivational interviewing: helping people change. New York, NY: Guilford Press; 2012.
7. Watson JM, Fairhurst C, Li J, Tober G, Crosby H, Lloyd C, ARiAS Research Group. ADAPTA: A pilot randomised controlled trial of an alcohol-focused intervention versus a healthy living intervention for problem drinkers identified in a general hospital setting. Drug Alcohol Depend 2015 Sep 01;154:117-124 [FREE Full text] [doi: 10.1016/j.drugalcdep.2015.06.030] [Medline: 26194955]
8. Schwarz A, Nielsen B, Søgaard J, Søgaard Nielsen A. Making a bridge between general hospital and specialised community-based treatment for alcohol use disorder-A pragmatic randomised controlled trial. Drug Alcohol Depend 2019 Mar 01;196:51-56. [doi: 10.1016/j.drugalcdep.2018.12.017] [Medline: 30665152]
9. Shourie S, Conigrave K, Proude E, Haber P. Detection of and intervention for excessive alcohol and tobacco use among adult hospital in-patients. Drug Alcohol Rev 2007 Mar;26(2):127-133. [doi: 10.1080/09595230601145175] [Medline: 17364847]
10. Hellum R, Bjerregaard L, Nielsen AS. Factors influencing whether nurses talk to somatic patients about their alcohol consumption. Nordic Studies on Alcohol and Drugs 2017 Feb 08;33(4):415-436. [doi: 10.1515/nsad-2016-0034]
11. Oxholm C, Christensen AS, Christiansen R, Nielsen AS. Can We Talk about Alcohol for a Minute? Thoughts and Opinions Expressed by Health Professionals and Patients at a Somatic Hospital. Alcoholism Treatment Quarterly 2020 Aug 12;39(1):119-130. [doi: 10.1080/07347324.2020.1803168]
12. Howard MO, Chung SS. Nurses' attitudes toward substance misusers. II. Experiments and studies comparing nurses to other groups. Subst Use Misuse 2000 Mar 03;35(4):503-532. [doi: 10.3109/10826080009147470] [Medline: 10741539]
13. McCormick KA, Cochran NE, Back AL, Merrill JO, Williams EC, Bradley KA. How primary care providers talk to patients about alcohol a qualitative study. J Gen Intern Med 2006 Sep;21(9):966-972. [doi: 10.1007/bf02743146]
14. Escobar GJ, LaGuardia JC, Turk BJ, Ragins A, Kipnis P, Draper D. Early detection of impending physiologic deterioration among patients who are not in intensive care: development of predictive models using data from an automated electronic medical record. J Hosp Med 2012 Mar 22;7(5):388-395. [doi: 10.1002/jhm.1929] [Medline: 22447632]
15. 15 EG, Ragins A, Scheirer P, Liu V, Robles J, Kipnis P. Nonelective rehospitalizations and postdischarge mortality: predictive models suitable for use in real time. Med Care 2015;53:916-923. [doi: 10.1097/mlr.0000000000000435]

XSL•FO
RenderX

16. Hackmann G, Chen M, Chipara O, Lu C, Chen Y, Kollef M, et al. Toward a two-tier clinical warning system for hospitalized patients. 2011 Presented at: AMIA Annu Symp Proc; 2011; Washington DC, USA p. 511-519.
17. Mishra NK, Son RY, Arnzen JJ. Towards Automatic Diabetes Case Detection and ABCS Protocol Compliance Assessment. Clinical Medicine & Research 2012 May 25;10(3):106-121. [doi: 10.3121/cmr.2012.1047]
18. Ebrahimi A, Nielsen AS, Will UK, Mansourvar M. The Prediction of Alcohol Use Disorder: A Scoping Review. 2019 Presented at: Proc IEEE Symp Comput Commun; 2019; Barcelona, Spain p. 1062-1067. [doi: 10.1109/iscc47284.2019.8969764]
19. Ebrahimi A, Wiil U, Andersen K, Mansourvar M, Nielsen A. A Predictive Machine Learning Model to Determine Alcohol Use Disorder. 2020 Presented at: Proc IEEE Symp Comput Commun 2020; 7-10 July 2020; Rennes, France p. 1-7. [doi: 10.1109/iscc50000.2020.9219685]
20. Anderson R. The collection, linking and use of data in biomedical research and healthcarethical issues. Nuffield: Nuffield Council on Bioethics; 2015:1-198.
21. Khairat S, Marc D, Crosby W, Al Sanousi A. Reasons For Physicians Not Adopting Clinical Decision Support Systems: Critical Analysis. JMIR Med Inform 2018 Apr 18;6(2):e24. [doi: 10.2196/medinform.8912]
22. Christensen AMS. The institutional framework of professional virtue. In: Carr D, editor. Cultivating moral character and virtue in professional practice. 1st ed. London: Routledge; 2018:124-134.
23. Kaba R, Sooriakumaran P. The evolution of the doctor-patient relationship. International Journal of Surgery 2007 Feb;5(1):57-65. [doi: 10.1016/j.ijsu.2006.01.005]
24. Mead N, Bower P. Patient-centredness: a conceptual framework and review of the empirical literature. Social Science & Medicine 2000 Oct;51(7):1087-1110. [doi: 10.1016/s0277-9536(00)00098-8]

Abbreviations

EHR: electronic health record
IT: information technology
SBIRT: screening, brief intervention, and referral to specialized treatment

Edited by G Eysenbach; submitted 25.01.20; peer-reviewed by Y Chu, C Jacob; comments to author 23.09.20; revised version received 03.11.20; accepted 08.05.21; published 09.08.21.

Please cite as:
Oxholm C, Christensen AMS, Christiansen R, Wiil UK, Nielsen AS
Attitudes of Patients and Health Professionals Regarding Screening Algorithms: Qualitative Study
JMIR Form Res 2021;5(8):e17971
URL: https://formative.jmir.org/2021/8/e17971
doi:10.2196/17971
PMID:34383666

©Christina Oxholm, Anne-Marie Soendergaard Christensen, Regina Christiansen, Uffe Kock Wiil, Anette Søgaard Nielsen. Originally published in JMIR Formative Research (https://formative.jmir.org), 09.08.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Estimation of Psychological Distress in Japanese Youth Through Narrative Writing: Text-Based Stylometric and Sentiment Analyses

Masae Manabe[1], MEd; Kongmeng Liew[1], PhD; Shuntaro Yada[1], PhD; Shoko Wakamiya[1], PhD; Eiji Aramaki[1], PhD

Nara Institute of Science and Technology, Nara, Japan

Corresponding Author:
Eiji Aramaki, PhD
Nara Institute of Science and Technology
8916-5, Takayama-cho, Ikoma
Nara
Japan
Phone: 81 743 72 5250
Email: aramaki@is.naist.jp

Abstract

Background: Internalizing mental illnesses associated with psychological distress are often underdetected. Text-based detection using natural language processing (NLP) methods is increasingly being used to complement conventional detection efforts. However, these approaches often rely on self-disclosure through autobiographical narratives that may not always be possible, especially in the context of the collectivistic Japanese culture.

Objective: We propose the use of narrative writing as an alternative resource for mental illness detection in youth. Accordingly, in this study, we investigated the textual characteristics of narratives written by youth with psychological distress; our research focuses on the detection of psychopathological tendencies in written imaginative narratives.

Methods: Using NLP tools such as stylometric measures and lexicon-based sentiment analysis, we examined short narratives from 52 Japanese youth (mean age 19.8 years, SD 3.1) obtained through crowdsourcing. Participants wrote a short narrative introduction to an imagined story before completing a questionnaire to quantify their tendencies toward psychological distress. Based on this score, participants were categorized into higher distress and lower distress groups. The written narratives were then analyzed using NLP tools and examined for between-group differences. Although outside the scope of this study, we also carried out a supplementary analysis of narratives written by adults using the same procedure.

Results: Youth demonstrating higher tendencies toward psychological distress used significantly more positive (happiness-related) words, revealing differences in valence of the narrative content. No other significant differences were observed between the high and low distress groups.

Conclusions: Youth with tendencies toward mental illness were found to write more positive stories that contained more happiness-related terms. These results may potentially have widespread implications on psychological distress screening on online platforms, particularly in cultures such as Japan that are not accustomed to self-disclosure. Although the mechanisms that we propose in explaining our results are speculative, we believe that this interpretation paves the way for future research in online surveillance and detection efforts.

(JMIR Form Res 2021;5(8):e29500) doi:10.2196/29500

KEYWORDS

psychological distress; youth; narratives; natural language processing; Japan; mental health; stress; distress; young adult; teenager; sentiment

Introduction

Background

Adolescents often display premonitory symptoms of mental illnesses arising from psychological distress (such as depression and anxiety-related disorders) [1-3]. Noticing the onset of such symptoms is important to facilitate effective treatment and diagnosis. Any delay between the onset of psychopathological symptoms to diagnosis and clinical treatment can result in a worsening of the condition, leading to complications not only for the individual's mental health but also with respect to their physical health and social relationships [4]. In particular, youth tend to report more symptoms of internalizing mental illnesses

XSL•FO
RenderX

that are related to psychological distress [5] and accompanying physical or vegetative changes (eg, weight and appetite changes) compared to adults. Consequently, as with any psychopathology, early access to medical and psychological assistance is important for the recovery process. Accordingly, the development of methods for the early detection of psychological distress in youth is of crucial importance.

Text-Based Screening

In recent years, the rapid development of natural language processing (NLP) technology has enabled screening for dementia or depression through examinations of language use in free-text tasks [6-8]. One proposal to increase the effectiveness of such examinations is to limit and control the topics of these free-text tasks. This would reduce the occurrence of situations where differences in language use are due to differences in writing tasks and topics, which may confound language use detection with psychopathology. For example, the written narrative in recollecting a painful experience may be very different from the narrative recounting an interesting episode heard from a friend. Regardless of the underlying mental health of both instances, it is highly likely that different language and terms would be used in the recounting of these memories.

In this study, we adopted NLP techniques for detecting psychological distress disorders in young Japanese writers, and then describe possible language indicators that may reflect these differences. Accordingly, we consider the characteristics of youth with mental illness, and propose that writing fictional, imaginative narratives may appropriately allow for the generation of content that facilitates the detection of psychological distress.

Psychopathology and Autobiographical Narratives

Prior research on narrative writing and psychopathology detection has focused on autobiographical narratives stemming from self-disclosure. For example, when a patient with depression (ie, major depressive disorder) recalls their own experience, they tend to rely on general or repeated memory (eg, I played a game with friends last week) rather than concrete memory (eg, I played volleyball with Sato last Thursday) or extended memory (eg, looking at this souvenir reminded me of my cruise trip last winter). This phenomenon is referred to as overgeneral autobiographical memory [9-11] and has been well-documented even in youth with depression [12].

However, detection of psychopathology in autobiographical self-disclosure–based narratives may not be as effective in Japan, given that the Japanese are generally less accustomed to self-disclosure [13,14]. Furthermore, self-disclosure on online platforms comes with risks such as cyberbullying and social isolation [15]. Thus, an alternative method of detection is needed. Toward this end, we propose the use of fictional, imaginative writing that does not require any form of self-disclosure. As a first step, such research would need to examine which part of the narrative affords detection of psychological distress in the writer in detail. We consider that story creation relies on one's past experiences and memory, but in a transformed, unrecognizable manner, resembling a generalized experience that can be shared with others.

Imaginative Narratives as Writing Topics

We propose that psychological distress in the writer may be observable through their written narratives. In other words, topic questions should encourage respondents to engage in creative, imaginative writing. Imagination, or fantasy, relies on stories and retrievals of one's previous experiences that are combined and reworked based on certain elements of these past experiences in a creative manner to generate new propositions and scenarios [16].

We consider that creativity in writing imaginative narratives is important for psychological distress detection. A meta-analysis on trait creativity and psychopathology showed that mild psychopathology is associated with increased creativity, to the extent that it does not impair day-to-day functioning [17]. Similarly, external ratings of creativity in visual art were found to be associated with dehydroepiandrosterone-sulfate (DHEA-S), an endocrine marker of depression. Specifically, lower levels of DHEA-S (indicating an increased risk of depression) were predictive of higher creativity ratings, and this relationship was significantly moderated by the individual's emotional vulnerability to social rejection or acceptance [18].

Accordingly, individuals' mental health states should be reflected in the creativity of their narrative content, and analysis of these narratives may function as effective early detectors of psychological distress. Furthermore, imaginative free-text writing tasks may be more suited for mental illness detection in children and adolescents. As creativity refers to the ability to imagine new and useful things or situations [19], it involves imagining something that does not presently exist in reality. Thus, we propose that by constraining free-text narrative tasks to imaginative writing, we can effectively detect adolescent writers with possible psychopathological tendencies through examining the creativity of the written narratives as quantified by the amount of variation in their vocabulary and language use.

Objectives

Based on the hypothesis outlined above, our research focuses on the detection of psychological distress in written imaginative narratives. To quantify these variations in vocabulary and language use, we applied NLP tools such as stylometric measures and lexicon-based sentiment analysis to analyze these texts. Although the act of analyzing stories itself is not a new concept, the use of NLP tools to quantify qualitative text in analyzing imaginative narratives remains a challenging approach. Nevertheless, we propose that written stories from youth at high and low risk of psychological distress can be differentiated through stylometric text analysis methods.

Methods

Participants

A total of 634 participants were initially recruited from Yahoo! Crowdsourcing, a Japanese online crowdsourcing platform. Participants were reimbursed 5 yen (~US $0.05) through Yahoo! Crowdsourcing. After excluding participants with meaningless words and reprints of copyrighted works in their responses, a total of 629 (267 males, 335 females, 27 no disclosed gender;

XSL•FO
RenderX

mean age 40.9 years, SD 12.3) participants were identified. Note that age was approximated from the mean of the class (group) data, since we only collected information on the age group. Of these participants, we further narrowed our focus to responses from young participants. Following the World Health Organization definition of "youth" as individuals aged between 15 and 25 years, we examined a final 52 participants (14 males, 29 females, 9 did not disclose gender; mean age 19.8 years, SD 3.1). Although the main focus of the analysis was on youth, we also examined supplementary data on 577 adults (253 males, 306 females, 18 did not disclose gender; mean age 42.4 years, SD 12.3) to allow for the possibility of identifying youth-specific characteristics. During the first 2 weeks, we limited recruitment to users under the age of 25 years, but had low participation rates. Therefore, for the next 2 weeks, we recruited users without any age restriction and then selected only users of the target age group. As a result, we collected data from the 52 youth and 577 adults. The adult data were not intended to be part of our research, but showed sufficient potential to provide a developmental perspective, and therefore were included for use in the supplemental analysis.

Materials

Participants were first instructed to "write an introduction to a story, with at least 200 (Japanese) characters." This was an open-ended response task, designed to let participants engage in narrative writing by creating the introduction to an imagined story. Subsequently, they were assessed for psychological distress using the Kessler Psychological Distress Scale (K-10), a brief 10-item test widely used in screening mental illness (eg, in the World Health Organization World Mental Health Survey [20]). The cutoff value for K-10 was set at 30, which is a reference value to indicate a respondent's state of severe psychological distress. Participants with a K-10 score ≥30 were assigned to the higher distress group and the remaining participants were assigned to the lower distress group.

NLP Measures

We adopted stylometric measures for authorship detection that have been shown to have relationships with the attitudes and psychological tendencies of authors. We utilized 12 types of stylometrics, as listed in Table 1, based on Japanese text metrics organized in Asaishi [21]. Additionally, we performed sentiment analyses to examine the ratio of specific emotion terms (eg, happiness, surprise, anger; see Table 1) to the total number of terms in the text.

To reduce confounds related to story length, we limited the open-ended responses to include only the first 200 characters while preserving the sentence unit.

That is, we kept the maximum number of the introductory sentences if the last sentence (a) did not exceed the 200-character limit or (b) exceeded the limit by fewer characters than the characters from the end of its previous sentence. During this process, sentences were separated by punctuation marks, except for those inside parentheses.

Because some metrics require grammatical information, we applied shallow natural language parsing techniques. Tokenization and part-of-speech tagging were processed through the morphological analyzer MeCab [22].

Note that Japanese text is not tokenized by authors with spaces. CaboCha [23] was used for syntax parsing.

XSL•FO
RenderX

Table 1. Stylometric measures (value format).

Stylometric	Description (value format)
Percentages of character types	The ratios (%) of hiragana, katakana, and kanji (Chinese characters) to the characters in the story
Type token ratio	The ratio (%) of different words to the total number of words in the story
Percentages of content words	The ratio (%) of content words (ie, nouns, verbs, adjectives, and adverbs) to the total number of words in the story
Modifying words and verb ratio	The ratio (%) of verbs to adjectives, adverbs, and conjunctions for the words in the story; this stylometric has been used as one of the indicators of author estimation [24]
Percentages of proper nouns	The ratio (%) of proper nouns (named entities) to all words in the story
Word abstraction	The abstraction degrees of the words in the story. The abstraction degrees were obtained from the Japanese word-abstraction dictionary AWD-J [25] (real number)
Ratios of emotional words	The ratios (%), relative to all words in the story, of words associated with each of the following seven categories of emotions: sadness, anxiety, anger, disgust, trust, surprise, and happiness. Weights are assigned such that each value spans between 0 and 1; the sum of all values is 1. The degree of association with emotion was determined according to the Japanese emotional-word dictionary JIWC [26]
Number of sentences	The total number of sentences that make up the story (integer)
Length of sentences	Descriptive statistics for the number of characters in each sentence that constitutes the story. In particular, the average sentence length has been suggested to be linked to the writer's creative attitude and personality [27] (real number)
Percentage of conversational sentences	Percentage of the total number of conversational sentences contained in the story
Depth of syntax tree	Descriptive statistics calculated for the depth of the dependency tree for each sentence in the story (real number)
Mean of the number of chunks per sentence	Descriptive statistics calculated for the average values of the number of chunks for each sentence in the story (real number)
Mean of the words per chunk	Descriptive statistics calculated for the average values of the number of words per chunk in the story (real number)

Results

The results of the comparison of language indicators between youth and adults are shown in Table 2. See Multimedia Appendix 1 for examples of narrative writing in youth and adults.

In contrast to our hypothesis, participants in the higher distress group did not show significant increases in word richness or diversity. Most of the stylometrics that examined variation in word use (eg, type token ratio) did not significantly differ between groups. In exploring additional differences in language content between the higher and lower distress groups for youth, significant differences were observed in emotion terms for happiness-related word ratios. This suggests that narratives written by participants in the higher distress group were more likely to use happiness-related phrases and words than those in the lower distress group. No other significant differences were observed.

Although outside the scope of this study, we also carried out a supplementary analysis of narratives written by adults using the same procedure. In the adult participants, there were significant differences between higher and lower distress groups in content words, sadness, and the mean number of words per sentence clause. The higher distress group used fewer content words, fewer emotion words about sadness, and fewer words per chunk than the lower distress group (Table 2).

XSL•FO
RenderX

Table 2. Comparison of mean values of stylometric measures.

Stylometric	Youth				Adults			
	Higher distress (n=21), mean (SD)	Lower distress (n=31), mean (SD)	t^a (*df*=50)	*P* value	Higher distress (n=94), mean (SD)	Lower distress (n=483), mean (SD)	t^a (*df*=575)	*P* value
Character types								
Hiragana	0.597 (0.079)	0.610 (0.075)	–0.591	.56	0.604 (0.087)	0.596 (0.084)	0.787	.43
Katakana	0.060 (0.058)	0.047 (0.036)	0.990	.33	0.0463 (0.053)	0.597 (0.084)	0.491	.62
Kanji (Chinese characters)	0.255 (0.053)	0.250 (0.062)	0.289	.77	0.268 (0.064)	0.276 (0.069)	–1.045	.30
Type token ratio	0.540 (0.075)	0.528 (0.058)	0.628	.53	0.550 (0.056)	0.542 (0.058)	1.168	.24
Content words	0.272 (0.068)	0.265 (0.053)	0.383	.70	2.490 (0.214)	2.548 (0.219)	–2.175	.03
Modifying words and verb ratio	0.394 (0.208)	0.424 (0.185)	–0.541	.59	0.406 (0.230)	0.393 (0.219)	0.504	.61
Proper nouns	0.010 (0.025)	0.009 (0.021)	0.152	.88	0.007 (0.018)	0.007 (0.013)	-0.056	.96
Word abstraction								
Maximum	3.133 (0.184)	3.071 (0.145)	1.364	.18	3.057 (0.181)	3.060 (0.183)	–0.153	.88
Average of the top 5 words	2.909 (0.105)	2.899 (0.091)	0.360	.72	2.878 (0.104)	2.886 (0.105)	0.509	.66
Emotional words								
Sadness	0.103 (0.010)	0.107 (0.012)	–1.223	.23	0.104 (0.016)	0.108 (0.020)	–2.051	.04
Anxiety	0.105 (0.017)	0.109 (0.023)	–0.690	.50	0.105 (0.024)	0.108 (0.020)	–1.402	.16
Anger	0.172 (0.021)	0.177 (0.018)	–0.956	.34	0.174 (0.024)	0.171 (0.021)	1.372	.17
Disgust	0.167 (0.036)	0.174 (0.039)	–0.731	.47	0.172 (0.041)	0.167 (0.034)	1.399	.16
Trust	0.165 (0.014)	0.160 (0.017)	1.071	.29	0.165 (0.026)	0.162 (0.022)	1.133	.26
Surprise	0.154 (0.019)	0.163 (0.015)	–1.967	.06	0.164 (0.028)	0.161 (0.022)	0.819	.41
Happiness	0.134 (0.049)	0.109 (0.018)	–2.657	.01	0.116 (0.041)	0.123 (0.039)	–1.384	.17
Number of sentences	6.619 (2.578)	7.419 (2.566)	–1.101	.28	7.011 (2.975)	6.778 (2.454)	0.809	.42
Length of sentences	33.740 (10.807)	29.544 (11.924)	1.292	.20	30.761 (10.898)	33.962 (22.711)	–1.336	.18
Conversational sentences	0.045 (0.082)	0.062 (0.107)	–0.606	.55	0.034 (0.121)	0.042 (0.114)	–0.298	.77
The number of chunks per sentence	8.194 (2.872)	8.167 (3.704)	0.028	.98	7.756 (3.400)	8.033 (4.302)	–0.546	.59
The words per chunk	2.586 (0.305)	2.561 (0.270)	0.305	.76	2.490 (0.214)	2.548 (0.219)	–2.175	.03

[a]Since equal variances were not assumed, Welch *t*-tests were used to examine between-group differences on the above measures.

Discussion

Narrative Language Features and Youth Mental Health

We found significant differences between the lower and higher distress groups, particularly with regard to the relative frequency of happiness-related terms. Happiness is commonly viewed as a positively valenced emotion across cultures [28]. At a glance, the increased usage of happiness-related words in the higher distress group suggests that youth with increased tendencies to psychological distress or illness may prefer writing more positive narratives. Considering that the concept of happiness is not typically associated with psychological distress, our result suggesting the stronger presence of happiness-related terms in the higher psychological distress group appears contradictory. Nevertheless, we posit two potential explanations for this result. First, these findings may reflect the possibility that youth with higher distress prefer happier stories. Alternatively, these findings may suggest that more frequent priming of happiness may induce psychological distress in youth. Although a causal relationship cannot be investigated from this single cross-sectional study, we speculate on several interpretations. Prior research on self-directed narratives has revealed an association with depression. For example, those who tended to imagine a positively inclined future for themselves had lower depression measures (eg, Center for Epidemiological Studies Depression Scale, Children's Depression Inventory, Beck Depression Inventory) at the time of their imagination, but in subsequent follow-up, they were found to have higher tendencies for depression than the group that had anticipated a less positive future [29].

XSL•FO
RenderX

By disambiguating these two concepts, one possibility could be that participants did not need to consider their "self" (current circumstances) in creative writing, and were able to imagine freely without these constraints. Thus, the increased positive affect (happiness scores) in their writing may suggest that this was more idealized or desired. Such an interpretation would be consistent with past research linking an increased desire for happiness with depression [30].

However, we reiterate that these explanations are speculations based on the pattern of results obtained from our data. More research is needed to confirm the applicability of these explanations in creative writing for individuals with tendencies toward psychological distress.

Type Token Ratio and Word Abstraction

Our results showed no significant difference in vocabulary usage, measured in this study through indicators such as the type token ratio, nor in the abstraction of the words for the narrative passages. One explanation could be that the memories of one's own experiences do not necessarily affect the generation of a story. Differentiation according to the level of word abstraction alone is insufficient, because narratives may be conceptualized separately from one's own experiences, referencing the collective knowledge from cultural media and the experiences of others.

Pros and Cons of Crowdsourcing

Several previous studies have shown that crowdsourcing is an appropriate tool for recruiting research participants [31,32]. Nevertheless, we acknowledge some problems with crowdsourcing data collection. For example, participants may attempt the same questionnaire twice, and incorrect comprehension of the instructional text may be problematic due to participation from nonnative speakers. Another problem we had to face was the phenomenon of "satisficing," in that participants may tend to conserve cognitive resources in survey research [33]. In crowdsourcing, people are often motivated to work on multiple tasks in as short a time as possible to increase their monetary reward because of low unit costs.

In this study, we took steps to reduce these problematic effects from crowdsourcing by using a service in which native speakers (in this case, Japanese speakers) form the vast majority of users. We also excluded duplicated IDs and data (narratives) that were directly lifted from copyrighted works or lists of meaningless words. Although we cannot definitively rule out any bias or duplicate participants in our research, precautions were taken to safeguard against these potential problems.

Limitations and Future Directions

We note several limitations of our study. First, we had limitations with regard to the age of our sample. The crowdsourcing platform had a minimum age of 15, meaning that adolescents under 15 years old were not included in this study. Second, we used the K-10 for screening purposes, which is a general questionnaire and lacks sensitivity to specific diagnoses. Thus, there is a possibility that different types of psychopathological tendencies may exhibit differentiated effects or underlying mechanisms. Future studies should consider the use of actual clinical diagnoses as the criteria for mental illness beyond crowdsourced convenience sampling.

Finally, the linguistic indicators used in this study are widely used for author estimation and are usually applied to large volumes of text. In this study, narrative sentences of about 200 characters each were used, which may be insufficient for detecting the presence of true effects. Consequently, we are unable to rule out the use of creativity measures in imaginative writing as a means for detecting psychological distress, as our lack of a significant finding in this aspect can also be explained by an insufficiency in the length of our narrative data.

Nevertheless, our study lays the groundwork for psychological distress surveillance programs in sensitive populations, especially in situations where recollection of personal, self-related narratives or self-disclosure may be problematic or risky. This may be more relevant in the Japanese context. Japanese undergraduate students were less likely to self-disclose experiences of bullying compared to a US sample, and this was fueled by a concern for disrupting social and relational harmony [34]. As such, surveillance methods that rely on fictional narrative content without reliance on self-disclosure may be more suitable in a collectivistic Japanese context. Our exploratory result identifies happiness and surprise-related content as potential indicators of psychological distress in the writer. Future research should confirm these findings on a larger scale, with a larger, preregistered study involving more diverse samples and cross-cultural comparisons. Narrative writing may even provide therapeutic benefits for individuals suffering from depression [35], and future studies can also quantify the effectiveness of such methods in randomized controlled trials.

Conclusions

Youth with tendencies toward mental illness were found to write more positive stories that contained more happiness-related terms. Although the mechanisms underlying these differences in frequency of happiness-related term usage are speculative at present, these results may potentially have more widespread implications on screening, particularly in cultures such as Japan that are not accustomed to self-disclosure. This is a preliminary finding and more confirmatory research is needed to establish the robustness of these results.

Acknowledgments

This research was supported by the Japan Society for the Promotion of Science KAKENHI Grant (16H06399).

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Examples of narrative writing. Shows the highest/lowest type token ratio and ratio of happiness-related words for youth and adults.
[XLSX File (Microsoft Excel File), 17 KB - formative_v5i8e29500_app1.xlsx]

References

1. Kessler RC, Amminger GP, Aguilar-Gaxiola S, Alonso J, Lee S, Ustün TB. Age of onset of mental disorders: a review of recent literature. Curr Opin Psychiatry 2007 Jul;20(4):359-364 [FREE Full text] [doi: 10.1097/YCO.0b013e32816ebc8c] [Medline: 17551351]
2. Stallman HM. Psychological distress in university students: A comparison with general population data. Austral Psychol 2010 Nov 06;45(4):249-257. [doi: 10.1080/00050067.2010.482109]
3. Moylan S, Maes M, Wray NR, Berk M. The neuroprogressive nature of major depressive disorder: pathways to disease evolution and resistance, and therapeutic implications. Mol Psychiatry 2013 May;18(5):595-606. [doi: 10.1038/mp.2012.33] [Medline: 22525486]
4. Rice F, Riglin L, Lomax T, Souter E, Potter R, Smith D, et al. Adolescent and adult differences in major depression symptom profiles. J Affect Disord 2019 Jan 15;243:175-181. [doi: 10.1016/j.jad.2018.09.015] [Medline: 30243197]
5. Wight RG, Sepúlveda JE, Aneshensel CS. Depressive symptoms: how do adolescents compare with adults? J Adolesc Health 2004 Apr;34(4):314-323. [doi: 10.1016/j.jadohealth.2003.05.003] [Medline: 15041001]
6. Jin H, Wu S. Text messaging as a screening tool for depression and related conditions in underserved, predominantly minority safety net primary care patients: validity study. J Med Internet Res 2020 Mar 26;22(3):e17282 [FREE Full text] [doi: 10.2196/17282] [Medline: 32213473]
7. Topaz M, Adams V, Wilson P, Woo K, Ryvicker M. Free-text documentation of dementia symptoms in home healthcare: a natural language processing study. Gerontol Geriatr Med 2020;6:2333721420959861 [FREE Full text] [doi: 10.1177/2333721420959861] [Medline: 33029550]
8. Morales M, Scherer S, Levitan R. Cross-modal review of indicators for depression detection systems. 2017 Presented at: Fourth Workshop on Computational Linguistics and Clinical Psychology - From Linguistic Signal to Clinical Reality; August 3, 2017; Vancouver, BC. [doi: 10.18653/v1/w17-3101]
9. Liu Y, Yu X, Yang B, Zhang F, Zou W, Na A, et al. Rumination mediates the relationship between overgeneral autobiographical memory and depression in patients with major depressive disorder. BMC Psychiatry 2017 Mar 21;17(1):103 [FREE Full text] [doi: 10.1186/s12888-017-1264-8] [Medline: 28327097]
10. Matsumoto N, Mochizuki S. Reduced autobiographical memory specificity due to depression: Assess the instructions for analog study. Japanese J Res Emotions 2013;21(1):11-18. [doi: 10.4092/jsre.21.11]
11. Williams JMG, Teasdale JD, Segal ZV, Soulsby J. Mindfulness-based cognitive therapy reduces overgeneral autobiographical memory in formerly depressed patients. J Abnorm Psychol 2000 Mar;109(1):150-155. [doi: 10.1037//0021-843x.109.1.150] [Medline: 10740947]
12. Kuyken W, Howell R, Dalgleish T. Overgeneral autobiographical memory in depressed adolescents with, versus without, a reported history of trauma. J Abnorm Psychol 2006 Aug;115(3):387-396. [doi: 10.1037/0021-843X.115.3.387] [Medline: 16866580]
13. Asai A, Barnlund D. Boundaries of the unconscious, private, and public self in Japanese and Americans: a cross-cultural comparison. Int J Intercult Relation 1998 Nov;22(4):431-452. [doi: 10.1016/s0147-1767(98)00017-0]
14. Schug J, Yuki M, Maddux W. Relational mobility explains between- and within-culture differences in self-disclosure to close friends. Psychol Sci 2010 Oct;21(10):1471-1478. [doi: 10.1177/0956797610382786] [Medline: 20817913]
15. Best P, Manktelow R, Taylor B. Online communication, social media and adolescent wellbeing: A systematic narrative review. Child Youth Serv Rev 2014 Jun;41:27-36. [doi: 10.1016/j.childyouth.2014.03.001]
16. Vygotsky LS. Imagination and creativity in childhood. J Russian East Eur Psychol 2014 Dec 08;42(1):7-97. [doi: 10.1080/10610405.2004.11059210]
17. Paek SH, Abdulla AM, Cramond B. A meta-analysis of the relationship between three common psychopathologies—ADHD, anxiety, and depression—and indicators of Little-c Creativity. Gift Child Quart 2016 Feb 04;60(2):117-133. [doi: 10.1177/0016986216630600]
18. Akinola M, Mendes WB. The dark side of creativity: biological vulnerability and negative emotions lead to greater artistic creativity. Pers Soc Psychol Bull 2008 Dec;34(12):1677-1686 [FREE Full text] [doi: 10.1177/0146167208323933] [Medline: 18832338]
19. Stein MI. Creativity and culture. J Psychol 1953 Oct;36(2):311-322. [doi: 10.1080/00223980.1953.9712897]

XSL•FO
RenderX

20. Kessler RC, Andrews G, Colpe LJ, Hiripi E, Mroczek DK, Normand SLT, et al. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. Psychol Med 2002 Aug;32(6):959-976. [doi: 10.1017/s0033291702006074] [Medline: 12214795]
21. Takuma A. An overview of indicators for quantifying text features [in Japanese]. J Japan Soc Library Inf Sci 2017;63(3):159-169. [doi: 10.20651/jslis.63.3_159]
22. MeCab: Yet Another Part-of-Speech and Morphological Analyzer. URL: https://taku910.github.io/mecab/ [accessed 2021-02-17]
23. CaboCha: Yet Another Japanese Dependency Structure Analyze. URL: https://taku910.github.io/cabocha/ [accessed 2021-02-17]
24. Watanabe M. Statistical analysis of the difference between original novels written by Haruki Murakami and foreign novels translated. Math Ling 2003;24(3):143-153.
25. AWD-J: Abstractness of Word Database for Japanese common words. URL: http://sociocom.jp/~data/2019-AWD-J/ [accessed 2021-02-03]
26. JIWC-Dictionary. GitHub. URL: https://github.com/sociocom/JIWC-Dictionary [accessed 2021-02-09]
27. Takeshige S, Inoue N. The writing style of Yasushi Inoue in his short novels. Dokusho Kagaku (The Science of Reading) 1986;30(2):51-62.
28. An S, Ji L, Marks M, Zhang Z. Corrigendum: Two sides of emotion: exploring positivity and negativity in six basic emotions across cultures. Front Psychol 2017;8:1467. [doi: 10.3389/fpsyg.2017.01467] [Medline: 28900408]
29. Oettingen G, Mayer D, Portnow S. Pleasure now, pain later: positive fantasies about the future predict symptoms of depression. Psychol Sci 2016 Mar;27(3):345-353. [doi: 10.1177/0956797615620783] [Medline: 26825106]
30. Mahmoodi Kahriz B, Bower J, Glover F, Vogt J. Wanting to be happy but not knowing how: poor attentional control and emotion-regulation abilities mediate the association between valuing happiness and depression. J Happiness Stud 2019 Oct 26;21(7):2583-2601. [doi: 10.1007/s10902-019-00193-9]
31. Clifford S, Jewell RM, Waggoner PD. Are samples drawn from Mechanical Turk valid for research on political ideology? Res Politics 2015 Dec 15;2(4):205316801562207. [doi: 10.1177/2053168015622072]
32. Majima Y. The feasibility of a Japanese crowdsourcing service for experimental research in psychology. SAGE Open 2017 Mar 23;7(1):215824401769873. [doi: 10.1177/2158244017698731]
33. Holland JL, Christian LM. The influence of topic interest and interactive probing on responses to open-ended questions in web surveys. Soc Sci Comput Rev 2008 Dec 29;27(2):196-212. [doi: 10.1177/0894439308327481]
34. Matsunaga M. Individual dispositions and interpersonal concerns underlying bullied victims' self-disclosure in Japan and the US. J Soc Person Relation 2010 Nov 09;27(8):1124-1148. [doi: 10.1177/0265407510380084]
35. Cooper P. Writing for depression in health care. Br J Occup Ther 2013 Apr 15;76(4):186-193. [doi: 10.4276/030802213x13651610908452]

Abbreviations

DHEA-S: dehydroepiandrosterone-sulfate
K-10: Kessler Psychological Distress Scale
NLP: natural language processing

Edited by G Eysenbach; submitted 13.04.21; peer-reviewed by S Matsuda; comments to author 04.05.21; revised version received 29.06.21; accepted 06.07.21; published 12.08.21.

Please cite as:
Manabe M, Liew K, Yada S, Wakamiya S, Aramaki E
Estimation of Psychological Distress in Japanese Youth Through Narrative Writing: Text-Based Stylometric and Sentiment Analyses
JMIR Form Res 2021;5(8):e29500
URL: https://formative.jmir.org/2021/8/e29500
doi:10.2196/29500
PMID:34387556

©Masae Manabe, Kongmeng Liew, Shuntaro Yada, Shoko Wakamiya, Eiji Aramaki. Originally published in JMIR Formative Research (https://formative.jmir.org), 12.08.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Screening Diabetic Retinopathy Using an Automated Retinal Image Analysis System in Independent and Assistive Use Cases in Mexico: Randomized Controlled Trial

Alejandro Noriega[1,2], PhD; Daniela Meizner[3], MD; Dalia Camacho[2,4], MSc; Jennifer Enciso[2,5], PhD; Hugo Quiroz-Mercado[3], MD; Virgilio Morales-Canton[3], MD; Abdullah Almaatouq[6], PhD; Alex Pentland[1], PhD

[1]MIT Media Laboratory, Massachusetts Institute of Technology, Cambridge, MA, United States
[2]Prosperia Salud, Mexico City, Mexico
[3]Retina Department, Asociación para Evitar la Ceguera en México, Mexico City, Mexico
[4]Engineering Academic Division, Instituto Tecnológico Autónomo de México, Mexico City, Mexico
[5]Posgrado de Ciencias Bioquímicas, Universidad Nacional Autónoma de México, Mexico City, Mexico
[6]Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA, United States

Corresponding Author:
Alejandro Noriega, PhD
Prosperia Salud
58D Secretaria de Marina 1206, Lomas del Chamizal
Mexico City, 05129
Mexico
Phone: 52 617 982 47
Email: noriega@mit.edu

Abstract

Background: The automated screening of patients at risk of developing diabetic retinopathy represents an opportunity to improve their midterm outcome and lower the public expenditure associated with direct and indirect costs of common sight-threatening complications of diabetes.

Objective: This study aimed to develop and evaluate the performance of an automated deep learning–based system to classify retinal fundus images as referable and nonreferable diabetic retinopathy cases, from international and Mexican patients. In particular, we aimed to evaluate the performance of the automated retina image analysis (ARIA) system under an independent scheme (ie, only ARIA screening) and 2 assistive schemes (ie, hybrid ARIA plus ophthalmologist screening), using a web-based platform for remote image analysis to determine and compare the sensibility and specificity of the 3 schemes.

Methods: A randomized controlled experiment was performed where 17 ophthalmologists were asked to classify a series of retinal fundus images under 3 different conditions. The conditions were to (1) screen the fundus image by themselves (solo); (2) screen the fundus image after exposure to the retina image classification of the ARIA system (ARIA answer); and (3) screen the fundus image after exposure to the classification of the ARIA system, as well as its level of confidence and an attention map highlighting the most important areas of interest in the image according to the ARIA system (ARIA explanation). The ophthalmologists' classification in each condition and the result from the ARIA system were compared against a gold standard generated by consulting and aggregating the opinion of 3 retina specialists for each fundus image.

Results: The ARIA system was able to classify referable vs nonreferable cases with an area under the receiver operating characteristic curve of 98%, a sensitivity of 95.1%, and a specificity of 91.5% for international patient cases. There was an area under the receiver operating characteristic curve of 98.3%, a sensitivity of 95.2%, and a specificity of 90% for Mexican patient cases. The ARIA system performance was more successful than the average performance of the 17 ophthalmologists enrolled in the study. Additionally, the results suggest that the ARIA system can be useful as an assistive tool, as sensitivity was significantly higher in the experimental condition where ophthalmologists were exposed to the ARIA system's answer prior to their own classification (93.3%), compared with the sensitivity of the condition where participants assessed the images independently (87.3%; P=.05).

XSL•FO
RenderX

Conclusions: These results demonstrate that both independent and assistive use cases of the ARIA system present, for Latin American countries such as Mexico, a substantial opportunity toward expanding the monitoring capacity for the early detection of diabetes-related blindness.

(JMIR Form Res 2021;5(8):e25290) doi:10.2196/25290

KEYWORDS

diabetic retinopathy; automated diagnosis; retina; fundus image analysis

Introduction

Impact of Diabetes

Diabetes is one of the most challenging health problems in the world, affecting more than 400 million people. Particularly, diabetes threatens the health care systems of low- and middle-income countries, where 80% of the world's diabetic population live [1,2] . Diabetes is a multifactorial and complex disease with a strong genetic component. In this regard, it has been demonstrated that Hispanic/Latino people have a greater susceptibility to develop type II diabetes, as well as diabetes-associated complications, including renal insufficiency and visual impairment [1-4].

In 2015, there were more than 41 million adults diagnosed with diabetes in Latin America and Caribbean countries, making it one of the major causes of premature death and disability in the region [5,6] . Particularly, Mexico ranked sixth among the world's diabetes prevalence in 2015 and second among Latin America, only after Brazil [7,8]. It is estimated that 26 million adults live in Mexico with diabetes or prediabetes, and only half of them have been diagnosed. Diabetes and its related complications are the first cause of disability and the third cause of death in the country, largely impacting productivity, life quality, and the economy [5].

Evolution and Treatment of Diabetic Retinopathy

Diabetic retinopathy (DR) is the most common complication in advanced or uncontrolled diabetic patients and is the leading cause of irreversible vision loss in working-age adults [9,10]. DR is a microvascular complication that emerges in diabetic patients as a consequence of chronic hyperglycemia that contributes to blood vessel damage in the retina, causing a combination of fluid leakage, swelling of the surrounding tissue, blood flow obstruction, and abnormal neovascularization [9,10].

DR progression is slow, gradual, and reversible in its first stage. However, if not treated promptly, it can lead to irreversible blindness . According to the International Clinical Diabetic Retinopathy Severity Scale, the first stage of DR is classified as mild nonproliferative diabetic retinopathy (NPDR), which is characterized by the presence of at least 1 microaneurysm and is highly reversible through blood pressure, cholesterol, and sugar level control. Only very rare cases that present macular edema (swelling of fluid and protein deposits on or under the macula) might require laser photocoagulation or intravitreal injections. Without adequate diabetic control, the disease advances to moderate and severe NPDR stages, which include the presence of hemorrhages, microaneurysms, hard exudates, venous beading, or intraretinal microvascular abnormalities. At these stages, metabolic control is not sufficient to stop the disease progression, and the patient will require invasive treatments such as photocoagulation and intravitreal antivascular endothelial growth factor agents or corticosteroids. The most advanced stage is proliferative DR and is characterized by neovascularization, preretinal hemorrhages, hemorrhages in the vitreous, traction retinal detachments, or macular edema. Proliferative DR is treated with the more aggressive laser therapy called scatter or pan-retinal photocoagulation; intravitreal injection; and, in some cases, vitreoretinal surgery, which removes scar tissue or blood from the vitreous cavity to repair retinal detachments or treat macular holes [10-13].

To increase early detection and prevent the progression of DR to advanced stages, diabetic patients are recommended to have annual or semiannual retinal screenings beginning at the moment when they are diagnosed with diabetes. However, according to data from the Diabetic Retinopathy Barometer, 27% of people living with diabetes declared that they never discussed eye complications with their doctors before the onset of complications, and only 13% of the diabetic population have visited an ophthalmologist after their diagnosis [4,14]. Through frequent, preventive screenings, 70% of the cases can be captured at the initial stages of the disease and treated with noninvasive strategies such as metabolic control or photocoagulation [15]. Unfortunately, in most developing countries, there is no ophthalmological attention at primary care clinics, and it is only when diabetic patients develop vision attenuation that they are referred to second- and third-level hospitals to be screened, diagnosed, and treated [16]. At this point, significant retinal damage has occurred, and, even with invasive vitreoretinal surgery or photocoagulation, vision cannot be restored.

The limited access to ophthalmologists and retina specialists at primary care clinics, due to financial and staff limitations at national health care institutions, precludes the continuous monitoring of diabetic patients in low- and middle-income countries such as Mexico.

Challenges of Diabetic Retinopathy Screening on a Large Scale

In Mexico, DR is a leading cause of irreversible blindness among the working-age population [4,13] . Approximately 30% of the patients diagnosed with diabetes develop DR, and, based on the predictions of diabetes increasing in prevalence, by 2045, there will be 245 million people with DR lesions and 77 million people with vision-threatening DR [17].

One of the main limitations for the establishment of a systematic eye-screening program is the limited availability of ophthalmologists and their unequal distribution around the country. Based on the 2013 registry of society-affiliated

XSL•FO
RenderX

ophthalmologists from the Mexican Society of Ophthalmology, the average number of ophthalmologists per 100,000 people is lower (2.68 per 100,000) than the average among Latin American countries (5.27 per 100,000). There is a particularly worrying distribution in rural areas, with 2 ophthalmologists per 100,000 people [18].

In particular, in low- and middle-income countries such as Costa Rica, Peru, and India, there have been several efforts to implement DR screening programs targeting the limitation of ophthalmologists with mobile screening units integrated with telemedicine [19-21]. In these contexts, 2 key factors were identified for achieving cost-effectiveness of these strategies: (1) accurate identification of the risk population and (2) optimization of the number of people screened per unit of time [21]. Notably, these 2 factors can be improved by leveraging automated retinal image analysis (ARIA) systems such as the one in this study.

ARIA for Diabetic Retinopathy Screening

In recent years, the combination of the development of advanced statistical methods, the greater availability of data, and the substantial increase in computing power has allowed for the application of advanced computational methodologies, including artificial intelligence (AI), in diverse social and medical domains. Among the use of AI for social welfare, AI applications in health care domains are one of the fastest growing sectors, with a compound annual growth rate above 40% during the period between 2014 and 2021 [22]. AI tools have been successfully applied to diagnostics, therapeutics, population health management, administration, and regulation, showing a capacity to augment societies' access to health care and improve the coverage and quality of the services provided.

Ultimately, AI applications in health care present opportunities to improve overall quality of life, patients' prognoses, and optimization of human and financial resources [23]. In particular, ARIA systems have emerged as a promising solution to increase early detection of DR at primary care clinics, particularly, in resource-constrained developing countries, thereby improving health outcomes, avoiding incapacitating complications, and reducing treatment costs.

ARIA systems analyze retinal fundus images by applying techniques such as deep learning (DL) to classify diabetic patients in (1) cases without retinal lesions associated to DR (nonreferable output) and (2) cases that need to undergo examination by an ophthalmologist to confirm diagnosis and define treatment (referable output) [24-28]. As of today, various analysis systems have been developed and implemented on the market in European countries, Canada, and the United States. However, very few have been tested in Latin America and Caribbean countries to evaluate their performance and usability in the particular resource-constrained settings of these countries [29]. To determine qualities of successful implementation in these countries, research must investigate patients' ethnicities, the training of health care personnel, community openness to new technologies, and hospital resources.

Aims and Key Findings of the Study

This study aimed to evaluate the performance of a DL-based ARIA system that classifies retinal fundus images in nonreferable or referable circumstances, based on the presence of DR damage, as well as the potential benefits of its use as an assistive tool for ophthalmic doctors. We also completed a randomized controlled trial where the performance of the ARIA system was compared with the accuracy of 17 ophthalmologists from one of the most reputable ophthalmic hospitals in Mexico, Hospital de la Ceguera, which is part of the "Association to Avoid Blindness in Mexico" (APEC). In particular, the performances of ophthalmologists in 3 experimental conditions were assessed: 1 independent condition, in which the ophthalmologists assessed the images independently from the ARIA system, and 2 assistive conditions, in which either ophthalmologists observed and were influenced by the ARIA system's classification and confidence or an ARIA system–generated, attention heatmap highlighted probable DR lesions in the retina.

The key findings were that the ARIA system developed using a DL strategy was able to classify referable vs nonreferable cases with an area under the receiver operating characteristic curve (AUROC) of 98%, a sensitivity of 95.1%, and a specificity of 91.5% for international patient cases. There was an AUROC of 98.3%, a sensitivity of 95.2%, and a specificity of 90% for Mexican patient cases. For Mexican patient cases, the ARIA system performance was more successful than the average performance of the 17 ophthalmologist participants in the study. Moreover, we found that the ARIA system can be useful as an assistive tool, as we found significant improvement in the specificity in the experimental condition where participants were able to consider the answer of the ARIA system as a second opinion (87.3%), compared with the specificity of the condition where participants assessed the images independently (93.3%; P=.05).

Hence, this study aimed to demonstrate the high potential value of the use of ARIA systems, in both independent and assistive schemes, toward the goal of effective mass screening for the early detection of DR in developing countries such as Mexico.

Methods

ARIA System

ARIA System Design

The ARIA system consists of an image preprocessing module and an image analysis module that returns a binary referable and nonreferable DR classification; the level of confidence of that classification; and an attention map that shows, pixel-wise, the indicative features for referable DR according to the model (Figure 1). The models constituting the ARIA system were implemented using the Keras library with the Tensorflow backend [30] in Python 3.5 [31].

Images from all datasets were annotated by ophthalmic specialists for 5-class identification according to the International Clinical Diabetic Retinopathy Severity Scales (ICDRSS) and subsequently labeled as nonreferable or referable

XSL•FO
RenderX

DR [32]. Table 1 describes the classification, and Figure 1A provides a graphical example. The gold standard classification used for the experimental phase of the study was provided by 3 retina specialists, as described in the following subsections.

Figure 1. Deep learning–based automated retinal image analysis system. (A) Example of classified retinal fundus images according to the International Clinical Diabetic Retinopathy Severity Scale used for the training data. (B) Flow chart describing the design of the automated retinal image analysis system; the data used for training, validation, and testing; and the algorithm's outputs. DR: diabetic retinopathy; NPDR: nonproliferative diabetic retinopathy; PDR: proliferative diabetic retinopathy.

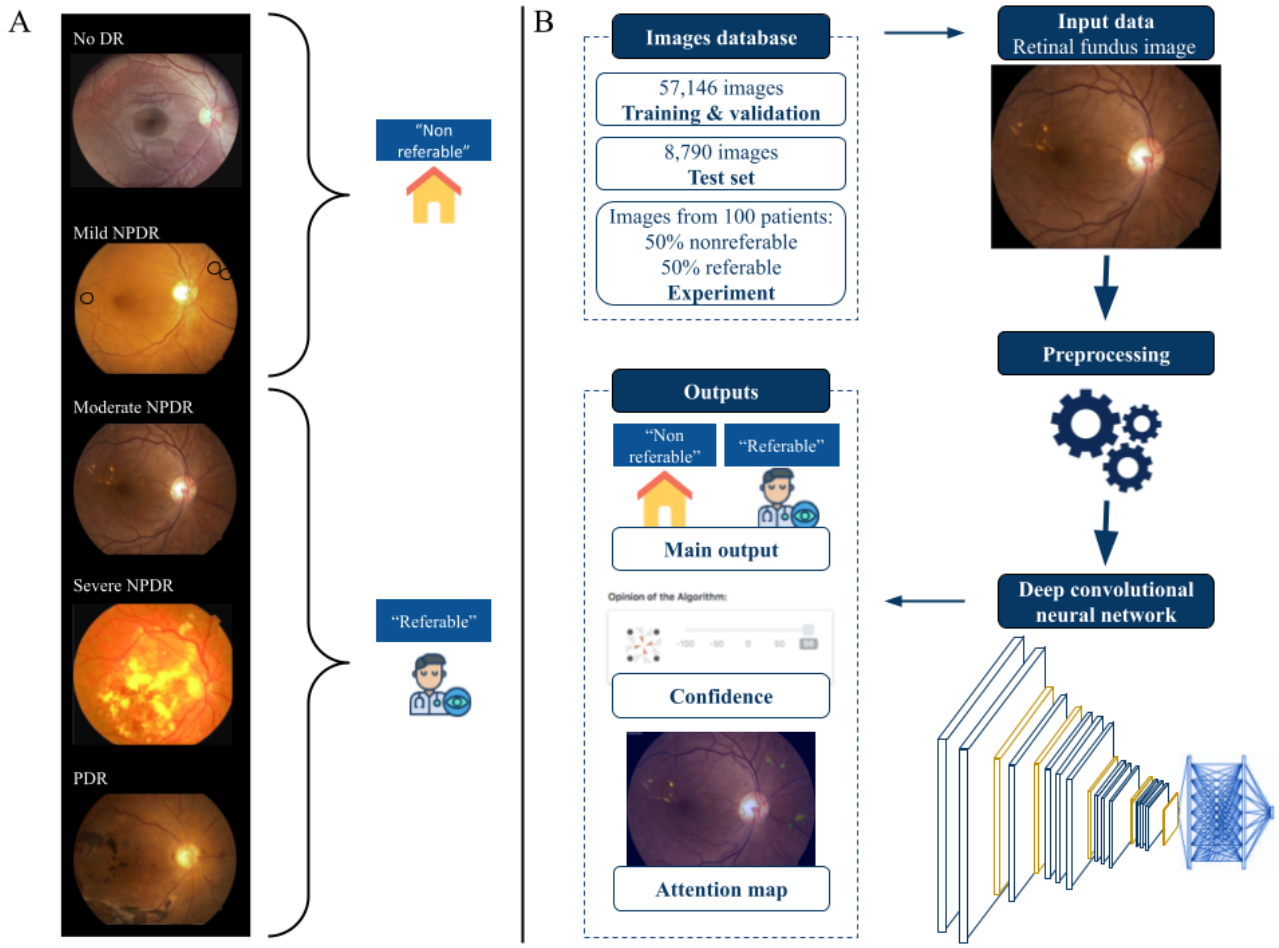


Table 1. International Clinical Diabetic Retinopathy Severity Scale and its classification for the automated retinal image analysis system [32].

ARIA[a] system classification	DR[b] severity scale	Ophthalmoscopy findings
Nonreferable	No apparent retinopathy (no DR)	No abnormalities
	Mild nonproliferative DR (mild DR)	Microaneurysms only
Referable	Moderate nonproliferative DR (moderate DR)	More than just microaneurysms but less than severe nonproliferative diabetic retinopathy
	Severe nonproliferative DR (severe DR)	≥20 intraretinal hemorrhages in each of 4 quadrants, definite venous beading in 2 quadrants, or prominent intraretinal microvascular abnormalities in 1 quadrant. No signs of proliferative retinopathy.
	Proliferative DR	Neovascularization or vitreous/preretinal hemorrhage.

[a]ARIA: automated retinal image analysis.

[b]DR: diabetic retinopathy.

Preprocessing

Before classifying the images and training the algorithms, a preprocessing procedure was applied. The procedure consisted of cropping the background to eliminate noninformative areas, padding the image to guarantee consistent squared image ratios, resizing the image to 224×224 pixels, and normalizing pixel values to the range 0-1.

Image Classification Model

The model used for image classification consisted of a deep convolutional neural network [33,34]. The network architecture

XSL•FO
RenderX

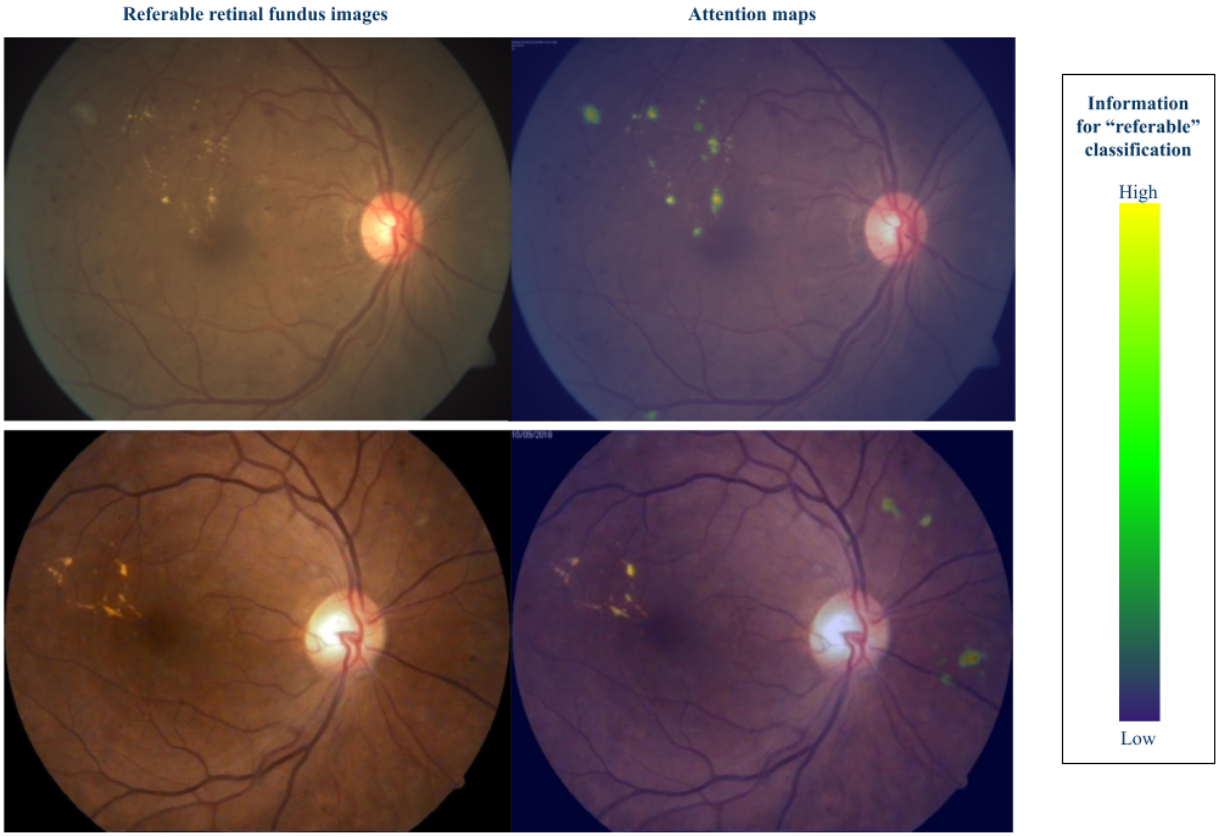
developed for this project consisted of 16 convolutional layers, a dense layer of 1024 neurons, 2 dropout layers to avoid overfitting, and a binary classification layer of a single unit with sigmoid activation. This architecture took the VGG model published by Simonyan and Zisserman [34] as a starting point. Hence, the model output is a value between 0 and 1, which may be interpreted as the confidence of the model regarding a referable DR classification. Lastly, a threshold of 0.5 was used to classify nonreferable (<0.5) and referable (≥ 0.5) DR.

The model was trained on an international dataset, of which most images were taken in primary care clinics in California, United States [35]. The training subset had 57,146 images (16,458/57,146, 28.80% with referable DR; 45,602/57,146, 79.80% gradable), and the evaluation subset had 8,790 images (694/8790, 7.90% with referable DR; 7067/8790, 80.40% gradable). The training and test subsets followed the same distribution used by Voets and colleagues [36]. Considering real-life scenarios, the training and validation datasets included images from different types of cameras and of different qualities (ie, with artifacts, out of focus, underexposed, or overexposed).

Attention Heatmaps

Attention heatmaps were developed to show lesion areas in the image by highlighting each pixel according to their importance to a referable DR classification, according to the model. These heatmaps were obtained by applying one of the most effective methods for building saliency maps on images, the layer-wise relevance propagation method, with an alpha-beta rule [37,38]. In essence, the layer-wise relevance propagation method redistributed the output value throughout the layers until the input layer (input image) was reached. Figure 2 shows examples of fundus images and the heatmaps generated using the methodology described.

Figure 2. Attention heatmaps for 2 referable images. Green and yellow colors indicate regions in the image that provide information to the algorithm to classify the image as referable.



Study Populations

We had 17 ophthalmologists from the Mexican ophthalmic hospital participating in the experimental study, and 3 retina specialists from the same institution participated in the generation of the gold standard. The 17 ophthalmologists evaluated 45° macula-centered fundus images from 100 Mexican patients, where 50% (50/100) had nonreferable DR and 50% (50/100) had referable DR levels. Each ophthalmologist evaluated 45 retinal images, in order for each image to be evaluated more than once. The ophthalmologists were retina specialization resident students, where 3 residents were in their second year, 12 were in their third year, and 2 were in their fourth year of residency.

Experimental Design

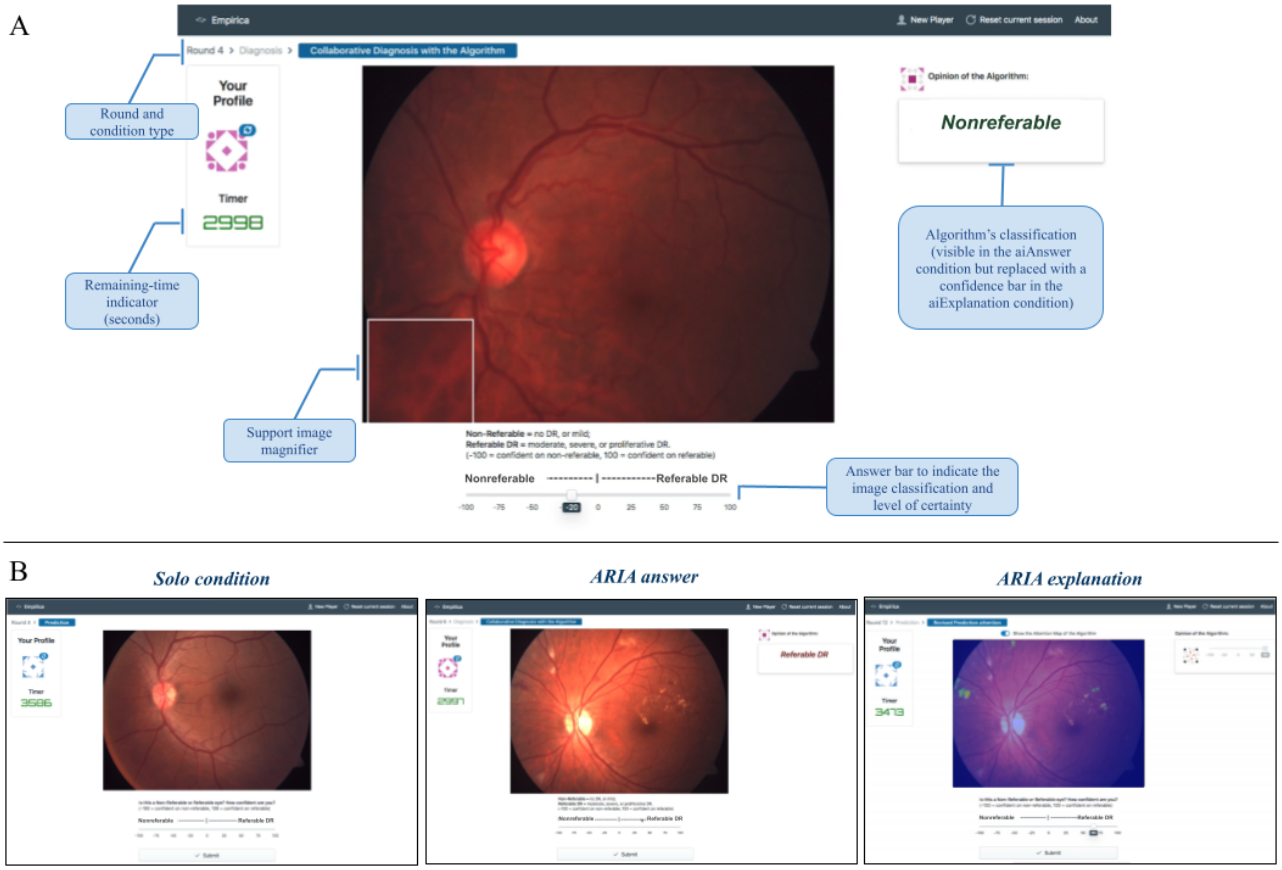
Overview of Study Design

We conducted a randomized controlled experiment to assess the performance of the ARIA system in comparison with ophthalmic doctors from the Mexican ophthalmic hospital and to evaluate the potential benefits of using the system as an assistive tool for doctors. To achieve this, a web-based experiment platform was developed where ophthalmologists

XSL•FO
RenderX

evaluated fundus retinal images under 3 different conditions—solo, ARIA answer, and ARIA explanation—described below. The platform was developed based on the Empirica framework [39]. Figure 3 displays the main screens of the web platform used in this experiment.

Figure 3. Web-platform design for patient-case classification. (A) Visual indicators and components of the classification window. (B) Visualization of the 3 experimental conditions. ARIA: automated retinal image analysis; DR: diabetic retinopathy.



Gold Standard and Image Quality

To generate a gold standard, the fundus images of all patient cases used in the experiment were graded by 3 retina specialists of the ophthalmic hospital, and a majority rule was used (ie, if there was a disagreement in the nonreferable/referable label, the label selected by 2 of 3 experts was considered the gold standard). We used the same web-based platform described in Figure 3 for image grading. The retina specialists also graded the image quality, and images graded as bad quality were not considered for the experiment. From the remaining images, 50 images from patients with referable DR and 50 images from patients with nonreferable DR were selected at random to be used for the study. According to the ICDRSS, the selected images had the following distribution: 49 with no apparent retinopathy, 1 with mild DR, 33 with moderate DR, 12 with severe nonproliferative DR, and 5 with proliferative DR. Since these images were taken at an ophthalmic hospital, most patients with DR were under treatment and therefore had more advanced DR stages (moderate, severe, and proliferative DR).

Experimental Conditions

The experiment followed a within-subjects design, where each ophthalmologist evaluated 45 randomly selected fundus images (from 45 different patients), 15 for each of the 3 treatment conditions: solo, ARIA answer, and ARIA explanation. The ophthalmologists were first asked to evaluate 15 fundus retinal images in the solo condition, followed by 30 images that randomly alternated between the ARIA answer and the ARIA explanation conditions. The 15 images in each condition subset were randomly selected for each participant without replacement from all images available for the experiment, generating a rough balance in the proportion of referable and nonreferable images across conditions. In particular, the average proportion of referable images was 49.8% (127/255) for the solo condition, 52.5% (134/255) for the ARIA answer condition, and 46.7% (119/255) for the ARIA explanation condition. In addition, Multimedia Appendix 1 reports the average number of observations of each ICDRSS class for each treatment condition.

In the solo condition, participants responded to the task in isolation, without any exposure to the ARIA system. In contrast, in the ARIA answer condition, participants were exposed to the binary answer of the ARIA system (ie, nonreferable or referable), as a second opinion, and then asked to submit their postexposure answer. The ARIA explanation condition was identical to the ARIA answer condition, with the exception that participants were shown not only the binary answer of the ARIA system but also its level of confidence and attention heatmap.

Finally, after completing all the classification tasks, the ophthalmologists were asked to submit an optional feedback survey about their experience.

XSL•FO
RenderX

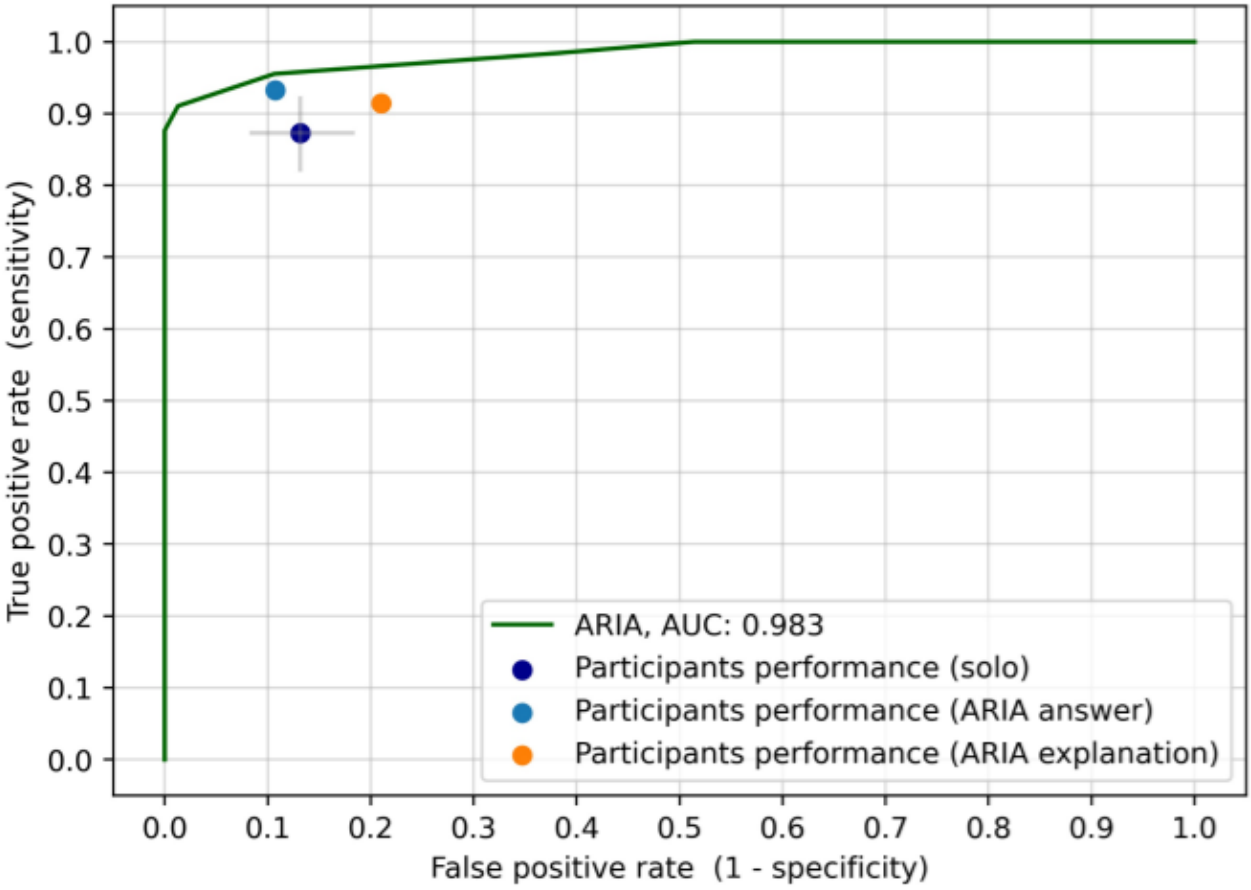
The study was reviewed and approved by the Committee on the Use of Humans as Experimental Subjects at the Massachusetts Institute of Technology, and all participants provided explicit consent prior to their participation.

Results

ARIA's Independent Performance

The ARIA system was first tested in a large dataset of international cases. It achieved an out-of-sample area under the receiver operating characteristic curve (AUROC) of 98% (Multimedia Appendix 1). In particular, using a given acceptance threshold, the ARIA system achieved a sensitivity of 95.1% and a specificity of 91.5%. Most importantly, the ARIA system also displayed high accuracy classifying images from patients from the Mexican ophthalmic hospital, where it had an AUROC of 98.3%, a sensitivity of 95.2%, and specificity of 90% (Figure 4).

Figure 4. Receiver operating characteristic curve of the ARIA system compared with the ophthalmologist's accuracy under the 3 experimental conditions (solo, ARIA answer, and ARIA explanation). Grey lines indicate 95% CIs for the solo condition. ARIA: automated retinal image analysis; AUC: area under the curve.



ARIA's Assistive Performance

Figure 4 shows the sensitivity and false positive rate (false positive rate = 1 – specificity) for each condition—solo, ARIA answer, and ARIA explanation—and compares them with the receiver operating characteristic curve of the ARIA system. The average sensitivity in the solo condition across the 17 participants was 87.3%, and the average specificity was 86.8%. In comparison, the average sensitivity and specificity across the 17 participants for the ARIA answer condition were 93.3% and 89.3%, respectively, and the average sensitivity and specificity across participants for the ARIA explanation condition were 91.5% and 79%, respectively.

The joint analysis of the ARIA system performance for Mexican patients, compared with the 3 experimental conditions involving ophthalmologist assessments, showed that the ARIA system is more accurate than the average accuracy of participants under any of the exposure conditions. In particular, the ARIA system increased sensitivity from 87.3% to 93.3% (P=.05; vertical movement between the dark blue dot and the green line in Figure 4) while maintaining participants' specificity at 86.8%. Compared with the solo condition, the ARIA system also increased specificity to 100% while maintaining participants' average sensitivity at 87.3% (horizontal movement from the dark blue dot leftwards to the green line in Figure 4).

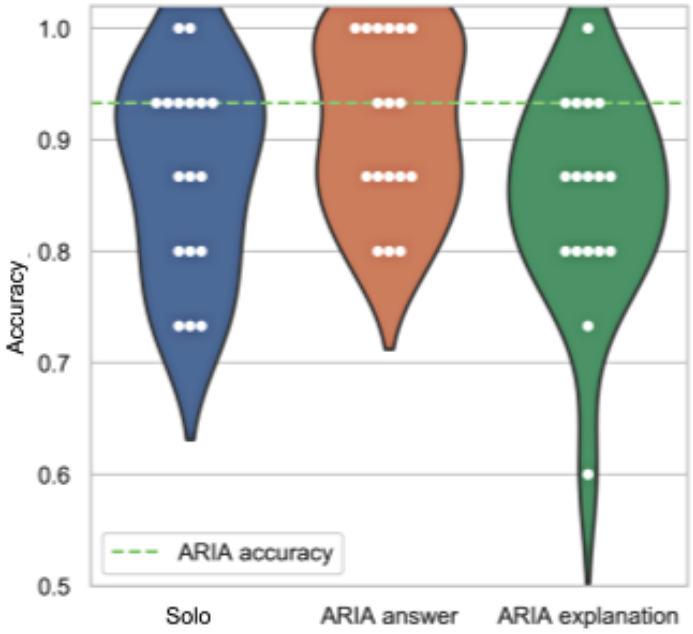
Most interestingly, Figure 4 shows that exposure to the ARIA system was able to improve the performance of human experts, particularly, in the ARIA answer condition, which significantly improved the sensitivity and specificity compared with the solo condition (distance between dark blue and light blue dots in Figure 4). However, performance in the ARIA explanation condition had mixed results, showing improved sensitivity but

XSL•FO
RenderX

worse specificity (distance between dark blue and orange dots in Figure 4).

Figure 5 provides more detail on the effect that exposure to information of the ARIA system had on the performance of ophthalmologists. In particular, it shows that the accuracy (% of correct answers) of the 17 experts consistently improved in the ARIA answer condition, shifting the distribution upwards and decreasing the variance across participants. For example, while only 2 participants had a perfect score in the solo condition, up to 6 participants had a perfect score in the ARIA answer condition. However, the ARIA explanation condition had mixed beneficial and detrimental effects on participants' accuracy and increased the variance of performance across participants compared with the solo condition.

Figure 5. Influence of the ARIA system on the ophthalmologists' decisions: ophthalmologists' performance after exposure to the ARIA answer or the ARIA explanation condition outputs. ARIA: automated retinal image analysis.



Discussion

Principal Findings

The number of people living with diabetes by 2045 is projected to reach 700 million people worldwide [7,40]. This means that routine eye screening might prevent vision loss in approximately 230 million patients. Just in Mexico, the prevention of DR would implicate savings of up to US $10 million for the 3 main public health care institutions [41]. The development of ARIA systems represents a possible solution to the increasing demand of eye screenings in health care systems, particularly, in limited-resource settings. However, it has been shown that acceptance of the human factors involved in the field processes are critical for the effective implementation of screening systems [42,43].

In this study, we successfully developed and evaluated a DL-based ARIA system to determine its performance as an independent decision-making system, as well as a supportive tool for health care professionals. As an independent decision-making tool, the ARIA system outperformed the average ophthalmologist participant in the experiment. On the other hand, as a supportive tool, the ARIA system exerted a strong influence on the opinion of human participants. However, its effect depended on the output's format, highlighting the importance of a well-designed platform that has been user-tested with final users.

ARIA's Independent Performance

The DL-based ARIA system presented in this work was evaluated with a subset of retinal images from international patient cases and an image set of patients from a Mexican ophthalmic hospital. In both datasets, the ARIA system outperformed the average sensitivity and specificity of 17 ophthalmology residents of retina specialty.

The sensitivities (95.1% and 95.2% for the international and Mexican datasets, respectively) are comparable to those reported for 7 other automated DR screening systems assessed in a systematic review, whose sensitivity values were between 87% and 95% [44]. On the other hand, the specificities reached by our ARIA system (91.5% and 90% for the international and Mexican datasets, respectively) were higher than the average specificity values of between 49% and 69% reported by Nørgaard and Grauslund [44]. Also, our system's sensitivity and specificity were comparable with those reported for commercial DR screening technologies with DL features, whose sensitivity and specificity values were 85%-99.3% and 68.8%-97.9%, respectively [45]. Compared with these commercial DR screening technologies, our ARIA system has one of the best balances between sensitivity and specificity, with both measurements above 90%.

ARIA Assistive Performance

Besides the sensitivity and specificity assessment, the ARIA system evaluation included 2 hybrid decision schemes, either assistive or a combination of human and AI. The experimental

XSL•FO
RenderX

design was developed to reflect that in real-world applications, results of an automated system are reviewed and confirmed by health care professionals to choose the most adequate therapeutic protocol for each patient. In these assistive evaluations, we confirmed the existence of significant synergies derived from the interaction among the human and AI dyads.

The ARIA output's influence on ophthalmologists' overall precision depended on its format. A simplified output (ie, nonreferable or referable classification) resulted in the most successful sensitivity and specificity for ophthalmologists' inputs. On the other hand, a more complex output (ie, with a confidence bar and attention map) partially improved ophthalmologists' decisions, increasing their sensitivity but also increasing the incidence of false positive classifications.

These results are coherent with some of the ophthalmologists' feedback submitted after the classification tasks, where some expressed that even when attention heatmaps were useful, the bar showing the confidence of the ARIA system was confusing.

Limitations

Future pilot studies with a larger number of patients and ophthalmologists will be useful to confirm the ARIA system's accuracy. Also, future studies might include direct ophthalmoscopy by retina specialists as the gold standard, in order to avoid errors related to image quality.

Additional experiments with alternative platform designs might be useful to generate a suitable screening tool that optimizes patient evaluations and referrals in 2 stages. In the first stage, an ARIA system might be useful to identify patients with a higher probability of developing DR. In the second stage, ophthalmologists would be able to evaluate the retinal images of high-risk patients, in combination with the ARIA system output, to make a first decision about the disease stage and treatment, sending referrals to retina specialists only for patients with an advanced disease.

Conclusions

The results of this study demonstrate a substantial opportunity for Latin American countries such as Mexico toward developing efficient mass screening systems for early detection of diabetes-related blindness, considering the short supply of ophthalmologists in their public health care system.

The web-based platform developed for this study was designed for the implementation of the ARIA system as an automatic screening tool and as a telemedicine platform to confirm or reject the ARIA system's output with assessment of an ophthalmologist or retina specialist. The platform was useful for this study and can be easily adapted for future studies that include the collection of additional information about other eye diseases detectable by image analysis (ie, glaucoma, age-related macular degeneration, or coat disease).

The conclusion of these results suggests the proposed ARIA system is valuable in an independent or assistive condition and can be useful to increase and improve DR diagnosis, as well as other ophthalmic diseases in the future. However, special attention to the design of an explanatory platform is required for successful implementation of the system.

Acknowledgments

The authors gratefully thank the retina specialists and the ophthalmologists from the APEC Hospital involved in this study for their evaluations of the retina fundus images. This project was carried out thanks to the fellowships received by individual members of the team, including fellowships of the Massachusetts Institute of Technology and the National Council of Science and Technology (CONACYT).

Authors' Contributions

AN conceived and designed the experiments, analyzed data, and contributed to the discussion and review of the paper. DC trained the models, performed the experiments, analyzed data, and contributed to the discussion of the paper. DM contributed to the experimental design, image classification, and discussion of the paper. JE contributed with data analysis and paper writing, including the discussion. HQM, VMC, AA, and AP contributed to various aspects of the paper, including experimental design, machine learning strategies, medical feedback, image evaluations, and the discussion.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Average number of observations according to the International Clinical Diabetic Retinopathy Severity Scales classification, for each set of 15 retina images in each treatment condition: solo, ARIA answer, and ARIA explanation.
[DOC File , 58 KB - formative_v5i8e25290_app1.doc]

Multimedia Appendix 2
CONSORT-EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 407 KB - formative_v5i8e25290_app2.pdf]

XSL•FO
RenderX

References

1. Zhang X, Saaddine JB, Chou C, Cotch MF, Cheng YJ, Geiss LS, et al. Prevalence of diabetic retinopathy in the United States, 2005-2008. JAMA 2010 Aug 11;304(6):649-656 [FREE Full text] [doi: 10.1001/jama.2010.1111] [Medline: 20699456]
2. SIGMA Type 2 Diabetes Consortium T, Williams AL, Jacobs SBR, Moreno-Macías H, Huerta-Chagoya A, Churchhouse C, et al. Sequence variants in SLC16A11 are a common risk factor for type 2 diabetes in Mexico. Nature 2014 Feb 06;506(7486):97-101 [FREE Full text] [doi: 10.1038/nature12828] [Medline: 24390345]
3. Caballero AE. Understanding the Hispanic/Latino patient. Am J Med 2011 Oct;124(10 Suppl):S10-S15. [doi: 10.1016/j.amjmed.2011.07.018] [Medline: 21939793]
4. Secretaría de Salud. 2016. Encuesta Nacional de Salud y Nutrición de Medio Camino 2016 (Ensanut MC 2016). Cuernavaca: INSP; 2016. URL: https://www.gob.mx/cms/uploads/attachment/file/209093/ENSANUT.pdf [accessed 2019-11-08]
5. Institute for Health Metrics and Evaluation. 2017. Country profiles (Mexico). URL: http://www.healthdata.org/results/country-profiles [accessed 2019-08-24]
6. Barcelo A, Arredondo A, Gordillo-Tobar A, Segovia J, Qiang A. The cost of diabetes in Latin America and the Caribbean in 2015: Evidence for decision and policy makers. Journal of Global Health 2017 Dec;7(2):020410 [FREE Full text] [doi: 10.7189/jogh.07.020410] [Medline: 29163935]
7. International Diabetes Federation. 2019. IDF Diabetes Atlas Ninth edition. 2019. URL: https://diabetesatlas.org/upload/resources/material/20200302_133352_2406-IDF-ATLAS-SPAN-BOOK.pdf [accessed 2019-11-15]
8. Gómez EJ. Political party ambitions and type-2 diabetes policy in Brazil and Mexico. Health Econ Policy Law 2020 Apr;15(2):261-276. [doi: 10.1017/S1744133118000415] [Medline: 30394254]
9. Deshpande A, Harris-Hayes M, Schootman M. Epidemiology of diabetes and diabetes-related complications. Phys Ther 2008 Nov;88(11):1254-1264 [FREE Full text] [doi: 10.2522/ptj.20080020] [Medline: 18801858]
10. Cheung N, Mitchell P, Wong TY. Diabetic retinopathy. The Lancet 2010 Jul;376(9735):124-136. [doi: 10.1016/S0140-6736(09)62124-3]
11. Secretaría de Salud. Diagnóstico y Tratamiento de Retinopatía Diabética. Evidencias y Recomendaciones. Catálogo Maestro. Guías de Práctica Clínica. México: CENETEC; 2011. URL: http://www.cenetec.salud.gob.mx/descargas/gpc/CatalogoMaestro/171_GPC_RETINOPATIA_DIABETICA/Imss_171RR.pdf [accessed 2019-07-19]
12. Jaime CL. RETINOPATÍA DIABÉTICA DESDE LA PREVENCIÓN. INTEGRAR LA PESQUISA EN LOS CENTROS DE DIABETES. Revista Médica Clínica Las Condes 2016 Mar;27(2):195-203. [doi: 10.1016/j.rmclc.2016.04.009]
13. Barría VF, Martínez CF. Clinical Practice Guide for Diabetic Retinopathy for Latin America for Ophthalmologists and Healthcare Professionals. 2011. URL: https://www.iapb.org/wp-content/uploads/2011-Clinical-Practice-Guide-for-DR-for-Latin-America.pdf [accessed 2020-08-21]
14. Cavan D, Makaroff L, da Rocha Fernandes J, Sylvanowicz M, Ackland P, Conlon J, et al. The Diabetic Retinopathy Barometer Study: Global perspectives on access to and experiences of diabetic retinopathy screening and treatment. Diabetes Res Clin Pract 2017 Jul;129:16-24. [doi: 10.1016/j.diabres.2017.03.023] [Medline: 28499163]
15. Jiménez Báez M, Márquez González H, Bárcenas Contreras R, Morales Montoya C, García Espinosa L. EARLY DIAGNOSIS OF DIABETIC RETINOPATHY IN PRIMARY CARE. Colombia Médica 2015 Jan 01;46(01):14-18 [FREE Full text] [doi: 10.25100/cm.v46i1.1681]
16. Carrillo-Alarcón L, Ávila-Pozos R, López LE, Cruz-Castillo R, Ocampo-Torres M, Alcalde-Rabanal J. Projection of Diabetic Patients Retinopathy in Hidalgo State-México, through 2030. EC Ophthalmology 2017;5(2):73-80.
17. Thomas RL, Halim S, Gurudas S, Sivaprasad S, Owens D. IDF Diabetes Atlas: A review of studies utilising retinal photography on the global prevalence of diabetes related retinopathy between 2015 and 2018. Diabetes Res Clin Pract 2019 Nov;157:107840. [doi: 10.1016/j.diabres.2019.107840] [Medline: 31733978]
18. Hong H, Mújica OJ, Anaya J, Lansingh VC, López E, Silva JC. The Challenge of Universal Eye Health in Latin America: distributive inequality of ophthalmologists in 14 countries. BMJ Open 2016 Nov 18;6(11):e012819 [FREE Full text] [doi: 10.1136/bmjopen-2016-012819] [Medline: 27864248]
19. Martinez J, Hernandez-Bogantes E, Wu L. Diabetic retinopathy screening using single-field digital fundus photography at a district level in Costa Rica: a pilot study. Int Ophthalmol 2011 Apr;31(2):83-88. [doi: 10.1007/s10792-010-9413-9] [Medline: 21274593]
20. Salamanca O, Geary A, Suárez N, Benavent S, Gonzalez M. Implementation of a diabetic retinopathy referral network, Peru. Bull World Health Organ 2018 Oct 01;96(10):674-681 [FREE Full text] [doi: 10.2471/BLT.18.212613] [Medline: 30455515]
21. Rachapelle S, Legood R, Alavi Y, Lindfield R, Sharma T, Kuper H, et al. The cost-utility of telemedicine to screen for diabetic retinopathy in India. Ophthalmology 2013 Mar;120(3):566-573. [doi: 10.1016/j.ophtha.2012.09.002] [Medline: 23211635]
22. 2016. From $600 M to $6 Billion, Artificial Intelligence Systems Poised for Dramatic Market Expansion in Healthcare. 2016. URL: https://ww2.frost.com/news/press-releases/600-m-6-billion-artificial-intelligence-systems-poised-dramatic-market-expansion-healthcare/ [accessed 2019-12-01]

XSL•FO
RenderX

23. He J, Baxter SL, Xu J, Xu J, Zhou X, Zhang K. The practical implementation of artificial intelligence technologies in medicine. Nat Med 2019 Jan;25(1):30-36. [doi: 10.1038/s41591-018-0307-0] [Medline: 30617336]
24. Arenas-Cavalli J, Ríos S, Pola M, Donoso R. A Web-based Platform for Automated Diabetic Retinopathy Screening. Procedia Computer Science 2015;60:557-563 [FREE Full text] [doi: 10.1016/j.procs.2015.08.179]
25. Yang Y, Li T, Li W, Wu H, Fan W, Zhang W. Lesion detection and Grading of Diabetic Retinopathy via Two-stages Deep Convolutional Neural Networks. In: Lecture Notes in Computer Science. Switzerland: Springer Nature; 2017 Presented at: 20th Medical Image Computing and Computer Assisted Intervention - MICCAI 2017; September 11-13; Quebec City, Canada p. 533-540 URL: https://doi.org/10.1007/978-3-319-66179-7_61 [doi: 10.1007/978-3-319-66179-7_61]
26. Bhaskaranand M, Ramachandra C, Bhat S, Cuadros J, Nittala MG, Sadda S, et al. Automated Diabetic Retinopathy Screening and Monitoring Using Retinal Fundus Image Analysis. Journal of Diabetes Science and Technology 2016 Feb 16;10(2):254-261 [FREE Full text] [doi: 10.1177/1932296816628546] [Medline: 26888972]
27. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. JAMA 2016 Dec 13;316(22):2402-2410. [doi: 10.1001/jama.2016.17216] [Medline: 27898976]
28. Tufail A, Rudisill C, Egan C, Kapetanakis VV, Salas-Vega S, Owen CG, et al. Automated Diabetic Retinopathy Image Assessment Software: Diagnostic Accuracy and Cost-Effectiveness Compared with Human Graders. Ophthalmology 2017 Mar;124(3):343-351 [FREE Full text] [doi: 10.1016/j.ophtha.2016.11.014] [Medline: 28024825]
29. Dutz M, Almeida R, Packard T. The Jobs of Tomorrow: Technology, Productivity, and Prosperity in Latin America and the Caribbean. Washington DC: World Bank Group; 2018:1-71.
30. Chollet F, others. Keras. 2015. URL: https://keras.io [accessed 2018-12-01]
31. Van RG, Drake JF. Python Tutorial Release 2. 2001. URL: https://docs.python.org/2.0/tut/tut.html [accessed 2018-12-01]
32. Wilkinson C, Ferris FL, Klein RE, Lee PP, Agardh CD, Davis M, Global Diabetic Retinopathy Project Group. Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales. Ophthalmology 2003 Sep;110(9):1677-1682. [doi: 10.1016/S0161-6420(03)00475-5] [Medline: 13129861]
33. LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, et al. Backpropagation Applied to Handwritten Zip Code Recognition. Neural Computation 1989 Dec;1(4):541-551. [doi: 10.1162/neco.1989.1.4.541]
34. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. ArXiv 2015;1409.1556v6:1-14 [FREE Full text]
35. Cuadros J, Bresnick G. EyePACS: an adaptable telemedicine system for diabetic retinopathy screening. Journal of Diabetes Science and Technology 2009 May 01;3(3):509-516 [FREE Full text] [doi: 10.1177/193229680900300315] [Medline: 20144289]
36. Voets M, Møllersen K, Bongo L. Reproduction study using public data of: Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. PLoS One 2019;14(6):e0217541 [FREE Full text] [doi: 10.1371/journal.pone.0217541] [Medline: 31170223]
37. Samek W, Montavon G, Binder A, La-puschkin S, Muller KR. Interpreting the Predictions of Complex ML Models by Layer-wise Relevance Propagation. ArXiv 2016;1611.08191v1:1-5 [FREE Full text]
38. Bach S, Binder A, Montavon G, Klauschen F, Müller K, Samek W. On Pixel-Wise Explanations for Non-Linear Classifier Decisions by Layer-Wise Relevance Propagation. PLoS One 2015 Jul;10(7):e0130140 [FREE Full text] [doi: 10.1371/journal.pone.0130140] [Medline: 26161953]
39. Almaatouq A, Becker J, Houghton JP, Paton N, Watts DJ, Whiting ME. Empirica: a virtual lab for high-throughput macro-level experiments. Behavior Research Methods 2021 Mar 29:1-14 [FREE Full text] [doi: 10.3758/s13428-020-01535-9] [Medline: 33782900]
40. Saeedi P, Petersohn I, Salpea P, Malanda B, Karuranga S, Unwin N, IDF Diabetes Atlas Committee. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9 edition. Diabetes Res Clin Pract 2019 Nov;157:107843 [FREE Full text] [doi: 10.1016/j.diabres.2019.107843] [Medline: 31518657]
41. Barquera S, Campos-Nonato I, Aguilar-Salinas C, Lopez-Ridaura R, Arredondo A, Rivera-Dommarco J. Diabetes in Mexico: cost and management of diabetes and its complications and challenges for health policy. Global Health 2013 Feb 02;9:3 [FREE Full text] [doi: 10.1186/1744-8603-9-3] [Medline: 23374611]
42. Raumviboonsuk P, Krause J, Chotcomwongse P, Sayres R, Raman R, Widner K, et al. Deep learning versus human graders for classifying diabetic retinopathy severity in a nationwide screening program. NPJ Digit Med 2019;2:25 [FREE Full text] [doi: 10.1038/s41746-019-0099-8] [Medline: 31304372]
43. Beede E, Baylor E, Hersch F, Iurchenko A, Wilcox L, Ruamviboonsuk P, et al. A Human-Centered Evaluation of a Deep Learning System Deployed in Clinics for the Detection of Diabetic Retinopathy. : Association for Computing Machinery; 2020 Apr Presented at: 2020 CHI Conference on Human Factors in Computing Systems (CHI '20); April 25-30; Honolulu p. 1-12 URL: https://doi.org/10.1145/3313831.3376718 [doi: 10.1145/3313831.3376718]
44. Nørgaard MF, Grauslund J. Automated Screening for Diabetic Retinopathy - A Systematic Review. Ophthalmic Res 2018;60(1):9-17 [FREE Full text] [doi: 10.1159/000486284] [Medline: 29339646]

XSL•FO
RenderX

45. Grzybowski A, Brona P, Lim G, Ruamviboonsuk P, Tan G, Abramoff M, et al. Artificial intelligence for diabetic retinopathy screening: a review. Eye (Lond) 2020 Mar;34(3):451-460 [FREE Full text] [doi: 10.1038/s41433-019-0566-0] [Medline: 31488886]

Abbreviations

AI: artificial intelligence
ARIA: automated retinal image analysis
AUROC: area under the receiver operating characteristic curve
DL: deep learning
DR: diabetic retinopathy
ICDRSS: International Clinical Diabetic Retinopathy Severity Scales
NPDR: nonproliferative diabetic retinopathy

Edited by G Eysenbach; submitted 18.11.20; peer-reviewed by G Lim, Y Li; comments to author 24.02.21; revised version received 12.04.21; accepted 19.05.21; published 26.08.21.

Please cite as:
Noriega A, Meizner D, Camacho D, Enciso J, Quiroz-Mercado H, Morales-Canton V, Almaatouq A, Pentland A
Screening Diabetic Retinopathy Using an Automated Retinal Image Analysis System in Independent and Assistive Use Cases in Mexico: Randomized Controlled Trial
JMIR Form Res 2021;5(8):e25290
URL: https://formative.jmir.org/2021/8/e25290
doi:10.2196/25290
PMID:34435963

©Alejandro Noriega, Daniela Meizner, Dalia Camacho, Jennifer Enciso, Hugo Quiroz-Mercado, Virgilio Morales-Canton, Abdullah Almaatouq, Alex Pentland. Originally published in JMIR Formative Research (https://formative.jmir.org), 26.08.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Preferences for Digital Smartphone Mental Health Apps Among Adolescents: Qualitative Interview Study

Robert Ribanszki[1], BSc, MSc; Jose Andres Saez Fonseca[1,2], MD, MSc; Joseph Matthew Barnby[1], BSc, MSc; Kimberly Jano[1], BSc, MSc; Fatima Osmani[1], BSc, MSc; Soma Almasi[1], BSc, MSc; Elias Tsakanikos[3,4], PhD

[1]Thrive: Therapeutic Software, London, United Kingdom

[2]Division of Psychiatry, Univeristy College London, London, United Kingdom

[3]Health Service and Population Research Department, Institute of Psychiatry, Psychology and Neuroscience, Kings' College London, UK, London, United Kingdom

[4]Department of Psychology, University of Crete, Crete, Greece

Corresponding Author:
Elias Tsakanikos, PhD
Health Service and Population Research Department
Institute of Psychiatry, Psychology and Neuroscience, Kings' College London, UK
16 De Crespigny Park
London, SE5 8AF
United Kingdom
Phone: 44 000 000 000
Email: elias.tsakanikos@kcl.ac.uk

Abstract

Background: Mental health digital apps hold promise for providing scalable solutions to individual self-care, education, and illness prevention. However, a problem with these apps is that they lack engaging user interfaces and experiences and thus potentially result in high attrition. Although guidelines for new digital interventions for adults have begun to examine engagement, there is a paucity of evidence on how to best address digital interventions for adolescents. As adolescence is a period of transition, during which the onset of many potentially lifelong mental health conditions frequently occurs, understanding how best to engage this population is crucial.

Objective: The study aims to detect potential barriers to engagement and to gather feedback on the current elements of app design regarding user experience, user interface, and content.

Methods: This study used a qualitative design. A sample of 14 adolescents was asked to use the app for 1 week and was interviewed using a semistructured interview schedule. The interviews were transcribed and analyzed using thematic analysis.

Results: Overall, 13 participants completed the interviews. The authors developed 6 main themes and 20 subthemes based on the data that influenced engagement with and the perceived usefulness of the app. Our main themes were *timing*, *stigma*, *perception*, *congruity*, *usefulness*, and *user experience*.

Conclusions: In line with previous research, we suggest how these aspects of app development should be considered for future apps that aim to prevent and manage mental health conditions.

(JMIR Form Res 2021;5(8):e14004) doi:10.2196/14004

KEYWORDS

qualitative; adolescents; mental health; digital smartphone app; digital mental health; mobile phone

Introduction

Background

The rise in common mental health issues among adolescents is a distressing trend. The World Health Organization has estimated that 20% of adolescents experience mental health conditions, and most of them do not receive or seek appropriate diagnosis and care [1]. Addressing this concern is an essential component of the current global mental health agenda [2]. Innovative solutions delivered by mental health apps (MHapps) could represent a feasible solution to tackle this issue.

There is already a plethora of mental health mobile apps available to adolescents, and the increasing use of smartphones

XSL•FO
RenderX

in this group might make these apps more acceptable and accessible [3]. Studies have also suggested that complex app-based mental health interventions for adolescents are feasible. For instance, in one study, cognitive behavioral therapy in the form of SMS text messages was considered useful by 75% of the participants [4]. Moreover, significant engagement with MHapps has been shown in the past, with nearly three-quarters of adolescents completing more than 80% diary entries over the 1-week intervention period [5]. MHapps also have the potential to reduce barriers to face-to-face help seeking, including stigma and distress about discussing one's own mental health [6]. This aspect of MHapps may be appealing to young people, given that most adolescents would not seek or pursue help with respect to mental health through traditional routes [7].

Adolescents' familiarity with mobile devices suggests that technology-based approaches would benefit them [8]; however, it is crucial to understand how best to tailor digital interventions to make them the most appealing. Tucker and Goodings [9] identified three themes that characterize most current MHapps, as follows: stress-inducing or stress-reducing apps, apps for configuring the body in space, and digital self-care apps. However, apps targeting adolescents likely need to expand into other areas such as positive focus, customizable features, human-human interaction, and easy access [3]. In terms of help-seeking preferences, adolescents have also expressed a desire for web-based, accessible information and health interventions, which are all technology-based needs rather than needs that can be met via in-person, telephone-based, or paper-based services [10]. When considering communication with providers, adolescents preferred email or text over video communication [11].

Despite their familiarity with digital technologies, engagement is generally low and evidence on the usefulness of concrete features is still scarce [12,13]. In addition, engagement with MHapps seems to vary independently from the presence of evidence-based features [14]. This raises questions about the clinical effectiveness and safety that undermine trust in both users and providers [15]. Furthermore, high dropout rates are generally associated with poor user experience (UX), whereas the specific components of engaging MHapp design are yet to be determined [16,17].

Objective

In this study, we seek to identify the key preferences and attitudes of adolescents that future digital mental health interventions may need to take into account to successfully reach this population. For this purpose, we used the *Thrive* mental health app, which has an established evidence base [18,19], to explore adolescents' perceptions of the potential usability of such a tool in their everyday life. Using a qualitative design, we gathered feedback from a sample of adolescents through face-to-face interviews. Specifically, our goal was to detect potential barriers to engagement and to gather feedback on the current elements of app design regarding UX, user interface (UI), and content. The long-term aim of this exploratory study is to provide the foundations for creating digital interventions for adolescents that are equally driven by clinical rigor and UX to help retain engagement.

Methods

Ethics and Preregistration

This research study was approved by the Roehampton University ethics board (reference number: PSYC 18/306) and preregistered as a qualitative protocol (reference number: TCYP171110). All methodologies adhered to the protocol unless stated otherwise.

Participants and Procedure

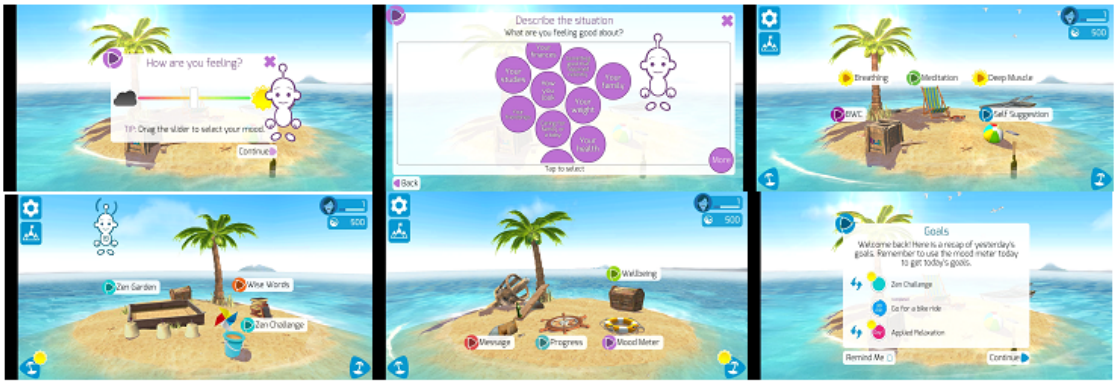
We recruited a total of 8 male and 6 female participants (N=14). Overall, 11 participants were recruited from a local secondary school attended by approximately 1300 students at the time of the study. Students were made aware of the study through combined advertisements in the school. For interested students, the researchers gave a talk that outlined the overall purpose of the app and explained that the study was trying to understand what features of digital MHapps were most useful to keep students engaged. Furthermore, 3 participants were recruited through their parents who were also users of the app and knew about the study. These 3 participants were enrolled after an introductory conversation with the researchers, explaining the details and the purpose of the study.

Participants in the school were asked to get in touch with the lead teacher to express their interest in participating in the study. The teachers assessed whether the students adhered to the inclusion criteria to be a part of the study. Researchers then got in touch with the lead teacher to schedule a meeting in person on school grounds to interview the participants. Participants who were recruited via end users were interviewed over the phone. Consent forms for participants, parents, and teachers were sent out and returned via email. The information sheets were sent via email. The participants were interviewed as soon as possible after they had given their consent forms and had used the app for at least 1 week.

All participants were asked to use the *Thrive: Feel Stress Free* app (Figure 1) for 1 week as much as they liked. Participants were asked to turn on their notifications in the app; however, this was not enforced as a strict inclusion criterion.

XSL•FO
RenderX

Figure 1. Screenshots of the Thrive: Feel Stress Free app. From top left to bottom right, this figure shows the mood meter, mood bubbles, first menu, second menu, third menu (all menus are accessible from the main screen and have an island theme), and goal menu.



Sample Size and Theme Saturation

Although our target sample size was 30, recruitment was stopped after 13 participants had been interviewed. The reason for this decision was twofold:

1. Owing to unforeseen barriers (school examinations, holidays, and teacher availability), our initial recruitment strategy resulted in 13 interviews, and another round of recruitment would have been needed to reach our intended sample size.
2. At this stage, we decided to review our sample size estimate by assessing theme saturation in our data.

We defined saturation as "the point during data analysis at which incoming data points (interviews) produce little or no new useful information relative to the study objectives" [20]. Using an approach recently refined by Guest et al [20], we set out to estimate how many additional interviews might be needed to reach theme saturation in our data. As most novel information is seen early in the coding process and follows an asymptotic curve in qualitative data sets [21], it is possible to use the occurrence of novel codes in each subsequent interview to estimate the slope of this curve and make an informed decision about a new recruitment target by systematically coding available interviews. However, after the coding of interview 9, it became apparent that saturation had already occurred and no further recruitment was necessary. To assess the saturation, interviews were coded by 2 of the researchers—JASF and JK—none of whom took part in interviewing participants or transcribing interviews.

The App

We chose the Thrive app for this study because we considered its design and UX elements as good examples of the state-of-the-art development principles in the digital mental health field. Moreover, one of the authors (JASF) was involved in the development of the Thrive app, which allowed us easy access to the app for the purpose of this study.

After opening the app, participants were guided through a short mood assessment and thought training exercise, which allowed the app to create a customized cognitive training plan for the user. The recommended exercise modules (and others) could be accessed after the assessment phase was complete. The included modules were a combination of guided relaxation techniques and guides that provided further background and recommendations related to the given feature, such as meditation, breathing, or self-suggestions. Exercises were explained in detail, and the users were guided through the entire process when practicing each module. Next to the cognitive training features, participants were able to play games, such as Zen Garden or word puzzle, aimed to provide a more engaging UX. When finished with the modules, the app also provided participants with an overview of their progress where they could track their mood, practice, and goals. Examples of these steps are presented in Figure 1.

Data Collection

Data were collected between January and July 2018. Owing to unforeseen delays, such as school examinations, holidays, and teacher availability, this was substantially longer than reported in the protocol. Eligible participants were between the ages of 11 and 18 years, owned a smartphone, used it frequently (more than 1 hour a day), and did not have an existing diagnosis of a mental health condition.

During the interview, no one else was present aside from the participant and the researcher, either in person or over the phone. We did not interview participants as a focus group because of the difficulty in organizing this at a convenient time.

The sessions were guided by a semistructured schedule (Multimedia Appendix 1). The discussion began with participants' overall impressions of the app. For example, "what bits of it [the app] were useful if any" or "which bits of it [the app] did you dislike?" The schedule then proceeded to more specific questions if not covered in the overall broad questions. For example, "what did you think of our avatar" or "what did you think of the journal?" This was not pilot-tested before commencing the interviews but was constructed by consensus between the authors.

Participants were asked when they would engage with the app during the day, if at all, and at which location this took place. We asked which barriers and facilitators led them to use the app less or more frequently. In addition, we asked participants to list what they would change about the app to make it more interesting. The participants led the conversation though the interviewers ensured that the participants were prompted to specific topics if previously missed. Each interview lasted approximately 20-40 minutes. No repeat interviews were

XSL•FO
RenderX

conducted. The interviews were audio-recorded and transcribed verbatim by SA, FO, RR, and JMB.

Analysis

The interview extracts were analyzed using inductive thematic analysis [22]. First, interview transcripts were read carefully by 3 researchers (RR, JASF, and KJ) to identify meaningful ideas relevant to the research topic. On the basis of this, an initial list of relevant concepts was generated. Second, short segments of the data, dealing with similar issues or concepts, were identified and grouped together using provisional codes. At this stage, researchers coded each transcript on their own and could use different codes for any single data segment. Third, codes were discussed and cross-referenced between the researchers and collated into a common framework that allowed for candidate themes to emerge. Finally, upon review, candidate themes were grouped into main (or meta) themes that formed coherent meaningful concepts across the texts. The main themes and subthemes were then reviewed and refined using the original transcripts and linked to participant quotes to ensure that the final themes indeed formed a coherent pattern across the whole data set. Our analysis resulted in 6 main themes and 20 subthemes (Textbox 1).

Textbox 1. Main themes and subthemes developed by the authors based on the data on the experience of using a digital mental health app.

- **Timing:** Daily schedule, exercise length
- **Stigma:** Family and friends, visibility, independence, framing, avoidance
- **Perception:** Trustworthiness, seriousness, skill versus quick fix
- **Congruity:** Design versus content, complexity, user journey
- **Usefulness:** Control, labelling, prompts, past experiences, valued features
- **User Experience:** Gamification, personalization

Results

Overview

One participant was not available for the interview, so our results included 13 interviews in total. Most participants were White and British (12/13, 92%), and 1 participant self-identified as being on Asian descent.

We labeled our main themes as follows: (1) timing, (2) stigma, (3) perception, (4) congruity, (5) usefulness, and (6) UX. Although our aim was to treat these categories as separate, they are nevertheless closely intertwined concepts with inevitable overlaps. The subthemes are referenced by their numbers in parentheses in Textbox 1.

Timing

All participants emphasized the importance of time constraints when engaging with the app. Participants often described their daily schedule as *busy* and *stressful* where they have to *get on with their work*. Under these circumstances, using the app often felt like an extra task where "you do have to make a conscious effort to go in" (Participant 5). This routine usually limited participants to only engage with the app at home, particularly before bed when there is nothing else left to do; however, this also led to its own set of issues. It was frequently mentioned that the app was *not being in sync* with this kind of schedule:

> *Generally, I was doing it more in the evenings and by then I couldn't, it was giving me like tasks to go out and go for a walk and things and I couldn't because it was dark outside.* [Participant 4]

The pressure of having a busy schedule also made exercise length an important question. In general, shorter and easily accessible exercises were more appealing, where users could simply log in and *check on themselves* by noting down their mood or completing a quick task. In this context, routine mood screening measures before accessing any particular exercise and longer tasks were usually seen as barriers to engagement. Participants were already conscious of these time restraints even before deciding to log in and knowing that they had enough time to complete a task was seen as an important determinant of engagement:

> *I didn't want to go on the app for a bit because I just thought 'this is going to take ages now, as if I knew I could do one that literally lasts like 4/5 minutes then I'd just click on that.* [Participant 6]

Stigma

Mental health–related stigma was one of the most frequent themes in our data, and it is likely to have a significant impact on engagement. Stigma primarily emerged through labels such as *weird*, *uncomfortable*, *private matter*, and *ashamed*, which referred to feelings of embarrassment and vulnerability associated with using a mental health app. Even though these issues were common to all, only 1 participant mentioned encountering any negative remark:

> *There was one goal to go out for a walk and going out randomly for a walk is a bit weird so you know I was explaining it to them and they found it a bit silly at first but then as I was going through it with them and explaining how it worked they found it more interesting...I mean there was one member of my family that still thought of it as still a bit gimmicky.* [Participant 4]

As this negative comment was unique in our data, we saw stigma as already internalized, which was brought to the surface by certain situations when using the app.

Most frequently, users were cautious about openly using and/or talking about the app among their family and friends and in public, such as on public transport where the *risk* of being seen

XSL•FO
RenderX

and labeled was highest. The presence of other people was also a barrier to performing various exercises, such as deep breathing or closing eyes, as these activities were seen as uncomfortable in public. In general, mental health was thought of as a *private matter* that belonged *behind a closed door.*

Thinking of mental health as a private issue also made the app more appealing in other ways. Participants perceived that it gave them more control over their issues without having to rely on other people:

> *I know a lot of teenagers who maybe wouldn't want to go to a counsellor or would be ashamed of going to a counsellor and in this way you're kind of helping yourself in your own way.* [Participant 12]

Participants also wanted to avoid being seen as a *downer* who complains about being stressed and preferred relying on an app rather than risking social rejection.

Participants were also concerned about their self-perception when using the app. Many of them took issue with the term *mental health app* and suggested other labels with less *loaded* connotations. Good examples of such alternative terms were stress reduction or well-being (app):

> *I wouldn't say I have something as like a mental health issue. I think it is more just being stressed, I think people wouldn't look at it the same way.* [Participant 13]

Some participants felt that having any issues is in itself problematic and question the utility of engaging with them:

> *I thought, like, sometimes stuff you didn't really realise you'd thought about was there and it's a bit depressing to think about that.* [Participant 6]

Perception

Engagement was also closely tied to the way participants perceived the app, most frequently, to whether it was perceived as trustworthy. However, it was not always clear what trustworthiness meant. Some participants defined it as *professional*:

> *It definitely seemed quite slick and professional rather than a new app. It seemed like it had a backing to it.* [Participant 6]

Although others described it in contrast to other apps on the market:

> *It didn't say from the very start that this is some meditation type hippie app...because you know there's other apps out there that come across from the very start as that kind of thing. And you just think 'they're not going to work.'* [Participant 7]

Participants also highlighted a fine middle ground between being too serious on one hand and childish on the other. Most of them picked up on *playful* design elements, which were generally considered childish, even though the app was designed with adults in mind. One participant even reflected on this sensitivity to being patronized:

> *Teenagers always get funny about things being childish. Especially like slightly younger 14-year olds. They wouldn't want to feel like it was for children at all.* [Participant 9]

Conversely, participants also did not like the *medical* label attached to the app, which may be perceived as too serious:

> *I think for lots of people who aren't necessarily that comfortable talking about mental health and all that would feel almost slightly uncomfortable by...when they are gonna download it and says it is medical. You don't like feeling like that; like they have got a problem.* [Participant 8]

Following design, a common element was the way participants viewed the role that an app like this should play in their daily life. When asked about how they would describe it to a friend, a "use it only when you feel stressed" approach was overwhelmingly popular:

> *I described it as like a self-help, meditation app that you could use in stressful situations.* [Participant 1]

Although the app itself was created with structured skill building in mind, the default perception was that it should be used as a quick check-in tool or quick intervention when one is anxious or needs immediate support.

Congruity

Congruity describes some of the key areas where the design of the app (both UX and UI) proved to be confusing or working against its intended purpose. A common experience that seemed off-putting to users was when the design of various features was not in line with its actual content. The most frequently mentioned examples were the relaxation exercises; a participant illustrated this as follows:

> *In comparison to the really slow breathing and stuff, a fast-moving thing (background) just seemed really big at the time.* [Participant 6]

Along the same lines, participants also found looking at a screen problematic right before going to sleep—one of the most popular times to use the app—as its negative effects may do more harm than good and missed the option of a voice-only session for this situation.

In addition, on the UX side, participants preferred more guidance from the app, as without it—or without preexisting knowledge of mental health and therapeutic techniques—some sections of the app proved too complex:

> *I found the meditation section kind of overwhelming at first. Because there are so many options. And I found I didn't know quite which one to start with and then carry on with.* [Participant 12]

These types of problems can be overcome by clear signposting and well-designed user journeys; however, for some participants, this aspect has proved confusing as well:

> *...when you log in it's not as straightforward to follow the instructions or like when you first log into it.* [Participant 1]

XSL•FO
RenderX

Usefulness

Going beyond content and design, the usefulness theme refers to common patterns in participants' subjective experiences that either contributed or hindered effective engagement with the app. Many pointed to the positive effects of having a sense of control when it comes to mental health. Therefore, the fact that they were able to *do something* about their problems was in itself beneficial:

> *So, it felt like I was like actively going out and helping myself...rather than me just thinking "oh my goodness I'm just swamped," I'm actually making an effort to climb out of this mess.* [Participant 4]

Along the same lines, exercises were also most beneficial when users clearly understood their purpose and saw their progress:

> *Because you can kind of start to see if anything is like changing. If you know what it's trying to achieve and how it's trying to achieve that, then you can kind of measure how it's working. So, yeah I think it was useful to know what it's doing and what it's trying to help you think about.* [Participant 2]

Although having control over a problem was generally seen as a good thing, predefined situation labels were generally seen as a step too far. Common complaints were that these labels were often simply *wrong*, too *restrictive*, or *not accurate enough*, and although predefined labels were also useful for many, the participants suggested adding the option of having their own labels:

> *I did like the idea, but I felt if we were allowed to write our own responses instead of choosing one it would be more powerful.* [Participant 9]

Having the choice of selecting from predefined automatic thought labels also proved problematic for some users, especially if they were already in a negative mood:

> *...when I was in a bad mood it would ask me what I was thinking before I hit that bad mood and some of them were quite extreme. So, I latched onto the quite extreme version of what I was thinking.* [Participant 4]

Others also described similar experiences related to mood questionnaires, pointing to a situation where the app might even cause extra distress, by highlighting existing problems:

> *I might not even realise I'd been thinking about, like worrying about doing this and that. Not really realising, more subconscious. And then I click on it and I realise I actually have been doing that, and it makes it worse as it explicitly says it.* [Participant 6]

Some users also developed an association between their low mood and engaging with the app, which over time even amplified certain negative states:

> *...when you're in a bad mood and you just kinda don't say explicitly you don't necessarily stay in that mood, but when you click on the app, like a dark cloud or whatever it is, I then just feel like down.* [Participant 5]

Users' past experiences with certain exercises, such as meditation, proved to be a strong contributor to the kind of features they liked and visited frequently. This also underlines the appeal of activities where users know what and why they are doing. Beyond meditation exercises, the most valued features were reminders, relaxation or sleep exercises, the mood tracker, and the peer support or communication function.

UX Theme

This final theme highlights some of the emergent contradictions and alignments between current trends in app design and participants' subjective experiences in 2 key areas: gamification and personalization.

With regard to gamification, users saw games in this context as either neutral or counterproductive, although there was one participant who suggested that games should be more competitive rather than *calming*. Outside of the concrete games, users also did not make much use of other *soft* gamification features, such as the point system, whereby users were able to unlock new achievements and earn credits. This progression system, without real tangible rewards, did not make sense to the participants. Conversely, specific features of progression systems, with which participants were able to unlock new content based on their achievements, were generally seen as useful in creating a sense of progress and contributing to engagement:

> *The games in the app are definitely more fun and enjoyable because I can feel better benefit from them whereas with other puzzle games I just get frustrated if I can't do it. Whereas with on the app, I think, if I can't do it, I just keep going. There is no frustration in it.* [Participant 2]

Participants valued the ability to personalize their own experiences by setting their own background to creating their own character. They also wanted the opportunity to send personal messages to one another and did not see much value in sending or receiving prepopulated messages:

> *So it would be quite nice to have that feature where you could talk to people. Sort of like a news column, where you could put your queries or ask the app and you could put in your questions and they answer back and you can do it vice versa.* [Participant 10]

Discussion

Principal Findings

We aimed to understand adolescents' preferences when engaging with digital mental health interventions using semistructured interviews. Using these data, we developed 6 main themes and 20 subthemes that captured distinct aspects of adolescents' experiences with the Thrive app. Overall, most themes corresponded well with previous research in the field [7]; however, we also gained new insights for further exploration.

Adolescents saw the app as helpful. Many expressed that simply having access to an MHapp was reassuring in itself and proved beneficial in increasing their sense of self-reliance and containing negative moods. Time was a major factor in terms

XSL•FO
RenderX

of engagement. Most participants reported having a busy schedule and preferred using the app in the evenings before bed or just for quick *check-ins* during the day. Preferred features also corresponded with evening use and underscored results from previous studies highlighting the need for brief and easy-to-access features [23].

In terms of specific features and design, participants highlighted the importance of clarity in both their user journey and available information, which was also emphasized in previous studies [7,23]. Features seemed the most popular and engaging when either users already had experience with similar exercises, such as meditation, or when the purpose and goals of a given feature were clearly defined. We believe these preferences suggest that clarity of information likely has a direct impact on effectiveness by providing users with a clear mental map through which they can progress.

In contrast to previous studies that suggest reward and progression systems as a way of facilitating engagement [23,24], we found a clear distinction between helpful and unhelpful progression systems. In our sample, simple leveling or point systems were not meaningful to participants if they were not tied to tangible rewards. Similarly, games did not facilitate engagement as users deemed them irrelevant in this context. This was unexpected given the age of the population and previous research indicating good acceptability of games in this context [25]. Most participants endorsed the ability to unlock new features and levels that provided them with access to new content and exercises.

Stigma emerged as a hidden but important barrier to engagement. Participants often expressed embarrassment and feelings of weakness related to mental health. This concern, although not surprising given the associated stigma in the field [26], led many participants to use the app only in private and question how the app was branded and framed. Participants' reluctance to use the app when traveling is especially problematic when we consider that this could be the most obvious opportunity to find the time to engage with a mental health app. As a solution, less conspicuous designs were suggested.

Other barriers mentioned by participants also converge with those of previous research pointing out the stress-inducing potential of these apps [9]. Specifically, some participants saw prepopulated questions about their feelings as stress inducing. Others also felt that engaging with negative thoughts, rather than ignoring them or simply letting them pass, was not desirable. However, this seems to contradict previous findings that praise apps for their ability to increase emotional awareness [27,28]. This avoidance may hint at a difficulty in enduring or accepting any amount of distress. This may stem from an assumption that one should not have to encounter anything that may be distressing or difficult in life. We believe that this potential unwanted effect of MHapps deserves further investigation, as bringing negative thoughts into awareness is a fundamental aspect of cognitive behavioral therapy and accepting this initial *dip* in mood is a prerequisite for effective engagement [29].

Along the same lines, the majority of participants used the app only for checkups and as a quick stress reduction tool in acutely stressful situations. This pattern of engagement is in stark contrast to the skill-based approach to improving mental health where users engage with MHapps to acquire the skill of managing their distress on their own and self-soothe. Although acute stress reduction can indeed be beneficial in certain situations, excessive reliance on this may increase dependence on an external source of soothing rather than reducing it. We see this mismatch in user attitudes and intended use as one of the key points to address in the future if MHapps are indeed to become scalable additions to therapy.

Finally, although previous studies often emphasized the importance of customizability [3] and this also emerged as a desirable feature in our sample, given that time seems to be one of the most important factors influencing engagement, this finding should be approached with caution. Participants may express their desire for customizability in an interview situation, but in reality, they may respond better to clearer design, short interactions with the app, and easy access.

Limitations

Participants were only provided with access to the app for 1 week before the interview, which might have influenced the depth and detail of their experience and limited our conclusions. However, given that our main goal was to detect immediate, noticeable features that were liked or disliked, and that we reached theme saturation in our analysis, we are confident that this time frame was sufficient to address our question.

Another potential source of bias is the small sample size and the convenience sample. This may mean that our participants may be overly similar in certain ways, having similar backgrounds and preferences, which could distort our findings. Collecting background information from participants would have enabled us to reflect on this aspect in more detail. Moreover, this information would have also helped us to adequately contextualize our results. Although our focus of this qualitative paper was to get a *sense* of how adolescents may view a mental health app, more background on demographics and other participant characteristics would have strengthened the interpretation of our findings.

Finally, 3 participants were known to the experimenters. This may have caused some bias in the study as participants may have been interested in the app from the start and had an incentive to speak favorably. Although it was made clear that they would be anonymous and their responses recorded by a member of the research team not known to them, it is still reasonable to assume that there may be some bias regarding their experience of the app. However, we did not observe any differences between the responses of the participants along these lines.

Conclusions

We identified 6 main themes and 20 subthemes that captured distinct aspects of adolescents' experiences with the Thrive app. Overall, participants preferred convenient, clear, and easy-to-access features that they could use on an ad hoc basis. They saw the app as a potential way of calming themselves

XSL•FO
RenderX

when needed rather than as a tool to learn to manage their mental health. Contrary to our initial assumptions, some specific design elements, such as the outlook or games, were seen as childish or useless. Furthermore, stigma emerged as a major barrier when engaging with the app. This again points to a mismatch between the intended purpose of the app and the way it is perceived by its potential users. Future studies with larger sample sizes are needed to dissect the specific preferences of specific user groups and to outline concrete ways to address these barriers in the digital mental health field.

Acknowledgments

The study was funded by the University of Roehampton.

Authors' Contributions

ET and JASF initially devised the study. JMB and RR collected the data. RR, JASF, and KJ coded the data. RR, JASF, and KJ developed the themes. JMB and RR wrote the initial manuscript. RR, KJ, JMB, FO, SA, ET, and JASF critically revised the manuscript.

Conflicts of Interest

During the write-up of the results, RR and JMB were both employed by Thrive Therapeutic Software (during the data collection and the designing of the study, they were employed by the University of Roehampton). KJ, FO, and SA were also working for Thrive Therapeutic Software on a voluntary basis. At the time of the study, JASF was the chief executive officer of Thrive Therapeutic Software but was not involved in data collection.

Multimedia Appendix 1
Semistructured interview schedule.
[DOCX File , 9 KB - formative_v5i8e14004_app1.docx]

References

1. Depression and Other Common Mental Disorders: Global Health Estimates. World Health Organization. 2017. URL: https://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf?sequence=1&isAllowed=y [accessed 2020-09-01]
2. Mental Health Action Plan 2013-2020. World Health Organization. URL: https://apps.who.int/iris/handle/10665/89966 [accessed 2020-09-01]
3. Kenny R, Dooley B, Fitzgerald A. Developing mental health mobile apps: exploring adolescents' perspectives. Health Informatics J 2016 Jun;22(2):265-275 [FREE Full text] [doi: 10.1177/1460458214555041] [Medline: 25385165]
4. Whittaker R, Merry S, Stasiak K, McDowell H, Doherty I, Shepherd M, et al. MEMO--a mobile phone depression prevention intervention for adolescents: development process and postprogram findings on acceptability from a randomized controlled trial. J Med Internet Res 2012 Jan 24;14(1):e13 [FREE Full text] [doi: 10.2196/jmir.1857] [Medline: 22278284]
5. Reid SC, Kauer SD, Dudgeon P, Sanci LA, Shrier LA, Patton GC. A mobile phone program to track young people's experiences of mood, stress and coping. Development and testing of the mobiletype program. Soc Psychiatry Psychiatr Epidemiol 2009 Jun;44(6):501-507. [doi: 10.1007/s00127-008-0455-5] [Medline: 19011723]
6. Schmoldt A, Benthe HF, Haberland G. Digitoxin metabolism by rat liver microsomes. Biochem Pharmacol 1975 Sep 1;24(17):1639-1641 [FREE Full text] [Medline: 10]
7. Grist R, Porter J, Stallard P. Mental health mobile apps for preadolescents and adolescents: a systematic review. J Med Internet Res 2017 May 25;19(5):e176 [FREE Full text] [doi: 10.2196/jmir.7332] [Medline: 28546138]
8. Bakker D, Kazantzis N, Rickwood D, Rickard N. Mental health smartphone apps: review and evidence-based recommendations for future developments. JMIR Ment Health 2016 Mar 1;3(1):e7 [FREE Full text] [doi: 10.2196/mental.4984] [Medline: 26932350]
9. Tucker I, Goodings L. Managing stress through the stress free app: practices of self-care in digitally mediated spaces. Digit Health 2015;1:2055207615580741 [FREE Full text] [doi: 10.1177/2055207615580741] [Medline: 29942538]
10. Radovic A, McCarty CA, Katzman K, Richardson LP. Adolescents' perspectives on using technology for health: qualitative study. JMIR Pediatr Parent 2018;1(1):e2 [FREE Full text] [doi: 10.2196/pediatrics.8677] [Medline: 30740590]
11. Ranney ML, Choo EK, Spirito A, Mello MJ. Adolescents' preference for technology-based emergency department behavioral interventions. Pediatric Emerg Care 2013;29(4):475-481. [doi: 10.1097/pec.0b013e31828a322f]
12. Lui JH, Marcus DK, Barry CT. Evidence-based apps? A review of mental health mobile applications in a psychotherapy context. Prof Psychol Res Pr 2017 Jun;48(3):199-210. [doi: 10.1037/pro0000122]
13. Huguet A, Rao S, McGrath PJ, Wozney L, Wheaton M, Conrod J, et al. A systematic review of cognitive behavioral therapy and behavioral activation apps for depression. PLoS One 2016;11(5):e0154248 [FREE Full text] [doi: 10.1371/journal.pone.0154248] [Medline: 27135410]

XSL•FO
RenderX

14. Kertz SJ, MacLaren Kelly J, Stevens KT, Schrock M, Danitz SB. A review of free iphone applications designed to target anxiety and worry. J Technol Behav Sci 2017 Jan 11;2(2):61-70. [doi: 10.1007/s41347-016-0006-y]
15. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. Evid Based Ment Health 2018 Aug;21(3):116-119. [doi: 10.1136/eb-2018-102891] [Medline: 29871870]
16. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res 2011 Aug 5;13(3):e52 [FREE Full text] [doi: 10.2196/jmir.1772] [Medline: 21821503]
17. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. J Med Internet Res 2009 Apr 24;11(2):e13 [FREE Full text] [doi: 10.2196/jmir.1194] [Medline: 19403466]
18. Christoforou M, Sáez Fonseca JA, Tsakanikos E. Two novel cognitive behavioral therapy-based mobile apps for agoraphobia: randomized controlled trial. J Med Internet Res 2017 Nov 24;19(11):e398 [FREE Full text] [doi: 10.2196/jmir.7747] [Medline: 29175809]
19. McCloud T, Jones R, Lewis G, Bell V, Tsakanikos E. Effectiveness of a mobile app intervention for anxiety and depression symptoms in university students: randomized controlled trial. JMIR Mhealth Uhealth 2020 Jul 31;8(7):e15418 [FREE Full text] [doi: 10.2196/15418] [Medline: 32735221]
20. Guest G, Namey E, Chen M. A simple method to assess and report thematic saturation in qualitative research. PLoS One 2020;15(5):e0232076 [FREE Full text] [doi: 10.1371/journal.pone.0232076] [Medline: 32369511]
21. Guest G, Namey E, McKenna K. How many focus groups are enough? Building an evidence base for nonprobability sample sizes. Field Methods 2016 Jul 24;29(1):3-22. [doi: 10.1177/1525822x16639015]
22. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
23. Werner-Seidler A, O'Dea B, Shand F, Johnston L, Frayne A, Fogarty AS, et al. A smartphone app for adolescents with sleep disturbance: development of the sleep ninja. JMIR Ment Health 2017 Jul 28;4(3):e28 [FREE Full text] [doi: 10.2196/mental.7614] [Medline: 28754651]
24. Matthews M, Doherty G, Coyle D, Sharry J. Designing mobile applications to support mental health interventions. In: Lumsden J, editor. Handbook of Research on User Interface Design and Evaluation for Mobile Technology. Hershey, Pennsylvania: IGI Global; 2008:635-656.
25. Brown M, O'Neill N, van Woerden H, Eslambolchilar P, Jones M, John A. Gamification and adherence to web-based mental health interventions: a systematic review. JMIR Ment Health 2016 Aug 24;3(3):e39 [FREE Full text] [doi: 10.2196/mental.5710] [Medline: 27558893]
26. Finzen A. Stigma and stigmatization within and beyond psychiatry. In: Gaebel W, Roessler W, Sartorius N, editors. The Stigma of Mental Illness - End of the Story?. Switzerland: Springer International Publishing; 2017:29-42.
27. Kauer SD, Reid SC, Crooke AH, Khor A, Hearps SJ, Jorm AF, et al. Self-monitoring using mobile phones in the early stages of adolescent depression: randomized controlled trial. J Med Internet Res 2012 Jun 25;14(3):e67 [FREE Full text] [doi: 10.2196/jmir.1858] [Medline: 22732135]
28. Morris ME, Kathawala Q, Leen TK, Gorenstein EE, Guilak F, Labhard M, et al. Mobile therapy: case study evaluations of a cell phone application for emotional self-awareness. J Med Internet Res 2010 Apr 30;12(2):e10 [FREE Full text] [doi: 10.2196/jmir.1371] [Medline: 20439251]
29. Clark DA. Cognitive Restructuring. The Wiley Handbook of Cognitive Behavioral Therapy. 2013. URL: https://doi.org/10.1002/9781118528563.wbcbt02 [accessed 2020-09-01]

Abbreviations

MHapps: mental health apps
UI: user interface
UX: user experience

Edited by G Eysenbach; submitted 13.03.19; peer-reviewed by I Tucker, K Stawarz, M DeJonckheere, A Bergin; comments to author 03.10.19; revised version received 11.11.20; accepted 06.05.21; published 27.08.21.

Please cite as:
Ribanszki R, Saez Fonseca JA, Barnby JM, Jano K, Osmani F, Almasi S, Tsakanikos E
Preferences for Digital Smartphone Mental Health Apps Among Adolescents: Qualitative Interview Study
JMIR Form Res 2021;5(8):e14004
URL: https://formative.jmir.org/2021/8/e14004
doi:10.2196/14004
PMID:34128814

XSL•FO
RenderX

©Robert Ribanszki, Jose Andres Saez Fonseca, Joseph Matthew Barnby, Kimberly Jano, Fatima Osmani, Soma Almasi, Elias Tsakanikos. Originally published in JMIR Formative Research (https://formative.jmir.org), 27.08.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Analysis of Hospital Quality Measures and Web-Based Chargemasters, 2019: Cross-sectional Study

Kunal N Patel[1*], DrPH, MPH, MDP; Olena Mazurenko[2*], MD, PhD, MS; Eric Ford[3*], MPH, PhD

[1]Northern Illinois University, DeKalb, IL, United States
[2]Indiana University, Indianapolis, IN, United States
[3]University of Alabama at Birmingham, Birmingham, AL, United States
[*]all authors contributed equally

Corresponding Author:
Kunal N Patel, DrPH, MPH, MDP
Northern Illinois University
1425 West Lincoln Highway
DeKalb, IL, 60115
United States
Phone: 1 815 753 1891
Email: kpatel27@niu.edu

Abstract

Background: The federal health care price transparency regulation from 2019 is aimed at bending the health care cost curve by increasing the availability of hospital pricing information for the public.

Objective: This study aims to examine the associations between publicly reported diagnosis-related group chargemaster prices on the internet and quality measures, process indicators, and patient-reported experience measures.

Methods: In this cross-sectional study, we collected and analyzed a random 5.02% (212/4221) stratified sample of US hospital prices in 2019 using descriptive statistics and multivariate analysis.

Results: We found extreme price variation in *shoppable* services and significantly greater price variation for medical versus surgical services (P=.006). In addition, we found that quality indicators were positively associated with standard charges, such as mortality (β=.929; P<.001) and readmissions (β=.514; P<.001). Other quality indicators, such as the effectiveness of care (β=−.919; P<.001), efficient use of medical imaging (β=−.458; P=.001), and patient recommendation scores (β=−.414; P<.001), were negatively associated with standard charges.

Conclusions: We found that hospital chargemasters display wide variations in prices for medical services and procedures and match variations in quality measures. Further work is required to investigate 100% of US hospital prices posted publicly on the internet and their relationship with quality measures.

(JMIR Form Res 2021;5(8):e26887) doi:10.2196/26887

KEYWORDS

chargemaster; standard charge; price transparency; health care; diagnosis-related group; DRG; quality measures; the Centers for Medicare and Medicaid Services regulation; CMS

Introduction

Background

Increases in health care expenditures have persisted throughout the years in the United States despite policy efforts to *bend the curve*. According to the Centers for Medicare and Medicaid Services (CMS), the US health care spending in 2018 increased 4.6% from the previous year and totaled US $3.6 trillion [1]. A contributing factor to the rise in health care expenditures comes from the fact that hospitals do not compete on price in the same way other efficient product markets do (such as the e-commerce sector). Currently, there are differences in what is charged by health care systems compared with what is paid by consumers [2]. Subsequently, consumers are, in effect, *price takers* accepting the hospital charges negotiated with their insurer [3].

As a result, historically, consumers have not been as price-sensitive toward making health care decisions when compared with consumer decision-making behaviors commonly observed in other economic sectors. With the continual increase in US health care spending, a widely held view is that greater

XSL•FO
RenderX

consumer engagement in health care will help hold prices down. In turn, greater consumer engagement will slow down the sector's expansion rate if (and when) consumers place a substantial emphasis on making price-sensitive decisions using pricing transparency information [4,5]. To that end, the CMS have issued 2 regulations that require hospitals to increase their *price transparency* [6,7]. The first regulation required hospitals to disclose their diagnosis-related group (DRG) chargemasters on the web publicly in a machine-readable form (such as a *Microsoft Excel* file) starting in 2019. Releasing the DRG chargemasters on the internet was met with little resistance from hospitals, as the information did not compromise revealing negotiated hospital pricing strategies vis-à-vis third-party payors or competitors. Although there was little resistance to the first federal regulation, previous literature has shown an abundance of nonprice-transparent and noncompliant hospitals and hospitals with inaccessible pricing information [8-10].

Nonetheless, understanding newly available US chargemaster information is vital to patients because American patients are sent a medical bill after receiving treatment. A medical bill will contain the patient's portion owed of hospital standard charges for medical services and procedures that were delivered net of any contractual allowances and third-party payments. Therefore, standard charges are relevant to the consumer, either directly by influencing their purchase decisions before receiving medical care or indirectly when they receive a medical bill after treatment.

Objective

This study aims to assess the variability of publicly available DRG chargemaster data and its relation to quality measures, process indicators, and patient experience measures as a source of information for consumer quality assessment and price-sensitive decision-making purposes. The research benefits three audiences. For policy makers, this study provides an early assessment of the pricing transparency regulation's utility. For researchers, being able to collect and compare hospitals' pricing data is an important task if they are to inform policy maker efforts on controlling health care spending. In addition, researchers can inform the public at large and assist other stakeholders, such as nongovernmental organizations, in providing an analysis of pricing information found on chargemasters that is understandable. Finally, for health care administrators, this research can shed new light on the importance of presenting standard charges to the public in compliance with the law.

Methods

Procedures

We conducted a cross-sectional study of web-based publicly available hospital chargemasters from 2019. First, we assessed the descriptive statistics and coefficients of variation (CVs) to describe the standard charges grouped by the DRG code. We aimed to describe the full extent of price variability in hospital standard charges.

We then performed 2 median chi-square tests on standard charges and type of service (either medical or surgical). Median chi-square tests were performed because the standard charges were not normally distributed, that is, standard charges were skewed to the right. The first test was for average standard charge (either above the median or below the median) by the type of service (either medical or surgical). Similarly, the second test was for the CV (either above the median or below the median) by the type of service (either medical or surgical).

Next, we performed a log-linear, ordinary least squares regression model to fit the natural log-transformed standard charges on hospital characteristics. We log-transformed the dependent variable (standard charges) owing to the right-skewness and lack of normal distribution. We removed outliers with residuals of IQR 1.5 below the first quartile or IQR 1.5 above the third quartile. β coefficients, *P* values, and robust SEs were presented as predictors. Robust SEs were clustered on hospital to correct for related observations. All analyses were conducted using *Microsoft Excel* and *Stata/SE 15.1*. The institutional review board of the University of Alabama at Birmingham exempted this study.

Data Source

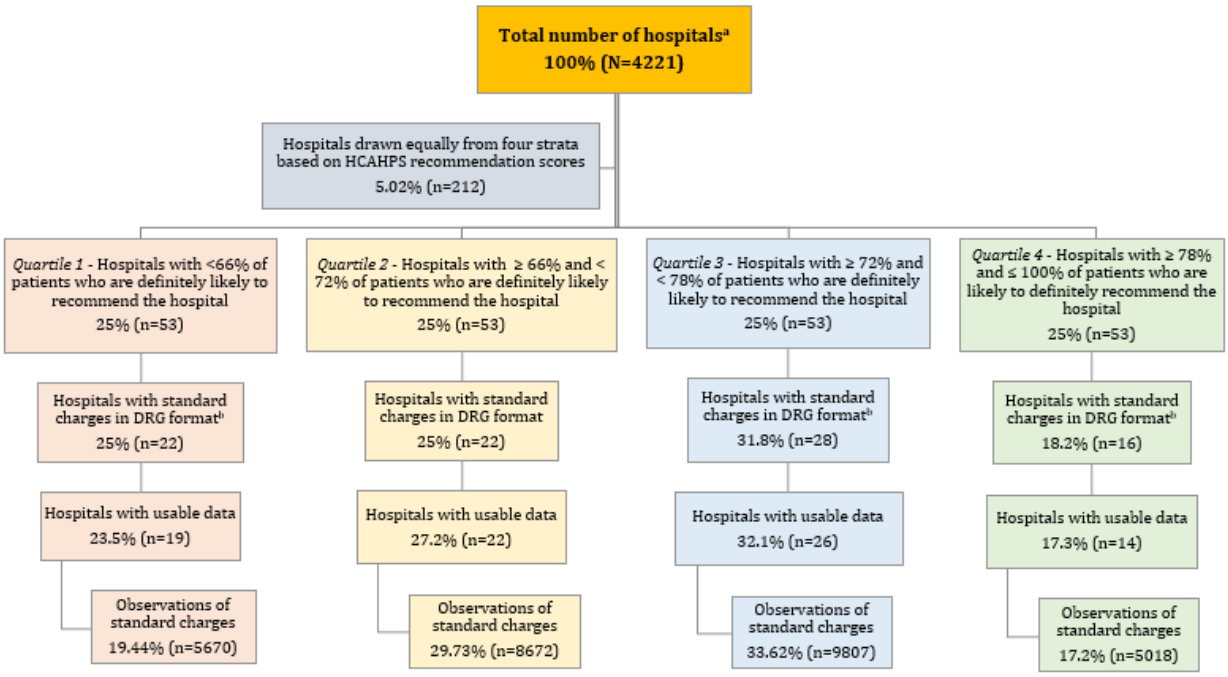
We retrieved chargemasters from hospital websites on the internet between August 25, 2019, and October 3, 2019, if they were formatted using DRG primary codes (eg, chargemasters in Healthcare Common Procedure Coding System or common procedural terminology primary code were excluded). In line with previous research, we obtained common hospital characteristic data from the Hospital Compare, CMS, American Hospital Association (AHA), and Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS).

Sampling Strategy

We constructed a random, stratified sample to assess US hospitals (Figure 1). It has been previously shown that hospital website quality is associated with HCAHPS recommendation scores [11]. Thus, to ensure an adequate variation of low- to high-quality websites, we stratified hospitals (n=4221) listed in HCAHPS data from October 1, 2017, to September 20, 2018, into 4 ranked quartiles based on the measure, "Patients who reported YES, they would definitely recommend the hospital." In total, 1.26% (53/4221) of hospitals were randomly selected from each of the 4 strata, representing a total of 5.02% (212/4221) of the hospitals in the HCAHPS data set. The sample size was restricted to maintain the feasibility of manual data collection and processing costs [12]. In sum, we had 29,167 observations of standard charges grouped by 81 different hospitals.

XSL•FO
RenderX

Figure 1. Data sampling strategy. [a]Total number of hospitals drawn from Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) survey from 3rd quarter of 2018. [b]Some hospitals provided data in an unusable format, such as in the All Patient Refined–diagnosis-related group coding format vs Medicare Severity–diagnosis-related group, providing maximum or minimum charges vs standard charges, etc. DRG: diagnosis-related group.



Predictors for Hospital Characteristics

Quality predictors included benchmark measures for the hospital's overall rating (hospital rating categories: 1 star, 2 stars, 3 stars, 4 stars, 5 stars, and missing), mortality rate, safety score, readmission rate, effectiveness of care score, efficient use of medical imaging score, and patient experience score. These measures and their categorical values (either below the national average, same as the national average, above the national average, or missing) were obtained using the Hospital General Information data set from the CMS. In addition, we included one additional patient experience measure: the likelihood of patients to recommend the hospital using the quartile categories described in the *Sampling Strategy* section (1=lowest quartile and 2, 3, and 4=highest quartile).

Controls included hospital ownership type (government—hospital district or authority, government—local, physician, proprietary, voluntary nonprofit—church, voluntary nonprofit—other, and voluntary nonprofit—private). Next, using data from the AHA Annual Survey, the hospital bed size (6-24 beds, 25-49 beds, 50-99 beds, 100-199 beds, 200-299 beds, 300-399 beds, 400-499 beds, and 500 or more beds) was controlled. Previous work has used the number of competitors in the market as a measure of competition (rather than the Herfindahl-Hirschman index) [13]. In line with these studies, we calculated a control measure for competition using the number of Medicare providers per 5-digit ZIP code (1=least and 2 and 3=most) from the HCAHPS data set. The DRG primary code was controlled using individual dummies for each of the DRG primary codes from the CMS data set for Medicare Severity–DRG version 36. Finally, 2 geographical control variables were included using the AHA Annual Survey data set. They were regions (New England, Mid Atlantic, South Atlantic, East North Central, East South Central, West North Central, West South Central, Mountain, and Pacific as defined by the AHA Annual Survey) and US states (individual dummies for each US state).

Results

Variance Analysis

CMS specifically defined 5 services using DRG primary codes to be *shoppable* in a forthcoming regulation on increasing health care price transparency (effective January 1, 2021). We present the price variability for these 5 *shoppable* medical services for our 5.02% (212/4221) sample of US hospitals in Table 1. The shoppable medical services included in our analysis were sorted from most to least variable, as measured by the CV. The maximum standard charge was frequently many orders of magnitude higher than the minimum. For example, the standard charge for DRG primary code 473 *cervical spinal fusion without comorbid conditions or major comorbid conditions or complications* had a mean (SD) value of US $89,302 (SD US $50,122), a CV of 0.561, and ranged from US $30,924 to US $249,283. The maximum standard charge for the procedure was over US $210,000 more than the minimum standard charge out of 44 hospitals that performed the service.

XSL•FO
RenderX

Table 1. Standard charges for Centers for Medicare and Medicaid Services–specified shoppable services[a].

Medicine and surgery services[b]	DRG[c] primary code	Hospitals observed (n=212), n (%)	Price variability (US $)		Coefficient of variation (SD divided by mean)
			Mean (SD)	Minimum-maximum, range	
Major joint replacement or reattachment of lower extremity without MCC[d]	470	75 (35.4)	68,329 (41,724)	26,401-224,052	0.611
Spinal fusion except cervical without MCC	460	50 (23.6)	123,744 (71,755)	30,995-427,374	0.580
Cervical spinal fusion without CC[e] or MCC	473	44 (20.8)	89,302 (50,122)	30,924-249,283	0.561
Cardiac valve and other major cardiothoracic procedures with cardiac catherization with MCC	216	28 (13.2)	430,274 (195,719)	139,460-912,194	0.455
Uterine and adnexa procedures for nonmalignancy without CC or MCC	743	62 (29.2)	41,338 (18,662)	11,863-87,981	0.451

[a]The table is sorted from most to least variable using the coefficient of variation.

[b]These are the only 5 selected services using the diagnosis-related group primary code (as opposed to the common procedural terminology or Healthcare Common Procedure Coding System) that the Centers for Medicare and Medicaid Services determined to include in the forthcoming regulation (effective date January 1, 2021), which mandates public disclosure of payer-specific negotiated charges, deidentified minimum and maximum negotiated charges, and discounted cash prices for at least 300 shoppable services, including 70 Centers for Medicare and Medicaid Services–specified shoppable services and 230 hospital-selected shoppable services.

[c]DRG: diagnosis-related group.

[d]MCC: major comorbid conditions or complications.

[e]CC: comorbid conditions.

Next, out of the set of 761 DRG primary codes, the data for the most and least variable services by type of service (either medical or surgical) with at least 30 observations are presented in Table 2. Most and least variable services were measured by the highest and lowest CVs, respectively. Noticeably, it appears that surgical procedures had higher means and lower CVs, which is investigated further in the *Standard Charge and Type of Service* section.

XSL•FO
RenderX

Table 2. Top 10 most and least variable services for diagnosis-related group primary codes[a].

Medicine and surgery services (rank)	DRG[b] primary code	Type	Hospitals (n=212), n (%)	Price variability (US $)			Coefficient of variation (SD divided by mean)	Magnitude of range (US $)
				Mean (SD)	Minimum-maximum, range			
Top 10 most variable services								
Normal newborn (1)	795	Medical	57 (26.9)	27,052 (167,415)	1005	1,268,646	6.189	1,267,641
Reticuloendothelial and immunity disorders with MCC[c] (2)	814	Medical	30 (14.2)	129,016 (376,201)	13,470	2,077,708	2.916	2,064,238
Skin ulcers with CC[d] (3)	593	Medical	43 (20.3)	44,730 (80,382)	6668	493,015	1.797	486,346
Other infectious and parasitic diseases diagnoses with MCC (4)	867	Medical	32 (15.1)	96,332 (170,037)	8608	962,984	1.765	954,377
Other respiratory system operating room procedures without CC or MCC (5)	168	Surgical	38 (17.9)	75,296 (109,621)	16,330	695,556	1.456	679,226
Neonates, died or transferred to another acute care facility (6)	789	Medical	50 (23.6)	91,278 (129,536)	1839	687,641	1.419	685,802
Other endocrine, nutritional, and metabolic operating room procedures with MCC (7)	628	Surgical	32 (15.1)	159,266 (209,529)	41,388	1,188,069	1.316	1,146,681
Other factors influencing health status (8)	951	Medical	52 (24.5)	19,934 (22,603)	32	109,262	1.134	109,230
Minor skin disorders without MCC (8)	607	Medical	57 (26.9)	28,153 (30,979)	5013	226,300	1.100	221,287
Depressive neuroses (10)	881	Medical	44 (20.8)	21,152 (21,902)	3144	140,972	1.035	137,829
Top 10 least variable services								
Percutaneous intracardiac procedures without MCC (1)	274	Surgical	36 (17)	112,404 (43,667)	46,738	255,453	0.388	208,715
Kidney and ureter procedures for neoplasm with CC (2)	657	Surgical	38 (17.9)	78,905 (31,783)	23,587	156,374	0.403	132,787
Other heart assist system implant (3)	215	Surgical	30 (14.2)	335,647 (139,272)	153,355	702,998	0.415	549,643
Ischemic stroke, precerebral occlusion, or transient ischemia with thrombolytic agent with CC (4)	62	Medical	39 (18.4)	84,117 (34,993)	27,457	187,137	0.416	159,680
Cardiac pacemaker revision except device replacement with CC (5)	261	Surgical	31 (14.6)	68,581 (30,005)	23,154	130,492	0.438	107,338
Vaginal delivery with sterilization and/or dilation and curettage (6)	767[e]	Surgical	50 (23.6)	24,912 (10,938)	8210	54,446	0.439	46,236
Cesarean section without CC or MCC (7)	766[e]	Surgical	56 (26.4)	24,106 (10,687)	9737	58,931	0.443	49,194
Disorders of the biliary tract without CC or MCC (8)	446	Medical	59 (27.8)	28,628 (12,805)	7936	72,899	0.447	64,963
Uterine and adnexa procedures for nonmalignancy without CC or MCC (9)	743	Surgical	62 (29.2)	41,338 (18,662)	11,863	87,981	0.451	76,118

XSL•FO
RenderX

Medicine and surgery services (rank)	DRG[b] primary code	Type	Hospitals (n=212), n (%)	Price variability (US $)			Coefficient of variation (SD divided by mean)	Magnitude of range (US $)
				Mean (SD)	Minimum-maximum, range			
Aortic and heart assist procedures except pulsation balloon with MCC (10)	268	Surgical	31 (14.6)	251,216 (113,721)	92,804	623,820	0.453	531,016

[a]Only services in the diagnosis-related group primary code with at least 30 observations are included. Most and least variable services are measured by the highest and lowest coefficients of variation, respectively.

[b]DRG: diagnosis-related group.

[c]MCC major comorbid conditions or complications.

[d]CC: comorbid conditions.

[e]Diagnosis-related group (DRG) codes 766 and 767 have been removed from Medicare Severity–DRG version 36.

Standard Charge and Type of Service

The relationship between standard charge and type of service (medical or surgical) was assessed for significant differences using 2 different median chi-square tests. The tables are presented in the top and bottom panels of Table 3. In both median chi-square tests, the cells represent the counts of individual DRG primary codes. The median chi-square test for average standard charge versus the type of service was significant (Pearson χ^2_1 [sample size=758]=284.1; P<.001). The observed number of average standard charges was significantly greater than the expected number for surgical services, with average standard charges greater than the median (observed=309 and expected=193). The median chi-square test for CV by type of service was also significant (Pearson χ^2_1 [sample size=758]=7.6; P=.006). However, in contrast to average standard charges, the observed number of CVs was significantly less than the expected number of CVs for surgical services, with CVs greater than the median. In summary, surgical services (as opposed to medical services) generally tended to have significantly more average standard charges and fewer CVs above the median.

Table 3. Contingency table for average standard charge versus the type of service and coefficient of variation for standard charge versus the type of service[a].

Standard Charge	Type of service						Total (n=758), n (%)
	Medical			Surgical			
	Observed (n=372), n (%)	Expected (n=372), n (%)	Chi-square contribution	Observed (n=386), n (%)	Expected (n=386), n (%)	Chi-square contribution	
Average							
Less than median	302 (79.7)	186 (50)	72.3	77 (20.3)	193 (50)	69.7	379 (50)
Greater than median	70 (18.5)	186 (50)	72.3	309 (81.5)	193 (50)	69.7	379 (50)
Coefficient of variation (SD divided by mean)							
Less than median	167 (44.1)	186 (50)	1.9	212 (55.9)	193 (50)	1.9	379 (50)
Greater than median	205 (54.1)	186 (50)	1.9	174 (45.9)	193 (50)	1.9	379 (50)

[a]Counts are individual diagnosis-related group primary codes, for example, 70 medical-type diagnosis-related group codes have averages greater than the median standard charge.

Hospital Characteristics of Standard Charges

We examined standard charges across hospital characteristics: hospital ownership, hospital rating, mortality, safety, readmission, effectiveness of care, patient experience, competition, efficient use of medical imaging, patient recommendation, region, bed size, US state, and DRG primary code (Table 4). Using multivariate regression modeling after removing outliers, we found that our model was able to explain nearly 90% of the variation in the randomized, stratified sample of standard charges in 2019 using categorical variables for the predictors (Table 5).

XSL•FO
RenderX

Table 4. Hospital characteristics of included standard charges (N=29,167)[a].

Variables	Hospital, n (%)
Hospital ownership	
Government	
Hospital district or authority	1534 (5.26)
Local	1197 (4.1)
Physician	150 (0.51)
Proprietary	9021 (30.93)
Voluntary nonprofit	
Church	1390 (4.77)
Other	4102 (14.06)
Private	11,773 (40.36)
Hospital rating	
1 star (worst)	2336 (8.01)
2 stars	9363 (32.1)
3 stars	6003 (20.58)
4 stars	9459 (32.43)
5 stars (best)	1903 (6.52)
Missing	103 (0.35)
Mortality	
Below the national average	4474 (15.34)
Same as the national average	17,696 (60.67)
Above the national average	5926 (20.32)
Missing	1071 (3.67)
Safety	
Below the national average	8619 (29.55)
Same as the national average	5695 (19.53)
Above the national average	13,633 (46.74)
Missing	1220 (4.18)
Readmission	
Below the national average	15,237 (52.24)
Same as the national average	2975 (10.20)
Above the national average	10,388 (35.62)
Missing	567 (1.94)
Effectiveness of care	
Below the national average	4383 (15.03)
Same as the national average	23,394 (80.21)
Above the national average	1287 (4.41)
Missing	103 (0.35)
Patient experience	
Below the national average	11,932 (40.91)
Same as the national average	8296 (28.44)
Above the national average	8537 (29.27)
Missing	402 (1.38)

XSL•FO
RenderX

Variables	Hospital, n (%)
Competition	
1 (least)	22,625 (77.57)
2	4360 (14.95)
3 (most)	2182 (7.48)
Efficient use of medical imaging	
Below the national average	4936 (16.92)
Same as the national average	16,343 (56.03)
Above the national average	6028 (20.67)
Missing	1860 (6.38)
Patient recommendation	
1 (lowest quartile)	5670 (19.44)
2	8672 (29.73)
3	9807 (33.62)
4 (highest quartile)	5018 (17.2)
Region	
New England	1384 (4.75)
Mid Atlantic	2845 (9.75)
South Atlantic	4616 (15.83)
East North Central	4359 (14.94)
East South Central	2713 (9.3)
West North Central	1083 (3.71)
West South Central	4797 (16.45)
Mountain	2766 (9.48)
Pacific	4604 (15.78)
Number of beds	
6-24	329 (1.13)
25-49	1017 (3.49)
50-99	3049 (10.45)
100-199	6017 (20.63)
200-299	4650 (15.94)
300-399	5249 (17.99)
400-499	4057 (13.91)
500 or more	4799 (16.45)

[a]The table does not show values for each category for US state and diagnosis-related group primary code.

XSL•FO
RenderX

Table 5. Regression results for standard charges in a sample of US hospitals (N=27,530)[a].

Variables	β (robust SE)	*P* value
Hospital ownership		
Government		
Hospital district or authority	.856 (0.104)	<.001
Local	.499 (0.159)	.002
Physician	−1.879 (0.257)	<.001
Proprietary	.828 (0.177)	<.001
Voluntary nonprofit		
Church	1.008 (0.089)	<.001
Other	Reference	N/A[b]
Private	.172 (0.111)	.13
Hospital rating		
1 star (worst)	.499 (0.212)	.02
2 star	.422 (0.076)	<.001
3 star	Reference	N/A
4 star	.133 (0.185)	.47
5 star (best)	.109 (0.207)	.60
Missing	−.042 (0.756)	.96
Mortality		
Below the national average	.514 (0.051)	<.001
Same as the national average	Reference	N/A
Above the national average	.244 (0.125)	.05
Missing	.961 (0.188)	<.001
Safety		
Below the national average	−.085 (0.046)	.06
Same as the national average	Reference	N/A
Above the national average	.102 (0.063)	.11
Missing	−.007 (0.187)	.97
Readmission		
Below the national average	.929 (0.158)	<.001
Same as the national average	Reference	N/A
Above the national average	.578 (0.147)	<.001
Missing	−1.023 (0.325)	.002
Patient experience		
Below the national average	−.046 (0.055)	.41
Same as the national average	Reference	N/A
Above the national average	.294 (0.075)	<.001
Missing	3.707 (0.346)	<.001
Effectiveness of care		
Below the national average	.081 (0.102)	.43
Same as the national average	Reference	N/A
Above the national average	−.919 (0.119)	<.001

XSL•FO
RenderX

Variables	β (robust SE)	*P* value
Efficient use of medical imaging		
Below the national average	−.277 (0.055)	<.001
Same as the national average	Reference	N/A
Above the national average	−.458 (0.128)	.001
Missing	.321 (0.084)	<.001
Patient recommendation		
1 (lowest quartile)	Reference	N/A
2	−.236 (0.066)	.001
3	−.169 (0.074)	.03
4 (highest quartile)	−.414 (0.072)	<.001
Competition		
1 (least)	Reference	N/A
2	.546 (0.065)	<.001
3 (most)	.552 (0.119)	<.001
Number of beds		
6-24	Reference	N/A
25-49	2.063 (0.287)	<.001
50-99	1.518 (0.261)	<.001
100-199	1.832 (0.32)	<.001
200-299	2.63 (0.341)	<.001
300-399	2.082 (0.286)	<.001
400-499	2.013 (0.294)	<.001
500 or more	2.075 (0.262)	<.001
Constant	10.004 (0.418)	<.001

[a]The table shows the results for a log-linear regression using the natural log function to transform the dependent variable, that is, standard charge. Other covariates for individual state code and diagnosis-related group Code Dummy Variables are not shown. Outliers with residuals IQR 1.5 below the first quartile or IQR 1.5 above the third quartile are omitted. Robust SEs are clustered on hospital to correct for related observations. Standard charges are in dollars. R^2=0.8955 and number of observations=27,530.

[b]N/A: not applicable.

All quality indicators were associated with standard charges at the statistically significant α=.05 level, except for the patient safety indicator. The 2 quality indicators associated with the largest significant *increases* in standard charges were below the national average mortality rate (β=.929; *P*<.001) and below the national average readmission rate (β=.514; *P*<.001); they were associated with 153% and 67% significantly higher standard charges on average, respectively, compared with the national average groups, holding other factors constant. On the contrary, the three quality indicators associated with the largest significant *decreases* in standard charges in our study were above the national average effectiveness of care (β=−.919; *P*<.001), above the national average efficient use of medical imaging (β=−.458; *P*=.001), and the highest quartile patient recommendation scores (β=−.414; *P*<.001); they were associated with 60%, 37%, and 34% significantly lower standard charges on average, respectively, than those of the reference groups, holding other factors constant.

Finally, for Table 5, please note that the interpretations of β coefficients were on average, while holding all else constant and using natural log-transformed standard charges as the outcome variable. In addition, the constant and β coefficients for the missing categories in the relevant variables were not described but can be found in the table. Robust SEs were clustered on hospital to correct for related observations. Outliers were removed as described in the *Methods* section, leaving 27,530 observations in the regression model. Overall, a large amount of variation in standard charges was explained by our regression model (R^2=89.55%).

Discussion

Principal Findings

Wide differences exist between hospital billed charges and the amount of money that hospitals expect to receive for services [14]. Our analysis found that chargemaster DRG prices on the internet varied greatly between facilities. At a minimum, the

XSL•FO
RenderX

web-based chargemaster data do not reflect the marginal cost of performing 1 instance of a procedure. Different hospitals have widely varying fixed costs that may drive the variance to some extent, but this is not sufficient to explain the differences observed [15]. A more plausible explanation is that there are systematic differences in the business strategies related to chargemaster construction, as found in our analysis.

Reviewing Table 2, even the procedures with *relatively* low SDs and CVs had wide enough ranges to indicate that there is little to no relation to the chargemaster's rates and actual underlying costs. For example, a previous study on Ohio state data from 2007 to 2012 showed that a hospital with the highest median charge for a normal newborn delivery (DRG primary code: 795) could be nearly 4 times as costly as the hospital with the lowest median charge despite no differences in length of stay (which typically is 2 days) [16]. We found even further drastic differences in our data set of the standard charge of a normal newborn delivery on a national level when compared with this study, where the maximum standard charge for the procedure was more than 1250 times greater than the minimum standard charge. Furthermore, our finding for the *Normal Delivery of a Newborn* service having the largest variation among our data was unusual for three reasons. First, the mean standard charge was relatively small, which usually leads to lower variances. Next, the upper bound of US $1,268,646 defied any reasonable expectations for this service. Finally, the minimum rate, US $1005, also defied logic. Even an uncomplicated delivery typically involves a 2-day stay with a per diem above US $1400, which would total more than US $2800 for the charge [13]. Additional examples of standard charges with wide ranges for the exact same service are commonly found in the literature [17-19].

Thereafter, we sought to test the differences in variability in the type of service (either medical or surgical). The estimated CVs for surgical-type DRGs were significantly smaller than those for medical DRGs using standard charge data for Maryland between 1979 and 1981 [20]. However, as the DRG patient diagnosis classifications are refined overtime, variation among medical-type DRGs could potentially converge toward the more favorable lower levels of variation of surgical-type DRGs [20]. However, we found that after 4 decades of revisions to DRG codes, where the number of unique codes increased from ≥400 in the 1980s to ≥700 in the 2020s, medical-type DRGs still had more variability than surgical-type DRGs. Our results may indicate challenges, as the results show that it is still increasingly more difficult to predict medical-type standard charges that have more variability when compared with surgical services. As a result, health care providers and other stakeholders will have to work increasingly harder to assist consumers in making informed decisions, especially for medical services.

Afterward, we sought to understand whether the wide variances observed were systematically related to hospital characteristics for quality performance indicators. A number of hospital characteristics were shown to be significantly associated with standard charges, including physical characteristics such as bed size or ownership structure, geographical characteristics, controls for the service or procedure code, competition, and quality indicators (such as patient recommendation scores or readmission rates).

Overall, our results were largely consistent with those of a previous study that found that standard charges in hospital chargemasters were well predicted using hospital characteristics [21]. However, the previous study did not find sufficient evidence that hospitals with higher prices also provided a higher quality of care [21]. In contrast to this finding, we found 2 key quality characteristics to be positively and significantly associated with standard charges (when controlling for market competition, physical characteristics, geographical differences, and DRG primary code in the multivariate analysis): mortality rates and readmission rates. Furthermore, these quality indicators are consistent with economic theory, where higher quality goods and services demand a higher price in the competitive market [22].

On the other hand, there is not a singular positive or negative relationship between price and quality, and at times, price and quality can either have a positive or negative relationship [23]. We found 3 health care quality indicators with contradictory results to standard economic theory: effectiveness of care, efficient use of medical imaging, and patient recommendation scores. In other words, as quality increases, the standard charge decreases, which is a contradictory pricing behavior.

Complexities exist in modern health care, which causes gaps in the ability of health care systems to deliver consistent, effective, and efficient care [24]. Therefore, significant undertreatment and overtreatment occur [24]. A possible explanation for the relationship between higher quality effectiveness of care and efficient use of medical imaging being associated with decreases in standard charges is that they lower waste, and thus, they reduce standard charges. Finally, a potential reason for higher patient recommendation scores being associated with lower standard charges is that the hospital may benefit from increased volume (or demand owing to more patient referrals) and, in turn, from economies of scale.

At this juncture, it is important to digress from the first phase of health price transparency regulation and discuss the implications of the second phase briefly to shed some light on other implications of this study in the context of present health care systems and policies. Although hospitals provide chargemaster data, the *standard charges* rarely provide information for patients to make informed health care decisions [25]. As a large number of patients are insured, they are more interested in cost sharing information and specific insurer-negotiated pricing rather than standard charges for health care services. Therefore, the second round of CMS transparency regulations more broadly requires hospitals to disclose the rates they have negotiated with third-party payers for service bundles starting in 2021 [7], including the following:

- Gross charge: the charge for an individual item or service that is reflected on a hospital's chargemaster, absent any discounts
- Discounted cash price: the charge that applies to an individual who pays cash or cash equivalent for a hospital item or service

XSL•FO
RenderX

- Payer-specific negotiated charge: the charge that a hospital has negotiated with a third-party payer for an item or service
- Deidentified minimum negotiated charges: the lowest charge that a hospital has negotiated with all third-party payers for an item or service
- Deidentified maximum negotiated charges: the highest charge that a hospital has negotiated with all third-party payers for an item or service.

Patients may use this additional information in 2021 to more accurately price-shop, insurers may use this information to bargain for better reimbursement rates, and other facilities may use this information to alter their pricing strategies and compete more effectively in the more transparent health care market. Therefore, this information is closely guarded by health plans [26]. Thus, hospitals, insurers, lobbying groups, and other stakeholders oppose this regulation because the negotiated prices have immeasurable proprietary strategic value, and disclosure thereof will have far-reaching implications on both price and quality competition. It remains to be seen if health systems will be able to block or alter the second phase of price transparency regulations before the scheduled implementation in 2021.

Limitations

While conducting this study on health care price transparency, there are 2 important limitations that need to be discussed. First, we did not analyze pricing information from other coding systems, such as common procedural terminology, Healthcare Common Procedure Coding System, or other proprietary formats. Some hospitals published chargemasters using other codes that were not mandated. Thus, the study results can only be generalized to the extent that DRG codes bundle services together correctly and correspond accurately to services rendered for patients. Some of these other coding systems rely on billing specialists to itemize services rendered, and they may or may not result in more accurate pricing, which could be higher or lower on average when compared with DRG-coded charges we analyzed in this study. However, the DRG coding system is one of the most widely used systems for preparing patient bills, and the results of this study are directly applicable to this most commonly used hospital pricing system in the United States.

Second, we did not follow up, investigate, or verify individual observations of standard charges. It is possible (and quite likely) that hospital chargemasters unintentionally contain outdated, erroneous, or inaccurate standard charges. These mistakes may have been published on the web for the public unbeknownst to hospital administrators. We mitigated these effects as much as possible by using statistical techniques where appropriate, such as analyzing median values and removing outliers.

Conclusions

Patients are not solely influenced by costs when making health care decisions; they base their decisions on several factors, including the opinions and information supplied by their health care providers and insurers. Moreover, previous literature has shown that patients just do not want to be a cog in the health care system, but in reality, they want to share in the decision-making processes regarding where to seek treatment with their health care providers [27,28]. Such health care–related decisions are commonly determined based on the quality of available medicals goods and services at a particular facility or by a specific provider. Therefore, patient decisions to seek treatment are being determined jointly by providers and consumers using both clinical quality and out-of-pocket cost information.

In summary, the results of this cross-sectional study, which analyzed the pricing behavior at hospitals in the first phase of the price transparency regulations, draw attention to the fact that policy makers, researchers, and health care administrators as well as, ultimately, consumers all need to be vigilant about health care price transparency and its relation to quality measures. There was extreme variation in *shoppable* services. Findings unearthed in this study include: one of the most commonly performed services (normal newborn delivery) had the most variation, significantly larger variation existed in medical services than surgical services, and quality variables were associated either positively or negatively with standard charges. It is ever more important for all the parties involved, such as researchers, policy makers, and health care administrators, to act in good faith and make the information as user-friendly and accessible as possible as well as use this information to the highest, fullest potential—bending the health care cost curve.

Future Directions

It is crucial for researchers, policy makers, and health care administrators to work together to design a holistic registry or database system to document these chargemasters. This study has demonstrated the potential value of such information using publicly available chargemaster data on the internet from a cross-sectional random, stratified sample of 5.02% of the US hospitals. This process can be scaled up to collect, clean, and document chargemasters for all US hospitals multiple times per year, such as quarterly or semiannually.

Conflicts of Interest

None declared.

References

1. Historical. Centers for Medicare & Medicaid Services (CMS). URL: https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NationalHealthAccountsHistorical [accessed 2021-08-03]
2. Babcock P. Paralyzed by prices: an analysis of price theory within the context of health care. Linacre Q 2019 Feb 24;86(1):89-102 [FREE Full text] [doi: 10.1177/0024363919837861] [Medline: 32431392]

XSL•FO
RenderX

3. Malakhov S. Moral hazard, optimal healthcare-seeking behavior, and competitive equilibrium. Expert J Econ 2017;5(2):71-79 [FREE Full text]
4. Kullgren JT, Duey KA, Werner RM. A census of state health care price transparency websites. J Am Med Assoc 2013 Jun 19;309(23):2437-2438. [doi: 10.1001/jama.2013.6557] [Medline: 23780454]
5. Gusmano MK, Maschke KJ, Solomon MZ. Patient-centered care, yes; patients as consumers, no. Health Aff (Millwood) 2019 Mar;38(3):368-373. [doi: 10.1377/hlthaff.2018.05019] [Medline: 30830817]
6. FY 2019 LTCH PPS final rule. Centers for Medicare & Medicaid Services (CMS). URL: https://www.cms.gov/newsroom/fact-sheets/fiscal-year-fy-2019-medicare-hospital-inpatient-prospective-payment-system-ipps-and-long-term-acute-0 [accessed 2021-07-29]
7. CY 2020 Hospital Outpatient Prospective Payment System (OPPS) policy changes: hospital price transparency requirements. Centers for Medicare & Medicaid Services (CMS). 2019. URL: https://www.cms.gov/newsroom/fact-sheets/cy-2020-hospital-outpatient-prospective-payment-system-opps-policy-changes-hospital-price [accessed 2021-07-29]
8. Mullens CL, Hernandez JA, Anderson ED, Allen L. Just because (most) hospitals are publishing charges does not mean prices are more transparent. JMIR Med Inform 2020 Feb 07;8(2):e14436 [FREE Full text] [doi: 10.2196/14436] [Medline: 32027308]
9. Lu A, Chen E, Vutam E, Brandt J, Sadda P. Price transparency implementation: accessibility of hospital chargemasters and variation in hospital pricing after CMS mandate. Healthc (Amst) 2020 Sep;8(3):100443. [doi: 10.1016/j.hjdsi.2020.100443] [Medline: 32919582]
10. Patel K, Rucks A, Ford E. Consumers' exposure to price transparency: compliance testing and sentiment analysis of US hospitals during 2019. In: Hefner JL, Al-Amin M, Huerta TR, Aldrich AM, Griesenbrock TE, editors. Transforming Health Care (Advances in Health Care Management, Vol. 19). Bingley: Emerald Publishing Limited; 2020:99-114. [doi: https://doi.org/10.1108/S1474-823120200000019010]
11. Ford EW, Huerta TR, Diana ML, Kazley AS, Menachemi N. Patient satisfaction scores and their relationship to hospital website quality measures. Health Mark Q 2013 Oct;30(4):334-348. [doi: 10.1080/07359683.2013.844041] [Medline: 24308412]
12. Zuckerman S, Hadley J, Iezzoni L. Measuring hospital efficiency with frontier cost functions. J Health Econ 1994 Oct;13(3):255-280. [doi: 10.1016/0167-6296(94)90027-2]
13. Ghiasi A, Zengul FD, Ozaydin B, Oner N, Breland BK. The impact of hospital competition on strategies and outcomes of hospitals: a systematic review of the U.S. Hospitals 1996-2016. J Healthc Fin 2017;44(2):142 [FREE Full text]
14. Muhlestein D. What types of hospitals have high charge-to-reimbursement ratios? Health Affairs Blog. 2013. URL: https://www.healthaffairs.org/do/10.1377/hblog20130715.033072/full/ [accessed 2021-07-29]
15. Boncheck L. Will more transparency help us lower the cost of health care? J Lancaster Gen Hosp 2019;14(2):33-34 [FREE Full text]
16. Hall ES, Lang MD, Levin L, Narendran V. Variation in neonatal inpatient charges at the state and local level. Hosp Top 2015 Jul 17;93(2):27-35. [doi: 10.1080/00185868.2015.1052273] [Medline: 26185931]
17. Agarwal A, Dayal A, Kircher SM, Chen RC, Royce TJ. Analysis of price transparency via National Cancer Institute-designated cancer centers' chargemasters for prostate cancer radiation therapy. JAMA Oncol 2020 Mar 01;6(3):409-412 [FREE Full text] [doi: 10.1001/jamaoncol.2019.5690] [Medline: 31944219]
18. Sprehe MR, Johnson R. Using chargemaster data to understand childhood leukemia costs. Blood 2016;128(22):5938. [doi: 10.1182/blood.V128.22.5938.5938]
19. White C, Whaley C. Prices paid to hospitals by private health plans are high relative to medicare and vary widely: findings from an employer-led transparency initiative. RAND Corporation 2019:1-60 [FREE Full text] [doi: 10.7249/RR3033]
20. Frank RG, Lave JR. The psychiatric DRGs. Are they different? Med Care 1985 Oct;23(10):1148-1155. [doi: 10.1097/00005650-198510000-00003] [Medline: 3932788]
21. Batty M, Ippolito B. Mystery of the chargemaster: examining the role of hospital list prices in what patients actually pay. Health Aff (Millwood) 2017 Apr 01;36(4):689-696. [doi: 10.1377/hlthaff.2016.0986] [Medline: 28373335]
22. Robinson JC. Hospital quality competition and the economics of imperfect information. Milbank Q 1988;66(3):465. [doi: 10.2307/3349965]
23. Fitzgerald MP, Yencha C. A test of policy makers' formal and lay theories regarding health care prices. J Public Pol Mark 2018 Dec 31;38(1):3-18. [doi: 10.1177/0743915618820573]
24. Oakes AH, Patel MS. A nudge towards increased experimentation to more rapidly improve healthcare. BMJ Qual Saf 2020 Mar 19;29(3):179-181. [doi: 10.1136/bmjqs-2019-009948] [Medline: 31744847]
25. Glover M, Whorms D, Singh R, Almeida RR, Prabhakar AM, Saini S, et al. A radiology-focused analysis of transparency and usability of top U.S. hospitals' chargemasters. Acad Radiol 2020 Nov;27(11):1603-1607. [doi: 10.1016/j.acra.2019.11.021] [Medline: 32014405]
26. Koller CF, Khullar D. The commercial differential for hospital prices: responses from states and employers. J Am Med Assoc 2019 Aug 27;322(8):723-724. [doi: 10.1001/jama.2019.9275] [Medline: 31329210]
27. Coulter A, Entwistle V, Gilbert D. Sharing decisions with patients: is the information good enough? Br Med J 1999 Jan 30;318(7179):318-322 [FREE Full text] [doi: 10.1136/bmj.318.7179.318] [Medline: 9924064]

XSL•FO
RenderX

28. Fowler Jr FJ, Levin CA, Sepucha KR. Informing and involving patients to improve the quality of medical decisions. Health Aff (Millwood) 2011 Apr;30(4):699-706. [doi: 10.1377/hlthaff.2011.0003] [Medline: 21471491]

Abbreviations

AHA: American Hospital Association
CMS: Centers for Medicare and Medicaid Services
CV: coefficient of variation
DRG: diagnosis-related group
HCAHPS: Hospital Consumer Assessment of Healthcare Providers and Systems

Edited by G Eysenbach; submitted 01.01.21; peer-reviewed by C Mullens, G Deckard; comments to author 18.04.21; revised version received 21.04.21; accepted 31.05.21; published 19.08.21.

Please cite as:
Patel KN, Mazurenko O, Ford E
Analysis of Hospital Quality Measures and Web-Based Chargemasters, 2019: Cross-sectional Study
JMIR Form Res 2021;5(8):e26887
URL: https://formative.jmir.org/2021/8/e26887
doi:10.2196/26887
PMID:34420914

©Kunal N Patel, Olena Mazurenko, Eric Ford. Originally published in JMIR Formative Research (https://formative.jmir.org), 19.08.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Redesigning a Web-Based Stakeholder Consensus Meeting About Core Outcomes for Clinical Trials: Formative Feedback Study

Roulla Katiri[1,2,3*], MSc; Deborah A Hall[1,2,4*], PhD; Derek J Hoare[1,2,5], PhD; Kathryn Fackrell[1,2,6], PhD; Adele Horobin[1,5], PhD; Nóra Buggy[2], MA; Nicholas Hogan[2]; Pádraig T Kitterick[1,2,5], PhD; Core Rehabilitation Outcome Set For Single-Sided Deafness (CROSSSD) Initiative[2]

[1]Hearing Sciences, Mental Health and Clinical Neurosciences, School of Medicine, University of Nottingham, Nottingham, United Kingdom
[2]National Institute for Health Research (NIHR) Nottingham Biomedical Research Centre (BRC), Ropewalk House, Nottingham, United Kingdom
[3]Audiology Department, Mater Misericordiae University Hospital, Dublin, Ireland
[4]Department of Psychology, School of Social Sciences, Heriot-Watt University Malaysia, Putrajaya, Malaysia
[5]Nottingham University Hospitals NHS Trust, Queen's Medical Centre, Nottingham, United Kingdom
[6]Wessex Institute, University of Southampton, University Road, Southampton, United Kingdom
[*]these authors contributed equally

Corresponding Author:
Pádraig T Kitterick, PhD
National Institute for Health Research (NIHR) Nottingham Biomedical Research Centre (BRC)
Ropewalk House
113 The Ropewalk
Nottingham, NG1 5DU
United Kingdom
Phone: 44 1158232626
Fax: 44 1158232615
Email: padraig.kitterick@nottingham.ac.uk

Abstract

Background: Clinical trials that assess the benefits and harms of an intervention do so by measuring and reporting *outcomes*. Inconsistent selection and diversity in the choice of outcomes make it challenging to directly compare interventions. To achieve an agreed core set of outcomes, a consensus methodology is recommended, comprising a web-based Delphi survey and a face-to-face consensus meeting. However, UK government regulations to control the pandemic prohibited plans for a face-to-face consensus meeting as part of the Core Rehabilitation Outcome Set for Single-Sided Deafness (CROSSSD) study.

Objective: This study aims to evaluate the modifications made by the CROSSSD study team to achieve consensus using web-based methods, but with minimal deviation from the original study protocol.

Methods: The study team worked with health care users and professionals to translate the planned face-to-face consensus meeting in a web-based format, preserving the key elements of the nominal group technique. A follow-up survey gathered evaluation feedback on the experiences of the 22 participating members. Feedback covered premeeting preparation, the process of facilitated discussions and voting, ability to contribute, and perceived fairness of the outcome.

Results: Overall, 98% (53/54) of feedback responses agreed or strongly agreed with the statements given, indicating that the web-based meeting achieved its original goals of open discussion, debate, and voting to agree with a core outcome set for single-sided deafness. Hearing-impaired participants were fully engaged, but there were some methodological challenges. For the participants, challenges included building rapport, understanding, and delivering the tasks in hand. For the study team, challenges included the need for thorough preparation and management of the unpredictability of tasks on the day.

Conclusions: Sharing our experiences and lessons learned can benefit future core outcome set developers. Overcoming the challenges of delivering a web-based consensus exercise in the face of the pandemic can be applied more generally to maximize inclusiveness, enhance geographical access, and reduce research costs.

(JMIR Form Res 2021;5(8):e28878) doi:10.2196/28878

XSL•FO
RenderX

KEYWORDS

COVID-19; nominal group technique; formative research; patient participation; persons with hearing impairments; mobile phone

Introduction

Background

When choosing a treatment for a disease or disorder, health care users, health care professionals, and other stakeholders need evidence of the benefits and harms of the treatments. Clinical trialists gather evidence by comparing and contrasting the benefits and harms (outcomes) of medical, surgical, or behavioral interventions. However, clinical trials evaluating interventions often measure and report different outcomes [1], making it challenging to synthesize evidence to inform recommendations on clinical management.

To address the inconsistency of outcome selection, clinical trialists recommend developing a *core outcome set* (COS). A COS prescribes the minimum set of outcomes that should be measured and reported when testing an intervention for a given health condition. The Core Outcome Measures in Effectiveness Trials (COMET) initiative has published a handbook to promote good practice in COS development methods [2]. The conventional process involves structured communication with patients and clinicians using a Delphi survey administered as a questionnaire [3], followed by a smaller scale consensus meeting [2]. Although the questionnaire can be administered on the internet, the consensus meeting is typically face-to-face [2,4], consistent with other applications of the nominal group technique (NGT) in the context of health care research [5].

Objectives

In line with the standard process, our project team (Core Rehabilitation Outcome Set for Single-Sided Deafness [CROSSSD]) had planned a web-based Delphi survey followed by a face-to-face consensus meeting [6]. However, we had to revise these meeting plans to comply with the travel and physical distancing restrictions imposed by the UK government in 2020 in response to the COVID-19 pandemic. There is limited information about web-based qualitative data gathering from groups in health care [7-10], and its evaluation from the participant perspective appears to be somewhat minimal [11]. Given that CROSSSD is about people with single-sided deafness (SSD), the web-based methods adopted had to be accessible to people with possible communication difficulties and suitable for data gathering, adding an extra layer of considerations.

The primary aims of this study are (1) to describe how we redesigned the consensus meeting format from face-to-face to web-based and (2) to evaluate stakeholder experiences.

Methods

Participants

A face-to-face consensus meeting was organized as per protocol [6] to take place in London, United Kingdom, on March 19, 2020. A total of 22 participants were invited (6/22, 27% health care users; 8/22, 36% health care professionals; 2/22, 9% public research partners, who had first-hand or lived experience of SSD; 1/22, 5% patient involvement manager; 2/22, 9% facilitators; 2/22, 9% members of the study management team; and 1/22, 5% observer). Overall, 68% (15/22) of participants traveled from within the United Kingdom and 32% (7/22) participants from Europe. On cancelation of the face-to-face meeting, we invited participants to continue their involvement. Methodological changes required only notification of a nonsubstantial amendment to the Nottingham 2 Research Ethics Committee, who approved the study. Examples of these changes included (1) amendment of the participant information leaflet to say *web-based consensus meeting*, (2) recording individual consent on the internet, and (3) extending the study end date.

Of the original group of participants, 1 health care user could not attend the rescheduled date (July 7, 2020) and 2 health care professionals did not respond to the invitation to the web-based meeting. A replacement health care user was recruited to maintain the balance across stakeholder groups. One additional facilitator was also recruited so that the web-based discussion groups were manageable.

A commercial representative based in Denmark and a US-based clinical researcher asked the CROSSSD team if they could join the meeting and so were invited to attend as nonparticipating observers. Therefore, the revised group of participants comprised 23 individuals, of which 12 (52%) were eligible to vote because they had completed both rounds of the Delphi survey. Participants consented to participate in the consensus meeting by completing a web-based form. Voting during the consensus meeting was conducted using hyperlinks to Jisc web-based surveys [12].

Redesigned Meeting

We used Microsoft Office Teams [13] for web-based discussions because (1) it was supported by the study sponsor; (2) it was freely available; and (3) it had desirable features, including a gallery view of all participants, a chat function, live caption ability, and audio recording. Optional one-to-one practical software tutorials were offered to all participants before the meeting to ensure that all necessary functionality was accessible and understood by participants. A discretionary virtual coffee morning was held the week before the meeting to enable participants to test the technology and meet each other socially. A total of 22 participants attended (1/22, 5% chairperson; 6/22, 27% health care users; 6/22, 27% health care professionals; 2/22, 9% public research partners; 1/22, 5% patient involvement manager; 2/22, 9% facilitators; 2/22, 9% other members of the study team; and 2/22, 9% observers). Overall, 3 health care professionals, 2 observers, 1 facilitator, and 1 study team member could not attend because of work commitments.

The public research partners, patient involvement manager, and facilitators with experience in conducting face-to-face COS consensus meetings and qualitative research contributed to the planning of the web-based meeting, including its structure, timing, preparatory activities, communication strategies, discussion points, and voting techniques. Public research partners helped to enhance accessibility for those with hearing

XSL•FO
RenderX

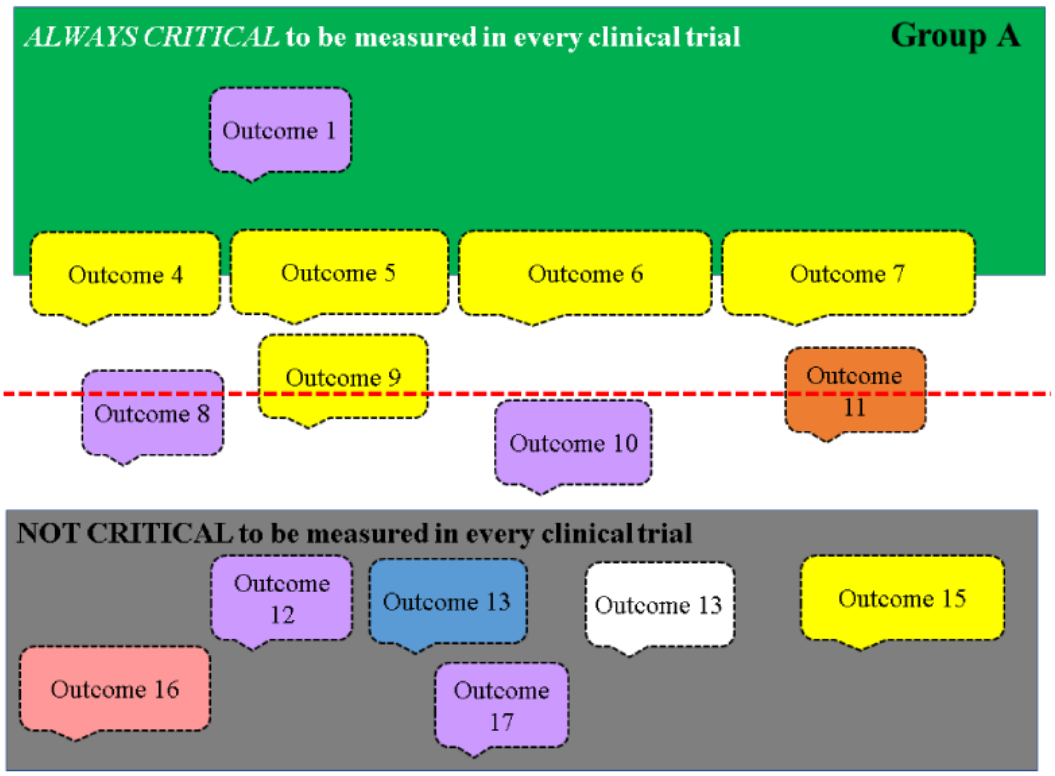
difficulties, drawing upon their own lived experience, as per recommendations when designing COS studies [2,14]. These enhancements included meeting etiquette (eg, use the *raise hand* function and wait your turn), chairing (eg, making the facilitators aware of their role, ways to resolve conflict, and adhering to the agenda), accessibility (eg, enabling the automatic captions), and troubleshooting (eg, use the chat function or exit and re-enter the software).

In line with the approach advocated by COMET, and to obtain qualitative information from our participant group in a structured manner [15], an NGT approach [16] was adopted. NGT allows groups to explore and thoroughly discuss issues in hand, identify, rank, and rate various problem dimensions with limited researcher influence or interference [5,7]. Conventionally, NGT comprises the following steps: (1) a chairperson introduces the group, sets ground rules, and explains the purpose of the meeting and procedures for the day; (2) the chairperson states the question and encourages each participant to individually reflect and brainstorm; (3) with the help of a facilitator, participants have an opportunity to discuss and clarify ideas; and (4) participants evaluate the ideas and vote anonymously for the *best ideas*. In CROSSSD, steps (3) and (4) were conducted in three parallel subgroups. Each facilitator presented the main discussion points using predetermined guidance (Multimedia Appendix 1) before voting. When consensus was required, an additional step (5) shared the voting results with the group and provided the opportunity to discuss and vote again. In this study, the results were presented using histograms embedded in PowerPoint slides.

CROSSSD extended the NGT by requesting participants to engage in certain activities in advance, namely, (1) inviting them to meet the group at the discretionary coffee morning; (2) introducing the meeting purpose, procedures, and Delphi survey results via an information pack (Multimedia Appendix 2), PowerPoint slides (Multimedia Appendix 3), and a prerecorded presentation (Multimedia Appendix 4); and (3) asking them to vote for three outcomes they considered crucial to include in the COS before the day of the consensus meeting. Further modifications were (4) introducing a structured ice-breaker activity and (5) providing subgroup support from a public research partner or patient involvement manager and facilitator. The subgroup composition was predetermined to achieve a balance of stakeholder perspectives and to facilitate the efficient organization of subgroup discussions on the day of the meeting.

During the consensus meeting, which was 7 hours long with three 30-minute-long breaks, a series of discussion and voting steps reduced the pool of candidate outcomes to a final COS. The first step was to present the results from the *top three outcomes* survey conducted before the meeting and asked whether participants agreed to exclude those outcomes that had not been selected by anyone to be in their top three. The voting options were *agree*, *disagree*, or *unsure*. Next, participants were asked to consider the remaining outcomes and to identify five outcomes that they considered critical to be measured in every clinical trial of interventions for SSD. During subgroup discussions, the facilitators moved the outcomes around a shared visual display to reflect discussions (Figure 1). The green zone included outcomes considered *always critical*, the gray zone included outcomes considered *not critical*, and the intermediate zone was for those with mixed opinions or not yet discussed.

Figure 1. The PowerPoint slide used to provide a visual display of the outcomes for consideration and to assist the facilitators when guiding discussions or summarizing subgroup discussions.



XSL•FO
RenderX

When participants returned to the full group, they were then asked to vote whether they would exclude those outcomes considered *not critical* (ie, in the gray zone). This process was repeated again after the whole group and subgroup discussions to reduce the list of candidate outcomes. Finally, participants considered the remaining outcomes that had not yet been voted *in* or *out* and voted on whether the *always critical* set should form the COS for SSD interventions. Applying the criterion of 70% agreement as per protocol [6], at least nine of the 12 participants agreed for any decision to be carried out.

For formative feedback, all 12 voting participants were asked to complete a web-based consensus meeting evaluation (Multimedia Appendix 5), adapted from the COMET Initiative [17]. Participants responded to six statements on the premeeting information, their experience of the consensus meeting, and fairness of the outcome using a 5-point Likert scale (strongly agree, agree, neither, disagree, or strongly disagree) and open text boxes for further comments. Each statement is described in the *Results* section. Two further open text boxes sought feedback on the practical arrangements for the meeting and suggestions for improvement. All other meeting participants were invited to respond to a modified version of the evaluation involving the open text comments only.

Results

Overview

Formative feedback was received from 75% (9/12) of voting participants (4 health care users and 5 health care professionals) and 40% (4/10) of the study team (2 public research partners, 1 patient involvement manager, and 1 facilitator). To illustrate the key points, many of the comments from one health care user (HU2) who was highly articulate are shared below. These views were confirmed by the study team's reflections.

Examples of Formative Feedback

Concerning whether the information provided in advance was helpful, all voting participants (9/9, 100%) agreed or strongly agreed. One said as follows:

> *Communication by the organisers with participants in advance of the meeting was absolutely first class with ample opportunity offered for consultation about any areas of concern and clarification when needed was always offered promptly and with considerable patience.* [HU2]

The facilitator commented as follows:

> *The pre-meeting information was very thorough. The Teams meeting was extremely valuable—contrary to my expectations. I expected this to be a confirmatory meeting; instead the facilitators highlighted aspects of the schedule which might not work so well, and everyone made contributive comments on how to make the online work. As a result some fundamental changes were made but we all felt we input into this process.*

The patient involvement manager agreed: "I had time at a prep-meeting to ask questions and to clarify the procedures for the day." Regarding whether the process used to agree with the COS was satisfactory, most participants (8/9, 89%) agreed or strongly agreed:

> *The process was particularly rigorous. The highest level of support was available from the leaders of the meeting but there was no heavy-handed intervention.* [HU2]

One public research partner highlighted the benefit of preparation:

> *The meeting had to be reconfigured to proceed remotely and this was handled exceptionally well [...] a lot of thought went into it and it showed, [...] technical support was provided promptly and without fuss or exasperation.*

Only one health care professional indicated that the process could have been improved by reorganizing the subgroups during the day:

> *I think the discussion in each group was influenced by the members, so some mixing would have helped [...] in the end there was a reasonably good outcome though.*

All participants (9/9, 100%) agreed or strongly agreed that meeting facilitation was satisfactory. Comments included were as follows:

> *The leaders were superb facilitators and every participant was made to feel as if their voice was important.* [HU2]
>
> *The facilitators were absolutely first class professionals and I felt privileged to have had the opportunity of working with them.* [HU2]

Again, all participants (9/9, 100%) agreed or strongly agreed that they felt able to contribute to the meeting. One supporting comment was as follows:

> *Everyone without exception was encouraged to participate fully at the meeting and the facilitators displayed great sensitivity to the needs of each individual contributor. From a personal point of view, I was concerned that the technology used for the meeting might impede successful and effective communication, but it didn't, thanks to the watchful eye of the leaders of the meeting who actively encouraged free expression from every participant while at the same time subtly guiding the proceedings to ensure maintenance of a structure which would lead to fulfilment of the consensus meeting's objectives. I would also like to add that a very fine rapport between participants was quickly established.* [HU2]

A facilitator indicated that past experience was important:

> *It had helped having been involved previously in facilitating three face-to-face COS consensus meetings. I drew heavily from that previous experience.*

XSL•FO
RenderX

Similarly, all participants (9/9, 100%) agreed or strongly agreed that they felt comfortable communicating their views. For example:

> *People taking part demonstrated great empathy for their fellows and there was a heart-warming sense of co-operation [...] delegates had ample opportunity to share their 'story' [...] I was made to feel like a person of value with something significant to contribute and I was particularly struck by the very high level of respect which people demonstrated for each other.* [HU2]

Finally, all participants (9/9, 100%) agreed or strongly agreed that the consensus meeting produced a fair result. One said:

> *There was at times quite heated debate, but I believe that a consensus was finally reached which reflected the majority view.* [HU2]

Subsequently, a Jisc survey of the wider stakeholder community confirmed 97% (89/92) agreement with the COS decisions made during the consensus meeting.

Participant Preferences

One of the major recurring themes was the preference for social interactions over web-based meetings. Two health care professionals said as follows:

> *Given the circumstances, this was a perfect solution, nevertheless I missed the social interactions.*
>
> *I personally don't like remote meetings. I feel they stifle free speech and the normal interactions and debate cannot happen in the same way.*
>
> *Nothing could have been better other than the face-to-face interaction [...] however, we enjoyed the benefits of the next best thing and there were also clearly some advantages in having a virtual workshop.* [HU2]

Another lesson concerned time management. At one point in the afternoon, there was some misunderstanding about the length of a break and when to reconvene, and this lost about 10-15 minutes of the schedule. Clear communication can avoid such issues. More generally, different stakeholders concurred that there was too little time for discussion. One health care professional said as follows:

> *I felt more time for each group to discuss the reasons behind their selected outcomes with the other groups, and to explain why they have selected one above another would have been useful...I enjoyed the in-group discussion, but felt the between-group discussions were a bit rushed/short.*

The patient involvement manager commented as follows:

> *We were a little rushed; not enough time for whole group discussions and voting.*

With regard to improvements to the web-based meeting, one public research partner recommended to plan more time at the end for discussion of the COS:

> *I felt that maybe a safety net or reserve of one hour might have been added to the end.*

Taking fatigue into account, the patient involvement manager suggested a debriefing might be deferred to a later date and be organized in the same way as the discretionary coffee morning to "allow participants to reflect with each other and to feel an appropriate 'closure', rather than a very intense day followed by a very quick 'goodbye'."

Discussion

Principal Findings

To the best of our knowledge, no previous COS development studies have adopted a fully web-based consensus methodology. On the basis of participant feedback, the premeeting preparation, choice of software, and approach to facilitation on the day proved effective in ensuring meaningful participant engagement. On balance, we conclude that the web-based method adopted for this meeting was successful and produced a result that was genuinely reflective of the group consensus.

We aim to create a safe environment with a *sense of belonging* to help participants feel valued, so they might share information more spontaneously. Although SSD is known to cause difficulty in following conversations in group situations, which can lead to listening fatigue and withdrawal [18], none of the participants mentioned such disadvantages. We suggest that the web-based meeting overcame these issues, as participants could access software features, including live subtitling on demand and *raise your hand*, which gives a clear right to conversation turn-taking that is not always achievable in face-to-face meetings [19]. Therefore, the structured discussions built into the web-based schedule were considered an advantage for this participant group.

Lessons Learned

Web-based meetings, unlike face-to-face meetings, can confer some advantages for participants to contribute effectively [8]. Three general methodological approaches have the greatest positive impact. The first approach concerned meeting planning and preparation. We followed recommendations to seek input from public research partners at all stages of COS development, drawing in perspectives based on the lived experience of SSD [2,14]. Their suggestions included lengthening the subgroup discussion time, screen sharing of the visual display, and sharing each subgroup's slide (Figure 1) with the meeting chair. Experienced facilitators pre-empted potential challenges. For example, a suggestion to add an activity prioritizing the outcomes within the COS was rejected after consultation with the facilitators because it was considered overambitious for the time scales. However, a suggestion to hold brief study team catch-ups during the breaks was endorsed, aiming to address any arising incidents and enhance participant contribution [8]. Detailed premeeting documentation informed participants about the process [7], setting clear expectations [20], and explaining the minimum participant requirements [11]. Participant feedback indicated that they appreciated this careful preparation. The second approach involved software training. Although all CROSSSD study facilitators and some participants had good

XSL•FO
RenderX

prior working knowledge of the audiovisual technology chosen to closely mirror a face-to-face environment, some participants had no previous experience. To optimize interactions, as recommended by Flynn et al [10], we ensured that all participants joined with a camera, either via a computer, smartphone, or tablet. Guidance notes, one-to-one tutorials, and the virtual coffee morning offered an opportunity to learn and practice using the technology. Finally, the third approach involved the study team taking a number of steps to ensure satisfaction with the meeting arrangements. Numerous authors have recommended offering participants a range of flexible times to allow for environment choice, for example, fitting around family timetables [8,11]. Although this was not feasible in this study because the NGT and voting had to be conducted in real time, the modifications we have described contributed to ensuring participants felt at ease and promoted positive group dynamics [10,11].

Limitations of the Evaluation

Although the response rate for the voting participants was acceptable, the majority of the open text comments came from one health care user (ie, HU2). The response rate for the study team was only 40% (4/10), with no responses received from the chairperson or observers. Furthermore, the COMET evaluation form was not tailored to the web-based meetings. To enrich the formative feedback and enhance the credibility of the present findings, the lead author (RK) sought an opportunity to triangulate our findings with 13 independent experts with experience in the planning and delivery of web-based consensus meetings. A meeting was convened in February 2021 by the Medical Research Council-National Institute for Health Research (NIHR) Trials Methodology Research Partnership Outcomes Working Group COS-subgroup, and experts joined virtually from the United Kingdom, Ireland, Amsterdam, the United States, Canada, and Australia. Agreed recommendations were directly relevant to many of our feedback findings, including the need for careful premeeting preparation, setting expectations to achieve less than what would be possible face-to-face, considering equity of engagement, ensuring the chairperson is strict with timings, and allowing time at the end for debriefing and reflection [21].

Conclusions

The COVID-19 pandemic presented a need and opportunity to introduce and evaluate a web-based consensus method involving hearing-impaired participants. Our findings indicate that it is feasible to conduct successful web-based consensus exercises with multistakeholder groups using audiovisual virtual meeting technology. We anticipate that the methodological changes made and the lessons learned are more widely applicable to other forms of research that require consensus-based decision-making and are not necessarily limited to COS development studies.

Acknowledgments

The study sponsor was the University of Nottingham. The CROSSSD study group would like to acknowledge the generous contributions of all health care users, health care professionals, manufacturers' representatives, and observers who attended and participated during the consensus meeting: Ad Snik, Carly Sygrove, Cherith Campbell-Bell, Christopher Parker, Daniel M Zeitler, Lewis Williams, Maxine Oxford, Patrick Boyle, Paul K James, Penelope R Hill-Feltham, Peter Toth, Richard Bowles, Richard Nicholson, Roger Bayston, and Tove Rosenbom. The authors would also like to thank Paul Van de Heyning (Department of Otorhinolaryngology, Head and Neck Surgery, Antwerp University Hospital, Antwerp, Belgium), Jill B Firszt (School of Medicine, Washington University in St Louis, St Louis, Missouri, USA), and Iain A Bruce (Manchester University Hospitals National Health Service Foundation Trust, Manchester Academic Health Science Centre, Manchester, United Kingdom) for their input as members of the CROSSSD study international experts steering group. The CROSSSD initiative also acknowledges the support of the NIHR Clinical Research Network in recruitment. The views expressed in this publication are those of the authors and not necessarily those of the National Health Service, the NIHR, or the Department of Health and Social Care. NIHR Nottingham Biomedical Research Centre, funding reference number BRC-1215-20003, provided funding. DAH is an NIHR senior investigator. KF was funded by the NIHR and Postdoctoral Fellowship PDF-2018-11-ST2-003. Additional grants obtained for the CROSSSD study are as follows: Graham Fraser Foundation Travel Grant to attend the 15th International Conference on Cochlear Implants and other Implantable Auditory Technology, where the study was first launched. Oticon Medical provided funding to purchase DelphiManager software from the COMET Initiative, University of Liverpool. The funding bodies had no role in the study design and implementation, writing the report, or decision to submit the report for publication.

Conflicts of Interest

PTK declares receiving grants and other support from Cochlear Europe Ltd and Cochlear Ltd (Australia) outside the submitted work. The rest of the authors declare no conflict of interest.

Multimedia Appendix 1
Core Rehabilitation Outcome Set for Single-Sided Deafness study facilitator pack.
[PDF File (Adobe PDF File), 717 KB - formative_v5i8e28878_app1.pdf]

Multimedia Appendix 2

XSL•FO
RenderX

Core Rehabilitation Outcome Set for Single-Sided Deafness participants plan and guide documents for virtual consensus meetings.
[PDF File (Adobe PDF File), 429 KB - formative_v5i8e28878_app2.pdf]

Multimedia Appendix 3
Core Rehabilitation Outcome Set for Single-Sided Deafness preconsensus meeting introductory presentation slides.
[PDF File (Adobe PDF File), 2759 KB - formative_v5i8e28878_app3.pdf]

Multimedia Appendix 4
Core Rehabilitation Outcome Set for Single-Sided Deafness preconsensus meeting introductory presentation recording.
[MOV File , 77222 KB - formative_v5i8e28878_app4.mov]

Multimedia Appendix 5
Core Rehabilitation Outcome Set for Single-Sided Deafness consensus meeting evaluation form.
[PDF File (Adobe PDF File), 183 KB - formative_v5i8e28878_app5.pdf]

References

1. Katiri R, Hall DA, Killan CF, Smith S, Prayuenyong P, Kitterick PT. Systematic review of outcome domains and instruments used in designs of clinical trials for interventions that seek to restore bilateral and binaural hearing in adults with unilateral severe to profound sensorineural hearing loss ('single-sided deafness'). Trials 2021 Mar 20;22(1):220 [FREE Full text] [doi: 10.1186/s13063-021-05160-5] [Medline: 33743802]
2. Williamson PR, Altman DG, Bagley H, Barnes KL, Blazeby JM, Brookes ST, et al. The COMET Handbook: version 1.0. Trials 2017 Jun 20;18(Suppl 3):280 [FREE Full text] [doi: 10.1186/s13063-017-1978-4] [Medline: 28681707]
3. Cole ZD, Donohoe HM, Stellefson ML. Internet-based Delphi research: case based discussion. Environ Manage 2013 Mar;51(3):511-523 [FREE Full text] [doi: 10.1007/s00267-012-0005-5] [Medline: 23288149]
4. Harvey N, Holmes CA. Nominal group technique: an effective method for obtaining group consensus. Int J Nurs Pract 2012 Apr;18(2):188-194. [doi: 10.1111/j.1440-172X.2012.02017.x] [Medline: 22435983]
5. Gladman T, Tylee G, Gallagher S, Mair J, Rennie SC, Grainger R. A tool for rRating the value of health education mobile apps to enhance student learning (MARuL): development and usability study. JMIR Mhealth Uhealth 2020 Jul 31;8(7):e18015 [FREE Full text] [doi: 10.2196/18015] [Medline: 32735228]
6. Katiri R, Hall DA, Buggy N, Hogan N, Horobin A, Van de Heyning P, et al. Core Rehabilitation Outcome Set for Single Sided Deafness (CROSSSD) study: protocol for an international consensus on outcome measures for single sided deafness interventions using a modified Delphi survey. Trials 2020 Mar 04;21(1):238 [FREE Full text] [doi: 10.1186/s13063-020-4094-9] [Medline: 32131880]
7. Olsen J. The Nominal Group Technique (NGT) as a tool for facilitating pan-disability focus groups and as a new method for quantifying changes in qualitative data. Int J Qual Methods 2019 Jul 26;18:1-10 [FREE Full text] [doi: 10.1177/1609406919866049]
8. Dodds S, Hess AC. Adapting research methodology during COVID-19: lessons for transformative service research. J Serv Manag 2020 Aug 10;32(2):203-217 [FREE Full text] [doi: 10.1108/josm-05-2020-0153]
9. Imlach F. Reflections on conducting research with healthcare users in a pandemic lockdown. N Z Med J 2020 Aug 21;133(1520):108-112. [Medline: 32994600]
10. Flynn R, Albrecht L, Scott SD. Two approaches to focus group data collection for qualitative health research. Int J Qual Methods 2018 Jan 17;17(1):160940691775078. [doi: 10.1177/1609406917750781]
11. Daniels N, Gillen P, Casson K, Wilson I. STEER: Factors to consider when designing online focus groups using audiovisual technology in health research. Int J Qual Methods 2019 Nov 19;18:160940691988578 [FREE Full text] [doi: 10.1177/1609406919885786]
12. Jisc Online Surveys. URL: https://www.onlinesurveys.ac.uk/ [accessed 2021-07-24]
13. Microsoft Teams. Microsoft Corporation Office 365. URL: https://www.microsoft.com/en/microsoft-365/microsoft-teams/group-chat-software [accessed 2021-07-24]
14. Young B, Bagley H. Including patients in core outcome set development: issues to consider based on three workshops with around 100 international delegates. Res Involv Engagem 2016;2:25 [FREE Full text] [doi: 10.1186/s40900-016-0039-6] [Medline: 29507761]
15. Azzari C, Baker SM. Ten lessons for qualitative transformative service researchers. J Serv Mark 2020 Jan 02;34(1):100-110 [FREE Full text] [doi: 10.1108/jsm-12-2018-0351]
16. Van de Ven AH, Delbecq AL. The nominal group as a research instrument for exploratory health studies. Am J Public Health 1972 Mar;62(3):337-342. [doi: 10.2105/ajph.62.3.337] [Medline: 5011164]
17. Consensus meeting evaluation. COMET Initiative. URL: https://www.comet-initiative.org/Patients/Consensus [accessed 2021-07-24]

XSL•FO
RenderX

18. Lucas L, Katiri R, Kitterick PT. The psychological and social consequences of single-sided deafness in adulthood. Int J Audiol 2018 Jan;57(1):21-30. [doi: 10.1080/14992027.2017.1398420] [Medline: 29132260]
19. Reid D, Reid FJ. Online Focus Groups: An in-depth comparison of computer-mediated and conventional focus group discussions. Int J Mark Res 2018 Jan 30;47(2):131-162 [FREE Full text] [doi: 10.1177/147078530504700204]
20. Smith H, Horobin A, Fackrell K, Colley V, Thacker B, Hall DA, Core Outcome Measures in Tinnitus (COMiT) initiative. Defining and evaluating novel procedures for involving patients in Core Outcome Set research: creating a meaningful long list of candidate outcome domains. Res Involv Engagem 2018;4:8 [FREE Full text] [doi: 10.1186/s40900-018-0091-5] [Medline: 29507772]
21. Gorst SL, Barrington H, Brookes ST, Chalmers JR, Devane D, Fledderus AC, et al. Online consensus meetings for COS development: issues to consider. COMET Online. URL: https://www.comet-initiative.org/Downloads/Issues%20to%20consider%20for%20online%20consensus%20meetings.pdf [accessed 2021-07-24]

Abbreviations

COMET: Core Outcome Measures in Effectiveness Trials
COS: core outcome set
CROSSSD: Core Rehabilitation Outcome Set for Single-Sided Deafness
NGT: nominal group technique
NIHR: National Institute for Health Research
SSD: single-sided deafness

Edited by G Eysenbach; submitted 17.03.21; peer-reviewed by N Daniels, N Harman; comments to author 07.05.21; revised version received 15.05.21; accepted 29.05.21; published 19.08.21.

Please cite as:
Katiri R, Hall DA, Hoare DJ, Fackrell K, Horobin A, Buggy N, Hogan N, Kitterick PT, Core Rehabilitation Outcome Set For Single-Sided Deafness (CROSSSD) Initiative
Redesigning a Web-Based Stakeholder Consensus Meeting About Core Outcomes for Clinical Trials: Formative Feedback Study
JMIR Form Res 2021;5(8):e28878
URL: https://formative.jmir.org/2021/8/e28878
doi:10.2196/28878
PMID:34420915

©Roulla Katiri, Deborah A Hall, Derek J Hoare, Kathryn Fackrell, Adele Horobin, Nóra Buggy, Nicholas Hogan, Pádraig T Kitterick, Core Rehabilitation Outcome Set For Single-Sided Deafness (CROSSSD) Initiative. Originally published in JMIR Formative Research (https://formative.jmir.org), 19.08.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Association of COVID-19 Risk Misperceptions With Household Isolation in the United States: Survey Study

Joseph A Ladapo[1], MD, PhD; Jonathan T Rothwell[2], PhD; Christina M Ramirez[3], PhD

[1]Division of General Internal Medicine and Health Services Research, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, United States

[2]Gallup, Washington, DC, United States

[3]Department of Biostatistics, Fielding School of Public Health, University of California, Los Angeles, Los Angeles, CA, United States

Corresponding Author:
Joseph A Ladapo, MD, PhD
Division of General Internal Medicine and Health Services Research
David Geffen School of Medicine
University of California, Los Angeles
1100 Glendon Ave, Suite 850
Los Angeles, CA, 90024
United States
Phone: 1 3107942728
Email: jladapo@mednet.ucla.edu

Abstract

Background: Adverse mental and emotional health outcomes are increasingly recognized as a public health challenge associated with the COVID-19 pandemic.

Objective: The goal of this study was to examine the association of COVID-19 risk misperceptions with self-reported household isolation, a potential risk factor for social isolation and loneliness.

Methods: We analyzed data from the Franklin Templeton-Gallup Economics of Recovery Study (July to December 2020) of 24,649 US adults. We also analyzed data from the Gallup Panel (March 2020 to February 2021), which included 123,516 observations about loneliness. The primary outcome was self-reported household isolation, which we defined as a respondent having no contact or very little contact with people outside their household, analogous to quarantining.

Results: From July to December 2020, 53% to 57% of respondents reported living in household isolation. Most participants reported beliefs about COVID-19 health risks that were inaccurate, and overestimation of health risk was most common. For example, while deaths in persons younger than 55 years old accounted for 7% of total US deaths, respondents estimated that this population represented 43% of deaths. Overestimating COVID-19 health risks was associated with increased self-reported household isolation, with percentage differences ranging from 5.6 to 11.8 ($P<.001$ at each time point). Characteristics associated with self-reported household isolation from the July and August 2020 surveys and persisting in the December 2020 survey included younger age (18 to 39 years), having a serious medical condition, having a household member with a serious medical condition, and identifying as a Democrat. In the Gallup Panel, self-reported household isolation was associated with a higher prevalence of loneliness.

Conclusions: Pandemic-related harms to emotional and mental well-being may be attenuated by reducing risk overestimation and household isolation preferences that exceed public health guidelines.

(JMIR Form Res 2021;5(8):e30164) doi:10.2196/30164

KEYWORDS

COVID-19; pandemic; mental health; public health; isolation; loneliness; guideline; risk; perception; United States; health risk; well-being

XSL•FO
RenderX

Introduction

Adverse mental and emotional health outcomes are increasingly recognized as a public health challenge associated with the COVID-19 pandemic. As early as March 2020, a national survey reported that 36% of US adults felt the pandemic would have a serious impact on their mental health [1]. In April 2020, another survey found that 14% of US adults reported serious psychological distress, compared to 4% during a similar time period in 2018 [2]. Rates of loneliness have also been high, with 36% of US adults—including 61% of adults aged 18 to 25 years—reporting significant loneliness in an October 2020 survey [3]. More recently, a March 2021 survey found that 48% of adults reported higher levels of stress in their lives compared to before the pandemic, and 61% reporting undesired weight changes [4].

These health sequalae of the COVID-19 pandemic are multifactorial, and social isolation is likely an important contributor [5,6]. Because of physical distancing mandates, quarantines, and fear of illness, a substantial proportion of Americans have limited their physical contact with others outside of their household. This trend has likely contributed to social isolation and loneliness. Household isolation is analogous to quarantining, and research has shown that quarantining is a risk factor for a variety of adverse mental and emotional health outcomes, including increased stress, anxiety, depression, fear, and detachment from other people [5,7].

The US Centers for Disease Control and Prevention (CDC) recently recommended that researchers examine drivers of adverse mental health during the COVID-19 pandemic [8]. One driver that has received little attention is the role that COVID-19 risk misperceptions may play in the behavioral decision to limit physical contact with others. While COVID-19 risk perceptions have been associated with protective health behaviors [9], they may lead to suboptimal behavioral choices, if individuals substantially overestimate or underestimate risk [10,11]. Overestimation, in particular, is of concern in the context of mental and emotional health and well-being because it tends to amplify social isolation and reduce contact with others. Using survey data from the Franklin Templeton-Gallup Economics of Recovery Study, we assessed the association of COVID-19 risk misperceptions with self-reported household isolation. Our findings are relevant to policy measures to reduce COVID-19–related social isolation and may inform the management of future epidemics and pandemics.

Methods

Data

We used data from the Franklin Templeton-Gallup Economics of Recovery Study, a self-administered web survey from an opt-in sample provided by Dynata of 24,649 US adults, aged 18 years and older, of whom 10,419 participated during more than one survey time point. The survey was conducted during the following time points: July 2 to 14, August 3 to 11, September 4 to 13, October 1 to 9, November 2 to 6, and December 1 to 7, 2020. Gallup weighted the obtained sample to correct for nonresponse and construct a nationally representative population. Nonresponse adjustments were made by adjusting the sample to match the national demographics of gender, age, race and ethnicity, region, educational level, marital status, and employment status. Demographic weighting targets were based on the Census Bureau's 2018 data release of the American Community Survey and the Current Population Survey (February 2020).

We also supplemented this survey with data from the Gallup Panel, a research panel that is representative of the US adult population and includes approximately 100,000 members. Gallup fielded the COVID-19 tracking survey on March 13, 2020, and collected approximately 1000 responses daily until April 26, 2020, when the sample declined to approximately 500 responses daily. The Gallup Panel's COVID-19 Tracking Survey includes information about self-reported household isolation and loneliness, with 123,516 observations from March 24, 2020, to February 21, 2021. This study was exempt from institutional review board review according to policies of the UCLA (University of California, Los Angeles) Office of the Human Research Protection Program.

Primary Measures

We assessed self-reported household isolation by asking participants about the degree of in-person contact outside their household that they had over the past 24 hours. We considered a participant to be isolated if they reported being *completely isolated* (no contact) or *mostly isolated* (very little contact) from people outside their household. We assessed loneliness in the Gallup Panel survey by asking participants, "Did you experience the following feelings during a lot of the day yesterday?" Loneliness and other emotional experiences were included as response options. The specific questions and respondent options are provided in Multimedia Appendix 1.

Perceptions of COVID-19 Health Risks

We evaluated perceptions of COVID-19 health risks using multiple questions. In July and August, respondents were asked about the percentage of all US COVID-19 deaths that fell into the following age strata: age 24 years and below, age 25 to 34 years, age 35 to 44 years, age 45 to 54 years, age 55 to 64 years, and age 65 years and older. We assessed misperceptions using the reported proportion of deaths attributable to persons under the age of 55 years because most deaths from COVID-19 have occurred in persons older than 55 years. Perceptions about age-related COVID-19 health risks were assessed in September and October 2020 with an analogous question about the age distribution of COVID-19 hospitalizations.

In the November 2020 survey, respondents were asked what percentage of patients hospitalized with COVID-19 died. In the December 2020 survey, respondents were asked what percentage of patients infected with COVID-19 required hospitalization.

Estimation of Actual COVID-19 Health Risks

CDC data were used to estimate the proportion of deaths from COVID-19 by age. Data from the COVID-19–Associated Hospitalization Surveillance Network (COVID-NET) was used to estimate hospitalizations by age [12]. Because COVID-NET data on hospitalizations were reported using different age strata

XSL•FO
RenderX

than those provided to survey respondents, we adjusted these data using simple proportional methods. Specifically, we multiplied hospitalizations reported in COVID-NET by the proportion of years in a corresponding age stratum in order to recategorize hospitalizations into different age strata. The likelihood of hospitalization after infection was estimated to be approximately 5% as reported by Reese and colleagues from the CDC estimate that there were 52.9 million infections from February to September 2020 and 2.4 million hospitalizations, implying a hospitalization rate of 4.5% [13]. Their method accounted for underreporting. We estimated the likelihood of death among patients hospitalized for COVID-19 to be 12% based on an analysis of 38,517 hospitalized patients from January 1, 2020, to June 30, 2020 [14].

Other Measures

In each wave, survey respondents were asked whether they or a household member had a comorbidity that increased the risk of severe COVID-19 illness. Respondents also reported sociodemographic characteristics, household income, and preferences for political parties. Per capita deaths from COVID-19 in each US county from March 1 until December 1 were assessed using CDC data [15].

Analyses of Survey Data

Descriptive analyses of respondents' characteristics were performed using data from July, August, and December 2020. The July and August surveys were combined for analyses because questions about risk perception were identical between those two time points. The September and October surveys were similarly combined. We used multivariable logistic regression analyses to examine the relationship between misperceptions about COVID-19 health risks and social isolation. Results from these models were presented as predictive margins, in which the regression models were used to estimate the marginal effect of risk overestimation, expressed as a proportion, while holding the distribution of all other covariates constant [16]. The adjusted association of respondent characteristics with social isolation was also presented using data from July and August as well as December in order to examine how behavioral patterns may have shifted over the course of the pandemic. In a secondary analysis, we used Gallup Panel data to assess the relationship between self-reported household isolation and loneliness.

Perceptions about risk were characterized as being overestimates, underestimates, or accurate estimates. To provide respondents with a reasonable degree of latitude and to account for any uncertainty in our reference estimates, we considered responses that were within 5 percentage points above or below the correct estimate as being accurate (eg, a response of 15% for estimated hospital mortality would be considered accurate because it fell within 5 percentage points of the actual rate of 12%) [14]. A range of 10 percentage points above or below was used for respondents' estimates of the proportion of hospitalizations occurring in persons younger than 55 years old, due to the larger proportion.

We performed mean imputation from the overall sample for education (missing <1%), income (missing <1%), and whether the respondent or their family had a serious medical condition (missing <1%). We used this method instead of a more robust multiple imputation model because of the low rate of missingness. We did not report percentages as n/N values because all reported percentages were estimated using analytic weights. All analyses were performed using Stata 14 (StataCorp LP) and incorporated analytic weights to account for the effects of nonresponse.

Data Availability

The data used in this study are available from the corresponding authors upon request and with the permission of Franklin Templeton and Gallup.

Results

Overview

We present descriptive characteristics of the respondents in Table 1. The mean age of the respondents was 47 years (SD 18), and 52.3% were female. The largest proportion of respondents had a household income that ranged from US $48,000 to US $89,999. Half of the respondents reported that they or a household member had a serious medical condition that increased their risk of serious illness from COVID-19.

XSL•FO
RenderX

Table 1. Characteristics of the study sample by time period.

Characteristics	July to August 2020: participants (N=15,014), n (%)[a]	September to October 2020: participants (N=10,019), n (%)[a]	November 2020: participants (N=5026), n (%)[a]	December 2020: participants (N=5009), n (%)[a]
Age group (years)				
18-39	5333 (37.5)	3665 (37.7)	1966 (38.7)	1749 (38.3)
40-64	6347 (41.3)	4140 (41.2)	1978 (40.4)	2148 (40.6)
≥65	3334 (21.2)	2214 (21.1)	1082 (20.9)	1112 (21.1)
Female				
Yes	8032 (52.3)	5430 (52.2)	2711 (51.9)	2733 (51.9)
No	6982 (47.8)	4589 (48.0)	2315 (48.2)	2276 (48.3)
Race or ethnic group				
White	9512 (63.4)	6465 (63.7)	3206 (63.7)	3228 (63.6)
Black	1909 (12.4)	1289 (12.3)	694 (12.3)	680 (12.4)
Hispanic	2491 (16.4)	1503 (16.0)	788 (16.0)	720 (15.8)
Other or unknown	1102 (7.8)	762 (8.1)	338 (8.1)	381 (8.2)
Educational level				
8th grade or some high school	328 (4.0)	274 (4.0)	165 (4.6)	106 (3.4)
High school graduate	2889 (36.7)	2268 (37.0)	1155 (36.5)	1173 (38.1)
Some college or college graduate	11,797 (59.5)	7477 (59.1)	3706 (59.0)	3730 (58.6)
Serious medical condition				
No	7708 (51.8)	5099 (51.0)	2427 (49.1)	2414 (48.2)
Yes, in respondent	3843 (26.1)	2465 (25.1)	1307 (26.4)	1211 (24.8)
Yes, in household member	2324 (15.3)	1652 (16.6)	844 (16.6)	890 (17.9)
Yes, in respondent and household member	1139 (7.1)	803 (7.5)	448 (8.3)	494 (9.4)
Political preference				
Democrat	5483 (36.1)	3763 (37.3)	1940 (37.9)	1990 (38.4)
Republican	4539 (31.4)	3185 (32.1)	1686 (34.2)	1545 (31.2)
Independent	4155 (27.1)	2594 (25.8)	1190 (23.6)	1232 (24.9)
Other or unknown party	837 (6.1)	477 (4.9)	210 (4.5)	242 (5.7)
Household income (US $)				
<24,000	2466 (19.1)	1922 (20.7)	1009 (20.6)	931 (20.4)
24,000-47,999	2825 (22.2)	2146 (24.4)	1078 (25.3)	1152 (26.1)
48,000-89,999	4519 (30.3)	2990 (29.5)	1456 (29.3)	1546 (30.4)
≥90,000	5204 (28.4)	2961 (25.4)	1483 (24.8)	1380 (23.1)
Married				
Yes	8018 (48.4)	5124 (48.0)	2652 (47.8)	2567 (47.6)
No	6996 (51.9)	4895 (52.2)	2374 (52.3)	2442 (52.5)
Live in rural area				
Yes	1262 (9.4)	982 (10.3)	527 (11.3)	474 (10.2)
No	13,752 (90.6)	9037 (89.8)	4499 (88.8)	4535 (89.9)

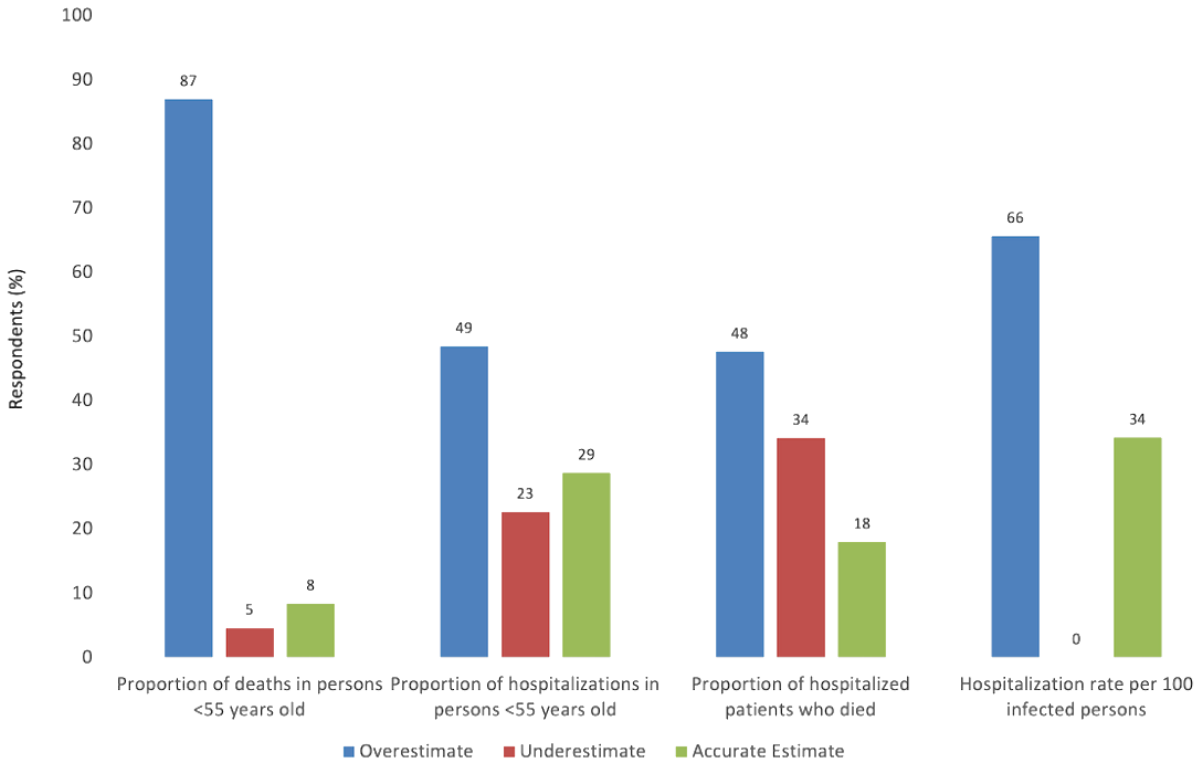
[a]Percentages are based on analytic weights.

XSL•FO
RenderX

Misperceptions About COVID-19 Health Risks

Most participants held beliefs about COVID-19 health risks that were inaccurate (Figure 1). Overestimation of health risk, rather than underestimation, was the most common type of inaccuracy at each survey time point. For example, while persons younger than 55 years old accounted for 7% of total US deaths at the time of the July and August surveys, respondents estimated that they accounted for 43% of total deaths. In addition, while the proportion of COVID-19 hospitalizations that occurred in persons younger than 55 years old was 38%, respondents in the September and October surveys reported that this population accounted for 46%. The mortality rate of patients hospitalized with COVID-19 was estimated by respondents in the November survey to be 25% compared to an actual rate of 12%. The proportion of patients hospitalized after being infected with COVID-19 was estimated by respondents in the December survey to be 34% compared to an actual proportion of 12%.

Figure 1. Comparison of respondents' perceived risk versus actual risk associated with COVID-19 illness.



Association Between COVID-19 Health Risk Perceptions and Household Isolation

The proportion of respondents living in self-reported household isolation did not vary substantially over time, ranging from 53% to 57% (Figure 2). Overestimating the proportion of death or hospitalizations from COVID-19 occurring in people under 55 years old was associated with a significantly increased likelihood of self-reported household isolation (Table 2). Overestimating the likelihood of death or hospitalization was also associated with a significantly increased likelihood of self-reported household isolation. Excluding respondents who underestimated risk modestly attenuated the results, but all associations between misperceptions and social isolation remained significant (Table S1 in Multimedia Appendix 2). In the Gallup Panel, adults living in self-reported household isolation reported higher rates of loneliness (Figure S1 in Multimedia Appendix 2).

XSL•FO
RenderX

Figure 2. Rates of household isolation during the COVID-19 pandemic.

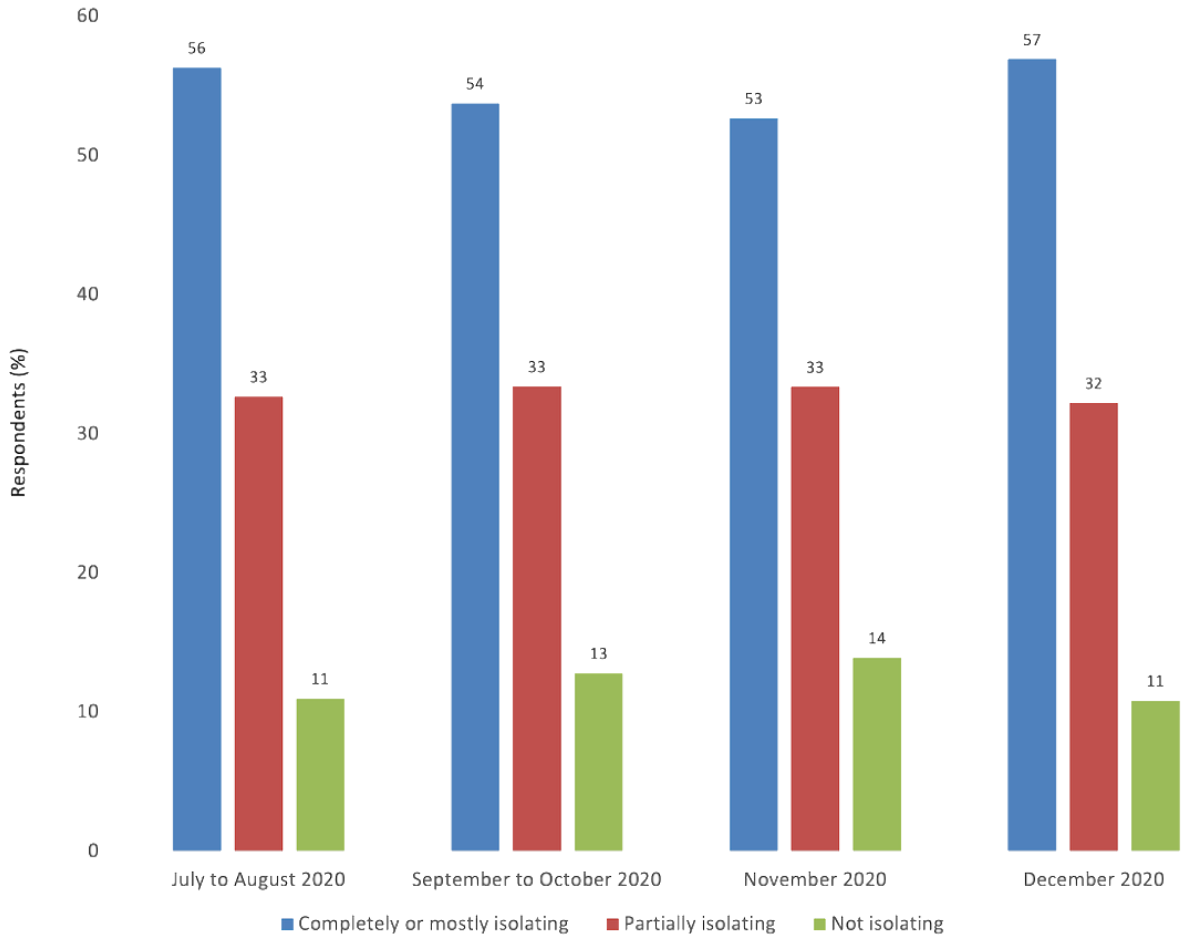


Table 2. Effect of misperceptions (ie, overestimation) of risk[a] on the likelihood of living in self-reported household isolation for each time period.

Misperceptions	Likelihood of living in social isolation							
	July to August 2020		September to October 2020		November 2020		December 2020	
	Difference in % (95% CI)	*P* value	Difference in % (95% CI)	*P* value	Difference in % (95% CI)	*P* value	Difference in % (95% CI)	*P* value
Misperception about deaths[b]	7.7 (5.3-10.1)	<.001	N/A[c]	N/A	N/A	N/A	N/A	N/A
Misperception about hospitalizations[d]	N/A	N/A	5.6 (3.6-7.6)	<.001	N/A	N/A	N/A	N/A
Misperception about hospital mortality[e]	N/A	N/A	N/A	N/A	9.9 (6.9-12.8)	<.001	N/A	N/A
Misperception about hospitalization risk[f]	N/A	N/A	N/A	N/A	N/A	N/A	11.8 (8.7-14.9)	<.001

[a]Adjusted marginal effect of misperception (ie, overestimation) of risk.

[b]Misperception about proportion of COVID-19 deaths attributable to persons younger than 55 years old.

[c]N/A: not applicable; questions about misperception were only asked at specific time points.

[d]Misperception about proportion of COVID-19 hospitalizations attributable to persons younger than 55 years old.

[e]Misperception about proportion of patients hospitalized with COVID-19 who die.

[f]Misperception about hospitalization risk if infected with COVID-19.

XSL•FO
RenderX

Other Characteristics Associated With Household Isolation

Characteristics associated with self-reported household isolation from the July and August 2020 surveys and persisting in the December 2020 survey included younger age (18 to 39 years), having a serious medical condition, having a household member with a serious medical condition, and identifying as a Democrat (Table 3). Being Black or Hispanic was associated with a higher likelihood of social isolation in July and August, but this relationship was not present in December. Reporting a higher income (>US $48,000) was associated with a lower likelihood of social isolation in July and August, but this relationship had largely waned by December.

XSL•FO
RenderX

Table 3. Association of misperceptions (ie, overestimation) about COVID-19 health risks with preferences for living in self-reported household isolation.

Characteristics	July to August 2020		September to October 2020		November 2020		December 2020	
	aOR[a] (95% CI)	*P* value	aOR (95% CI)	*P* value	aOR (95% CI)	*P* value	aOR (95% CI)	*P* value
Misperceptions about COVID-19 health risks								
Misperception about deaths[b]	1.38 (1.25-1.53)	<.001	N/A[c]	N/A	N/A	N/A	N/A	N/A
Misperception about hospitalizations[d]	N/A	N/A	1.27 (1.17-1.38)	.001	N/A	N/A	N/A	N/A
Misperception about hospital mortality[e]	N/A	N/A	N/A	N/A	1.52 (1.34-1.72)	.001	N/A	N/A
Misperception about hospitalization risk[f]	N/A	N/A	N/A	N/A	N/A	N/A	1.65 (1.45-1.88)	<.001
Age group (years)								
18-39 (reference)	1.00	__[g]	1.00	—	1.00	—	1.00	—
40-64	0.73 (0.67-0.79)	.001	0.89 (0.80-0.98)	.01	0.76 (0.66-0.88)	<.001	0.84 (0.73-0.96)	.01
≥65	0.88 (0.79-0.97)	.01	0.93 (0.83-1.05)	.25	0.76 (0.64-0.90)	.002	0.98 (0.82-1.17)	.83
Female								
No (reference)	1.00	—	1.00	—	1.00	—	1.00	—
Yes	0.96 (0.90-1.04)	.32	0.96 (0.88-1.04)	.33	0.86 (0.76-0.98)	.02	1.02 (0.90-1.15)	.78
Race or ethnic group								
White (reference)	1.00	—	1.00	—	1.00	—	1.00	—
Black	1.13 (1.00-1.26)	.04	1.44 (1.25-1.66)	<.001	1.33 (1.09-1.62)	.006	0.97 (0.79-1.18)	.74
Hispanic	1.35 (1.22-1.49)	<.001	1.27 (1.12-1.43)	<.001	1.29 (1.09-1.54)	.004	0.97 (0.82-1.16)	.76
Other	1.34 (1.17-1.53)	<.001	1.50 (1.27-1.75)	<.001	1.67 (1.33-2.10)	<.001	1.24 (0.99-1.56)	.06
Educational level								
8th grade or some high school (reference)	1.00	—	1.00	—	1.00	—	1.00	—
High school graduate	0.95 (0.79-1.15)	.59	1.00 (0.81-1.25)	.97	0.55 (0.41-0.75)	<.001	0.70 (0.50-0.99)	.04
Some college or college graduate	1.00 (0.83-1.21)	.99	1.01 (0.81-1.26)	.95	0.64 (0.47-0.87)	.004	0.84 (0.60-1.19)	.32
Serious medical condition								
No (reference)	1.00	—	1.00	—	1.00	—	1.00	—
Yes, in respondent	2.15 (1.98-2.35)	<.001	2.26 (2.04-2.50)	<.001	1.78 (1.53-2.06)	<.001	1.85 (1.60-2.15)	<.001
Yes, in household member	1.49 (1.35-1.64)	<.001	1.70 (1.51-1.91)	<.001	1.26 (1.07-1.50)	.007	1.34 (1.14-1.58)	<.001
Yes, in respondent and household member	1.66 (1.45-1.91)	<.001	1.89 (1.61-2.23)	<.001	1.43 (1.14-1.79)	.002	1.73 (1.39-2.15)	<.001
Political preference								
Democrat (reference)	1.00	—	1.00	—	1.00	—	1.00	—
Republican	0.64 (0.59-0.70)	<.001	0.68 (0.62-0.76)	<.001	0.57 (0.50-0.67)	<.001	0.54 (0.47-0.63)	<.001
Independent	0.72 (0.66-0.78)	<.001	0.79 (0.71-0.88)	<.001	0.75 (0.64-0.88)	<.001	0.76 (0.65-0.89)	<.001
Other party	0.72 (0.61-0.84)	<.001	0.68 (0.56-0.84)	<.001	0.70 (0.52-0.95)	.02	0.67 (0.51-0.87)	.004
Household income (US $)								
24,000 (reference)	1.00	—	1.00	—	1.00	—	1.00	—
24,000-47,999	0.90 (0.81-1.01)	.07	1.02 (0.90-1.16)	.72	0.85 (0.71-1.01)	.07	0.96 (0.80-1.15)	.65
48,000-89,999	0.80 (0.72-0.89)	<.001	0.92 (0.81-1.05)	.21	0.90 (0.75-1.08)	.24	0.98 (0.82-1.18)	.85

XSL•FO
RenderX

Characteristics	July to August 2020		September to October 2020		November 2020		December 2020	
	aOR[a] (95% CI)	*P* value	aOR (95% CI)	*P* value	aOR (95% CI)	*P* value	aOR (95% CI)	*P* value
≥90,000	0.89 (0.79-1.00)	.04	1.09 (0.95-1.26)	.23	1.11 (0.90-1.37)	.31	1.24 (1.01-1.53)	.04
Married								
No (reference)	1.00	—	1.00	—	1.00	—	1.00	—
Yes	0.99 (0.92-1.07)	.80	0.97 (0.89-1.06)	.54	1.04 (0.91-1.18)	.61	1.01 (0.88-1.15)	.92
Live in rural area								
No (reference)	1.00	—	1.00	—	1.00	—	1.00	—
Yes	0.93 (0.82-1.04)	.20	0.86 (0.75-0.99)	.04	0.95 (0.78-1.15)	.61	0.93 (0.76-1.13)	.45
Deaths per capita due to COVID-19	1.00 (0.99-1.00)	.79	1.01 (1.00-1.02)	.18	1.00 (0.99-1.02)	.54	1.00 (0.99-1.01)	.92

[a]aOR: adjusted odds ratio.

[b]Misperception about proportion of COVID-19 deaths attributable to persons younger than 55 years old.

[c]N/A: not applicable; questions about misperception were only asked at specific time points.

[d]Misperception about proportion of COVID-19 hospitalizations attributable to persons younger than 55 years old.

[e]Misperception about proportion of patients hospitalized with COVID-19 who die.

[f]Misperception about hospitalization risk if infected with COVID-19.

[g]*P* values cannot be calculated for reference values.

Discussion

Adverse mental and emotional health effects of the COVID-19 pandemic are an increasingly recognized public health challenge. Risk misperceptions about COVID-19 may be exacerbating this challenge. Using Franklin Templeton-Gallup Economics of Recovery Study surveys from July to December 2020, we found that respondents consistently overestimated health risks associated with COVID-19, as measured by four different questions assessing COVID-19 morbidity and mortality. Overestimation of risk was consistently associated with greater self-reported household isolation, which may have adverse emotional and mental effects similar to quarantining [5,7]. These findings are relevant to policy interventions for social isolation and loneliness because they suggest that more accurate public understanding of risk would yield an optimal balance between health precautions and healthy social interactions.

While social isolation and loneliness have often been considered health risks for older adults, prior research has shown that COVID-19–related emotional and mental health harms are disproportionately borne by younger adults [3,17]. A June 2020 CDC survey reported that approximately twice as many respondents seriously considered suicide in the previous 30 days compared to US adults in 2018 when asked about the previous 12 months (10.7% versus 4.3%) [8]. The highest rates of suicidal ideation were reported by persons aged 18 to 24 years. We found that self-reported household isolation was most common among persons aged 18 to 39 years, a finding that likely contributes to the high rates of emotional distress reported in this population during the pandemic. The analysis of the Gallup Panel, which included detailed questions about emotional health, demonstrated that US adults reporting household isolation also reported higher rates of loneliness. The disproportionate burden on young people also raises concerns about long-term health and economic consequences.

Suicidal ideation among young adults is a concern because it is associated with a markedly increased risk of suicide plan and attempt, particularly during the first year after onset of ideation [18]. Among persons aged 18 to 25 years who participated in the 2009 to 2015 National Surveys on Drug Use and Health, the 12-month prevalence of suicidal ideation increased from 6.1% to 8.3% [19]. However, over this time period, receipt of mental health care was unchanged for most suicidal young adults and declined slightly among young adults without health insurance. This combination of trends may exacerbate the effects of worsening mental and emotional health during the COVID-19 pandemic.

Respondents who had a serious medical condition or lived with a household member with a serious medical condition were more likely to engage in self-reported household isolation. Based on public health messages about risk factors for an adverse COVID-19 outcome, this finding was anticipated. Our finding that political party was associated with differences in self-reported preventive health behavior for COVID-19 has been confirmed in other work, including a survey study of 3000 American adults performed in March 2020 [20,21]. The partisan differences appear to also extend to policy preferences in response to COVID-19 [20]. Younger adults also reported higher rates of household isolation compared to older adults. Because these characteristics were independently associated with self-reported household isolation, they identify groups that may benefit from targeted public health messaging, in instances when the anticipated benefits of stricter household isolation due to reduction in likelihood of transmission may be outweighed by the mental and emotional health costs. The corollary is that there may also be populations who would benefit from greater engagement in household isolation to reduce the risk of infection.

Our study has limitations. We asked respondents about household isolation over the previous 24 hours rather than over

XSL•FO
RenderX

a longer period of time, which may have led to inaccuracies. However, the relatively stable distribution of self-reported household isolation from month to month suggests that a 24-hour recall period was informative. Although we measured self-reported household isolation, we could not quantify social isolation or loneliness for survey respondents in the Franklin Templeton-Gallup Economics of Recovery Study because we did not collect information about participation in group activities, social engagement with friends or relatives, or subjective experience of loneliness [22,23]. It is possible that some individuals who strictly avoided contact with people outside of their household experienced low levels of social isolation and loneliness, while others who did not isolate themselves experienced high levels of social isolation and loneliness. However, our analysis of Gallup Panel data showed that US adults who avoided contact with people outside their household also reported higher rates of loneliness. Furthermore, household isolation is analogous to quarantining, and research has shown that quarantining is associated with increased stress, anxiety, depression, fear, and detachment from other people [5,7].

Another limitation is that self-reporting bias may have affected the accuracy of our household isolation measure. However, the relatively stable distribution of self-reported household isolation across study periods suggests that this bias was minimal. In addition, the questions we used to assess risk perceptions changed over time, which precluded direct comparisons of risk perception between time periods. Another factor that further complicated measurement of COVID-19 health risks and the likelihood of reporting household isolation is that COVID-19 case levels varied during our study period. Our regression models adjusted for per capita COVID-19 cases at the county level, but there could be confounding effects from other pandemic factors that varied over time. Furthermore, despite differences in our survey questions related to COVID-19 risk perception, these questions consistently probed beliefs about hospitalization and mortality risk, and our finding of an association between risk overestimation and self-reported household isolation was consistent, despite the changing questions.

In conclusion, survey respondents overestimated several health risks associated with COVID-19, and this overestimation was associated with a respondent's decision to avoid contact with people outside the household. This relationship was consistent from July to December. Harms to emotional and mental well-being experienced by US adults during the COVID-19 pandemic may be mitigated by addressing risk misconceptions and attenuating household isolation preferences that exceed public health guidelines.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Study survey.
[DOCX File , 17 KB - formative_v5i8e30164_app1.docx]

Multimedia Appendix 2
Effect of misperceptions on the likelihood of living in household isolation and emotional distress by degree of household isolation and age.
[DOCX File , 126 KB - formative_v5i8e30164_app2.docx]

References

1. New poll: COVID-19 impacting mental well-being: Americans feeling anxious, especially for loved ones; older adults are less anxious. American Psychiatric Association. 2020 Mar 25. URL: https://www.psychiatry.org/newsroom/news-releases/new-poll-covid-19-impacting-mental-well-being-americans-feeling-anxious-especially-for-loved-ones-older-adults-are-less-anxious [accessed 2021-03-25]
2. McGinty EE, Presskreischer R, Han H, Barry CL. Psychological distress and loneliness reported by US adults in 2018 and April 2020. JAMA 2020 Jul 07;324(1):93-94 [FREE Full text] [doi: 10.1001/jama.2020.9740] [Medline: 32492088]
3. Loneliness in America: How the pandemic has deepened an epidemic of loneliness and what we can do about it. Making Caring Common Project, Harvard Graduate School of Education. 2021 Feb. URL: https://mcc.gse.harvard.edu/reports/loneliness-in-america [accessed 2021-03-28]
4. One year later, a new wave of pandemic health concerns. American Psychological Association. 2021 Mar 11. URL: https://www.apa.org/news/press/releases/stress/2021/one-year-pandemic-stress [accessed 2021-03-29]
5. Pfefferbaum B, North CS. Mental health and the Covid-19 pandemic. N Engl J Med 2020 Aug 06;383(6):510-512. [doi: 10.1056/NEJMp2008017] [Medline: 32283003]
6. Reger MA, Stanley IH, Joiner TE. Suicide mortality and coronavirus disease 2019-A perfect storm? JAMA Psychiatry 2020 Nov 01;77(11):1093-1094. [doi: 10.1001/jamapsychiatry.2020.1060] [Medline: 32275300]

XSL•FO
RenderX

7. Brooks SK, Webster RK, Smith LE, Woodland L, Wessely S, Greenberg N, et al. The psychological impact of quarantine and how to reduce it: Rapid review of the evidence. Lancet 2020 Mar 14;395(10227):912-920 [FREE Full text] [doi: 10.1016/S0140-6736(20)30460-8] [Medline: 32112714]
8. Czeisler M, Lane RI, Petrosky E, Wiley JF, Christensen A, Njai R, et al. Mental health, substance use, and suicidal ideation during the COVID-19 pandemic - United States, June 24-30, 2020. MMWR Morb Mortal Wkly Rep 2020 Aug 14;69(32):1049-1057 [FREE Full text] [doi: 10.15585/mmwr.mm6932a1] [Medline: 32790653]
9. Bruine de Bruin W, Bennett D. Relationships between initial COVID-19 risk perceptions and protective health behaviors: A national survey. Am J Prev Med 2020 Aug;59(2):157-167 [FREE Full text] [doi: 10.1016/j.amepre.2020.05.001] [Medline: 32576418]
10. Zeckhauser RJ, Viscusi WK. Risk within reason. Science 1990 May 04;248(4955):559-564. [doi: 10.1126/science.2333509] [Medline: 2333509]
11. Zeckhauser RJ, Viscusi WK. The risk management dilemma. Ann Am Acad Pol Soc Sci 2016 Sep 08;545(1):144-155. [doi: 10.1177/0002716296545001015]
12. Coronavirus Disease 2019 (COVID-19)-Associated Hospitalization Surveillance Network (COVID-NET). Centers for Disease Control and Prevention. 2020. URL: https://www.cdc.gov/coronavirus/2019-ncov/covid-data/covid-net/purpose-methods.html [accessed 2021-01-12]
13. Reese H, Iuliano AD, Patel NN, Garg S, Kim L, Silk BJ, et al. Estimated incidence of coronavirus disease 2019 (COVID-19) illness and hospitalization-United States, February-September 2020. Clin Infect Dis 2021 Jun 15;72(12):e1010-e1017 [FREE Full text] [doi: 10.1093/cid/ciaa1780] [Medline: 33237993]
14. Asch DA, Sheils NE, Islam MN, Chen Y, Werner RM, Buresh J, et al. Variation in US hospital mortality rates for patients admitted with COVID-19 during the first 6 months of the pandemic. JAMA Intern Med 2021 Apr 01;181(4):471-478 [FREE Full text] [doi: 10.1001/jamainternmed.2020.8193] [Medline: 33351068]
15. US COVID-19 cases and deaths by state. USAFacts. 2021. URL: https://usafacts.org/visualizations/coronavirus-covid-19-spread-map/ [accessed 2021-01-31]
16. Graubard BI, Korn EL. Predictive margins with survey data. Biometrics 1999 Jun;55(2):652-659. [doi: 10.1111/j.0006-341x.1999.00652.x] [Medline: 11318229]
17. Shovestul B, Han J, Germine L, Dodell-Feder D. Risk factors for loneliness: The high relative importance of age versus other factors. PLoS One 2020;15(2):e0229087 [FREE Full text] [doi: 10.1371/journal.pone.0229087] [Medline: 32045467]
18. Nock MK, Borges G, Bromet EJ, Alonso J, Angermeyer M, Beautrais A, et al. Cross-national prevalence and risk factors for suicidal ideation, plans and attempts. Br J Psychiatry 2008 Feb;192(2):98-105 [FREE Full text] [doi: 10.1192/bjp.bp.107.040113] [Medline: 18245022]
19. Han B, Compton WM, Blanco C, Colpe L, Huang L, McKeon R. National trends in the prevalence of suicidal ideation and behavior among young adults and receipt of mental health care among suicidal young adults. J Am Acad Child Adolesc Psychiatry 2018 Jan;57(1):20-27.e2. [doi: 10.1016/j.jaac.2017.10.013] [Medline: 29301664]
20. Gadarian SK, Goodman SW, Pepinsky TB. Partisanship, health behavior, and policy attitudes in the early stages of the COVID-19 pandemic. PLoS One 2021;16(4):e0249596 [FREE Full text] [doi: 10.1371/journal.pone.0249596] [Medline: 33826646]
21. Gollwitzer A, Martel C, Brady WJ, Pärnamets P, Freedman IG, Knowles ED, et al. Partisan differences in physical distancing are linked to health outcomes during the COVID-19 pandemic. Nat Hum Behav 2020 Nov;4(11):1186-1197. [doi: 10.1038/s41562-020-00977-7] [Medline: 33139897]
22. Alcaraz KI, Eddens KS, Blase JL, Diver WR, Patel AV, Teras LR, et al. Social isolation and mortality in US Black and White men and women. Am J Epidemiol 2019 Jan 01;188(1):102-109 [FREE Full text] [doi: 10.1093/aje/kwy231] [Medline: 30325407]
23. Holt-Lunstad J, Smith TB, Baker M, Harris T, Stephenson D. Loneliness and social isolation as risk factors for mortality: A meta-analytic review. Perspect Psychol Sci 2015 Mar;10(2):227-237. [doi: 10.1177/1745691614568352] [Medline: 25910392]

Abbreviations

CDC: Centers for Disease Control and Prevention
COVID-NET: COVID-19–Associated Hospitalization Surveillance Network
UCLA: University of California, Los Angeles

XSL•FO
RenderX

Edited by J Torous, G Eysenbach; submitted 03.05.21; peer-reviewed by S Six, Y Yu; comments to author 30.06.21; revised version received 05.07.21; accepted 06.07.21; published 30.08.21.

Please cite as:
Ladapo JA, Rothwell JT, Ramirez CM
Association of COVID-19 Risk Misperceptions With Household Isolation in the United States: Survey Study
JMIR Form Res 2021;5(8):e30164
URL: https://formative.jmir.org/2021/8/e30164
doi:10.2196/30164
PMID:34253507

©Joseph A Ladapo, Jonathan T Rothwell, Christina M Ramirez. Originally published in JMIR Formative Research (https://formative.jmir.org), 30.08.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Online Search Trends Influencing Anticoagulation in Patients With COVID-19: Observational Study

Amy P Worrall[1*], BA, BSc, MB BAO MCh, MSc; Claire Kelly[2*], MB BAO BCh, MSc; Aine O'Neill[1], BSc, MB BAO BCh; Murray O'Doherty[1], BA, MB BAO BCh; Eoin Kelleher[3], MB BAO BCh; Anne Marie Cushen[4], BSc, MSc; Cora McNally[1], MB BAO BCh, MRCPI; Samuel McConkey[1,5], MB BAO, MD, DMTH, MRCPI; Siobhan Glavey[2], BSc, MB BAO BCh, PhD, MRCPI; Michelle Lavin[2,6], MB BAO BCh, PhD, MRCPI; Eoghan de Barra[1,5], BSc, MB BAO BCh, MD, MRCPI

[1]Department of Infectious Diseases, Beaumont Hospital, Dublin, Ireland
[2]Department of Haematology, Beaumont Hospital, Dublin, Ireland
[3]Department of Anaesthesiology, Beaumont Hospital, Dublin, Ireland
[4]Department of Pharmacy, Beaumont Hospital, Dublin, Ireland
[5]Department of International Health and Tropical Medicine, Royal College of Surgeons Ireland, Dublin, Ireland
[6]Irish Centre for Vascular Biology, School of Pharmacy & Biomedical Sciences, Royal College of Surgeons Ireland, Dublin, Ireland
[*]these authors contributed equally

Corresponding Author:
Amy P Worrall, BA, BSc, MB BAO MCh, MSc
Department of Infectious Diseases
Beaumont Hospital
P.O. Box 1297
Beaumont Road
Dublin, 9
Ireland
Phone: 353 1 8093000
Email: worralap@tcd.ie

Abstract

Background: Early evidence of COVID-19–associated coagulopathy disseminated rapidly online during the first months of 2020, followed by clinical debate about how best to manage thrombotic risks in these patients. The rapid online spread of case reports was followed by online interim guidelines, discussions, and worldwide online searches for further information. The impact of global online search trends and online discussion on local approaches to coagulopathy in patients with COVID-19 has not been studied.

Objective: The goal of this study was to investigate the relationship between online search trends using Google Trends and the rate of appropriate venous thromboembolism (VTE) prophylaxis and anticoagulation therapy in a cohort of patients with COVID-19 admitted to a tertiary hospital in Ireland.

Methods: A retrospective audit of anticoagulation therapy and VTE prophylaxis among patients with COVID-19 who were admitted to a tertiary hospital was conducted between February 29 and May 31, 2020. Worldwide Google search trends of the term “COVID-19” and anticoagulation synonyms during this time period were determined and correlated against one another using a Spearman correlation. A P value of <.05 was considered significant, and analysis was completed using Prism, version 8 (GraphPad).

Results: A statistically significant Spearman correlation (P<.001, r=0.71) was found between the two data sets, showing an increase in VTE prophylaxis in patients with COVID-19 with increasing online searches worldwide. This represents a proxy for online searches and discussion, dissemination of information, and Google search trends relating to COVID-19 and clotting risk, in particular, which correlated with an increasing trend of providing thromboprophylaxis and anticoagulation therapy to patients with COVID-19 in our tertiary center.

Conclusions: We described a correlation of local change in clinical practice with worldwide online dialogue and digital search trends that influenced individual clinicians, prior to the publication of formal guidelines or a local quality-improvement intervention.

XSL•FO
RenderX

(JMIR Form Res 2021;5(8):e21817) doi:10.2196/21817

KEYWORDS

COVID-19; coronavirus; online search engines; anticoagulation; thrombosis; online influence; health information dissemination

Introduction

Since late 2019, the knowledge of the clinical sequelae of COVID-19 has increased largely through the rapid dissemination of information through various information platforms. In addition to asymptomatic infection, SARS-CoV-2 can cause a broad range of symptoms, from mild coryzal symptoms to neurological and gastrointestinal presentations and, most worryingly, severe acute respiratory failure. In the early months of 2020, reports of a high incidence of COVID-19–related coagulopathy circulated online, with anecdotal evidence and case reports emerging initially. This was soon followed by marked increases in clinical thrombosis, including deep vein thromboses, pulmonary emboli [1,2], and microthrombi in pulmonary vasculature in postmortem pathological studies [3]. Clinical debate ensued surrounding the best strategy to prevent and treat COVID-19–associated coagulopathy. By the middle of April 2020, an international position paper recommended the provision of adequate thromboprophylaxis in hospitalized patients with COVID-19 [4]. Debate began on the use of therapeutic anticoagulation in the absence of confirmed thrombosis, or an intermediary dose of low-molecular-weight heparin as a thromboprophylaxis in this cohort. The risk-benefit ratio of bleeding versus clotting had to be taken into consideration, especially as some patients displayed signs of disseminated intravascular coagulopathy. However, the role of online search trends and online health information tools in the spread of this information has yet to be assessed.

The COVID-19 pandemic has rapidly evolved, spreading throughout the globe from its origin in China. The growing need to quickly report clinical findings, guidance, and recommendations became paramount as the seriousness of the situation unfolded. Chinese reports and studies were soon followed by those of our Italian colleagues. Italy became the first country to provide a European angle on treatment and management of COVID-19. Dramatic changes occurred worldwide, forcing the public into the safety of their homes. Social distancing measures, including quarantine, business closures, and travel bans, meant the ability for academic, clinical, and public discourse was limited. In order to gain information on COVID-19, the most recent updates, and suggested management, clinicians began engaging with online and social media platforms, ones that have already been described in the literature [5].

Infodemiology studies have already shown that search trends are useful parameters in the measurement of this pandemic; for example, search trends for symptoms correlating with disease outbreak [6]. Similarly, infodemiology metrics for search trends on Google across European countries showed a strong correlation between COVID-19 cases locally and worldwide, identifying new avenues for online disease surveillance and response efforts [7].

Here, we investigate the power of search trends as representative of discussion and dissemination of clinical information and recommendations online as well as how trending themes and discussion have influenced clinical practice in our local center.

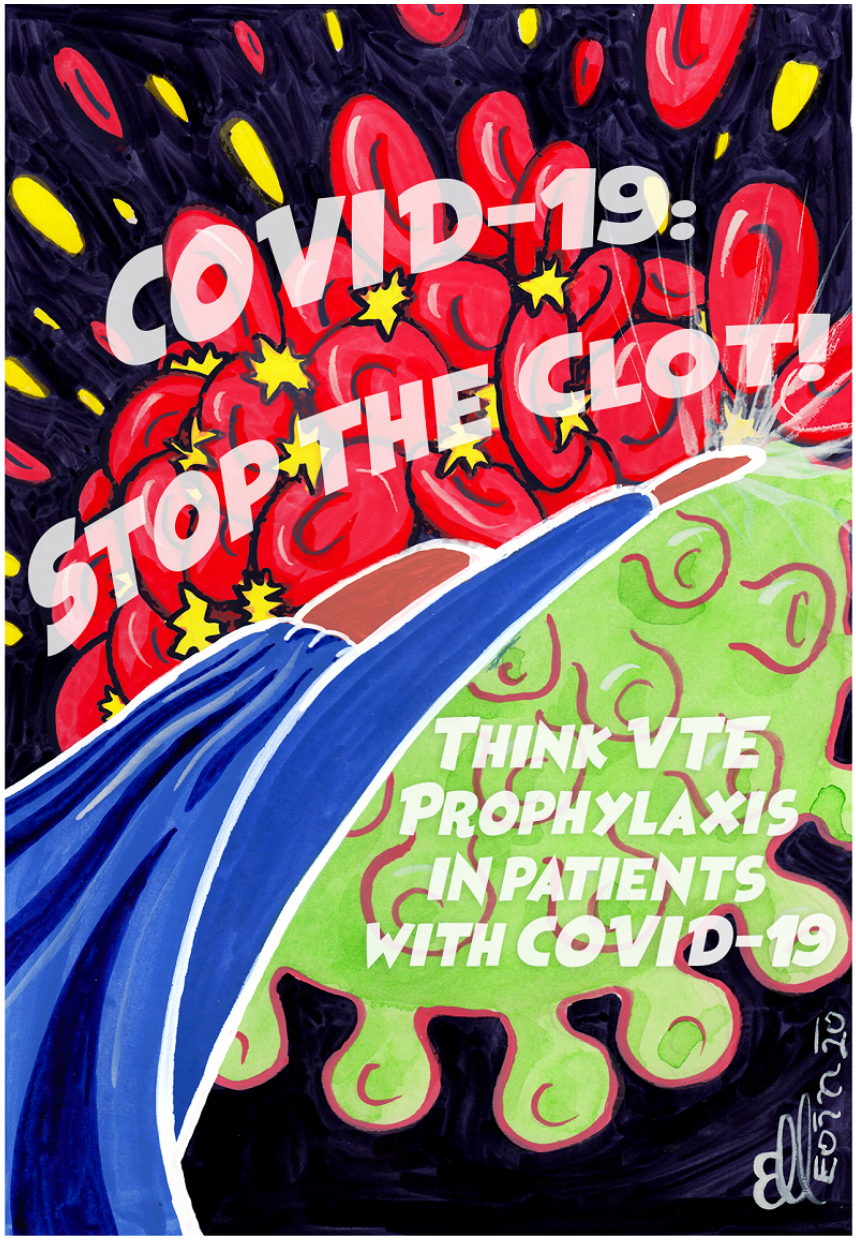
Methods

We conducted a retrospective audit of patients who tested positive for COVID-19 by real-time reverse transcription polymerase chain reaction at Beaumont Hospital in Dublin, Ireland, a tertiary 820-bed hospital, from March 1 to May 31, 2020. The audit received approval from the audit department and the research ethics committee of Beaumont Hospital. All patients were followed until the June 20, 2020. Data regarding venous thromboembolism (VTE) risk was assessed using the Padua Prediction Score system. The appropriate prescribing of thromboprophylaxis or anticoagulation therapy within 24 hours of admission or 24 hours of a positive COVID-19 result were collected. Three local interventions in the hospital were noted as key events that influenced anticoagulation and thromboprophylaxis guidelines for patients with COVID-19. These three interventions included the following: a COVID-19 teaching session by the coagulation specialist, a consultant hematologist (attending); a hospital-wide email with VTE prophylaxis guidelines; and hospital-wide circulation of an infographic about VTE prophylaxis in patients with COVID-19 (Figure 1).

Data from Google Trends were retrieved online in comma-separated values (CSV) format and used to compile a representative trend of worldwide searches for “COVID-19” and synonyms (eg, “coronavirus” and “covid”), together with the following terms: “anticoagulation,” “VTE prophylaxis,” “thrombosis,” “clots,” and “clotting.” The search results were then summated. Data collection included results from February 29 until May 31, 2020. The data were worldwide data, rather than European or Irish Google Trends data, as Google search results at the time were very scarce in individual countries. The Google Trends data were being used to infer online activity on numerous social media platforms during the early period of the COVID-19 pandemic. Data were analyzed using Prism, version 8 (GraphPad), according to nonparametric Spearman correlations. A P value of <.05 was considered statistically significant.

XSL•FO
RenderX

Figure 1. Local infographic used to encourage venous thromboembolism (VTE) prophylaxis in patients with COVID-19, designed by Dr Eoin Kelleher (@eoinkr) and circulated in print and on digital platforms throughout our tertiary center.



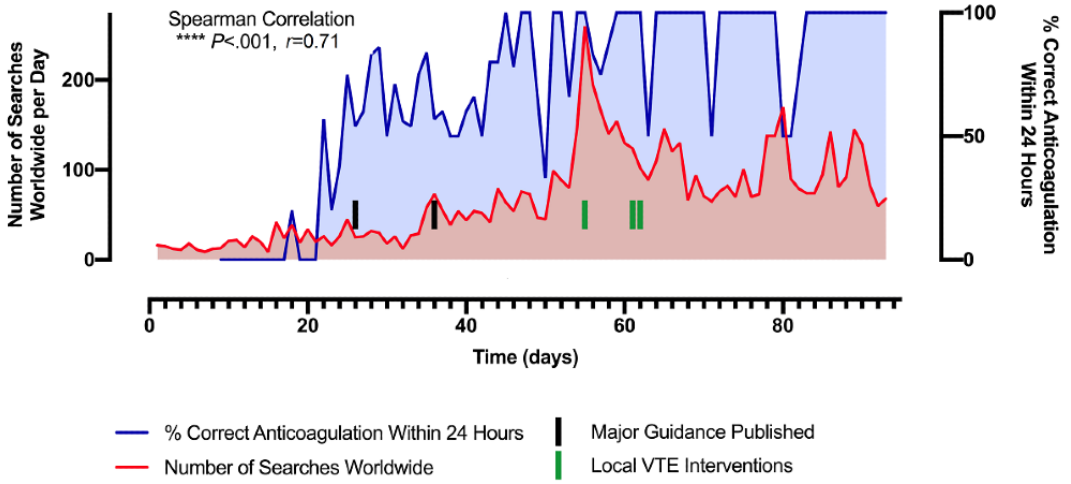
Results

A total of 399 patients consecutively diagnosed with COVID-19 were reviewed during the study period. The median age of the patients was 70 years (IQR 27, range 21-99), the majority were male (247/399, 61.9%) and Caucasian (360/399, 90.2%), and 81.5% (325/399) had underlying comorbidities. On admission, patients were assessed for thrombotic and bleeding risks. The median Padua Prediction Score was 4 (IQR 4, range 1-12). A total of 14.3% of patients (57/399) were on anticoagulation therapy for thrombotic or cardiac disorders on admission. A total of 55.1% of patients (220/399) had commenced standard thromboprophylaxis doses of heparin within the first 24 hours of admission or within 24 hours of a positive COVID-19 result. A total of 89.1% of anticoagulation prescriptions (196/220) were correctly adjusted for BMI or renal function within the same admission time frame.

A total of 6361 individual searches worldwide were collated from Google Trends over the study period relating to the COVID-19 pandemic and thrombosis. The percentage of patients with COVID-19 who were on appropriate anticoagulation therapy was graphed as a percentage of the total number of patients during the admission period (Figure 2 [4,8]). A statistically significant Spearman correlation ($P<.001$, $r=0.71$) was found between the two data sets. Our study demonstrated that online searches and discussion, dissemination of information, and Google search trends relating to COVID-19 and clotting risk, in particular, correlated with an increasing trend of providing thromboprophylaxis and anticoagulation therapy to patients with COVID-19 in our tertiary center. Following the publication of two major guideline papers [4,8] (Figure 2) and three local VTE interventions, there was significant improvement in VTE prophylaxis among our cohort of patients with COVID-19 infection (Figure 2).

XSL•FO
RenderX

Figure 2. Run chart of daily online searches relating to COVID-19 and anticoagulation (blue, left axis) and the percentage of correct anticoagulation medications administered to the cohort of patients with COVID-19 (red, right axis). Black dashes indicate the publication of two major international, interim, guidance papers for patients with COVID-19 [4,8], and green dashes indicate three local venous thromboembolism (VTE) interventions in our tertiary center.



Discussion

Principal Findings

Mavragani and colleagues demonstrated that Google Trends search queries correlated strongly with COVID-19 cases and COVID-19 deaths, but also that their correlations were most accurate during the initial months of a region's outbreak [7]. Our correlation similarly reflected search trends that correlated with a change in clinical practice locally. However, in 2015, Narayanaswami et al published a longitudinal study of the effects of traditional versus novel information dissemination (online resources, social media platforms, etc) on implementing and disseminating clinical guidelines [9]. They found no additional benefit from online resources in increasing awareness and implementation of clinical guidelines, though we hypothesize that this question may need revisiting if considering the developments in online resources since 2015 and the unique global impact and increased reliance on online assets during the COVID-19 pandemic.

Online learning platforms, information and news sources, as well as social media do impact and influence a subgroup of physicians that use those resources [5], and it is not outlandish to suggest that the number of clinicians engaging in online health information discussions has increased during the COVID-19 pandemic. The consequences of this can be positive; for example, the use of our local VTE prophylaxis increased significantly within our cohort prior to the three local VTE anticoagulation interventions, most likely secondary to two major interim guidance documents being published and significant international online discussion surrounding COVID-19 coagulopathy. However, as seen with the rapid and somewhat premature use of hydroxychloroquine in patients with COVID-19 or the hasty alarm surrounding ibuprofen use in COVID-19 cohorts, in an evolving pandemic, knee-jerk medical management and interventions can occur without stringent scientific or medical evidence to back up these actions. These actions are often a result of the rapid dissemination of data and research that were in the preprint stage and had not undergone full peer review [10,11]. We must acknowledge that the influence from online search trends, online social media platform discussions, and dissemination of information is not always reliable and can also be harmful [10].

While social media and online health information sources have almost always universally been appreciated for their potential, they have often not been successfully implemented as key components of public health strategy [12]. Only a small minority of medical and public health researchers would have actively engaged in this space in a professional capacity, as social media has been seen as a means of disseminating information rather than obtaining it [12]. The COVID-19 pandemic has reshaped the use of online platforms from online tools, search engines, and social media sites as potential tools of public health strategy, but also for disease surveillance and monitoring and, indeed, for more immediate transmission of acute changes in disease management [7]. A balance must be struck between the lethargic delay in translational medicine and the impulsive dissemination of as yet–unproven research and information dissemination, in particular during a global health crisis.

Limitations

This study is limited by the single local center analyzed in Ireland. The search terms and trends utilized were from one social media platform and can only provide a tentative inference of impact on practice. The Google Trends data were only collected on a worldwide sample, as the number of local and regional searches were too small for appropriate analysis. We also note that values from Google Trends are not absolute values of search results, but rather normalized values. However, the use of Google Trends in this study was used to represent a general trend of online discussion that occurred online during the early months of the COVID-19 pandemic, whereby online platforms, such as Google, Twitter, and Facebook, provided platforms for swift data dissemination for health practitioners [13].

Conclusions

In this paper, we described a phenomenon of local change in clinical practice following worldwide online conversation and digital search trends that influenced individual clinicians before

XSL•FO
RenderX

the formation of formal clinical guidelines, with tentative natural improvement and significant improvement following a quality-improvement intervention.

Authors' Contributions

APW, CK, AO'N, MO'D, and ML collected the data. APW and CK analyzed the data. All authors were involved in writing and reviewing the paper. EK designed and kindly provided the graphic design for the local "Think VTE Prophylaxis in Patients With COVID-19" intervention in Beaumont Hospital.

Conflicts of Interest

None declared.

References

1. Middeldorp S, Coppens M, van Haaps TF, Foppen M, Vlaar AP, Müller MCA, et al. Incidence of venous thromboembolism in hospitalized patients with COVID-19. J Thromb Haemost 2020 Aug;18(8):1995-2002 [FREE Full text] [doi: 10.1111/jth.14888] [Medline: 32369666]
2. Helms J, Tacquard C, Severac F, Leonard-Lorant I, Ohana M, Delabranche X, CRICS TRIGGERSEP Group (Clinical Research in Intensive Care and Sepsis Trial Group for Global Evaluation and Research in Sepsis). High risk of thrombosis in patients with severe SARS-CoV-2 infection: A multicenter prospective cohort study. Intensive Care Med 2020 Jun;46(6):1089-1098 [FREE Full text] [doi: 10.1007/s00134-020-06062-x] [Medline: 32367170]
3. Fox SE, Akmatbekov A, Harbert JL, Li G, Quincy Brown J, Vander Heide RS. Pulmonary and cardiac pathology in African American patients with COVID-19: An autopsy series from New Orleans. Lancet Respir Med 2020 Jul;8(7):681-686 [FREE Full text] [doi: 10.1016/S2213-2600(20)30243-5] [Medline: 32473124]
4. Spyropoulos AC, Levy JH, Ageno W, Connors JM, Hunt BJ, Iba T, Subcommittee on Perioperative, Critical Care Thrombosis, Haemostasis of the Scientific, Standardization Committee of the International Society on Thrombosis and Haemostasis. Scientific and Standardization Committee communication: Clinical guidance on the diagnosis, prevention, and treatment of venous thromboembolism in hospitalized patients with COVID-19. J Thromb Haemost 2020 Aug;18(8):1859-1865 [FREE Full text] [doi: 10.1111/jth.14929] [Medline: 32459046]
5. Chan T, Trueger NS, Roland D, Thoma B. Evidence-based medicine in the era of social media: Scholarly engagement through participation and online interaction. CJEM 2018 Jan;20(1):3-8. [doi: 10.1017/cem.2016.407] [Medline: 28077195]
6. Higgins TS, Wu AW, Sharma D, Illing EA, Rubel K, Ting JY, Snot Force Alliance. Correlations of online search engine trends with coronavirus disease (COVID-19) incidence: Infodemiology study. JMIR Public Health Surveill 2020 May 21;6(2):e19702 [FREE Full text] [doi: 10.2196/19702] [Medline: 32401211]
7. Mavragani A. Tracking COVID-19 in Europe: Infodemiology approach. JMIR Public Health Surveill 2020 Apr 20;6(2):e18941 [FREE Full text] [doi: 10.2196/18941] [Medline: 32250957]
8. Barrett CD, Moore HB, Yaffe MB, Moore EE. ISTH interim guidance on recognition and management of coagulopathy in COVID-19: A comment. J Thromb Haemost 2020 Aug;18(8):2060-2063. [doi: 10.1111/jth.14860] [Medline: 32302462]
9. Narayanaswami P, Gronseth G, Dubinsky R, Penfold-Murray R, Cox J, Bever C, et al. The impact of social media on dissemination and implementation of clinical practice guidelines: A longitudinal observational study. J Med Internet Res 2015 Aug 13;17(8):e193 [FREE Full text] [doi: 10.2196/jmir.4414] [Medline: 26272267]
10. Ferner RE, Aronson JK. Chloroquine and hydroxychloroquine in COVID-19. BMJ 2020 Apr 08;369:m1432. [doi: 10.1136/bmj.m1432] [Medline: 32269046]
11. Carius BM, Schauer SG. Ibuprofen during the COVID-19 pandemic: Social media precautions and implications. West J Emerg Med 2020 Apr 24;21(3):497-498 [FREE Full text] [doi: 10.5811/westjem.2020.4.47686] [Medline: 32421496]
12. Keller B, Labrique A, Jain KM, Pekosz A, Levine O. Mind the gap: Social media engagement by public health researchers. J Med Internet Res 2014 Jan 14;16(1):e8 [FREE Full text] [doi: 10.2196/jmir.2982] [Medline: 24425670]
13. Chan AKM, Nickson CP, Rudolph JW, Lee A, Joynt GM. Social media for rapid knowledge dissemination: Early experience from the COVID-19 pandemic. Anaesthesia 2020 Dec;75(12):1579-1582 [FREE Full text] [doi: 10.1111/anae.15057] [Medline: 32227594]

Abbreviations

CSV: comma-separated values
VTE: venous thromboembolism

XSL•FO
RenderX

Edited by G Eysenbach; submitted 13.08.20; peer-reviewed by T Chan, A Mavragani; comments to author 19.09.20; revised version received 11.04.21; accepted 11.04.21; published 31.08.21.

Please cite as:
Worrall AP, Kelly C, O'Neill A, O'Doherty M, Kelleher E, Cushen AM, McNally C, McConkey S, Glavey S, Lavin M, de Barra E
Online Search Trends Influencing Anticoagulation in Patients With COVID-19: Observational Study
JMIR Form Res 2021;5(8):e21817
URL: https://formative.jmir.org/2021/8/e21817
doi:10.2196/21817
PMID:

©Amy P Worrall, Claire Kelly, Aine O'Neill, Murray O'Doherty, Eoin Kelleher, Anne Marie Cushen, Cora McNally, Samuel McConkey, Siobhan Glavey, Michelle Lavin, Eoghan de Barra. Originally published in JMIR Formative Research (https://formative.jmir.org), 31.08.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

XSL•FO
RenderX