JMIR Formative Research

Impact Factor (2023): 2.0
Volume 6 (2022), Issue 11 ISSN 2561-326X Editor in Chief: Amaryllis Mavragani, PhDc

Contents

Original Papers

Viewpoint

Review

Original Paper

Web-Based System Navigation Database to Support Equitable Access to Assistive Technology: Usability Testing Study

Tamika Jarvis[1*], MPH; Allison M L Mah[2*], MScOT; Rosalie H Wang[2*], BScOT, PhD; Michael G Wilson[1,3,4*], PhD

[1]Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, ON, Canada

[2]Department of Occupational Science & Occupational Therapy, University of Toronto, Toronto, ON, Canada

[3]McMaster Health Forum, McMaster University, Hamilton, ON, Canada

[4]Centre for Health Economics and Policy Analysis, Hamilton, ON, Canada

[*]all authors contributed equally

Corresponding Author:
Michael G Wilson, PhD
McMaster Health Forum
McMaster University
1280 Main Street West
Hamilton, ON, MML-417
Canada
Phone: 1 905 525 9140 ext 22121
Email: wilsom2@mcmaster.ca

Abstract

Background: Assistive technology (AT) can contribute to how individuals participate and engage in everyday activities, such as communication and mobility, and facilitates access to the services they require. Navigating Canada's AT system has been described as fragmented and complex, presenting barriers for individuals who require AT, caregivers, and health service providers. AccessATCanada was developed as a centralized web-based resource to help support access to AT by providing information about the existing jurisdictional funding programs and services.

Objective: This study aimed to evaluate the usability of AccessATCanada by gathering feedback about its features, functionality, and areas of strength and opportunity from potential end users.

Methods: A usability testing study using a think-aloud approach and semistructured interviews was conducted to measure the effectiveness and efficiency of and user satisfaction with AccessATCanada and to identify issues with the interface during end-user interaction. A qualitative thematic analysis was used to generate insights into and core themes about user experiences. User feedback was used to inform subsequent updates of the database with the goal of enhancing website friendliness and functionality before its official launch.

Results: A total of 10 participants (6 consumers, 1 caregiver, and 3 providers) participated in the usability testing study. The usability performance and scores tended to improve between the 2 testing cycles. Most participants were able to successfully complete all the tasks independently. The efficiency scores tended to improve as the users continued to engage with the interface. The website received an overall System Usability Score of 62.22, which was ranked as "OK/fair to good." The users provided an overall positive evaluation of the beta version of the web-based resource tested over 2 cycles and helped to identify areas for improvement. They commented on the functionality and added value of the website, discovery of new programs and resources, and design aesthetics. Most usability issues were reported as minor challenges related to presentation, functionality, and language, and feedback was adopted into later iterations of the website.

Conclusions: This study provides reflections on the value of usability testing and elements that are key to the creation of user-centered resources, such as the inclusion of participants with various abilities and considerations regarding website design and accessibility in an increasingly web-based world. AccessATCanada is now part of a growing global response to expand the reach of AT programs and services, improve the equity of access to AT, and reduce the complexity of navigating AT systems.

(JMIR Form Res 2022;6(11):e36949) doi:10.2196/36949

KEYWORDS

assistive technology; program funding; usability testing; internet; web-based database; health services

Introduction

Background

Assistive technology (AT) contributes to how an individual participates and engages in their everyday activities, such as communication and mobility, and facilitates access to the services that they require [1]. Access to AT, which promotes the inclusion, engagement, and participation of the world's growing aging population and population of persons with disabilities, remains challenging. According to the World Health Organization, an estimated 90% of individuals who are in need of AT do not have access to it [2]. As noted in the "Global Report on Assistive Technology," there are 3 phases in the pathway to access AT: seeking, obtaining, and realizing [3]. Seeking encompasses the first steps taken to enhance access to AT and necessitates that consumers, caregivers, and providers are informed about the available AT and can find and obtain related information. Successful access to information provides a foundation for individuals to continue accessing appropriate AT, regardless of whether the need is acute or long term.

More than 220 government and charitable organizational programs that provide funding and services for AT exist across Canada [1]. Access to AT programs has been described as fragmented and complex to navigate and uncoordinated between national, subnational, and local levels [4]. Adding to this complexity is the fact that programs are highly variable regarding the range of ATs that can be covered, eligibility criteria, and service-delivery mechanisms used [5]. Some Canadian jurisdictions, such as British Columbia, have introduced "one-stop" approaches or single-entry point systems, which are typically organizations that perform a range of activities, such as assessments, training, and AT, as well as manage access to funding sources [6]. However, these systems are far from being universal. Despite the many programs available for AT, poor "consistency in the quality and quantity of AT information," the high cost of AT, and the lack of governmental funding support create difficulties with navigating programs and make the acquisition of AT challenging for those who need them. In addition, the lack of training presents significant barriers for health service providers in recommending appropriate AT [7].

The lack of a user-friendly system is a challenge that needs to be addressed in a way that is efficient, effective, and satisfactory to help reduce the impact of inequitable access to needed supports. These findings motivated the development of AccessATCanada, a web-based resource designed to be easily searchable by various users with different abilities. Program information for the database was initially gathered through a jurisdictional scan to identify the types of AT covered under jurisdictional programs and funding, eligibility criteria for AT programs, and currently available AT funding and services [1]. The database contents and website were updated before launch in January 2021. The website was created in compliance with Web Content Accessibility Guidelines (WCAG) 2.0 by a professional website development company.

The Home page of the website allows users to search by keywords and has filters to search by jurisdiction, and AT types. It also features an interactive map of Canada, which allows users to search by province or territory (Figure 1). In addition, a feedback form is provided at the bottom of the page to allow users to report errors. The menu at the top of the Home page directs users to information about the website and collaborators (About Us page); a Programs page for users to find AT programs; a Resources page, which provides reports and publications related to AT in Canada; and a Contact Us page, where users are able to contact the project leads and provide feedback about the website.

Figure 1. AccessATCanada Home page.

Objectives

The objective of this study was to examine the usability of the website and the interaction of a representative set of users with the website to further develop and refine a user-friendly resource that meets the needs of the end users. The process used to gather feedback, identify elements considered important to users, and refine the AccessATCanada website can help further enhance this resource and help guide those looking to create similar resources [8].

Methods

This usability testing study used a think-aloud approach to gather feedback from prospective users related to effectiveness, user efficiency, and satisfaction with the website and to identify any potential issues with the interface [9].

Recruitment

A purposive sample of participants from 3 groups of anticipated users were recruited: (1) consumers, defined as individuals who self-identified as living with at least one type of disability (eg, a physical, sensory, cognitive, or mental health limitation) and who may benefit from AT use; (2) caregivers, who identified as individuals who cared for someone who required or used AT; and (3) health care providers, community social services providers, and industry vendors, who identified as individuals who assist others in accessing AT. Although the aim was to recruit approximately 5 individuals from each user group, the authors found that there was a sufficient amount of information gathered to make immediate updates to the website with a smaller number of participants. The participants were required to have internet access via a desktop or laptop computer with screen sharing to capture immediate responses when using the website and for the researchers to assist with troubleshooting as needed. The participants were recruited through project partners who shared recruitment advertisements with those who used their services and their networks and informal networks of service providers.

Ethics Approval

The study was approved by the Hamilton Integrated Research Ethics Board (REB-8325) and the University of Toronto Research Ethics Board (REB-38715). Informed consent was provided by all the participants.

Data Collection

Overall, 2 iterative cycles of task-based usability testing evaluations were conducted using an unreleased beta version of the AccessATCanada website. A total of 21 tasks were given to the participants, across 4 mock case scenarios (Multimedia Appendix 1). The usability testing sessions were facilitated by 2 members of the research team (AMLM and TJ). The usability testing sessions were conducted individually with each participant, lasting approximately 60 to 90 minutes in duration and audio-visually recorded through institutionally licensed WebEx (Cisco Systems, Inc) and Zoom (Zoom Video Communications) videoconferencing software. Figure 2

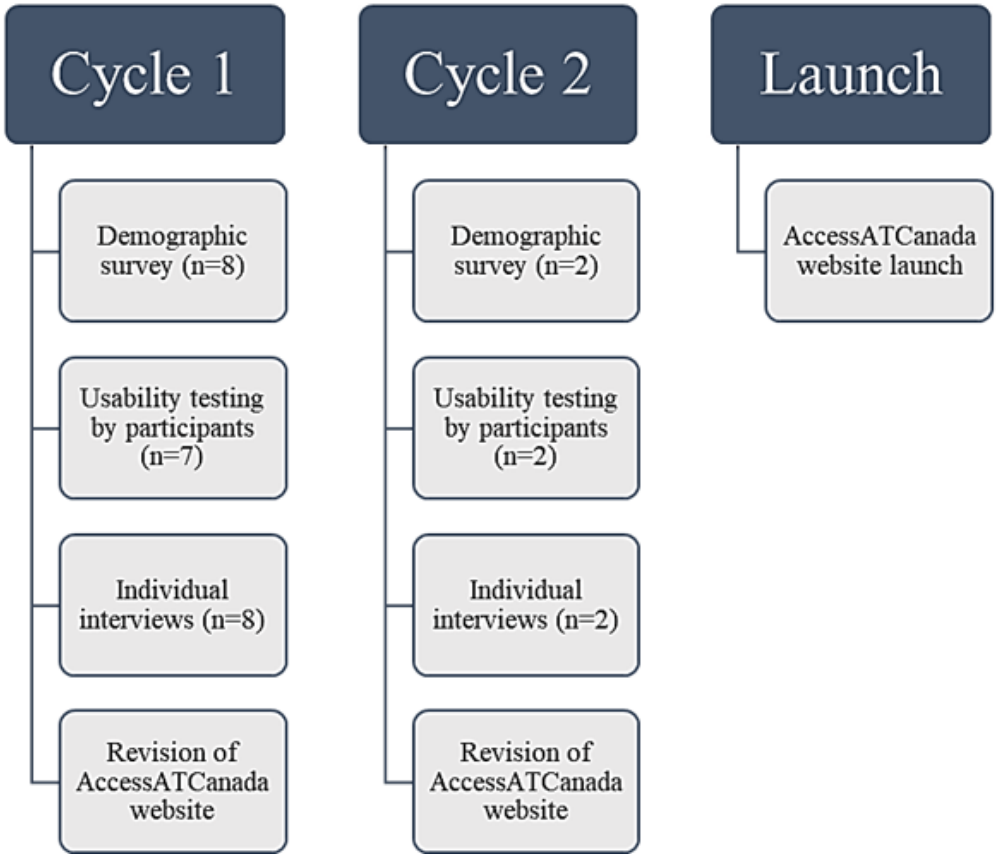
illustrates the cycles of usability testing and iterative refinement of the website. Usability testing included the completion of mock case scenarios, System Usability Scale (SUS), and a brief semistructured interview [10].

Before usability testing, the participants were asked to complete a brief demographic survey and to self-identify as a consumer, caregiver, or health care provider. Mock case scenarios were pilot tested with a member of the research team before data collection. They were written with the goal of obtaining information from diverse content searches within the website and required the participants to explore and use different website functions. The participants were asked to explore different provincial and territorial AT programs, navigate between government and charitable programs, find eligibility criteria for a funding service, and locate programs that funded specific AT needs. The scenarios and tasks were read aloud by the researcher in the same sequence for each participant. Notably, some tasks were performed out of order or were completed simultaneously by some participants unprompted; therefore, these tasks were skipped to avoid task repetition. The participants were prompted to think aloud and verbalize their actions and thoughts as they interacted with the system, which allowed for observation and real-time feedback [11]. To help identify areas for improvement, usability issues and errors (eg, issues related to functionality, presentation, and language and events that impacted the ability to use the website effectively and efficiently) were noted during testing sessions [12,13]. Approximately 60 minutes were allotted to completing the 4 mock scenarios; however, not all participants completed the tasks because of time constraints or technological challenges that arose during user testing sessions.

Following the completion of the mock scenarios, each participant was asked to complete the SUS. The SUS is a widely used and highly rated user-centered questionnaire that includes questions related to the learnability and complexity of and satisfaction with website use [10,14]. The SUS measures user comfort, satisfaction, and perception of usability of the website by verbally ranking their agreement with 10 statements on a 5-point Likert scale (Multimedia Appendix 2). After the completion of the SUS, a brief semistructured individual interview was conducted to elicit further clarification and elaboration of participants' responses and experience while using the website, elaborate on their SUS ratings, and describe the website features that they liked the best and least. The interview guide can be found in Multimedia Appendix 3.

Participants' observations and comments were compiled into interim summary reports between each usability testing cycle, which informed and optimized the next website iteration. A final report was then compiled based on the feedback provided by the second cycle of participants, which informed the recommended changes before the official launch of the website.

Figure 2. Diagram depicting the cycles of the usability testing procedure and website development.



Data Analysis

The assessment parameters and metrics of effectiveness, efficiency, and satisfaction were guided by the International Organization for Standardization 9241-11 [9]. All the tasks attempted were included in the data analysis.

Effectiveness was calculated and defined as the successful completion of tasks and scenarios [15]. The tasks that the

participants were unable to complete or were completed with the assistance of the researcher were graded as "failure (not completed)" and given a score of 0, and the tasks that were completed with ease or with no assistance were graded as "success" and given a score of 1. The average effectiveness was calculated based on task failures or full successes among the attempted tasks. Tasks that were not attempted by the user were excluded from the data analysis. Usability problems were defined if less than 70% of the participants were able to complete the tasks successfully [15,16]. The number of errors was recorded and counted. Errors were defined as unintended actions or mistakes made while attempting tasks [17]. These included, for example, the number of times the participants used the "back" button on their browsers or modified searches such as by removing keywords or filters.

The overall relative efficiency was assessed by the length of time (seconds) taken by the participants who successfully completed tasks and compared with the total time taken by all the participants [16,17]. A baseline target time was not established, as we expect users with a range of characteristics, such as familiarity and web-based comfort levels, to access this resource.

Satisfaction is a subjective measure of user attitudes and comfort while using a system. An overall usability problem was identified when the average SUS score was <68 or lower than the 50th percentile range [10]. To help provide a meaningful interpretation of individual and overall SUS ratings, an adjective rating scale was also used [18,19]. The interpretations of the SUS ratings are provided in the subsequent section and Multimedia Appendix 2. Semistructured interview questions were used to enhance the understanding of the participants' SUS ratings.

The severity of website usability issues and errors was recorded as users thought aloud. Severity was defined as the impact of a problem during website navigation that prevented users from completing tasks successfully or efficiently [16]. Three levels of severity were reported: minor issues that caused annoyance, but the task could still be completed; serious issues that caused frustrations and may have caused users to abandon tasks; and critical issues that, when not fixed, impeded the ability to complete tasks.

Usability testing sessions and interviews were transcribed verbatim. Using an inductive approach, audio-visual recordings from the testing sessions and facilitator notes were reviewed to extract feedback on usability and experience. NVivo 12 Pro (QSR International) was used for coding the data. Test facilitators (AMLM and TJ) read and independently coded transcripts using an iterative open coding process. The results were compared to discuss codes and patterns, and the differences were resolved by discussion.

Think-aloud and semistructured interview data were analyzed through thematic analysis [20]. Thematic analysis is a useful and flexible approach for identifying, analyzing, and reporting patterns of meaning in qualitative data. The analysis began with the review of raw transcripts, facilitator notes, and semistructured interviews. Data were inductively and iteratively analyzed and coded using open coding to *chunk* data and apply descriptors, and axial coding was used to categorize codes and develop themes from categories [21]. A preliminary list of initial codes was developed and sorted into themes and supporting subthemes. To help form patterns and further conceptualize the subject, a constant comparison of the data was performed throughout the data analysis [22]. Themes were discussed with all the members of the research team and iteratively refined. Quotations and dialogues were extracted from the participant transcripts to illustrate core themes.

Results

Overview

A total of 10 participants were recruited into the study, of which 9 (90%) participants completed both the user testing session and semistructured interview, and 1 (10%) participant was unable to complete the testing session because of technological issues but was able to provide feedback about the website through an interview. One of the users completed the user testing session by telephone after attempting and noting technical difficulties with a tablet computer, and the other users participated through a laptop or desktop computer.

Overall, 2 cycles of usability tests were conducted by 2 facilitators to capture major challenges with usability. The first cycle of user tests was conducted with 7 participants, and the second cycle was conducted with 2 participants.

Participant Characteristics

The participant demographics are presented in Table 1 (N=10). The sample consisted of 40% (4/10) of male participants and 60% (6/10) of female participants. Participants' age ranged from 25 to 74 years, with 50% (5/10) aged 35 to 44 years, representing the largest age group. A total of 6 participants were identified as consumers, 1 as a caregiver, and 3 as health care providers. The participants reported that they experienced either one or a combination of visual, hearing, touch, physical (eg, mobility limitation and fatigue), and cognitive challenges. A total of 9 participants reported that they used a computer at home, and all the participants reported feeling very comfortable with using the internet. The participants reported preferring to seek health and service information from health care professionals (8/10, 80%), the internet (6/10, 60%), support agencies (4/10, 40%), and other sources (1/10, 10%); 30% (3/10) of participants reported having used websites or databases related to AT in the previous 3 months, such as an assistive device funding program within a province.

Table 1. Demographic characteristics of the study participants (N=10).

Characteristics	Values, n (%)
Sex	
Male	4 (40)
Female	6 (60)
Other (please specify)	0 (0)
Prefer not to say	0 (0)
Age (years)	
18-24	0 (0)
25-34	2 (20)
35-44	5 (50)
45-54	1 (10)
55-64	0 (0)
65-74	2 (20)
>74	0 (0)
User type	
Consumer	6 (60)
Caregiver	1 (10)
Health care provider	3 (30)
Type of disability	
Visual	2 (20)
Hearing	2 (20)
Touch	1 (10)
Physical (eg, mobility limitation and fatigue)	2 (20)
Cognitive	1 (10)
Mental health	0 (0)
None reported	2 (20)
Use of a computer at home	
Yes	9 (90)
No	1 (10)
Comfort level with using the internet	
Not at all comfortable	0 (0)
A little comfortable	0 (0)
Comfortable	0 (0)
Very comfortable	10 (100)
Preferred methods for seeking health or service information	
Health care professional	8 (80)
Internet	6 (60)
Support agencies (eg, governmental and nonprofit)	4 (40)
Other (please specify)	1 (10)
Used web-based health resources related to AT[a]<3 months	
Yes	3 (30)
No	7 (70)

[a]AT: assistive technology.

Usability Evaluation Findings

Task Performance Measures of Effectiveness

We measured task performance based on the ease of navigating the site for the participants unfamiliar with the website and the number of errors made (Multimedia Appendix 4). Participants attempted 164 tasks across all scenarios, with 120 tasks being successfully completed. A total of 175 errors were recorded. Errors included the number of times the participants restarted their search, removed keywords or filters after a search returned 0 results, and used the "back" button to return to the previous webpage.

In summary, 9 tasks were completed easily by the participants and had low error rates. These were tasks that asked participants to find information on government AT programs, government-funded programs, and charitable funding programs. The participants had the highest number of errors when performing the first task within each set of scenarios and when identifying charity programs within provinces. In general, some features were not initially obvious to the participants. For example, a participant was not aware that they could conduct a keyword search on the "Home" page and suggested including text to indicate this feature. Some participants reported that because this feature did not operate as expected (eg, did not return any matching programs when using keyword searching as they would using an internet search engine), they did not feel confident using the keyword search feature and that their frustration would likely result in them abandoning the use of the database and returning to methods that they were already familiar with; for example, using other search engines.

Task Performance Measures of Efficiency

Efficiency scores ranged from 29% (2/7) of participants successfully completing tasks with ease, to 100% (8/8) of participants completing tasks with ease. Of all tasks successfully attempted, only 4 tasks had an efficiency score of 100%. The efficiency scores tended to improve as the users continued to engage with the interface. As reflected by the overall scores and comments provided, the facilitators observed that the first few tasks across each scenario took participants longer to successfully complete than other tasks. Across scenarios, the participants were less efficient at finding charitable AT programs. For example, only 29% (2/7) of participants were successful in finding a charity program related to hearing impairment services in the Yukon Territory, without assistance. In another example, when asked to find a charity program in Ontario, only 63% (5/8) of participants were able to complete this task without assistance. Searching presented a challenge for the participants because they did not know what search terms to use or which filters to use or were confused by the language used on the website. Facilitators observed that the keyword search functions on the "Home" and main "Programs" pages were a source of frustration for many participants, who commented that this feature appeared to be less integrated and inconsistent in returning results.

Task Performance Measures of Perceived Satisfaction

The SUS scores from both the testing cycles are presented in Table 2. Scores >68 indicate average usability. The overall SUS score for this study was 62.22 or "OK/fair", as described by Bangor et al [18,19]. Most participants provided an SUS rating described as fair or higher. The first cycle received an average SUS score of 59.3 (SD 17.48; OK or fair), with 29% (2/7) of participants providing a rating described as poor or lower. After cycle 1, revisions addressing usability issues were made to the website, specifically addressing areas related to presentation, functionality, and language. Furthermore, 22% (2/9) of additional participants were involved in user testing after these updates and provided an average SUS score of 72.5 (SD 3.54; good), indicating an improvement in comfort with the website.

During the 2 cycles of user testing, the severity of website usability issues and errors were reported, and feedback was summarized into 3 areas: presentation (eg, visual difficulties or issues when navigating the website), functionality (eg, issues impacting the ability to use or navigate the website), and language (eg, messages or meanings that users had a difficult time understanding or interpreting). Textbox 1 provides a summary of the major modifications made because of the user observations and feedback.

Several users provided suggestions for refining the design elements of the website to enhance intuitiveness and accessibility, which were adopted into subsequent iterations. This feedback was related to textual spacing, font size, and color contrast in accordance with the Canadian National Institute for the Blind and WCAG 2.0 accessibility guidelines, placement of the search filters, and refinement of the list of displayed results. A participant described that they sometimes felt overwhelmed by the amount of information presented on the website and returned searches; for example, the formatting of the "Programs" page and the result list was commented on the most by several users. Specifically, users provided suggestions for where to place prompts to reduce scrolling, such as "displaying number of matching results," whereas other participants suggested additional text to help instruct users to locate returned results. Other suggestions to improve the result list included adding additional cues, such as indents and borders between results to indicate the separation of results, particularly as the length of some program names, while detailed, could present issues with readability, and improving the color contrast on the filter function to indicate when filters and subfilters have been selected. A user stated the following about overlapping text and illustrations:

> *It is not user friendly...because it forces you to distinguish, to make effort in distinguishing between various layers. And if you have vision loss, that is an additional effort that you shouldn't have to make.*

Most challenges with the functionality of the website were considered minor issues largely related to (1) the use of the filters (eg, leading to increased scrolling in the menu on the "Programs" page), (2) the inability to select a filter if a specific area was not selected, and (3) the unreliable functionality of the

"clear all" feature. Feedback on some of the language or terminology used on the website was also provided, such as recommending using less technical language. For example, one of the participants said the following:

> *I question if everyone would understand what the "program," "assistive technology" mean. [M]aybe it would be better to have a description on that.*

Table 2. System usability scale (SUS) scores and corresponding grades.

SUS scores	Percentile range	Adjective
Cycle 1: individual scores		
87.5	96-100	Best imaginable
50	N/A[a]	Poor
62.5	15-34	OK or fair
65	41-59	OK or fair
55	N/A	OK or fair
30	N/A	Worst imaginable
65	41-59	OK or fair
Cycle 1 average score, mean (SD)		
59.3 (17.48)	15-34	OK or fair
Cycle 2: individual scores		
75	70-79	Good
70	41-59	Good
Cycle 2 average score, mean (SD)		
72.5 (3.54)	60-64	Good
Overall score (cycles 1 and 2), mean (SD)		
62.2 (16.27)	15-34	OK or fair

[a]N/A: not applicable.

Textbox 1. Summary of major modifications made to AccessATCanada.

Presentation

- Included instructions on how to conduct a search on the "Home" page
- Improved font contrast and size across the website
- Minimized the instances where text overlapped graphics to improve readability
- Included additional labels and content description for the filters on the "Programs" page
- Made the selected filters more apparent by improving highlighting
- Improved readability by increasing spacing and adding dividers between search results
- Improved the presentation of the search results to reduce scrolling
- Made the search button more intuitive by repositioning it under the search bar rather than at the end of the filters list
- Displayed the total number of matching results at the top of the search results list

Functionality

- Addressed the issues with the *clear all* and *search* functionalities on the "Programs" page
- Made the bars of the filters one selection area to enable users to click anywhere to make their selection
- Improved the query process to fix the issue of high filter sensitivity

Language

- Website reviewed for the use of lay language
- Continual efforts made to review the resource pages and make information less ambiguous

Themes

The key usability findings from the think-aloud and semistructured interviews were organized into the following themes: functionality and added value, the discovery of new programs and resources, and design aesthetics.

Functionality and Added Value

Overall, the participants commented that the website had features that were positive, "straightforward," "intuitive," and "easy to learn." The filters were stated to be one of the most useful features on the website with respect to clear filter choices and navigation, clear understanding of drop-down menus, the population of the selected filters and search criteria in a single area, and the ability to clear filters or search criteria easily. For example, a user commented that the use of filters made searching more effective:

> *The thing I liked about it best though was [...] just being able to say, ok lets take these [filters] out [...] and see what it looks like. That was really well done. I don't think I've seen that before, so either I'm really ignorant or you've got a really neat tool to use.*

Some participants also found the interactive map of Canada on the "Home" page to be an interesting feature. Several participants provided positive feedback about the icons indicating the type of AT available from the program on their information pages, stating that they were helpful.

Several participants provided suggestions that might further the beneficial impact of the website. For example, a participant observed that local programs within their province were not included in the database and suggested that smaller, "municipally based" programs may be more beneficial for clients to be connected to, as they "are more likely to get funding, because they are so local." Moreover, 2 participants suggested expanding the eligible population groups (eg, youth students), and adding filters that allow users to find community-based health centers and assessment locations and programs that provide equipment rentals. One of the participants explained that one of the biggest barriers to accessing ATs was the financial barrier and finding information related to funding and suggested filling this gap by including funding applications in a simplified way.

Discovery of New Programs and Resources

Many participants described the website as a valuable resource for centralizing information about supports and programs related to AT and discovering new programs. Across the consumer, caregiver, and health care provider groups, the participants largely described the website as a "one-stop shop" that provides access to accurate information. For example, a participant stated the following:

> *I think the best feature is that it has all the information you need that's all funding service programs, instead of going to multiple websites to find information, it has it down in one spot and you can filter for what you need to.*

The participants stated that this website would be useful for individuals, particularly consumers and health care providers, as an additional tool to search for funding opportunities across different provinces and refer clients to. A participant, who was very familiar with their provincial resources because of their specific area of practice in AT, stated that the website would be helpful for new health care professionals and providers in general areas of practice for finding new resources and improving familiarity with the available programs:

> *I think if I was more of a generalist, like a community [occupational therapist] or maybe if I was working in a hospital and kind of supporting people and connecting to people to resources before going back into the community, I would probably use it.*

Design Aesthetics

Most challenges with presentation, such as those related to website layout and general appearance, were reported as "minor issues" that caused annoyance but were not severe enough to hinder users from completing their tasks. The users identified areas to improve the formatting of the content for the ease of navigation and readability and ways to minimize the number of actions required to obtain a search result and improve readability by altering the color and contrast of the on-screen text. For example, the participants commented on the intuitiveness of where URL program links and the number of returned results were placed, spacing between the text of search results, amount of information presented, and the need to better differentiate between the chosen filters.

Discussion

Principal Findings

This study describes the methods and results of the usability testing of AccessATCanada, a web-based resource for improving access to information about AT programs and services in Canada. The usability of AccessATCanada was evaluated to understand how participants of different abilities would interact with the website and to identify and address major usability problems through navigating case scenarios. Although the overall SUS score was lower than the 68th percentile benchmark, the scores are subjective to the participants' prior history of navigating AT programs, knowledge of services, and previous resources that they have used. In addition, it is important to note that because of the small sample size, these results are likely skewed. For example, the first participant in the study gave a high SUS score, which was noted as an outlier. The participants may have also given a lower score because of challenges experienced during usability testing, such as inconsistencies in the filter results. The overall evaluation of the beta version of the web-based resource tested over 2 cycles revealed positive experiences, such as the ease of navigation, clean layout, and value, which validated the objective of AccessATCanada, and opportunities to incorporate feedback to improve user experience and usability, such as searching, terminology, and accessibility challenges.

The current state of equitable access to available resources and funding opportunities in Canada and the complexity of navigating the AT system were catalysts for the creation of this resource. Our project fits into the growing response by providing

a central, easy-to-use resource for various regional and national programs. This resource has the potential to identify and highlight areas where gaps exist, which might further reduce inequitable access to AT, specifically with respect to eligibility age, the types of AT and type of AT programs available, sources of funding (eg, charitable programs, government insurance programs, and other resource programs), conditions, and target population groups (eg, programs specific to Indigenous peoples, Veterans, refugees, and people who are incarcerated) within each province and territory. The ability to clearly identify the needs in these areas will be useful for priority planning for government programs and leaders working in this space.

An important area of consideration pertains to database maintenance and updating resource pages with the most recent and relevant information, particularly with resources that capture programs that are subject to changes in government regulations, such as AT. Two participants expressed concern about the long-term maintenance and sustainability of the database. For example, one of them described frustration with previous experiences with national databases running out of funding and failing to update their information. The other participant stated that keeping the website updated would encourage clinicians and end users to continue using the website. In recognition of this, the quality improvement of AccessATCanada has been planned as an ongoing process to ensure that the website is continuously updated and improved to provide up-to-date and accurate information.

It is also worth noting that outside the Canadian context, other countries, including those in less resourced settings, have also begun to respond to the need to improve access to AT by developing similar mobile apps and web-based resources [23,24]. Similar country-level or regional information search engines have been developed to provide publicly available information for end users and providers, such as Europe's Global Assistive Technology Information Network, Denmark's Assistive Technology Data-Denmark, Australia's National Equipment Database, and South Africa's Assistive Technology Database, which provides AT-related information for 10 other African countries.

Considering the aim of this website and that similar resources are to be usable by a broad range of people with and without different types of disabilities, user testing revealed the importance of including the ultimate end users during website development. As consumers and health information increasingly move to the web, previous studies have also suggested that usability testing is an important consideration for designers and developers [25,26]. Similar to other studies, we found that user experiences were reported more positively when website presentation and layout were considered and met WCAG 2.0 standards, such as the amount of text on a page, contrast, and reducing overlap between text and graphics [25,27,28]. The COVID-19 pandemic has illustrated just how critical web-based and digital experience has become and has especially highlighted the need to consider accessibility as an ongoing effort. For example, although AccessATCanada was built according to the WCAG 2.0 guidelines, the participants were still able to identify areas where meeting these guidelines could be improved.

We were unable to include users with significant technological challenges that might have been resolved if testing was conducted in person by providing appropriate technical support or guidance. This included engaging users with different disabilities in usability testing, for example, users who used eye-tracking technology, such as Tobii devices. This highlights an area where usability testing approaches can be improved to make products more user centered and inclusive for people of all abilities. For example, Asan and Yang [29] noted that there were few real-world usability evaluations that included eye-tracking technologies. Researchers, particularly those who conduct research in the areas of AT and disability supports, can more thoughtfully consider approaches for collecting data in time-sensitive environments from participants with various abilities.

Outside this resource, a lesson learned relates to centering website features on user needs, literacy, and common language. For example, although the participants expressed that the filter option was one of the best features, our early iteration revealed that when participants applied filters or used keywords that felt intuitive to them, the number of results returned was severely limited. During testing, the participants noted that keyword searching was not inclusive of different word variations, for example "wheelchair" and "wheel chair." Likewise, there remains an issue regarding terminology that creates challenges in finding AT funding, as identified by participants and highlighted in previous studies [30]. The participants identified that some of the terminology used within the website was unclear, for example, "jurisdictions" or "government-legislated insurance programs," and although we addressed this to the best of our ability, some language used was maintained to provide information continuity. The users emphasized the importance of using a common lay language that is easy for people to understand. However, inconsistent language may then present an accessibility challenge and could lead to avoidance of participating and applying for the funding service they may need. Language consistency is also important in current and future policy and program creations, as it could ultimately impact who can understand and access the available AT services within Canada. Although these issues were fixed in later iterations, these observations may benefit those interested in developing similar resources before launch.

Limitations

Limitations were considered within the context of the study. First, recruitment was conducted during the early phases of the COVID-19 pandemic. A key recruitment strategy involved engaging project partners to share recruitment material with those who used their services; however, these partners were understandably prioritizing the urgent needs of their clients during this time. Although the authors had the intention of recruiting a larger number of individuals for the study, another cycle of usability testing could not be conducted, which may have impacted the evaluation of the website. However, despite its small sample size, in combination with the "think-aloud" technique, this study was able to identify major areas of improvement that were valuable in improving and directing updates for the website regarding its usability and functionality [31].

Second, because of the remote nature of the study, some participants experienced technological challenges that were difficult for researchers to address remotely, which may have impacted their engagement in usability testing. The participants may not have been able to complete the scenario tasks because of these difficulties, as technology may have contributed to potential feelings of frustration. In addition, all the participants in the study reported that they were "very comfortable" with internet use, which may have impacted their experience of navigating the website compared with those who may have lower comfort. Other considerations were that the participants were able to self-select browsers and devices to use while participating in the usability testing study, which could have led to differences in the website layout. Website usability was also not explicitly tested with participants who used AT devices to navigate the internet, such as eye-tracking or text-to-speech technology.

Considering these limitations, future research may recruit a greater variety of stakeholders and users with different abilities and levels of knowledge about AT. Further iterations of the website based on participant feedback will positively enhance the website's usability and functionality.

Conclusions

A critical lack of information about available types, programs, and funding opportunities is a significant barrier to accessing AT. AccessATCanada was developed as a first attempt at creating a resource to map and centralize information on AT programs and funding organizations in Canada. This study used an iterative approach to the usability testing of an innovative digital resource involving people with different disabilities to evaluate its effectiveness, efficiency, and end-user satisfaction and experiences. Usability testing is useful for incorporating user perspectives in the design process, assessing satisfaction, and identifying areas for iterative refinement of technology among a wider range of users [27]. This study highlights the value and elements that are key to the creation of user-centered resources. The goal of creating an easily searchable and functional website was supported by the results of usability testing metrics and feedback, which were used to develop and enhance the website. Although information provision and enhancing awareness about the types of available AT programs are essential steps to improve access to AT, equitable access remains a key policy issue in Canada and abroad, and further efforts are required to meet the needs of end users and caregivers who rely on AT the most.

Acknowledgments

The project was led by the McMaster Health Forum and University of Toronto and was funded by AGE-WELL NCE Inc (Aging Gracefully across Environments using Technology to Support Wellness, Engagement and Long Life, Networks of Centres of Excellence Inc). The authors would like to thank the participants for their valuable feedback.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Scenarios and tasks.
[DOCX File , 14 KB - formative_v6i11e36949_app1.docx]

Multimedia Appendix 2
System Usability Scale interpretation tables.
[DOCX File , 14 KB - formative_v6i11e36949_app2.docx]

Multimedia Appendix 3
Semistructured qualitative interview questions.
[DOCX File , 12 KB - formative_v6i11e36949_app3.docx]

Multimedia Appendix 4
Task performance effectiveness findings.
[DOCX File , 16 KB - formative_v6i11e36949_app4.docx]

References

1. Schreiber D, Wang RH, Durocher E, Wilson MG. Access to assistive technology in Canada: a jurisdictional scan of programs. University of Toronto. 2017. URL: https://tinyurl.com/3p3r47p5 [accessed 2021-10-23]
2. World health assembly resolution 71.8 improving access to assistive technology. World Health Assembly. 2018 May 26. URL: https://apps.who.int/gb/ebwha/pdf_files/WHA71/A71_R8-en.pdf?ua=1 [accessed 2022-07-11]
3. World Health Organization, United Nations Children's Fund. Global Report on Assistive Technology. World Health Organization. 2022 May 15. URL: https://www.who.int/publications-detail-redirect/9789240049451 [accessed 2022-07-11]

4. Wang R, Wilson M. Access to assistive technology in Canada. In: Putnam M, Bigby C, editors. Handbook on Ageing with Disability. New York, NY, USA: Routledge; 2021.
5. Mattison CA, Wilson MG, Wang RH, Waddell K. Enhancing equitable access to assistive technologies in Canada: insights from citizens and stakeholders. Can J Aging 2020 Mar;39(1):69-88. [doi: 10.1017/S0714980819000187] [Medline: 30968788]
6. Penton V, Gustafson DL. Access to assistive technology and single entry point programs. Can J Disabil Stud 2014 Feb 16;3(1):93-120. [doi: 10.15353/cjds.v3i1.90]
7. Wang RH, Zdaniuk N, Durocher E, Wilson MG. Policymaker and stakeholder perspectives on access to assistive technologies in Canada: challenges and proposed solutions for enhancing equitable access. Disabil Rehabil Assist Technol 2022 Jan;17(1):61-73. [doi: 10.1080/17483107.2020.1765033] [Medline: 32489125]
8. Assistive Technology Funding and Service Programs in Canada. AccessATCanada. 2019. URL: https://www.accessassistivetech.ca/home [accessed 2021-11-17]
9. Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts - ISO 9241-11:2018. International Organization for Standardization. 2018. URL: https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en [accessed 2019-09-18]
10. Brooke J. SUS: a 'quick and dirty' usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation in Industry. London, UK: CRC Press; 1996:189-194.
11. Rubin J, Chisnell D. Chapter 9: Conduct the test sessions. In: Rubin J, Chisnell D, Spool J, editors. Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests. 2nd edition. Indianapolis, IN, USA: Wiley; 2008:201-228.
12. Keenan SL, Hartson HR, Kafura DG, Schulman RS. The usability problem taxonomy: a framework for classification and analysis. Empir Softw Eng 1999;4(1):71-104. [doi: 10.1023/A:1009855231530]
13. Khajouei R, Peute LW, Hasman A, Jaspers MW. Classification and prioritization of usability problems using an augmented classification scheme. J Biomed Inform 2011 Dec;44(6):948-957 [FREE Full text] [doi: 10.1016/j.jbi.2011.07.002] [Medline: 21782036]
14. Sousa VE, Dunn Lopez KD. Towards usable e-health. A systematic review of usability questionnaires. Appl Clin Inform 2017 May 10;8(2):470-490 [FREE Full text] [doi: 10.4338/ACI-2016-10-R-0170] [Medline: 28487932]
15. Neilsen J, Budiu R. Success rate: the simplest usability metric. Nielsen Norman Group. 2001 Feb 17. URL: https://www.nngroup.com/articles/success-rate-the-simplest-usability-metric/ [accessed 2019-09-23]
16. Rubin J, Chisnell D. Chapter 11: Analyze data and observations. In: Rubin J, Chisnell D, Spool J, editors. Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests. 2nd edition. Indianapolis, IN, USA: Wiley; 2008:245-268.
17. Mifsud J. Usability Metrics - A Guide To Quantify The Usability Of Any System. Usability Geek. 2015. URL: https://usabilitygeek.com/usability-metrics-a-guide-to-quantify-system-usability/ [accessed 2022-07-13]
18. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009;4(3):114-123 [FREE Full text]
19. Sauro J. A Practical Guide to the System Usability Scale: Background, Benchmarks & Best Practices. Denver, CO, USA: Measuring Usability LLC; 2011.
20. Fereday J, Muir-Cochrane E. Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. Int J Qual Methods 2006 Mar;5(1):80-92. [doi: 10.1177/160940690600500107]
21. Bradshaw C, Atkinson S, Doody O. Employing a qualitative description approach in health care research. Glob Qual Nurs Res 2017 Nov 24;4:2333393617742282 [FREE Full text] [doi: 10.1177/2333393617742282] [Medline: 29204457]
22. Boeije H. A purposeful approach to the constant comparative method in the analysis of qualitative interviews. Qual Quant 2002 Nov;36:391-409. [doi: 10.1023/A:1020909529486]
23. Visagie S, Matter R, Kayange G, Chiwaula M, Harniss M, Kahonde C. Perspectives on a mobile application that maps assistive technology resources in Africa. Afr J Disabil 2019 Aug 22;8:567 [FREE Full text] [doi: 10.4102/ajod.v8i0.567] [Medline: 31534918]
24. Karki J, Rushton S, Bhattarai S, De Witte L. Access to assistive technology for persons with disabilities: a critical review from Nepal, India and Bangladesh. Disabil Rehabil Assist Technol (forthcoming) 2021 Mar 02:1-9. [doi: 10.1080/17483107.2021.1892843] [Medline: 33651968]
25. Foley A. Exploring the design, development and use of websites through accessibility and usability studies. J Educ Multimed Hypermedia 2011 Dec;20(4):361-385.
26. Goldberg L, Lide B, Lowry S, Massett HA, O'Connell T, Preece J, et al. Usability and accessibility in consumer health informatics current trends and future challenges. Am J Prev Med 2011 May;40(5 Suppl 2):S187-S197. [doi: 10.1016/j.amepre.2011.01.009] [Medline: 21521594]
27. Barbara AM, Dobbins M, Haynes RB, Iorio A, Lavis JN, Raina P, et al. The McMaster optimal aging portal: usability evaluation of a unique evidence-based health information website. JMIR Hum Factors 2016 May 11;3(1):e14 [FREE Full text] [doi: 10.2196/humanfactors.4800] [Medline: 27170443]
28. (WCAG) Overview. Web Content Accessibility Guidelines. URL: https://www.w3.org/WAI/standards-guidelines/wcag/ [accessed 2019-06-30]

29. Asan O, Yang Y. Using eye trackers for usability evaluation of health information technology: a systematic literature review. JMIR Hum Factors 2015 Apr 14;2(1):e5 [FREE Full text] [doi: 10.2196/humanfactors.4062] [Medline: 27026079]
30. Jarvis T, Wang RH, Wilson MG. Dialogue summary: implementing a policy vision for enhancing equitable access to assistive technologies in Canada. McMaster Health Forum, McMaster University. 2020 Feb 26. URL: https://tinyurl.com/5n93asv5 [accessed 2021-12-19]
31. Nielsen J. Why you only need to test with 5 users. Nielsen Norman Group. 2000 Mar 18. URL: https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/ [accessed 2019-06-30]

Abbreviations

AT: assistive technology
SUS: System Usability Scale
WCAG: Web Content Accessibility Guidelines

Edited by A Mavragani; submitted 31.01.22; peer-reviewed by L Alabood, L Escobedo, P Worthy, C Or; comments to author 16.06.22; revised version received 30.08.22; accepted 31.08.22; published 03.11.22.

Please cite as:
Jarvis T, Mah AML, Wang RH, Wilson MG
Web-Based System Navigation Database to Support Equitable Access to Assistive Technology: Usability Testing Study
JMIR Form Res 2022;6(11):e36949
URL: https://formative.jmir.org/2022/11/e36949
doi:10.2196/36949
PMID:36326813

Original Paper

Importance of Patient Involvement in Creating Content for eHealth Interventions: Qualitative Case Report in Orthopedics

Thomas Timmers[1,2*], PhD; Walter van der Weegen[3*], PhD; Loes Janssen[4], PhD; Jan Kremer[1*], MD, PhD; Rudolf Bertijn Kool[1], MD, PhD

[1]IQ healthcare, Radboud Institute for Health Sciences, Radboud University Medical Center, Nijmegen, Netherlands

[2]Interactive Studios, Den Bosch, Netherlands

[3]Anna Hospital, Geldrop, Netherlands

[4]Maxima Medical Center, Veldhoven, Netherlands

[*]these authors contributed equally

Corresponding Author:
Thomas Timmers, PhD
IQ healthcare
Radboud Institute for Health Sciences
Radboud University Medical Center
Geert Grooteplein Zuid 10
Nijmegen, 6525GA
Netherlands
Phone: 31 746446069
Email: thomas@interactivestudios.nl

Abstract

Background: In many industries, collaboration with end users is a standard practice when developing or improving a product or service. This process aims for a much better understanding of who the end user is and how the product or service could be of added value to them. Although patient (end user) involvement in the development of eHealth apps is increasing, this involvement has mainly focused on the design, functionalities, usability, and readability of its content thus far. Although this is very important, it does not ensure that the content provided aligns with patients' priorities.

Objective: In this study, we aimed to explore the added value of patient involvement in developing the content for an eHealth app. By comparing the findings from this study with the existing app, we aimed to identify the additional informational needs of patients. In addition, we aimed to help improve the content of apps that are already available for patients with knee replacements, including the app our group studied in 2019.

Methods: Patients from a large Dutch orthopedic clinic participated in semistructured one-on-one interviews and a focus group session. All the patients had undergone knee replacement surgery in the months before the interviews, had used the app, and were therefore capable of discussing what information they missed or wished for before and after the surgery. The output was inductively organized into larger themes and an overview of suggestions for improvement.

Results: The interviews and focus group session with 11 patients identified 6 major themes and 30 suggestions for improvement, ranging from information for better management of expectations to various practical needs during each stage of the treatment. The outcomes were discussed with the medical staff for learning purposes and properly translated into an improved version of the app's content.

Conclusions: In this study, patients identified many suggestions for improvement, demonstrating the added value of involving patients when creating the content of eHealth interventions. In addition, our study demonstrates that a relatively small group of patients can contribute to improving an app's content from the patient's perspective. Given the growing emphasis on patients' self-management, it is crucial that the information they receive is not only relevant from a health care provider's perspective but also aligns with what really matters to patients.

Trial Registration: Netherlands Trial Register NL8295; https://trialsearch.who.int/Trial2.aspx?TrialID=NL8295

(JMIR Form Res 2022;6(11):e39637) doi:10.2196/39637

KEYWORDS

eHealth; qualitative research; qualitative; focus group; knee; surgery; feedback; user need; patient need; user centered; content codevelopment; patient involvement; co-design; participatory; app design; mobile health; mHealth; health app; orthopedic; mobile phone

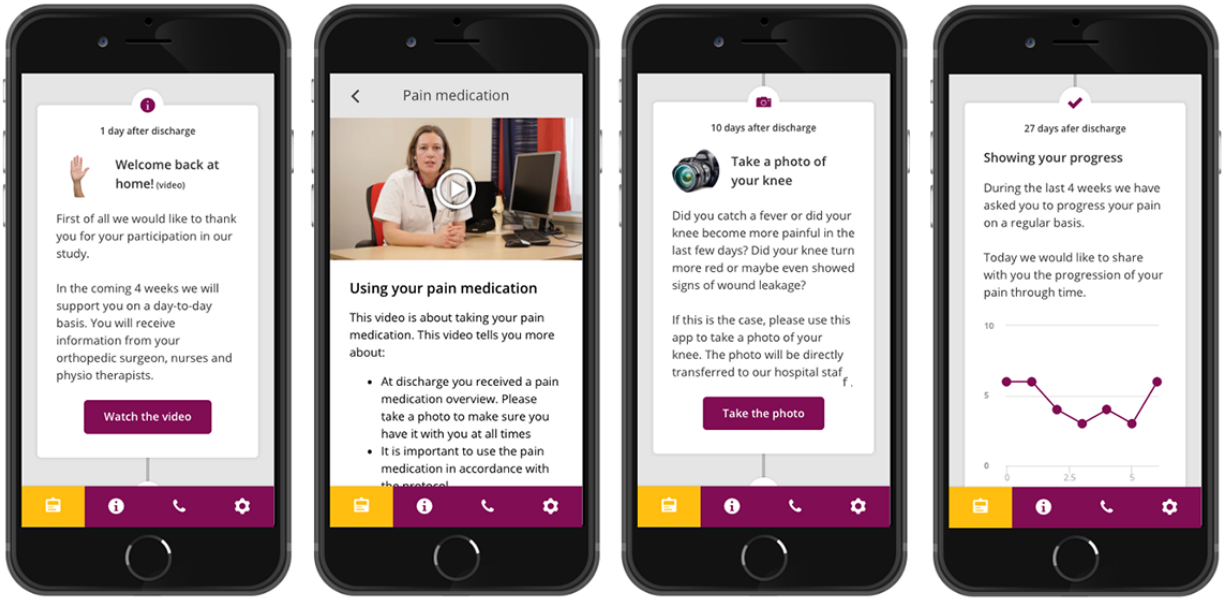
Introduction

Background

In many industries, collaboration with end users is a standard practice when developing or improving a product or service. This process aims to obtain a better understanding of who the end user is and how the product or service could be of added value to him or her. End-user involvement can also be applied in the development of eHealth apps by inviting patients, in addition to health care providers and software developers, to share their thoughts and ideas. Recent research has demonstrated the importance and effectiveness of user involvement, focusing on the design, functionality, usability, security, and privacy of eHealth apps [1-4]. In addition, users are increasingly involved in assessing the readability of the content that has already been created by health care providers and communication professionals [5,6]. Although having readable content is of great importance, it does not ensure that the content aligns with what really matters to the patients. In other words, a patient might easily understand what they are reading, but it still does not inform them about the topics that are important to them.

In 2019, our group studied the effectiveness of an app that provided patients with timely information: small pieces of information that were actively delivered to patients through an app via push notifications when the information was most relevant to patients (Figure 1) [7].

Figure 1. Examples of the interactive app used as an intervention, translated from Dutch (language used in the study) to English. From left to right, (A) the welcoming of patients to the app, (B) video and text information about medication use, (C) an invitation to send a photo of the wound (in case of fever, increased levels of pain, or wound leakage), and (D) a patient-reported pain score–progress tracker.



The app we used for the study was developed to support patients undergoing knee replacement surgery, one of the most commonly performed orthopedic procedures worldwide [8-11]. Even though the treatment is highly effective, a 2018 systematic review, including 95,560 patients, indicated that approximately 15% of the patients were unsatisfied [12]. Patients indicate that unfulfilled expectations are the most important reason for this and that improved pre- and postoperative information is needed [13-18]. Previous studies have indicated that the need for improved pre- and postoperative information is not just typical for orthopedic patients [17,19-21]. To develop the content in the 2019 version of the app, we collaborated with orthopedic surgeons, specialized nurses, nurses, and physiotherapists from 5 different Dutch hospitals. In total, the patients received 30 pieces of information and 16 unique videos in a timely manner during the initial months after surgery. Patients were notified of the newly available information through push notifications. Although the results demonstrated significant improvements in outcomes such as pain management, knee function, quality of life, and satisfaction, it is unknown whether the information presented in the app aligns with what really matters to patients, as it was developed solely by health care providers.

To gain an in-depth understanding of additional or different informational needs of patients, firsthand perspectives are required. To obtain these insights, a qualitative research approach is most appropriate, as it allows access to the thoughts and feelings of patients' understanding of how they have experienced their treatment and how the information that they were offered supported them [22].

Objectives

In this study, we aimed to explore the added value of patient involvement in developing the content for an eHealth app. By comparing the findings from this study with the existing app, we aimed to identify the additional informational needs of patients. In addition, we aimed to help improve the content of the apps that are already available for patients with knee replacements, including the app our group studied in 2019.

Methods

Study Design

This qualitative, mixed methods, prospective cohort study was conducted at the orthopedic department of the St. Anna General Hospital (Geldrop, the Netherlands). The hospital participated in the 2019 study and currently offers this app as the standard of care to patients undergoing knee replacement surgery. As of April 20, 2022, the app was downloaded 2923 times and used more than 120,000 times by patients during various stages of their treatment.

To identify the themes that are significant to most patients, we used a 2-step approach. First, we conducted face-to-face, semistructured interviews with individual patients. After transcribing, coding, and thematizing the results, we invited all the patients to discuss the themes that were identified and the suggestions for improvement that were derived from the interviews during a focus group session. In the final overview of suggestions, we added the element of timing to ensure that the patients receive information when it is most relevant to them. For the duration of the study, the patients were not informed that the members of the research team were involved in the development of the app and its content, as this might have impacted the thoughts and feelings the patients wanted to share, especially with regard to questions about communication from the hospital.

The study was registered at the Netherlands Trial Register (NL8295). Patients were asked to consider study participation by their physiotherapists, specialized nurses, or orthopedic surgeons. Typical sampling was used to ensure that the study population had characteristics similar to those of the general patient population (with respect to age, sex, and marital status). The senior researcher from the hospital (WvdW) contacted the patients who considered participating by phone in the days following surgery. After providing information about the study, a consent form was sent by email. After a few days, the researcher again contacted the patients to answer any remaining questions and to schedule a face-to-face interview with those who were interested in participation. The patients were informed that they would be invited to a focus group session at a later stage of the study. Participants could choose to have their first interview at home or at the hospital. After interviewing a patient at home and a patient at the hospital in February 2020, we were restricted to meetings on the web because of the COVID-19 pandemic. WhatsApp Messenger (Meta) was used to conduct the remaining interviews via video.

We followed the consolidated criteria for reporting qualitative research guidelines to obtain results [23].

Ethics Approval

Ethics approval was obtained from the Radboud Academic Medical Center Regional Ethics Board (reference 2020.6087) as well as from the St. Anna Local Ethics Board (reference 2020.004).

Participant Selection

Patients who underwent knee replacement surgery were invited to participate in the study. To limit the risk of recall bias, patients were only eligible for inclusion if they were able to schedule the interview within a maximum of 20 weeks after they had undergone their surgery.

Interviews

Patient characteristics were collected at the beginning of the interview. Next, an interview topic guide was used to address 3 specific aspects of the perioperative and early postoperative periods (Multimedia Appendix 1). The first specific aspect was general recovery. For this, the topic guide began with open, broad questions about how the patient had experienced recovery so far and which experiences were positive or negative. For each experience recalled, participants were asked to elaborate in as much detail as possible. Specifically, the interviewer asked how the experience was related to what they expected and what the basis was for these expectations. The second aspect was the information they received from their health care providers about their knee replacement recovery and how they aligned with their lived experiences. When a difference or gap was identified between the patients' expectation and experience, the patients were asked to elaborate on this in as much detail as possible. The third and final aspect of the interview focused on the information and education patients obtained for themselves (eg, consultation with health care providers, brochures, websites or apps, search engines, and social media). When all the 3 aspects were covered, patients were invited to follow-up with information about specific topics or add anything else that they felt was valuable.

A trained qualitative interviewer (TT) conducted the interviews together with a senior orthopedic researcher (WvdW). When the interviewers presumed data saturation and no new insights were gained, 2 additional interviews were conducted to finally evaluate the interviews. All the interviews were audio recorded and professionally transcribed verbatim.

Focus Groups

After transcribing, coding, and assigning themes to the results, patients were invited to a focus group session on September 24, 2020. The session was held at St. Anna Hospital and lasted from 7:30 PM to 9:15 PM. The session was hosted by TT and WvdW, who were accompanied by an orthopedic physician's assistant from the hospital (Ellis Bos) to answer the medical questions. A slide presentation served as the topic guide for the session (Multimedia Appendix 2).

To ensure that the session was open to active participation, the hosts began by asking the patients which themes they thought would (or should) be discussed. Next, the research team shared some of the quotes provided during the interviews, allowing participants to define the underlying theme themselves. After

participants shared their thoughts and feelings related to the theme, the researchers summarized their findings and provided an overview of the suggestions for improvement that came from the interviews. Patients were invited to provide feedback and add other elements to the overview. Next, a summary was provided, and conclusions were drawn from the results of the interviews and focus group session. Finally, patients were thanked for their participation and received a gift card.

Data Analysis

The transcription and coding of the interviews commenced after data saturation was reached. All interview transcripts were anonymized and uploaded to Atlas.ti (version 8.4.4; ATLAS.ti GmbH). Two members of the research team (TT and WvdW) independently read the transcripts from the first 3 interviews to identify codes. The researchers then compared their findings and agreed on a coding framework that could be used to code the remaining data. After all the interviews were coded by both researchers, they inductively organized the codes into larger categories using thematic analysis [24]. Finally, the relationships among the themes were identified, and the order in which they would be presented was determined.

Results

Study Sample

A total of 11 participants were interviewed between February 2020 and April 2020. The mean age of the patients was 66 (range 57-74) years (Multimedia Appendix 3). In total, 64% (7/11) of women and 36% (4/11) of men were interviewed. The mean duration of the interviews was 13 weeks after surgery (range 10-17 weeks). Before the interview, 18% (2/11) of patients underwent knee replacement surgery of the other knee. An interview was conducted at a patient's home and another at the hospital, and WhatsApp Messenger was used in 82% (9/11) of cases. The mean duration of the interviews was 40 minutes (range 25-57 minutes).

The following paragraphs provide an overview of the themes that emerged from the interviews. A narrative synthesis and an overview of suggestions for improvement are provided for each theme. In addition, a narrative synthesis is provided of the outcome of the focus group session, as well as the feedback that health care providers provided after the research team had shared their findings with them. Finally, examples of newly implemented information in the app are presented.

Theme 1: Expectations Versus Reality

A variety of answers were provided when the participants were asked how they experienced their recovery. In most cases, the recovery took longer than expected. Patients often referred to "the first 6 weeks" because they understood that most of the troubles and difficulties with pain, sleeping, and moving around would cease or be minimized by then. These sentiments were evident from the following excerpts from interviews:

> *My orthopedic surgeon told me that I would be able to get rid of the first crutch after 4 weeks and the second one after 6 weeks. Other people I talked to confirm this. I have to be honest, I guess I also heard the things I wanted to hear.* [PT08, female, aged 72 years]

> *I wouldn't be able to tell you what went well so far. At the moment, my knee hurts more than before the surgery, which was performed 10 weeks ago. I would have thought to have gained quite a lot in the first 6 weeks already.* [PT07, female, aged 66 years]

> *It was really 200% better than I expected... I started working behind the bar at the elderly center again about 8 weeks after surgery. I actually wanted to start again after 4 weeks, but my son and my GP wouldn't allow me. Unfortunately, the whole COVID-19 situation required us to close the bar, but I was up for it.* [PT04, female, aged 74 years]

Two suggestions for improvements were identified for this theme (Table 1).

Table 1. Improvement of information and timing for the theme "Expectations and Recovery."

Information	Timing
Duration and intensity of the recovery. It takes months, not weeks.	During decision-making for TKR[a] treatment and as a reminder in the weeks before and the weeks after surgery
Recovery differs greatly from patient to patient. There is no definition or standard for patients to compare themselves with in terms of speed and of duration of recovery.	First week after surgery or repeat 1 or 2 times in the following weeks

[a]TKR: total knee replacement.

Theme 2: Postoperative Pain

Patients identified postoperative pain as the central theme of early postoperative recovery period. In most patients, the severity and duration of the pain were unexpected. Pain during the night was an unpleasant surprise to many patients, as they expected that pain medication and lying at rest in their bed would provide them with a good night's sleep. Finally, some patients mentioned being somewhat surprised by the existence and duration of neuropathic pain, ranging from shooting pain throughout the leg to the feeling of a very tight band around the knee as seen in the following quotes:

> *My orthopedic surgeon had clearly told me that after the surgery it would start to hurt even more and then get better, he was completely right.* [PT05, female, aged 61 years]

> *Sufficient attention is paid to the pain in the first few days, but the long-term pain rarely mentioned.* [PT08, female, aged 72 years]

The first 5 weeks after the operation I slept really bad. It is as if you don't really feel the pain during the day, but you do at night. [PT11, male, aged 72 years]

Six suggestions for improvements were identified for this theme (Table 2).

Table 2. Improvement of information and timing for the theme "Postoperative Pain."

Information	Timing
In the initial weeks or months after surgery, patients will possibly experience more (but different) pain than what they did before the surgery.	After surgery (repeat several times)
Pain during the night is common in patients with knee replacements. This is unexpected, as patients expect the combination of pain medication and lying at rest to be beneficial. Patients might sleep poorly for weeks or even months. This is caused not only by unexpected turning while sleeping but also because of jolting neuropathic pain, the feeling of a tight bandage being wrapped around the knee, or the knee becoming suddenly very warm. Multiple nights of poor sleep can have a serious impact on a patient's mental and physical well-being, negatively influencing their recovery.	After surgery (repeat several times)
It is advised to use the pain medication as prescribed by the patient's health care professional. There is a very low risk of addiction to high-dose (rescue) pain medication even when it is only used for a short period.	Initial weeks after surgery
Some patients report side effects from the pain medication. If this is the case, patients should contact the hospital to see if the medication can be changed.	Initial weeks after surgery
For unknown reasons, pain medication is not always effective. Patients should contact the hospital when they feel the medication is not effective, leaving them with unmanageable postoperative pain. The hospital staff might be able to prescribe different, more effective medication.	Initial weeks after surgery
Neuropathy, or nerve pain, is caused by tiny nerves that were damaged or pinched during the surgery. This can lead to a constant or jolting pain around the knee or in the entire leg in the weeks following surgery. Although it might take a long time, in most cases the neuropathy disappears over time.	Initial weeks after surgery; reminder after 1 and 2 months

Theme 3: Information and Educational Materials

There were various patient needs concerning the type and amount of information they wanted or expected to receive about the treatment. Some patients wanted to know all the details about the surgery and prosthesis, whereas others preferred not to know about the details at all. Patients said that regardless of the type or amount of information they received, it was still very difficult to prepare for something they had never previously undergone. Remarkably, this was also reported by 2 patients who underwent surgery on their other knee <2 years before.

In preparation for their surgery, the patients were advised and supported by the hospital staff to download the hospital's app for knee replacement surgery. All the patients reported that they had downloaded and used them. Approximately half of the patients enabled the push notification feature from the app and were satisfied push notification feature. The other patients said they were not aware of this feature but still used the app multiple times during their treatment. The videos in the app were often mentioned as a great service, making the information easily available and digestible. Participants felt that because the videos were made by hospital staff, they were a trustworthy source of information. In addition to the app, the patients were offered hospital brochures. Most patients reported that they used the brochures at least once.

Some patients reported that the app should not only focus on complications, pain, and functional outcomes such as bending and stretching of the knee but also practical information, including details about "normal" recovery trajectory as well as the fact that recovery experience differs from patient to patient. The timing of the information in the app did not always align with the patients' actual recovery time, which unsettled the patients who recovered more slowly.

Patients did not spend much time searching for additional information on the internet. The following excerpts reveal the patient response to educational materials:

I read the app completely, I really liked that. When I received a push notification, I read the information right away. [PT03, female, aged 67 years]

You don't really know what's normal. You won't find those feelings anywhere. What am I supposed to feel now? Is this normal? That is of course something very personal, I understand that. [PT09, male, aged 57 years]

Four suggestions for improvements were identified for this theme (Table 3).

Table 3. Improvement of information and timing for the theme "Informational and Educational Materials."

Information	Timing
Provide comprehensive, though subdivided information about the anatomy of the knee, origin of complaints, knee replacement components, and the surgery. Ensure patients have a choice for the level of detail, for instance through "read more" links.	At the start of using the app
The app uses push notifications to actively inform patients about newly available information. Advise them to check whether they have enabled this functionality by going to the settings screen of their smartphone or tablet.	At the start of using the app
The recovery trajectory differs from person to person. There is no "normal" recovery or a graph that patients can or have to compare themselves with.	Initial weeks after surgery
The timing of the information in this app might differ from an individual patient's recovery time; state clearly this is not a cause for concern	Initial weeks after surgery

Theme 4: Physiotherapy Exercises

The patients unanimously agreed on the importance of physiotherapy sessions after discharge from the hospital. Being with their physiotherapist motivated them to perform the exercises and persuaded them to comfortably bend and stretch their knees just a little more than they would do at home by themselves. In addition, many considered the physiotherapist to be a personal coach. However, some found it confusing that sometimes, even within a single practice, the therapist's advice on the type of exercises or how to execute them differed from or even contradicted the previous advice. Several patients reported that they performed specific knee strengthening exercises in preparation for their surgery and felt that this benefited them during recovery.

Patients reported various physiotherapy rehabilitation approaches. There were differences not only between group sessions and individual coaching but also in the therapy itself, ranging from "fixing the knee to a bench and applying brute force to it for a short period of time" to sessions where patients would just come in during the day (unscheduled) and spend up to 2 or 3 hours doing semisupervised exercises. When asked about the physiotherapy exercises performed in the hospital directly after surgery, more than one participant mentioned that they did not feel ready for it, stating that it felt a bit hurried and rather strange doing exercises so soon after surgery with a "fresh, 30-cm-long wound" in their knee. Patients voiced their opinion as follows:

> *I have to say I really need the physiotherapist, because he keeps on saying everything will be fine.* [PT07, female, aged 66 years]

> *What is actually rarely mentioned, is the importance of the therapy. The entire course of your rehabilitation depends on how good the therapy and the therapist are.* [PT09, male, aged 57 years]

> *What amazes me a bit, and I'm sure they all mean it well, is that one therapist says: "You have to jump to the left" and then the other comes in and says: "What are you doing? You have to jump to the right."* [PT10, male, aged 57 years]

Eight suggestions for improvements were identified for this theme (Table 4).

Table 4. Improvement of information and timing for the theme "Physiotherapy Exercises."

Information	Timing
Performing exercises in the 4 to 8 weeks before surgery could strengthen the muscles around the knee and increase the flexibility of the knee.	4 to 8 weeks before surgery
Physiotherapy is offered many ways, ranging from individual to group therapy sessions, and from high-intensity shorter training sessions to longer semisupervised sessions. Regardless of these different forms of physiotherapy, the physiotherapist is seen as someone who motivates and persuades patients to do their exercises (to the full extent). In addition, patients regard their therapist as their personal coach, which is important in difficult times during recovery.	4 to 8 weeks before surgery
It is important for patients to find a physiotherapy practice that matches their personal preferences. Visiting clinics before the surgery could help in finding the right practice.	4 to 8 weeks before surgery
Before surgery, patients can start practicing walking, getting in and out of a chair, and climbing stairs with crutches. This will make the patient more comfortable using crutches when they return home from the hospital.	2 to 4 weeks before surgery
Patients should be aware of the fact that almost directly after the surgery, they will start to perform physiotherapy exercises with one of the hospital's physiotherapists. This might feel as something impossible to do right after surgery, but many patients are not aware that rapid mobilization of the knee has a positive impact on the progress of the recovery.	1 week before surgery
Patients should be aware of the importance of moving around and performing exercises. Even though it might sound or feel strange after undergoing major surgery, mobilization of the knee is of vital importance for a successful recovery. It is recommended to share the rationale behind early mobilization with patients after surgery and before performing the exercises.	Initial weeks after surgery
Patients should contact their physiotherapist when the instructions they receive are unclear or seem contradictory.	Initial weeks after surgery
There are major differences in recovery among patients. For example, the distance people can walk or cycle without too much pain, stiffness, or swelling of the knee varies widely. Patients should bear in mind that each recovery trajectory is very unique and that patients should not always compare their experience with others' or even with their own previous joint replacement surgery experience if they had one.	Initial weeks after surgery
Severe postoperative stiffness of the knee is, to a certain extent, something that can be treated through physiotherapy exercises. Only in rare cases patients are admitted to the hospital again, where the orthopedic surgeon manipulates the knee under local anesthesia to improve the knee's range of motion. This procedure is effective but adds weeks to months to the recovery trajectory.	4 to 8 weeks after surgery

Theme 5: Activities of Daily Living

Major differences in activities of daily living were noted among the patients. A patient stated that he was disappointed that 6 weeks after surgery, he was able to walk for only 200 m. Another patient reported being able to walk for approximately an hour, 6 weeks after surgery but was still rather disappointed because she expected to be able to do more. In addition, most patients mentioned having trouble cycling and walking. In many cases, this led to increased knee stiffness and swelling as did walking stairs, and one patient described it as, "dragging yourself up the stairs and jumping down the stairs, just to prevent the knee from bending." Taking a shower was another example in which clinical guidelines and patient experiences contradicted. From a clinical perspective, it was safe to take a shower, but from a practical perspective, it took patients about an hour and a half to do so, exhausting them. Three additional reported issues were sitting on the toilet (and getting up), difficulty walking with crutches, and the use of a walker or rollator to support in-house activities such as getting a cup of coffee and carrying small items through the house. A patient wore an Apple Watch (Apple) for fall detection because she lived alone. This technology allowed her neighbors to be notified if she had fallen. The following quotes highlight the reported issues:

> *Yes, I can cycle. Distances of just 20 or 30 km because my physical condition has deteriorated a bit.* [PT04, female, aged 73 years]

> *When I get on the exercise bike in the morning, I do a few cycles forwards and backwards. Then, I can cycle for 10 to 12 mins. The second time it takes a little longer to prepare, and in the evening, my knee is really stiff and thick.* [PT06, female, aged 72 years]

Nine suggestions for improvements were identified for this theme (Table 5).

Table 5. Improvement of information and timing for the theme "Activities of Daily Living."

Information	Timing
By using a toilet seat riser, patients will not have to bend their knees to the extent they would with a standard, much lower, toilet seat. This improves patient comfort in the postoperative phase.	2 to 4 weeks before surgery
Even though using crutches in the initials weeks after surgery enable mobility, carrying things such as coffee or lunch from one place to another with crutches is almost impossible. Using a walker or rollator may help patients manage their own health.	2 to 4 weeks before surgery
Instant discoloring of the knee or entire leg may appear in the initial weeks after surgery. This is something that can happen overnight, ranging from yellow green to dark blue and purple. This is something patients should not worry about, and it often disappears within several weeks.	Initial weeks after surgery
For some patients, it might be more comfortable to use a patch or band-aid on the knee to prevent clothing from continuously rubbing over the wound on the knee.	Initial weeks after surgery
Wearing soft clothing, such as sweat pants or training pants, protects the knee and the wound from uncomfortable pressure and rubbing.	Initial weeks after surgery
The knee might stiffen after surgery, which can last for weeks or even months. The knee can be stiff when getting out of bed in the morning or become stiff (again) during the day after activity. Specific exercises or training on a home trainer for 10 to 15 minutes helps in the prevention of stiffness.	Initial weeks after surgery
Swelling of the knee is reported by many patients, and it can last for weeks or even months. Swelling of the knee often occurs during or after activities such as walking or cycling. When the knee is swollen, patients are advised to elevate the leg (place pillow under the calf, not the knee itself) and apply ice packs on the knee.	Initial weeks after surgery
Taking a shower is often allowed from the moment patients get home. However, patients should be told explicitly that oftentimes, going up and down the stairs, undressing and dressing, and taking a shower are not very easy after knee replacement surgery. It will take a very long time and can be exhausting. It is advised to place a high stool in the shower to rest upon. Ensure the stool is stable and cannot slip or slide away.	Initial weeks after surgery
Patients indicate that in many cases, ointment with vitamin E helps to alleviate knee pain or itchiness.	From 2 weeks after surgery onwards

Theme 6: Hospitalization and Aftercare

All the patients reported that they felt very welcome at the hospital and that they were supported from arrival onwards by very friendly and helpful hospital staff, who had time to talk to them and answer their questions. For some patients, the discharge process felt a bit hurried, such as finishing a checklist, but no one felt as if they left the hospital without having enough information to take care of themselves. The fact that most patients left the hospital after only 24 hours still amazed some of them, mainly because they had undergone a major surgery. Other patients were more than happy to go home and were confident that being at home would be beneficial for their recovery.

Approximately 1 week after the surgery, all the patients received a phone call from one of the specialized orthopedic nurses, regardless of whether or not they had contacted the hospital themselves. Although clearly stated in the information patients received from the hospital, the phone call came as a pleasant surprise to patients, as it allowed them to ask questions or just briefly share their experiences. Patients who initiated phone calls or extra hospital visits were also very positive about the experience. This was not always the case with the formal 8-week follow-up consultation, which was reported by some patients to be rather clinical as it focused on the x-ray results. The consultation focused on discussing the x-ray results, that pain is normal, and that performing exercises benefits the recovery. The patients reported that they expected to have more meaningful and practical information about their recovery during this consultation as follows:

> *I thought it was great from the moment of admission onwards. In the operating room and the recovery room too, that all went perfectly.* [PT02, male, aged 62 years]

> *After about a week, I also got a call (from one of the nurses). I could talk a little about medication. The person you are talking to is really someone who understands what you are going through.* [PT03, female, aged 67 years]

Two suggestions for improvements were identified for this theme (Table 6).

Table 6. Improvement of information and timing for the theme "Hospitalization and Aftercare."

Information	Timing
When patients are informed about receiving a phone call from a health care provider, they are advised to write down their questions in advance.	Initial days after surgery (depending on when the call is)
Patients might be informed that the follow-up consultation with their orthopedic surgeon is focused on the alignment of the knee prosthesis, the pain, and importance of exercises in the coming months. Patients might be advised to write down their questions in advance.	Approximately a week before the follow-up consult

Results From the Focus Group Session

All 11 patients were invited to participate in the focus group session; 7 (64%) of them accepted the invitation. Owing to mandatory self-quarantine and COVID-19–related symptoms, 29% (2/7) of patients had to cancel their participation, and 71% (5/7) of participants were finally included. The focus group session led to new or additional participant insights in some cases; however, in most cases, the findings of the research team were confirmed. One of the main discussion items was why, from a medical perspective, there were differences among patients regarding pain management, duration of recovery, or recommendations for returning to sports, hobbies, or work. Ellis Bos and WvdW responded that both from a clinical and scientific perspective, these differences had been reported before and that it was still unclear why they occurred. The differences in recovery reported by the 2 patients who underwent total knee replacement in both knees were exemplary for this phenomenon.

At the end of the focus group session, the research team provided an overview of the suggestions for each theme to improve the current educational materials for patients undergoing knee replacement surgery. Feedback on these suggestions and additional ideas were discussed, leading to a final overview of the improvements.

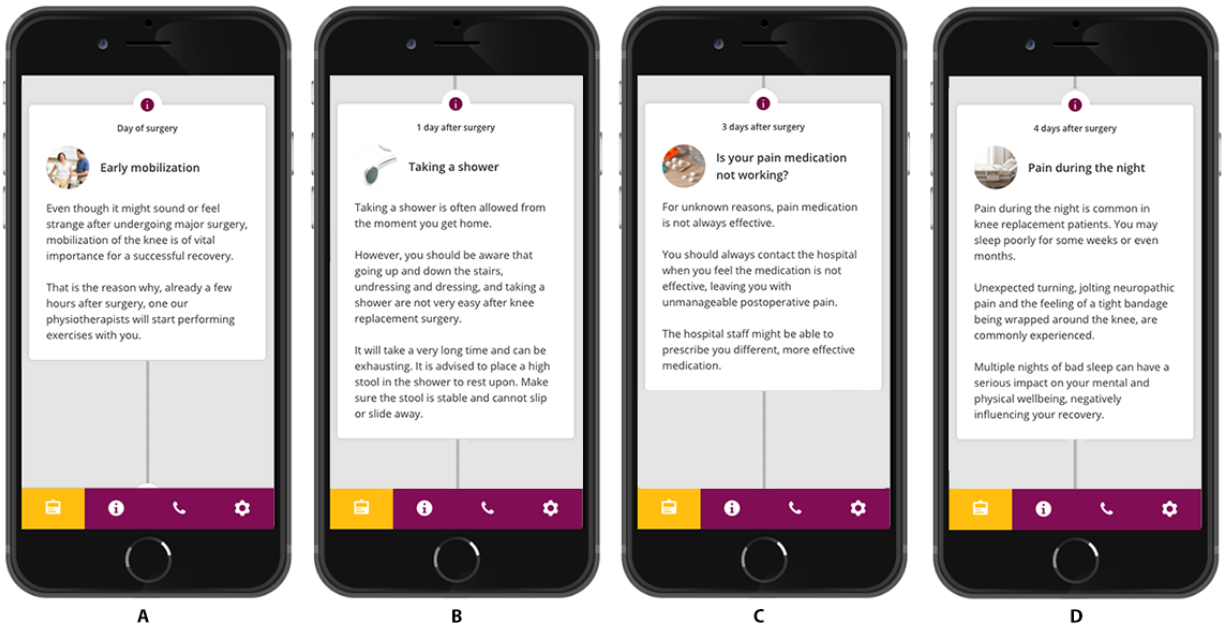
Health Care Providers' Response to the Feedback From Patients

Findings from the interviews and focus group session were shared with the health care providers at the orthopedic department. Their first response was a mixture of being glad that the study was performed because it provided new insights into what matters to patients and being surprised that so much of the information that they normally shared with patients did not resonate. This ranged from complex topics such as the duration of the recovery phase and pain management in the early postoperative phase to practical measures such as using toilet seat raiser or knowing how to act when the knee starts to swell. The difference between clinically oriented information and practical patient perspectives was best described by the example of "taking a shower." which is okay from a clinical perspective but considered a difficult endeavor from a patient perspective. The same goes for climbing up and down the stairs and getting dressed. Finally, the team was positive about the fact that all the patients had arranged their physiotherapy sessions but at the same time, was surprised to hear about the many different types of therapies patients described and the differences in approach, even within the same practice.

Examples of the Newly Implemented Information in the App

Some of the suggestions from the interviews and focus group session have already been included in the app that is currently in use at the hospital (Figure 2).

Figure 2. Examples of the information that has been added to the app (translated from Dutch to English). From left to right, (A) early mobilization after surgery, (B) taking a shower, (C) ineffective pain medication, and (D) pain at night.



Discussion

Principal Findings

The results of our study demonstrate the added value of patient involvement in developing patient-centered content for an eHealth app. From the interviews and focus group session, we learned that health care providers tended to focus on the clinical aspects of recovery such as pain, range of motion, complications, and wound care. However, patients are more interested in practical matters such as the (unexpected) intensity and duration of pain, taking a shower, dealing with swelling, going for a walk, riding a bike, or driving a car. In addition, they want to know "what is normal" in the abnormal situation in which they undergo surgery. Compared with the health care

provider–developed content of the existing app for knee replacement patients, more than 30 suggestions for improvement came from patients.

Previous studies, both in orthopedics and other medical fields, have identified patients' need to improve education and information delivery, to set realistic expectations about recovery period, to improve satisfaction after the treatment, and to increase self-management of pain [3,13-15,17,19,25-28]. However, most studies have not provided practical information on how to achieve and implement these improvements. We aimed to identify these gaps in information for patients, presented them to health care providers, jointly discussed them during a focus group session, and processed the outcomes into a ready-to-use overview and an update of the existing app.

Limitations

A limitation of our study is the relatively small number of patients who were interviewed. However, even this small group of patients provided valuable insights to improve the existing health care provider–developed content. Moreover, both interviewers agreed that data saturation had already been reached after interviewing 9 patients after which they decided to interview 2 more patients to confirm this. In addition, the inclusion of patients from a single hospital might limit the generalizability of our results. Given that we found more than 30 suggestions for improvement while the participating hospital already offers patients an app, a brochure, a website, and a complimentary postoperative phone call, we are confident that patient involvement in content creation can be of added value for other hospitals as well.

Clinical Implications and Future Research

Our study demonstrates that patients' practical informational needs can differ substantially from the clinically oriented information that health care providers want to offer them. The significant question is not whether health care providers offer this information, patients are overwhelmed with information overload, or the timing or format of content delivery was incorrect, but rather what if patients overlooked the information and were therefore unable to use it to set realistic expectations or manage their own health.

When patients and providers collaborate on the development of content, they can learn from each other's perspectives and create a blend of clinically and practically relevant information. In this process, we suggest inviting patients who having undergone the procedure, provide valuable opinion. Inviting patients from, for instance, a general hospital panel or patient advocate group will ensure the readability of the content and usability of the app, but they cannot determine whether the information itself fits the treatment-specific needs.

Finally, from an implication and implementation perspective, we suggest inviting the (software) developer or developers to participate in the interviews and focus group sessions. The firsthand patient perspectives will teach and inspire them to consider input from end users even more when building new content and software. For future research, we suggest focusing on the effects of patient-centered approach on patient-reported outcomes, patient expectations, contact between the patient and the hospital, and the level of satisfaction with the procedure and recovery.

Conclusions

This study demonstrates the added value of involving patients when creating the content of eHealth interventions. In total, more than 30 suggestions for improvement were identified, ranging from information to better manage expectations to different kinds of practical needs during various stages of the treatment. In addition, our study demonstrated that a relatively small group of patients can contribute to the improvement of an app's content. Given the growing emphasis on patients' self-management, it is crucial that the information they receive is not only relevant from a health care provider's perspective but also aligns with patient's priorities.

Acknowledgments

The authors would like to thank all the patients who participated in the interviews and focus group discussions and also thank Ellis Bos, a physician assistant at the St. Anna Hospital, for her contributions to both the focus group session and the process of translating all the patient perspectives into the patient education handout in this paper. Finally, the authors would like to thank Dr Anke Oerlemans, assistant professor at Radboud University Medical Center, for her support during the initiation of the study.

Data Availability

Extracts from the data sets used and analyzed during this study are available from the corresponding author upon reasonable request. For confidentiality reasons, the full transcripts cannot be provided.

Authors' Contributions

TT, RBK, JK, and LJ conceived the study and designed the trial. TT and WvdW performed the interviews and coded the transcriptions. TT drafted the manuscript. All authors contributed to revision of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Semistructured topic guide used for the interviews.
[PDF File (Adobe PDF File), 91 KB - formative_v6i11e39637_app1.pdf]

Multimedia Appendix 2
PowerPoint presentation for the focus group.
[PDF File (Adobe PDF File), 230 KB - formative_v6i11e39637_app2.pdf]

Multimedia Appendix 3
Participant characteristics.
[PDF File (Adobe PDF File), 77 KB - formative_v6i11e39637_app3.pdf]

References

1. Song T, Yu P, Bliokas V, Probst Y, Peoples GE, Qian S, et al. A clinician-led, experience-based co-design approach for developing mHealth services to support the patient self-management of chronic conditions: development study and design case. JMIR Mhealth Uhealth 2021 Jul 20;9(7):e20650 [FREE Full text] [doi: 10.2196/20650] [Medline: 34283030]
2. Chudyk AM, Ragheb S, Kent D, Duhamel TA, Hyra C, Dave MG, et al. Patient engagement in the design of a mobile health app that supports enhanced recovery protocols for cardiac surgery: development study. JMIR Perioper Med 2021 Nov 30;4(2):e26597 [FREE Full text] [doi: 10.2196/26597] [Medline: 34851299]
3. Barber T, Sharif B, Teare S, Miller J, Shewchuk B, Green LA, et al. Qualitative study to elicit patients' and primary care physicians' perspectives on the use of a self-management mobile health application for knee osteoarthritis. BMJ Open 2019 Feb 01;9(1):e024016 [FREE Full text] [doi: 10.1136/bmjopen-2018-024016] [Medline: 30782723]
4. Vaisson G, Provencher T, Dugas M, Trottier M, Chipenda Dansokho S, Colquhoun H, et al. User involvement in the design and development of patient decision aids and other personal health tools: a systematic review. Med Decis Making 2021 Mar 03;41(3):261-274. [doi: 10.1177/0272989x20984134]
5. Mrklas KJ, Barber T, Campbell-Scherer D, Green LA, Li LC, Marlett N, et al. Co-design in the development of a mobile health app for the management of knee osteoarthritis by patients and physicians: qualitative study. JMIR Mhealth Uhealth 2020 Jul 10;8(7):e17893 [FREE Full text] [doi: 10.2196/17893] [Medline: 32673245]
6. Ward R, Taha KM. Patient involvement as experts in the development and assessment of a smartphone app as a patient education tool for the management of thalassemia and iron overload syndromes. Hemoglobin 2016 Sep 22;40(5):323-329. [doi: 10.1080/03630269.2016.1217875] [Medline: 27535692]
7. Timmers T, Janssen L, van der Weegen W, Das D, Marijnissen W, Hannink G, et al. The effect of an app for day-to-day postoperative care education on patients with total knee replacement: randomized controlled trial. JMIR Mhealth Uhealth 2019 Oct 21;7(10):e15323 [FREE Full text] [doi: 10.2196/15323] [Medline: 31638594]
8. Choi Y, Ra HJ. Patient satisfaction after total knee arthroplasty. Knee Surg Relat Res 2016 Mar 30;28(1):1-15 [FREE Full text] [doi: 10.5792/ksrr.2016.28.1.1] [Medline: 26955608]
9. Maradit Kremers H, Larson DR, Crowson CS, Kremers WK, Washington RE, Steiner CA, et al. Prevalence of total hip and knee replacement in the United States. J Bone Joint Surg Am 2015 Sep 02;97(17):1386-1397 [FREE Full text] [doi: 10.2106/JBJS.N.01141] [Medline: 26333733]
10. Kurtz S, Ong K, Lau E, Mowat F, Halpern M. Projections of primary and revision hip and knee arthroplasty in the United States from 2005 to 2030. J Bone Joint Surg 2007;89(4):780-785. [doi: 10.2106/jbjs.f.00222]
11. Turkiewicz A, Petersson IF, Björk J, Hawker G, Dahlberg LE, Lohmander LS, et al. Current and future impact of osteoarthritis on health care: a population-based study with projections to year 2032. Osteoarthritis Cartilage 2014 Nov;22(11):1826-1832 [FREE Full text] [doi: 10.1016/j.joca.2014.07.015] [Medline: 25084132]
12. Kahlenberg CA, Nwachukwu BU, McLawhorn AS, Cross MB, Cornell CN, Padgett DE. Patient satisfaction after total knee replacement: a systematic review. HSS J 2018 Jul 5;14(2):192-201 [FREE Full text] [doi: 10.1007/s11420-018-9614-8] [Medline: 29983663]
13. Goldsmith LJ, Suryaprakash N, Randall E, Shum J, MacDonald V, Sawatzky R, et al. The importance of informational, clinical and personal support in patient experience with total knee replacement: a qualitative investigation. BMC Musculoskelet Disord 2017 Mar 24;18(1):127 [FREE Full text] [doi: 10.1186/s12891-017-1474-8] [Medline: 28340610]
14. Kennedy D, Wainwright A, Pereira L, Robarts S, Dickson P, Christian J, et al. A qualitative study of patient education needs for hip and knee replacement. BMC Musculoskelet Disord 2017 Oct 12;18(1):413 [FREE Full text] [doi: 10.1186/s12891-017-1769-9] [Medline: 29025397]
15. Specht K, Agerskov H, Kjaersgaard-Andersen P, Jester R, Pedersen BD. Patients' experiences during the first 12 weeks after discharge in fast-track hip and knee arthroplasty - a qualitative study. Int J Orthop Trauma Nurs 2018 Nov;31:13-19. [doi: 10.1016/j.ijotn.2018.08.002] [Medline: 30297138]
16. Bourne R, Chesworth BM, Davis AM, Mahomed NN, Charron KD. Patient satisfaction after total knee arthroplasty: who is satisfied and who is not? Clin Orthop Relat Res 2010 Jan;468(1):57-63 [FREE Full text] [doi: 10.1007/s11999-009-1119-9] [Medline: 19844772]

17. Hamilton DF, Lane JV, Gaston P, Patton JT, Macdonald D, Simpson AH, et al. What determines patient satisfaction with surgery? A prospective cohort study of 4709 patients following total joint replacement. BMJ Open 2013;3(4):e002525 [FREE Full text] [doi: 10.1136/bmjopen-2012-002525] [Medline: 23575998]
18. Tolk JJ, Janssen RP, Haanstra TM, Bierma-Zeinstra SM, Reijman M. The EKSPECT study: the influence of Expectation modification in Knee arthroplasty on Satisfaction of PatiEnts: study protocol for a randomized Controlled Trial. Trials 2018 Aug 14;19(1):437 [FREE Full text] [doi: 10.1186/s13063-018-2821-2] [Medline: 30107814]
19. Myles PS, Williams DL, Hendrata M, Anderson H, Weeks AM. Patient satisfaction after anaesthesia and surgery: results of a prospective survey of 10,811 patients. Br J Anaesth 2000 Jan;84(1):6-10 [FREE Full text] [doi: 10.1093/oxfordjournals.bja.a013383] [Medline: 10740539]
20. Lemos P, Pinto A, Morais G, Pereira J, Loureiro R, Teixeira S, et al. Patient satisfaction following day surgery. J Clin Anesth 2009 May;21(3):200-205. [doi: 10.1016/j.jclinane.2008.08.016] [Medline: 19464614]
21. Sivaganesan A, Khan I, Pennings JS, Roth SG, Nolan ER, Oleisky ER, et al. Why are patients dissatisfied after spine surgery when improvements in disability and pain are clinically meaningful? Spine J 2020 Oct;20(10):1535-1543. [doi: 10.1016/j.spinee.2020.06.008] [Medline: 32544721]
22. Austin Z, Sutton J. Qualitative research: getting started. Can J Hosp Pharm 2014 Nov 22;67(6):436-440 [FREE Full text] [doi: 10.4212/cjhp.v67i6.1406] [Medline: 25548401]
23. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
24. 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]
25. Mahdi A, Svantesson M, Wretenberg P, Hälleberg-Nyman M. Patients' experiences of discontentment one year after total knee arthroplasty- a qualitative study. BMC Musculoskelet Disord 2020 Jan 14;21(1):29 [FREE Full text] [doi: 10.1186/s12891-020-3041-y] [Medline: 31937282]
26. Høvik LH, Aglen B, Husby VS. Patient experience with early discharge after total knee arthroplasty: a focus group study. Scand J Caring Sci 2018 Jun 23;32(2):833-842. [doi: 10.1111/scs.12514] [Medline: 28833302]
27. Jaensson M, Dahlberg K, Nilsson U. Factors influencing day surgery patients' quality of postoperative recovery and satisfaction with recovery: a narrative review. Perioper Med (Lond) 2019;8:3 [FREE Full text] [doi: 10.1186/s13741-019-0115-1] [Medline: 31139359]
28. Gunaratne R, Pratt DN, Banda J, Fick DP, Khan RJ, Robertson BW. Patient dissatisfaction following total knee arthroplasty: a systematic review of the literature. J Arthroplasty 2017 Dec;32(12):3854-3860. [doi: 10.1016/j.arth.2017.07.021] [Medline: 28844632]

Edited by A Mavragani; submitted 16.05.22; peer-reviewed by S Joshi, F Medina; comments to author 07.09.22; revised version received 27.09.22; accepted 29.09.22; published 03.11.22.

Please cite as:
Timmers T, van der Weegen W, Janssen L, Kremer J, Kool RB
Importance of Patient Involvement in Creating Content for eHealth Interventions: Qualitative Case Report in Orthopedics
JMIR Form Res 2022;6(11):e39637
URL: https://formative.jmir.org/2022/11/e39637
doi:10.2196/39637
PMID:36326799

Original Paper

Enabling Rural Telehealth for Older Adults in Underserved Rural Communities: Focus Group Study

Inga Hunter[1], MBChB, MPhil; Caroline Lockhart[1], RN, MBS; Vasudha Rao[1], PhD; Beth Tootell[1], PGDipTT, MCom; Samuel Wong[2], PGDipSci, BSc

[1]School of Management, Massey Business School, Massey University, Palmerston North, New Zealand

[2]Vensa Health, Auckland, New Zealand

Corresponding Author:
Inga Hunter, MBChB, MPhil
School of Management
Massey Business School
Massey University
BSC, 3rd floor
Tennent Drive
Palmerston North, 4410
New Zealand
Phone: 64 6 356 9099 ext 84913
Email: I.Hunter@massey.ac.nz

Abstract

Background: Telehealth is often suggested to improve access to health care and has had significant publicity worldwide during the COVID-19 pandemic. However, limited studies have examined the telehealth needs of underserved populations such as rural communities.

Objective: This study aims to investigate enablers for telehealth use in underserved rural populations to improve access to health care for rural older adults.

Methods: In total, 7 focus group discussions and 13 individual interviews were held across 4 diverse underserved rural communities. A total of 98 adults aged ≥55 years participated. The participants were asked whether they had used telehealth, how they saw their community's health service needs evolving, how telehealth might help provide these services, and how they perceived barriers to and enablers of telehealth for older adults in rural communities. Focus group transcripts were thematically analyzed.

Results: The term *telehealth* was not initially understood by many participants and required an explanation. Those who had used telehealth reported positive experiences (time and cost savings) and were likely to use telehealth again. A total of 2 main themes were identified through an equity lens. The first theme was *trust*, with 3 subthemes—trust in the telehealth technology, trust in the user (consumer and health provider), and trust in the health system. Having access to reliable and affordable internet connectivity and digital devices was a key enabler for telehealth use. Most rural areas had intermittent and unreliable internet connectivity. Another key enabler is easy access to user support. Trust in the health system focused on waiting times, lack of and/or delayed communication and coordination, and cost. The second theme was *choice*, with 3 subthemes—health service access, consultation type, and telehealth deployment. Access to health services through telehealth needs to be culturally appropriate and enable access to currently limited or absent services such as mental health and specialist services. Accessing specialist care through telehealth was extremely popular, although some participants preferred to be seen in person. A major enabler for telehealth was telehealth deployment by a fixed community *hub* or on a mobile bus, with support available, particularly when combined with non–health-related services such as internet banking.

Conclusions: Overall, participants were keen on the idea of telehealth. Several barriers and enablers were identified, particularly trust and choice. The term *telehealth* is not well understood. The unreliable and expensive connectivity options available to rural communities have limited telehealth experience to phone or patient portal use for those with connectivity. Having the opportunity to try telehealth, particularly by using video, would increase the understanding and acceptance of telehealth. This study highlights that local rural communities need to be involved in designing telehealth services within their communities.

(JMIR Form Res 2022;6(11):e35864) doi:10.2196/35864

KEYWORDS

access; choice; trust; telehealth; rural; barriers; enablers; underserved populations; underserved; equity; elder; older adult; eHealth; telemedicine; barrier; enabler; facilitator; focus group

Introduction

Background

Telehealth is defined as "healthcare delivered using digital technology where participants may be separated by time and/or distance" [1]. Telehealth has been available for >50 years [2] but had not been widely adopted in New Zealand before the COVID-19 pandemic [3-5], despite its known benefits [6-8]. The use of telehealth has increased during the pandemic [4]. Although telehealth use is now higher than that before the COVID-19 pandemic, it is mainly by phone, and telehealth has not yet been embedded as a *business as usual* option for access to health care [4,8-10]. Considerable work has been undertaken in New Zealand to address this problem, particularly regarding the use of video consultations [11-13].

Rural Underserved Populations

The term *underserved* (also known as underresourced) population addresses situations where health care inequity exists because of system failures in health care delivery [14]. A major advantage of telehealth is increased access to health care for underserved populations, such as rural communities [15,16]; however, telehealth use during the COVID-19 pandemic was significantly lower in rural than in urban areas [17]. Rural areas are typified by low population density and less infrastructure than urban areas, with greater distances between services and people making it harder to deliver health services in rural areas [18]. In New Zealand, 1 in 4 people live in rural and semirural areas, with children, older adults, and Māori (the indigenous people of New Zealand) forming the greater proportion [19]. Owing to rural downturns and urban migration, the proportion of older adults in rural populations has been increasing at a faster rate than in urban areas [18]. The most deprived areas of New Zealand are rural [20], and ensuring equity in health care for rural New Zealanders is a priority [19,21].

Older Adults

Similar to many high-income countries, the population of New Zealand is aging, and the proportion of older adults aged ≥65 years is expected to increase to 22% by 2031, accounting for approximately 50% of government health expenditure [21]. Older adults have higher rates of long-term comorbidities and disabilities [21] and can benefit from telehealth [10].

Thus, the overarching research question for this project was *how can telehealth systems be designed and implemented in rural underserved populations to improve access to health care services in New Zealand*.

This paper presents the findings of this qualitative and exploratory project. This work adds to the academic body of knowledge by examining key barriers to and enablers of telehealth adoption within an underserved population—older adults in rural communities.

Methods

Overview

The project adopted a sociotechnical systems perspective for the development of health care technologies [22], recognizing that the adoption of new technologies involves an interaction between complex infrastructures and human behavior. A qualitative approach was used to determine the barriers to and enablers of telehealth in rural communities. Qualitative research techniques use interactive methods [23] and an approach that assumes that individuals see their reality from a set of values, attitudes, and beliefs that reflect their life experiences [24]. Group methods are an effective means of gaining such insights [25].

Ethics Approval

The New Zealand Health and Disability Ethics Committee operating procedures did not require a Health and Disability Ethics Committee review for this project. Following Massey University Human Ethics Committee processes, this project was evaluated by peer review, including suitability from a Māori research perspective, and judged to be low risk. Therefore, a low-risk notification was made to, and recorded by, the Massey University Human Ethics Committee as per the university's process [26]. Participation in the project was voluntary, and informed consent was obtained from all participants.

Eligibility

Criteria for selecting the rural areas in which to hold focus groups were identified using a modified Delphi process with the authors and the project advisory group (Textbox 1). Four geographically rural regions that met these criteria were identified: regions 1 to 3, aligning with the top, middle, and lower areas of the North Island, and region 4, aligning with the top of the South Island of New Zealand. In addition, the selected areas aligned with 1 district health board per region; for a map of New Zealand health regions and district health boards, refer to the New Zealand Health Partnerships website [27]. New Zealand underwent significant health reforms from July 1, 2022, although the regions and districts remained the same [28].

Eligibility criteria for the participants are listed in Textbox 2. The participants were recruited through purposive convenience sampling, which aims to gather a range of perspectives from diverse rural communities. The age for participation was ≥55 years, rather than ≥65 years, to allow for the fact that Māori tend to experience higher morbidity and mortality at a young age than non-Māori [21].

Textbox 1. Criteria for region selection.

Rural community criteria

- Geographic spread (three-quarters of New Zealanders live on the North Island)
- Collaborative rural community
- Existing researcher networks in the community
- Travel >30 minutes to the nearest permanent primary health care center

Textbox 2. Criteria for participant selection eligibility.

Participant criteria

- Normally residing in the selected rural communities
- Aged ≥55 years
- Willing to participate and able to consent
- Could have used or had not used telehealth before

Significance of Māori (Indigenous People) in the Research

In New Zealand, 1 in 4 people live in rural and semirural areas; children, older adults, and Māori contribute to the greater proportion of those who live rurally [19]. Māori are the indigenous population of New Zealand, accounting for 16.7% of the total population [29], with a higher proportion living in underserved areas and experiencing poorer health outcomes [19].

The New Zealand health and disability system, which includes health research [26], has obligations within its relationship with Māori under *Te Tiriti o Waitangi* (The Treaty of Waitangi). These obligations are contained in 3 principles—partnership, participation, and protection. Equitable access to health care and Māori self-determination with health and disability services form part of these principles; hence, health services must work together with Māori in the governance, design, delivery, and monitoring of health and disability services. Māori must be co-designers of the health system for Māori [25,30].

Purposive sampling was undertaken to ensure representation of different types of population groups within rural areas and, in particular, to ensure that the Māori voice was heard as per the Treaty of Waitangi's obligations. In addition to having Māori present in the community focus groups, 1 focus group was conducted on a *marae* (Māori meeting place), at the invitation of the local *iwi* (tribe), who had connections with participants from other Māori in region 1 (Northern Region). The success of this sampling strategy is reflected in the number of Māori respondents (29/98, 29.5% of the research population) exceeding the proportion of Māori in the general New Zealand population (16.7%) [29].

Recruitment

Participants could bring a support person (of any age), and a translator was available if required. Individuals wishing to take part in the study but unable to attend the focus group in their region had the option of an individual interview via Zoom videoconferencing or phone. Focus group participation was voluntary, and informed consent was obtained from all participants. A thank you gift in the form of chocolates, petrol vouchers, or supermarket vouchers was offered. The focus groups were conducted by IH, CL, and SW in person (on location) from June to July 2021, and pivoted interviews were conducted by CL by phone during August to September 2021. The focus group discussions and interviews were transcribed in full.

The process of prior engagement provided an opportunity to discuss the purpose of the research and time for the participant to get to know the interviewer [31] and assist researchers in gathering thick, rich data for analysis [32]. Most participants who registered an interest in participating in the focus groups or phone interviews were contacted by phone or email by CL before receiving the participant information sheet, signing the consent form, and conducting the focus group or phone interview. Before the focus group was held on the *marae*, IH, CL, and SW were welcomed with a *pōwhiri* (a formal Māori welcome) before entering the *whare* (Māori meeting house), thereby providing the *tikanga* or the general foundation for the context of the work and observing the cultural norms of the Māori participants [33]. The importance of respectfully engaging with Māori in *their space* is integral to establishing a relationship of trust and acceptance of the researchers and for acknowledging the principles within the Treaty of Waitangi [33].

Focus Groups

In total, 7 in-person focus groups were held (June to July) with older adults living in the 4 rural regions of New Zealand. The eighth focus group pivoted to interviews by phone (August to September) because of the COVID-19 national *lockdown*. Furthermore, 1 to 3 focus groups were held in 3 of the 4 rural regions, and the fourth region had interviews by phone. The focus groups were conducted at local community halls, with 1 being conducted on a *marae*, and lasted 1 to 1.5 hours.

Each focus group was split into 2—a table for those who had used telehealth and a table for those who had not used telehealth—and the tables were run simultaneously. The project

funder required data on the number of participants who had and had not used telehealth within the regions. The research team ran the 2 tables separately to allow older adults who were less confident with technology, or had had less contact with the health system, to have a different level of facilitated discussion than those who were familiar with telehealth and were more comfortable using the technology. Participants could choose the table they were most comfortable joining, irrespective of telehealth use; however, whether the participant had prior telehealth use experience was noted within the information in the demographics questions in the survey attached to the consent form. Despite the separation into 2 groups (used or had not used), when the discussions on the separate tables were reviewed after the focus group, they were remarkedly similar, with no themes identified by 1 group only.

The focus group semistructured interview guide of 4 questions with prompts was loosely based on the modification of the Penchansky and Thomas [34] access dimensions by Saurman [35], with adaptation for telehealth. Respondents could deviate from the interview guide, provided the discussion or interview remained relevant to the research question of the study. The same set of questions was used for the focus groups and phone interviews. After each event, the researchers debriefed and iteratively reviewed the process of conducting the focus group, which was recorded using memos. IH and CL alternated between tables of those participants who had and had not used telehealth with each new focus group. SW joined the table with the most participants.

Analysis

Both deductive and inductive approaches were used, with deductive coding drawing from the Saurman [35] modification of the access dimensions by Penchansky and Thomas [34]. Following Braun and Clarke [36], these concepts were used as an initial coding device to attract analysts' attention to relevant aspects of the data and understand users' ideas of telehealth use, not to test any framework.

The data set was analyzed as a whole, as well as according to those who had had a teleconsultation or not had a teleconsultation. An inductive analysis following the qualitative thematic analysis procedure of Strauss and Corbin [37] was then performed to identify feelings, attitudes, and perceptions and to understand the participants' experiences of rural telehealth. A total of 2 members of the research team coded the study independently, and the authors coded responses with an intercoder agreement of 94% [38]. In the first pass—open coding—the authors holistically read each response. In the second pass—axial coding—the 2 authors jointly identified subthemes within the larger categories. Finally, in the third pass—selective coding—the authors searched the data for specific responses that illustrated the subthemes. Any issues concerning the identification of subthemes or the coding of an individual response were discussed and resolved by the 2 authors and by reference to the whole team.

Results

Overview

In total, 98 adults aged ≥55 years took part across 7 in-person focus groups from 3 geographical regions, with 1 focus group pivoting to 10 individual phone interviews. In addition, 3 phone interviews were held with participants who could not attend the focus group in their area but wished to participate in the project. Phone interviews via Zoom videoconferencing were attempted; however, the unstable connectivity did not allow for an uninterrupted interview with any of the participants where Zoom videoconferencing could be used.

Demographics

General demographics are shown in Table 1, including those who had experienced telehealth consultations as per the selection criteria. The age range of the patients was 55 to 92 years. More women than men participated (68/98, 69%), similar to other web-based surveys, in which women tended to be more likely to self-select to participate [39]. No attempt was made to attract more participants who had used telehealth than those who had not used telehealth previously. The distribution of those who had used telehealth was almost 60% (58/98), and those who had not used telehealth comprised almost 40% (40/98) of the participant population. Māori accounted for 30% (29/98) of the participants.

Table 1. Participants by region and total.

Characteristics	Region 1: Northern Region[a]	Region 2: Midlands Region[b]	Region 3: Central Region[c]	Region 4: Southern Region[c]	Total
Age (years), range	55-82	56-82	55-82	60-92	55-92
Gender, n (%)					
Men	10 (27)	2 (20)	6 (31)	12 (38)	30 (31)
Women	27 (73)	8 (80)	12 (69)	21 (62)	68 (69)
Ethnicity, n (%)					
New Zealand European	10 (27)	6 (60)	15 (89)	30 (94)	61 (64)
Māori	25 (68)	4 (40)	0 (0)	0 (0)	29 (30)
Other	2 (5)	0 (0)	2 (11)	2 (6)	6 (6)
Has used telehealth, n (%)	24 (65)	7 (70)	12 (67)	15 (45)	58 (59)
Not used telehealth, n (%)	13 (35)	3 (30)	6 (33)	18 (54)	40 (41)

[a]A total of 3 focus groups; 2 community focus groups and 1 focus group held on a *marae*.

[b]A total of 10 individual phone interviews (because of the COVID-19 pandemic national lockdown).

[c]A total of 2 community focus groups.

Thematic Analysis

Overview

Data were entered into NVivo (version 1.6.1; QSR International). The grounded theory process was not used for this research; however, the grounded theory approach for the data analysis was used as it offered a systematic method of thematic data analysis, ideal for smaller data sets and generating rich descriptions and exhaustive coverage [40]. Comparative notes (memos) were used by the researchers as they conducted an iterative review after each of the focus groups and some of the phone interviews.

The overarching theme was enthusiasm and willingness to use telehealth from all communities who are keen to be involved in further research and implementation of telehealth systems, as evidenced by the following:

> *I'd love it if I could see him [health provider] by telehealth...and it would be much easier than having to drive for 3-4 hours* [Region 4 participant]

Each focus group included participants who had and had not used telehealth. The level of enthusiasm varied between those who had and had not used telehealth. It took time and discussion for some people to understand the potential benefits of telehealth for them. A few participants said that they would not use telehealth, either because of their overall health or disability or because they would always want to be seen in person. Very few participants who had experienced a telehealth consultation said that they were not keen to use it again; however, they all suggested areas of improvement. Therefore, participants fell into 3 groups: those who would not use telehealth, those who would probably use telehealth, and those who would use telehealth. The key message was that even those who had not tried telehealth would be willing to use it, although having had the experience of using telehealth resulted in a much more positive attitude toward telehealth and subsequent use. Therefore, providing opportunities to use telehealth in its broadest form would increase the success of telehealth systems in rural communities.

Two major themes emerged from a thematic analysis of the data, namely, trust and choice, each with further subthemes connected by equity (Textbox 3).

Textbox 3. Themes and subthemes connected to equity.

Trust

- Technology and telehealth
- Ability to use telehealth
- Health care system

Choice

- Health service access
- Consultation type
- Telehealth deployment

Trust

Three subthemes were associated with participant discussions around trust: trust in the technology and telehealth system using that technology; trust in a person's ability to use that technology; and, finally, trust in the health care system and its provision of care regardless of the mode of delivery.

Trust in the Technology and Telehealth System

This subtheme highlighted the need for end users to be able to trust that the technology used in telehealth systems would work when needed and with expected outcomes. A participant reinforced the need for consistency and reliability by saying the following:

> *The telehealth system has to work well otherwise people will start using it and you'll lose them straight away [if it doesn't work]* [Region 4 participant]

Connectivity was a major barrier or enabler for trusting telehealth. The issues reported with rural networks included a lack of connection to networks or unreliable and unstable networks that disconnect without warning. Participants expressed skepticism about dependable connectivity and the expectation that they would consistently be able to access health care via telehealth:

> *I don't think we're going to be able to do a lot of these things [telehealth] that we've just been talking about, until there is a vast improvement in the internet services,...If I leave this house, I can't get any reception if I'm at the back of the farm, and once I leave this house, there'll be no reception for about half an hour* [Region 2 participant]

Web-based banking and emailing were used in the discussions as examples of providing services on the web as participants were familiar with using them through rural network connectivity in their own areas. Alternative options to access web-based banking services (internet banking) were raised in each focus group and during many phone interviews. There was considerable uncertainty associated with managing financial transactions on the web. Connectivity was described as follows:

> *...it's intermittent...I think I've sent that email and no,...I haven't or, I thought I had paid that bill, and no I hadn't paid it [because the connection dropped off]* [Region 4 participant]

Frustration was expressed because of issues specific to rural settings, such as the pending removal of copper wire (landline phone) in New Zealand leaving some people with no connection, frequent power outages that cut off cell phone towers, and preventing the recharging of mobile phones. The lack of a collective community approach to cell tower installation, focusing instead on individual and tourist connections, was a cause for concern in 1 region. The participants in 1 focus group explained this as follows:

> *...we had flooding for 4 days. Electricity went [power outage] straight away and that's why a lot of people have got generators for freezers. During the flood, after 3 days the copper lines went down too, and there were no batteries in the tower to charge our phones* [Region 3 participant]

Most participants used mobile data or Wi-Fi by repeaters. Many had cheaper, older, basic mobile phones, PCs, tablets, or routers without the capacity for high-quality video calls or were on limited prepaid phone plans; therefore, the quality and speed of the video, audio, and text communications were poor. Moreover, a lack of planning by mobile and Wi-Fi networks when providing connectivity meant that network access was inconsistent, patchy, and completely absent, even on the same road. A typical example of issues that those living in rural areas experienced were expressed as follows:

> *...I've got 40 minutes to do my [online] banking on some days where it should take five [minutes],...and the presumption is that we can all be online to get all these good services, it is not realistic because of the [variable] connectivity* [Region 4 participant]

Access to rural connectivity in New Zealand is primarily through 2 formats: cellular technology through the use of mobile cell towers or base transceiver stations and wireless internet offered through a combination of wireless broadband linked (often by repeaters or network extenders) to a home-based Wi-Fi router as part of a local area network. Sometimes, the network backhaul (transmitting a signal from a remote site or network to another site) is supported by satellite and rural fiber broadband. The strength of the signal, either emitted by mobile cell towers or wireless broadband, is influenced by several factors. These factors include environmental conditions such as distance, weather, and obstacles; the technology applied, including antenna design, capacity, and frequency type; and the position where signal transmission towers or access points are placed. Owing to these factors, a person with a device capable of receiving signals from cellular or home-based Wi-Fi may experience significant fluctuations in signal strength when they move between rooms or locations on their rural property.

Similar international studies [41,42] have reported on broadband access challenges with telehealth programs for both rural and underserved populations. The shift to internet-based health consultations and associated increased reliance on internet connections because of the impact of the COVID-19 pandemic have further negatively affected telehealth use for those with existing health disparities.

The broadband initiative released by the Ministry of Business, Innovation, and Employment addresses some of these problems with both the Ultra-Fast Broadband Programme and the Rural Broadband Initiative. These 2 initiatives are currently on track to having 80% of New Zealanders with access to ultrafast broadband and improved rural coverage to 90% by 2025 [43].

Cost was another major issue, both to upgrade devices and access the network, with many participants, particularly from lower socioeconomic communities, reporting not being able to afford these costs and participants saying that their options were few:

> *It's too expensive [upgrading to get better network access]* [Region 1 participant]

The final issue for trusting telehealth was security. The overall impression was that the participants trusted the system to be secure. However, a participant stated the following:

> *...You need to be wary of a lot of stuff going on the internet, such as scammers...[however] when you need expert medical advice,...don't worry about who's dealing with your security* [Region 1 participant]

Security features such as 2-factor authentication and maintaining updated software were a cause for frustration for some families as many rural families shared email addresses and devices, and updates and changes to the software itself were difficult. A gentleman related the following:

> *...I was fortunate because I registered for a patient portal first, and they won't accept my wife on the same email address* [Region 4 participant]

Trust in a Person's Ability to Use Telehealth

The second subtheme highlighted issues related to the ability to use telehealth systems by both consumers and providers. Age was not an indicator of digital capability, and participants reported varying comfort levels with different digital technologies across all age groups, not only for themselves but also for the need for health providers:

> *...to come up to speed [learn how to use technology]* [Region 1 participant]

That said, people with disabilities experienced greater difficulty in using telehealth than others. However, at the same time, technology was also credited with increasing access for some people living with disabilities. Texting was one such example:

> *[technology] has been an amazing thing for the deaf community* [Region 3 participant]

Participants reported a widespread lack of knowledge on both how to use digital technology to access health care and what digital technology is available to be used. For example, when discussing a patient portal, one of the participants was excited to learn that there were other options available to access health care:

> *So, I just need to contact the doctor and ask them about a portal, then I can see all my medical things on the report? This is just the best bloody thing that ever happened because I'm sick of ringing [them] back having just missed a call from the nurse* [Region 1 participant]

There were several ways in which participants addressed their lack of knowledge about using digital technology. Writing down instructions or using teaching videos were some ways in which they coped with their lack of trust in being able to use technology and telehealth:

> *I see these video clips a lot on YouTube. I have done quite a bit of learning online.* [Region 3 participant]

Others took help from younger family members or partners, which was not always ideal:

> *...When they finally come and visit, my children,...they take the phone off you, and do it for you...[which meant a lost opportunity to upskill with the use of a digital device]* [Region 1 participant]

Ongoing training was another suggestion. It was preferred to be provided collectively in the community rather than individually, although some participants had undertaken individual web-based courses. It was felt that having group training enabled people to support each other:

> *...And that's where community education kicks in..., bringing the people to a central hub and actually educating them,..., in a community, the buy-in comes from the community, the capacity comes from the community...* [Region 1 participant]

Finally, participants proposed having ongoing support available, for example, a support person in a local hub, community center, or health center or in a mobile van:

> *If the health centre had a little workshop, they [older adults] could do it along with a medical visit, where they can sit down with someone who can help [to learn about patient portals]* [Region 4 participant]

Trust in the Health Care System

The third subtheme highlighted concerns raised regarding trust in the health system. Some comments were positive, particularly the provision of emergency care (first responders and rescue helicopters), with participants supporting more funding for helicopters, in particular. However, others were negative, particularly regarding chronic care management and referrals between services. Waiting times and the lack of communication and coordination between providers such as general practitioners (family physicians and primary care providers) and community pharmacists were frustrating for many. One of the participants went further with their experience:

> *...To get the same GP you have to book up to one month ahead...and the last appointment I had with the GP, I got down there, and they hadn't let me know that she [the GP] wasn't going to be there* [Region 4 participant]

Establishing and maintaining long-term relationships between consumers and providers were highlighted, with participants indicating that they, or members of their family, would rather travel a long distance to see their regular physician for consistent treatment rather than see a locum for the management of chronic conditions:

> *...They (family) want to have a face to face [in-person consultation]. They don't like having to see different people because you get given the wrong medication* [Region 1 participant]

Finally, the cost of accessing health care in general made telehealth a preferred option for some. One of the suggestions was as follows:

> *A number of health programmes could be provided online, and could be publicly funded...* [Region 2 participant]

This comment raised the need to discuss the funding of telehealth services, both at provider and consumer levels; telehealth may not necessarily be a cheaper option for the

consumer once telehealth is established as a *business-as-usual* model of care. During the COVID-19 pandemic, and to date, the cost of telehealth services in New Zealand has varied throughout the country but has often been free to the consumer, giving a false impression of the true cost of accessing health care through telehealth.

Choice

Choice was the second major theme that emerged from the focus groups and interviews, with 3 further associated subthemes: health service access, consultation type, and telehealth deployment.

Choice in Health Service Access

This theme highlighted the considerations raised by participants in accessing health care services:

> *...Choice is important, cheques [banking checks] were discontinued, we had no choice* [Region 3 participant]

New Zealand banks stopped using checks after May 31, 2021, a contentious mandated decision that required a move to digital payment options, with which many older adults were unfamiliar.

Cultural appropriateness (safety) was a key idea that emerged, particularly for Māori. Participants indicated that health care services needed to be places where people felt culturally safe, supported, and with people with whom they have a good relationship:

> *They [patients] haven't even opened the door [to access the health service] because they feel uncomfortable. The result being that family members did not access the health services.* [Region 1 participant]

Therefore, some participants traveled more than an hour to see a preferred primary health care provider rather than one closer. One of the participants explained as follows:

> *...I just keep my GP, I drive two and a half hours each way [to see the same GP]...* [Region 1 participant]

Privacy was an important part of choice, irrespective of cultural norms and expectations. Having the option to choose who was present during a consultation was part of feeling safe with a health service, particularly for Māori, a collective-based society [25], who, in general, prefer having support people from their *whānau* (family) accompanying them to a consultation; however, some Māori participants indicated that they would prefer to be unaccompanied for privacy reasons. Hence, having the choice of a support person being present is important to meet Māori cultural needs and enhance their engagement with telehealth. However, some non-Māori participants indicated that they would also like to be able to choose whether and whom they could have present at a consultation, hence making this a choice option for anyone:

> *I don't mind a [family] member there just helping and setting everything up [telehealth consultation]...[when talking with a doctor], but I really don't need some of my whānau [family] knowing that I have a problem somewhere else.* [Region 1 participant]

Travel was another important determinant for accessing health care services. Some participants liked to combine a visit to health care services in the town with other activities such as shopping, visiting the library, performing social activities, picking up medications from the pharmacy, or having a blood test. The preferred mode of transport varied: some participants favored a rural bus service that went from their rural community to the local town, particularly if they were unable to drive, did not have a driving license, or were unable or unwilling to ask for a lift to town:

> *...bus options were also limited...it's just the one bus [available bus service] but think of all the people who might potentially use the xxx bus, they're mainly using friends and family. So, if you're [living] out here and you don't have a car, you would be having someone take you..., but how often can you ask your friends [to drive you]?* [Region 4 participant]

However, other participants did not like traveling on such a bus as it took up most of the day, and users found it very tiring. Some rural regions did not have a rural bus service, and without other options for transport, people stayed home and did not access any health care. One of the participants clearly stated the outcome:

> *...but I can't make it [the travel], so I go without [healthcare]. A lot of us go without [health care]...they die!* [Region 1 participant]

Accessing specialist care was a major issue that was raised, and travel times could be up to 6 to 8 hours 1 way to see a specialist depending on road conditions, often requiring an overnight stay. The ability to have a specialist consultation via telehealth would mitigate the anxiety associated with driving long distances and navigating large cities. Telehealth was a popular option for accessing specialist services for almost all participants. One of the participants said the following:

> *If I've got to go [to see a specialist], I go the day before because I suffer really badly from anxiety. I can't do long trips and I drive myself, so I've got to take someone [with me]...* [Region 1 participant]

Accessing mental health services was highlighted as another service in which telehealth could have a large impact:

> *Accessing mental health services...there is a level of stigma if you live in a small community—if you walk into a counsellor's office, you feel exposed, so having the session from home means you have the comfort, and it may mean that treatment is sought, rather than not...and you can have loved ones with you sharing that time* [Region 2 participant]

Combining other services such as point-of-care testing, blood tests, and hearing and vision testing were suggestions for telehealth hubs, as well as medication delivery to home or to the site of telehealth services for collection. Banking services, where many are promoted to be accessed on the web or via regional branches with limited operating days and hours, and taxation services, where the main access is via email or automated phone answering services, are known to be problematic for older adults living in rural communities,

particularly those with sensory loss or mobility issues. Options for a combination of mobile telehealth and non–health care services, with an accompanying technical support person to assist older adults with technical aspects coming out to rural communities, were well received.

An older adult participant shared with us why they would support mobile services:

> *I'm not game [confident] to do online banking, because all these years we've had a cheque book, [checking account] well, they're gone now. I haven't gone on online, but I've gone on to phone banking which, I don't know how much safer it is* [Region 2 participant]

Choice in Consultation Type

The second subtheme was having a choice on how to have a consultation, for example, by video, phone, patient portals, text, or email. Having a choice of device —mobile phone, tablet, PC, and laptop—was also important:

> *I had a phone consultation with our pharmacist, it was absolutely fabulous* [Region 3 participant]

In fact, some participants were annoyed when they realized what telehealth was and that they had not been offered this as an option by their health care provider:

> *Out of the four consultations I've had since last March, one has been a phone [consultation], and the other three I've gone in [in person to see the GP], and I didn't need to go in for any of those four [consultations]* [Region 1 participant]

Using telehealth for regular reviews was raised by many participants:

> *I would be happy to check in to the medical centre via video to make the check-ins 6 monthly rather than 3 monthly* [Region 4 participant]

Using telehealth to see whether an in-person visit was needed was also raised:

> *...if you had video, you could talk with the GP, you could show your husband's swollen legs, or that he couldn't move his arms,...or that doctor could look at his legs and say "take more frusemide" because he knows what medications he is on* [Region 3 participant]

Furthermore, using telehealth to provide access to services otherwise unavailable or inaccessible was raised:

> *I think it's unreasonable for us to expect to have services on tap [instant access]. I think it all gets back to connectivity, and if I can talk to the physio on the phone or zoom and she can see how I'm going in my own house, I think that's what we need* [Region 4 participant]

Choice in Telehealth Deployment

The final subtheme considered the deployment of telehealth services. For participants who were very comfortable using digital technology, there was a desire to use it from home. One of the participants said the following:

> *All the way through my [treatment], I have gone to only a couple of appointments, I never went to my GP here, ever,...I don't want to be driving. Something clicked when this happened, and I thought I'm not going to use my energy for all of that [travelling]...I did it all on my phone, I wasn't concerned about seeing their visual* [Region 4 participant]

However, most indicated that they would prefer some sort of *hub*, with the required secure technology and a technology support person available. Some communities indicated a preference for a fixed telehealth hub located at a community hall, primary care provider building, or local rural hospital. The repurposing of existing facilities was supported:

> *We have WiFi here at the meeting hall, we have set it up and that was one of the reasons to push on our side, was so that if we needed to have the doctor on [present on a telehealth call], he could link up [with us]* [Region 1 participant]

However, others suggested a mobile option for a telehealth hub, with a bus that travels around a set schedule of rural communities with a technology support person and maybe a nurse or health care assistant:

> *...my grandson had some teeth work done in the dental caravan, she [the dental nurse in the caravan] was able to be do a zoom conference to the dentist in the hospital* [Region 1 participant]

The idea of mobile health services is not new to New Zealand; there is a mammogram bus, surgery bus, and dental bus that travel around the country to different rural areas, and additional services such as a mobile echocardiogram service in areas with high rates of heart disease are planned. Thus, the idea of mobile telehealth services, either separate from or incorporated into existing mobile services is not unreasonable. The idea of staffing the bus with someone who could assist in the primary purpose of the bus in addition to providing digital technology support for the consumer was popular. For example, the mobile service could enable a telehealth consultation with a specialist located in another region and support the consumer through experience with digital education so that they may choose to undertake telehealth consultations in the future.

Barriers and Enablers

Mapping these themes and subthemes to barriers and enablers is shown in Textbox 4. Often, an enabler is the opposite of a barrier; for example, one of the enablers is reliable connectivity—the opposite of this barrier is unreliable connectivity.

Textbox 4. Barriers and enablers.

Themes and subthemes mapped to barriers and enablers

- Barriers
  - Unreliable connectivity
  - Cost—network, devices, and data
  - Lack of access to devices
  - Security or privacy concerns
  - Low technology comfort level
  - Low digital literacy
  - Health service waiting times
  - Poor communication
  - Poor service coordination
  - Lack of services
  - Cost of health services
- Enablers
  - Reliable connectivity
  - Trust
  - Having choice
  - Flexible to individual or community needs
  - Easy to use
  - Support
  - Training
  - Reduction in travel
  - Culturally safe
  - Variety of deployment
  - Access to a wide range of services (health and nonhealth)

Discussion

Principal Findings

The work reported in this paper is part of an exploratory project that investigates how underserved rural communities would like to use telehealth to improve their access to health services. In total, 7 focus groups and 10 interviews (pivoted from focus groups because of the COVID-19 lockdown) comprising 98 adults aged ≥55 years from 4 rural areas discussed their future needs for health services, how telehealth could improve access to these health services, and the barriers and enablers to using telehealth for their rural communities. Diversity was evident within the chosen rural communities, and all participants faced multiple challenges related to their access to health care services.

Rural communities are keen to adopt telehealth; therefore, the time is right to deliver health services by telehealth, although it needs to be implemented correctly the first time.

Although telehealth has been used in the health sector for decades, albeit in a limited capacity, the term *telehealth* is not well understood by consumers, and there is a lack of consumer awareness of the availability, benefits, and device options for telehealth. However, those who have used telehealth to access health services find it extremely helpful and would willingly use it again. Thus, providing opportunities for consumers to see telehealth in action and use it with appropriate support and digital literacy training would increase their understanding and awareness of telehealth and subsequently increase telehealth adoption.

The findings from this study align with the concept of co-design, identifying benefits for the user as opposed to the provider organization, leading to increased alignment with user requirements and user acceptance [44]. Partnership, participation, and active protection, the 3 principles within the Treaty of Waitangi [31], which includes the Māori voice in the design of telehealth systems *with rural communities* rather than *for rural communities*, can increase the successful use of telehealth to improve access to health care for different communities. The inclusion of the rural consumer voice will go somewhat toward addressing some of the inequity that exists with access to digital options for underserved communities.

The best choice for the mode of telehealth deployment—mobile digital technology bus, fixed venue, hybrid approach, or extending rural hospital capability—will vary among different rural communities, and there may be further solutions that have not been identified in this research. Combining telehealth services (either via mobile bus or fixed) at community venues with point-of-care testing or blood tests, hearing and vision testing, and social and health-based activities (such as mum and baby sessions, coffee or tea, wellness checks, and vaccinations), as well as allowing the use of telehealth technology to deliver access to nonhealth services such as web-based banking, would add value to telehealth systems and thus increase telehealth acceptance and use.

Barriers and Enablers: Trust and Choice

The barriers and enablers shown in Textbox 4 align with other studies [4,45,46] and with the findings of previous studies by 2 of the authors who explored using sensor technology to support aging in place [47,48].

However, the themes of trust and choice have not been identified as enablers of telehealth. These themes, and their associated subthemes, align with the access dimensions discussed by Penchansky and Thomas [34], particularly those relating to connectivity and device availability, affordability, and accessibility. The addition of awareness by Saurman [35] is also clearly shown in this study, as many participants were unaware of both available health services and telehealth and of how to use digital technology that forms the basis of telehealth.

Trust is central to health care [49], easy to lose, and very hard to regain [50]. Implementing telehealth in an ad hoc manner that is difficult for users to connect to, with poor experiences, will not engender the needed sense of trust in the telehealth system. Hence, it is critical that telehealth implementations are culturally appropriate, well planned, sufficiently resourced, and have end user support. An example is the unreliable or absent connectivity experienced by the participants. The selected locations had some connectivity [51]; however, the participants' experience was poor. This raises the issue that it is not just the existence of broadband coverage but also the quality of that coverage that must be considered to determine whether it is sufficient to sustain a telehealth consultation.

Choice requires having the option to select between ≥2 possibilities and, thus, that >1 possibility exists and is known. Individuals and rural communities face similar but different challenges to urban dwellers, and they already have a limited choice with broadband provider options and data speeds; hence, having a choice over how to access health services will enable them to better engage with telehealth. For example, having a choice over devices that fit their income or that can connect to the available broadband, or over telehealth services that are delivered in a way that makes them feel safe and welcomed, will enable people in underserved rural communities to engage with telehealth.

Importantly, building trust and choice into telehealth system design would result in telehealth systems that are culturally appropriate for First Nations' people and indigenous populations who already experience significant inequity in access to health care and health outcomes [30,52,53].

Limitations

Beyond the limitations inherent to the nature of conducting focus groups [54], although rich data in consistent themes were obtained, care should be taken with reproducibility and transferability of the findings drawn from the study and beyond the study locations and these rural underserved communities. As the number of focus groups and phone interviews grew, discussions with participants became transferable, and fewer new themes were raised; however, data saturation and the end of data collection are contentious and much debated topics with different definitions [55]. It is possible that other researchers may interpret the data into different themes or that further analysis might identify additional or different codes, as the nature of qualitative research is that it is subjective and therefore influenced by researchers' personal biases [56].

Two main aspects determined the end of data collection for this project; available resources and the collection of sufficient data to enable meaningful, albeit subjective, analysis and inferences to be drawn by the researchers. The participants were not drawn from a random sample of individuals; rather, they actively volunteered to be involved in the study. The research was constrained in terms of the time to complete the project, with the added complication of needing to cancel 2 scheduled focus groups because of the second national COVID-19 pandemic lockdown and reverting from these last 2 focus groups to scheduling 10 separate phone interviews in an entirely new region. It could be argued that reverting to 10 separate phone interviews provided ≥8 hours to the transcription content and, therefore, even more credibility in obtaining thick, rich data from participants [32]. The team also had a cap on funding, limiting further exploration.

Finally, the impact that the COVID-19 pandemic may have had on recruitment, mode of engagement, and consumer attitudes toward telehealth must be considered.

Conclusions

The participants from the underserved rural communities were keen to use telehealth to access health services but wanted more information about, and support to use, telehealth systems. Rural communities want to be involved in designing telehealth services available in their communities. Maintaining trust and supporting choice in the use of telehealth to access health care are key enablers of telehealth acceptance.

Acknowledgments

This project was funded by Vensa Health as a contract awarded to IH.

The research team would like to acknowledge and thank the New Zealand Telehealth Forum for serving as our advisory group; Taria Tane from Ora Project Solutions for her assistance in connecting with communities in Northland; and Dr Dick Whiddett, Professor Hans Guesgen, and Dr Farah Palmer, Massey University, for their academic support and ethical peer review. The authors would especially like to thank all the participants for their enthusiasm in supporting the focus groups and their willingness to participate.

Authors' Contributions

IH conceptualized, designed, and wrote the first draft of this manuscript. CL led consumer engagement, recruitment, and project management. IH, CL, and SW were responsible for data collection. SW attended the focus groups and assisted with technical knowledge of the research. VR and BT led the data analysis. All authors were involved in the final interpretation of the results and read, commented on, approved, and signed the final version of this manuscript.

Conflicts of Interest

SW was employed by Vensa during the study period. Complete academic freedom with the design and presentation of the findings was preserved.
Reviewer Karen Day collaborates on telehealth research with some of the authors.
Reviewer David Parry declares being acquainted with the authors, but does not have any recent publications with them.

References

1. About telehealth. NZ Telehealth Resource Centre. URL: https://www.telehealth.org.nz/health-provider/what-is-telehealth/ [accessed 2022-08-20]
2. Institute of Medicine, Board on Health Care Services. The Role of Telehealth in an Evolving Health Care Environment Workshop Summary. Washington (DC): National Academies Press (US); 2012.
3. Telehealth around the world: a global guide. DLA Piper. 2020 Nov. URL: https://www.dlapiper.com/~/media/files/insights/publications/2020/12/dla-piper-global-telehealth-guide-december-2020.pdf [accessed 2022-09-06]
4. Kichloo A, Albosta M, Dettloff K, Wani F, El-Amir Z, Singh J, et al. Telemedicine, the current COVID-19 pandemic and the future: a narrative review and perspectives moving forward in the USA. Fam Med Community Health 2020 Aug;8(3):e000530 [FREE Full text] [doi: 10.1136/fmch-2020-000530] [Medline: 32816942]
5. New Zealand telehealth survey 2019: District Health Boards. NZ Telehealth Forum & Resource Centre. 2019 Nov. URL: https://www.telehealth.org.nz/telehealth-forum/submissions/dhbsurvey/ [accessed 2020-09-06]
6. 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]
7. Hallam L, Joyner S. Navigating telehealth: the patient's perspective. Health Consumers NSW. URL: https://www.hcnsw.org.au/navigating-telehealth/ [accessed 2022-09-06]
8. Imlach F, McKinlay E, Middleton L, Kennedy J, Pledger M, Russell L, et al. Telehealth consultations in general practice during a pandemic lockdown: survey and interviews on patient experiences and preferences. BMC Fam Pract 2020 Dec 13;21(1):269 [FREE Full text] [doi: 10.1186/s12875-020-01336-1] [Medline: 33308161]
9. Demeke HB, Merali S, Marks S, Pao LZ, Romero L, Sandhu P, et al. Trends in use of telehealth among health centers during the COVID-19 pandemic - United States, June 26-November 6, 2020. MMWR Morb Mortal Wkly Rep 2021 Feb 19;70(7):240-244 [FREE Full text] [doi: 10.15585/mmwr.mm7007a3] [Medline: 33600385]
10. Health and Disability System Review – Final Report. Wellington: HDSR; 2020.
11. Webinars. NZ Telehealth Forum and Resource Centre. URL: https://www.telehealth.org.nz/webinars/ [accessed 2022-09-06]
12. Large R, Rao V, Hunter I, Webb S, Bycroft J, Zuleta N, et al. PASE whitepaper: delivering health care from the cloud. NZ Doctor. 2022 May. URL: https://www.nzdoctor.co.nz/document/download/telehealth_white_paper-patients-anywhere-specialists-elsewhere-pase [accessed 2022-09-06]
13. National telehealth register. NZ Telehealth Forum and Resource Centre. URL: https://www.telehealth.org.nz/resources/nationaltelehealthregister/ [accessed 2022-09-06]
14. Piggott T, Orkin A. Deconstructing the concept of special populations for health care, research, and policy. In: Under-Served Health Determinants of Indigenous, Inner-City, and Migrant Populations in Canada. Toronto, Canada: Canadian Scholars; 2018.
15. Armfield NR, Edirippulige SK, Bradford N, Smith AC. Telemedicine--is the cart being put before the horse? Med J Aust 2014 May 19;200(9):530-533. [doi: 10.5694/mja13.11101] [Medline: 24835716]
16. Kim T, Zuckerman J. Realizing the potential of telemedicine in global health. J Glob Health 2019 Dec;9(2):020307 [FREE Full text] [doi: 10.7189/jogh.09.020307] [Medline: 31656598]
17. Chu C, Cram P, Pang A, Stamenova V, Tadrous M, Bhatia RS. Rural telemedicine use before and during the COVID-19 pandemic: repeated cross-sectional study. J Med Internet Res 2021 Apr 05;23(4):e26960 [FREE Full text] [doi: 10.2196/26960] [Medline: 33769942]

18. Bidwell S. Successful Models of Rural Health Service Delivery and Community Involvement in Rural Health: International Literature Review. Christchurch, New Zealand: Centre for Rural Health; 2001.
19. Rural health. Ministry of Health. 2019. URL: https://www.health.govt.nz/our-work/populations/rural-health [accessed 2022-09-06]
20. New Zealand Index of Deprivation (NZDep). Ehinz. URL: https://ehinz.ac.nz/indicators/population-vulnerability/socioeconomic-deprivation-profile/#new-zealand-index-of-deprivation-nzdep [accessed 2022-09-06]
21. Associate Minister of Health. Healthy Ageing Strategy. Wellington: Ministry of Health; 2016.
22. Coiera E. Putting the technical back into socio-technical systems research. Int J Med Inform 2007 Jun;76 Suppl 1:S98-103. [doi: 10.1016/j.ijmedinf.2006.05.026] [Medline: 16807084]
23. Rossman G, Rallis S. An Introduction to Qualitative Research Learning in the Field. Thousand Oaks, California: SAGE Publications; 2016.
24. Denzin NK, Lincoln YS, editors. The SAGE Handbook of Qualitative Research Fifth Edition. Thousand Oaks, California: SAGE Publications; Feb 2017.
25. Boulton A, Gifford H, Potaka-Osborne M. Realising whānau ora through community action: the role of Māori community health workers. Educ Health (Abingdon) 2009 Aug;22(2):188. [Medline: 20029746]
26. Code of ethical conduct for research, teaching and evaluations involving human participants. Massey University. 2017. URL: https://www.massey.ac.nz/massey/research/research-ethics/human-ethics/code-ethical-conduct.cfm [accessed 2022-09-06]
27. DHB Regions. NZ Health Partnerships. 2017. URL: http://www.nzhealthpartnerships.co.nz/wp-content/uploads/2017/08/DHB-Map-Aug-2017-1.pdf [accessed 2022-09-06]
28. About the health reforms / Mō ngā whakahounga hauora. Future of Health Te Anamata o Te Oranga. 2022 Aug. URL: https://www.futureofhealth.govt.nz/about-the-reforms/ [accessed 2022-09-06]
29. Māori population estimates: at 30 June 2020. NZ Stats. 2020 Nov 17. URL: https://www.stats.govt.nz/information-releases/maori-population-estimates-at-30-june-2020 [accessed 2022-09-06]
30. Waitangi Tribunal. Hauora Report On Stage One Of The Health Services And Outcomes Kaupapa Inquiry. New Zealand: Waitangi Tribunal; 2019.
31. Corbin J, Morse JM. The unstructured interactive interview: issues of reciprocity and risks when dealing with sensitive topics. Qual Inquiry 2016 Jun 29;9(3):335-354. [doi: 10.1177/1077800403009003001]
32. Morse JM. Critical analysis of strategies for determining rigor in qualitative inquiry. Qual Health Res 2015 Sep 16;25(9):1212-1222. [doi: 10.1177/1049732315588501] [Medline: 26184336]
33. Hudson M, Milne M, Reynolds P, Russell K, Smith B. TE ARA TIKA Guidelines for Māori Research Ethics: a framework for researchers and ethics committee members. Pūtaiora. 2010 Feb 9. URL: https://www.fmhs.auckland.ac.nz/assets/fmhs/faculty/tkhm/tumuaki/docs/teara.pdf [accessed 2022-09-06]
34. Penchansky R, Thomas J. The concept of access: definition and relationship to consumer satisfaction. Med Care 1981 Feb;19(2):127-140. [doi: 10.1097/00005650-198102000-00001] [Medline: 7206846]
35. Saurman E. Improving access: modifying Penchansky and Thomas's Theory of Access. J Health Serv Res Policy 2016 Jan;21(1):36-39. [doi: 10.1177/1355819615600001] [Medline: 26377728]
36. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
37. Strauss A, Corbin J. Basics of Qualitative Research (3rd ed.): Techniques and Procedures for Developing Grounded Theory. Thousand Oaks, California: SAGE Publications; 2008.
38. Holsti O. Content Analysis for the Social Sciences and Humanities. Harlow: Longman Higher Education; 1969.
39. Glass DC, Kelsall HL, Slegers C, Forbes AB, Loff B, Zion D, et al. A telephone survey of factors affecting willingness to participate in health research surveys. BMC Public Health 2015 Oct 05;15:1017 [FREE Full text] [doi: 10.1186/s12889-015-2350-9] [Medline: 26438148]
40. Guest G, MacQueen K, Namey E. Applied Thematic Analysis. Thousand Oaks, California: SAGE Publications; 2012.
41. Hirko K, Kerver J, Ford S, Szafranski C, Beckett J, Kitchen C, et al. Telehealth in response to the COVID-19 pandemic: implications for rural health disparities. J Am Med Inform Assoc 2020 Nov 01;27(11):1816-1818 [FREE Full text] [doi: 10.1093/jamia/ocaa156] [Medline: 32589735]
42. Zahnd WE, Bell N, Larson AE. Geographic, racial/ethnic, and socioeconomic inequities in broadband access. J Rural Health 2022 Jun 18;38(3):519-526. [doi: 10.1111/jrh.12635] [Medline: 34792815]
43. Broadband and mobile programmes. Ministry of Business, Innovation & Employment New Zealand Government. URL: https://www.mbie.govt.nz/science-and-technology/it-communications-and-broadband/fast-broadband/broadband-and-mobile-programmes/ [accessed 2020-07-18]
44. Steen M, Manschot M, De Koning N. Benefits of co-design in service design projects. Int J Design 2011;5(2):53-60.
45. Fernandes LG, Devan H, Fioratti I, Kamper SJ, Williams CM, Saragiotto BT. At my own pace, space, and place: a systematic review of qualitative studies of enablers and barriers to telehealth interventions for people with chronic pain. Pain 2022 Feb 01;163(2):e165-e181. [doi: 10.1097/j.pain.0000000000002364] [Medline: 34433776]

46. 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]
47. Elers P, Hunter I, Whiddett D, Lockhart C, Guesgen H, Singh A. User requirements for technology to assist aging in place: qualitative study of older people and their informal support networks. JMIR Mhealth Uhealth 2018 Jun 06;6(6):e10741 [FREE Full text] [doi: 10.2196/10741] [Medline: 29875083]
48. Hunter I, Elers P, Lockhart C, Guesgen H, Singh A, Whiddett D. Issues associated with the management and governance of sensor data and information to assist aging in place: focus group study with health care professionals. JMIR Mhealth Uhealth 2020 Dec 02;8(12):e24157 [FREE Full text] [doi: 10.2196/24157] [Medline: 33263551]
49. Goold SD. Trust, distrust and trustworthiness. J Gen Intern Med 2002 Jan;17(1):79-81 [FREE Full text] [doi: 10.1046/j.1525-1497.2002.11132.x] [Medline: 11903779]
50. Read L, Nelson H, Korenda L. Rebuilding trust in health care What do consumers want—and need—organizations to do? Deloitte Insights. 2021 Aug 5. URL: https://www2.deloitte.com/us/en/insights/industry/health-care/trust-in-health-care-system.html [accessed 2022-09-06]
51. Can I get faster broadband? Crown Infrastructure Partners. URL: https://www.crowninfrastructure.govt.nz/ [accessed 2022-09-06]
52. Graham R, Masters-Awatere B. Experiences of Māori of Aotearoa New Zealand's public health system: a systematic review of two decades of published qualitative research. Aust N Z J Public Health 2020 Jun;44(3):193-200 [FREE Full text] [doi: 10.1111/1753-6405.12971] [Medline: 32311187]
53. Palmer S, Gray H, Huria T, Lacey C, Beckert L, Pitama SG. Reported Māori consumer experiences of health systems and programs in qualitative research: a systematic review with meta-synthesis. Int J Equity Health 2019 Oct 28;18(1):163 [FREE Full text] [doi: 10.1186/s12939-019-1057-4] [Medline: 31660988]
54. Short S. Focus group interviews. In: A Handbook for Social Science Field Research: Essays & Bibliographic Sources on Research Design and Methods. Thousand Oaks, California: SAGE Publications; 2006.
55. Braun V, Clarke V. To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. Qual Res Sport Exercise Health 2019 Dec 26;13(2):201-216. [doi: 10.1080/2159676x.2019.1704846]
56. Anderson C. Presenting and evaluating qualitative research. Am J Pharm Educ 2010 Oct 11;74(8):141 [FREE Full text] [doi: 10.5688/aj7408141] [Medline: 21179252]

Edited by A Mavragani; submitted 20.12.21; peer-reviewed by K Day, D Parry, L Friesen, E van der Velde; comments to author 26.04.22; revised version received 18.07.22; accepted 10.08.22; published 04.11.22.

Please cite as:
Hunter I, Lockhart C, Rao V, Tootell B, Wong S
Enabling Rural Telehealth for Older Adults in Underserved Rural Communities: Focus Group Study
JMIR Form Res 2022;6(11):e35864
URL: https://formative.jmir.org/2022/11/e35864
doi:10.2196/35864
PMID:36331533

Original Paper

Atrial Fibrillation Detection With an Analog Smartwatch: Prospective Clinical Study and Algorithm Validation

David Campo[1], PhD; Valery Elie[1], PhD; Tristan de Gallard[1], MSc; Pierre Bartet[1], PhD; Tristan Morichau-Beauchant[2], MD; Nicolas Genain[1], MSc; Antoine Fayol[3], MD; David Fouassier[4], MD; Adrien Pasteur-Rousseau[5], MD; Etienne Puymirat[3], MD, PhD; Julien Nahum[2], MD

[1]Withings, Issy Les Moulineaux, France

[2]Intensive Care Unit, Centre Cardiologique du Nord, Sainte-Denis, France

[3]Cardiology Intensive Care Unit, Hopital Europeen Georges Pompidou, Paris, France

[4]Institut Coeur Paris Centre Turin, Paris, France

[5]Institut Coeur Paris Centre Floréal, Bagnolet, France

Corresponding Author:
Valery Elie, PhD
Withings
2 rue Maurice Hartmann
Issy Les Moulineaux, 92130
France
Phone: 33 637430705
Email: valery.elie@withings.com

Abstract

Background: Atrial fibrillation affects approximately 4% of the world's population and is one of the major causes of stroke, heart failure, sudden death, and cardiovascular morbidity. It can be difficult to diagnose when asymptomatic or in the paroxysmal stage, and its natural history is not well understood. New wearables and connected devices offer an opportunity to improve on this situation.

Objective: We aimed to validate an algorithm for the automatic detection of atrial fibrillation from a single-lead electrocardiogram taken with a smartwatch.

Methods: Eligible patients were recruited from 4 sites in Paris, France. Electrocardiograms (12-lead reference and single lead) were captured simultaneously. The electrocardiograms were reviewed by independent, blinded board-certified cardiologists. The sensitivity and specificity of the algorithm to detect atrial fibrillation and normal sinus rhythm were calculated. The quality of single-lead electrocardiograms (visibility and polarity of waves, interval durations, heart rate) was assessed in comparison with the gold standard (12-lead electrocardiogram).

Results: A total of 262 patients (atrial fibrillation: n=100, age: mean 74.3 years, SD 12.3; normal sinus rhythm: n=113, age: 61.8 years, SD 14.3; other arrhythmia: n=45, 66.9 years, SD 15.2; unreadable electrocardiograms: n=4) were included in the final analysis; 6.9% (18/262) were classified as Noise by the algorithm. Excluding other arrhythmias and Noise, the sensitivity for atrial fibrillation detection was 0.963 (95% CI lower bound 0.894), and the specificity was 1.000 (95% CI lower bound 0.967). Visibility and polarity accuracies were similar (1-lead electrocardiogram: P waves: 96.9%, QRS complexes: 99.2%, T waves: 91.2%; 12-lead electrocardiogram: P waves: 100%, QRS complexes: 98.8%, T waves: 99.5%). P-wave visibility accuracy was 99% (99/100) for patients with atrial fibrillation and 95.7% (155/162) for patients with normal sinus rhythm, other arrhythmias, and unreadable electrocardiograms. The absolute values of the mean differences in PR duration and QRS width were <3 ms, and more than 97% were <40 ms. The mean difference between the heart rates from the 1-lead electrocardiogram calculated by the algorithm and those calculated by cardiologists was 0.55 bpm.

Conclusions: The algorithm demonstrated great diagnostic performance for atrial fibrillation detection. The smartwatch's single-lead electrocardiogram also demonstrated good quality for physician use in daily routine care.

Trial Registration: ClinicalTrials.gov NCT04351386; http://clinicaltrials.gov/ct2/show/NCT04351386

(JMIR Form Res 2022;6(11):e37280) doi:10.2196/37280

KEYWORDS

atrial fibrillation; mobile health; mHealth; diagnosis; electrocardiogram; ECG; smartwatch; smart technology; wearable; cardiology; cardiac; heart failure; heart disease; cardiovascular; morbidity; automatic detection; algorithm; physician; sensor; digital health

Introduction

Atrial fibrillation is the most common form of arrhythmia in the world—it affects 8 million people in Europe and 5 million people in the United States [1]. Experts believe that the number of patients with atrial fibrillation will increase in the next few years to one-quarter of middle-aged adults in the United States and in Europe [2,3]. Despite improvements in its management, atrial fibrillation remains one of the major causes of stroke, heart failure, sudden death, and cardiovascular morbidity in the world [3,4].

Atrial fibrillation is associated with a 5-fold increase in the risk of stroke [5-9]. Large randomized controlled trials [10,11] are underway to evaluate stroke prevention by using anticoagulation treatment in patients with subclinical atrial fibrillation. Asymptomatic patients represent 32.6% to 39.4% of patients in large international registries [12,13]. Thus subclinical atrial fibrillation could represent approximately one-third of the atrial fibrillation population [14] and is admittedly found frequently among older adults [15,16].

Traditionally, atrial fibrillation is diagnosed using an electrocardiogram (ECG), a Holter monitor worn for 24 to 48 hours, an event recorder monitoring heart activity for several weeks, or implanted pacemakers or defibrillators with an atrial lead. There is trade-off in efficacy of detection of paroxysmal atrial fibrillation between short- and long-term monitoring.

Given the increasing number of patients with asymptomatic atrial fibrillation, new simple and efficient diagnostic devices are important supplements to traditional methods to allow early diagnosis [17], screening [18], or management [19]. It is therefore crucial that these devices be evaluated in clinical studies in comparison with standard 12-lead electrocardiography [20].

We aimed to validate the diagnostic performance (ie, classification into atrial fibrillation or normal sinus rhythm) and safety of a single-channel ECG smartwatch (ScanWatch, Withings) and its associated software (Scan Monitor, Withings Inc), in the detection of atrial fibrillation in comparison with reference diagnoses made by independent blind cardiologists using simultaneously recorded 12-lead ECG.

Methods

Study Design and Population

We conducted a prospective nonrandomized open-label comparative multicenter study. Cardiology in-patients and out-patients were consecutively screened from 4 sites in Paris, France. Inclusion criteria were male or female patients aged 18 years or older with atrial fibrillation or sinus rhythm, recruited with a 1:1 ratio. Patients with pacemakers, who were physically incapable of wearing a watch on their wrist, with linguistic or mental incapacity that precluded signing a written informed consent form, or vulnerable individuals (as defined by French regulation) were not included. Eligible patients were informed and provided consent prior to any study-related procedure.

Study Procedures

Data Collection

For each patient, simultaneous 30-second single-lead ECGs were recorded with ScanWatch with embedded software (Scan Monitor) and 12-lead ECGs recorded with Schiller Cardiovit FT1 electrocardiograph (CE marked and FDA cleared [21]).

ScanWatch uses 3 dry electrodes to record a 30-second single-lead ECG that is similar to lead I of a traditional 12-lead ECG. A real-time signal captured with the watch is streamed to a smartphone app (Health Mate, Withings; for Android and iOS), stored, exported in PDF format, and classified into 4 categories—atrial fibrillation, normal sinus rhythm, noise, or other—by the proprietary algorithm. The classification is performed on 30-second recordings using features related to the visibility of P waves and the irregularity of R-R intervals. Algorithm classifications were kept on Withings servers, and investigators were blinded from it. Data from the 12-lead ECG were exported using DICOM/HL7 ECG Waveform Export software (SemaServer, version 19.02; Schiller) in accordance with manufacturer's instructions. Patient information and data were collected and reported by site staff on the study case report form.

Data Review

The 12-lead reference ECGs and single-lead smartwatch ECGs were independently reviewed by blinded, board-certified cardiologists. Each recording was reviewed by 3 reviewers. The reviewers were instructed to classify each recording into one of the following categories: (1) normal sinus rhythm, (2) atrial fibrillation, (3) supraventricular tachycardia, (4) abnormal rhythm, such as frequent premature atrial contractions, frequent premature ventricular contractions, atrial flutter, ventricular tachycardia, ventricular fibrillation, second-degree atrioventricular block type I, second-degree atrioventricular block type II, third-degree atrioventricular block, and other, or (5) uninterpretable (ie, a classification cannot be made as the strip is not adequate for reading). If there were multiple rhythms, reviewers reported all the rhythms, then classified the recording in one of the above classes with justification. If there was a discrepancy between reviewers' classifications, the diagnosis of the majority was retained. For the 3 primary rhythms (normal sinus rhythm, atrial fibrillation, and supraventricular tachycardia), if classifications differed, the final classification was determined by consensus.

Reviewers' classifications were compared to the algorithm's automatic classification of the single-channel strips collected with the smartwatch to assess atrial fibrillation detection performance. The software classified the smartwatch strips as normal sinus rhythm, atrial fibrillation, Noise, or other

arrhythmia (supraventricular tachycardia and other abnormal rhythm reference diagnoses were pooled).

To assess the quality of the single-lead ECG signals generated by the smartwatch and whether the smartwatch could be used by cardiologists in clinical practice, cardiologists assessed the visibility and polarity of P, QRS, and T waves; measured the durations of PR, QRS, and QT intervals; and measured heart rate. An additional reviewer selected a well-defined beat to be later used by the reviewers for secondary outcome measures.

Diagnoses made by the reviewers from the single-channel strips generated by the smartwatch were compared with those from the 12-lead ECG.

Statistical Analysis

The primary outcomes were sensitivity and specificity of atrial fibrillation detection. Sensitivity and specificity were calculated, first, by considering all available categories, and second, by excluding Other and Noise. The reason for the second calculation is that the device is primarily intended to discriminate between atrial fibrillation and normal sinus rhythm. Calculated sensitivities and specificities are reported with their lower confidence interval bounds; their positive and negative likelihood ratios were also calculated.

Continuous variables were expressed as mean and standard deviation, or median and range, while categorical variables were expressed as numbers and proportions. Exact 95% confidence intervals of proportions were calculated with the Clopper-Pearson method. Results for the 4 classes (atrial fibrillation, normal sinus rhythm, Noise, and Other) are presented with a 4×4 confusion matrix.

Baseline characteristics were compared using the *t* test for sample means for normal distributions, Mann-Whitney test for sample medians otherwise, and Fisher exact test for proportions. A Shapiro-Wilk test was used to test normality.

Cardiologist agreement was measured using Cohen κ and average accuracy in comparison with the consensus diagnosis.

Sample size was calculated for sensitivities and specificities >0.9 and a statistical power >90%. All statistical tests were 2-sided with a statistical significance threshold at .05. Analyses were performed with Python (version 3.6.8) scikit-learn (version 0.23.2), statsmodels (version 0.12.2), and scipy (version 1.5.2) toolkits.

Ethics

This study was conducted in compliance with Good Clinical Practice and 1964 Declaration of Helsinki and subsequent amendments [22]. The study was approved by the French *Comité de Protection des PersonnesCPP Sud-Est IV* (19.06.28.65727) and registered (ClinicalTrials.gov, NCT04351386).

Results

Population Characteristics

Between December 2019 to April 2021, 283 patients were enrolled in the study; however, 10 patients prematurely discontinued their participation, and 11 patients were excluded, which resulted in an analysis set of 262 patients (Figure 1). Patients characteristics (men: n=160, women: n=102; age: mean 67.7 years, SD 14.8, BMI: mean 27.5, SD 5.7 kg/m^2) are presented in Table 1.

Figure 1. Study flowchart. ECG: electrocardiogram.

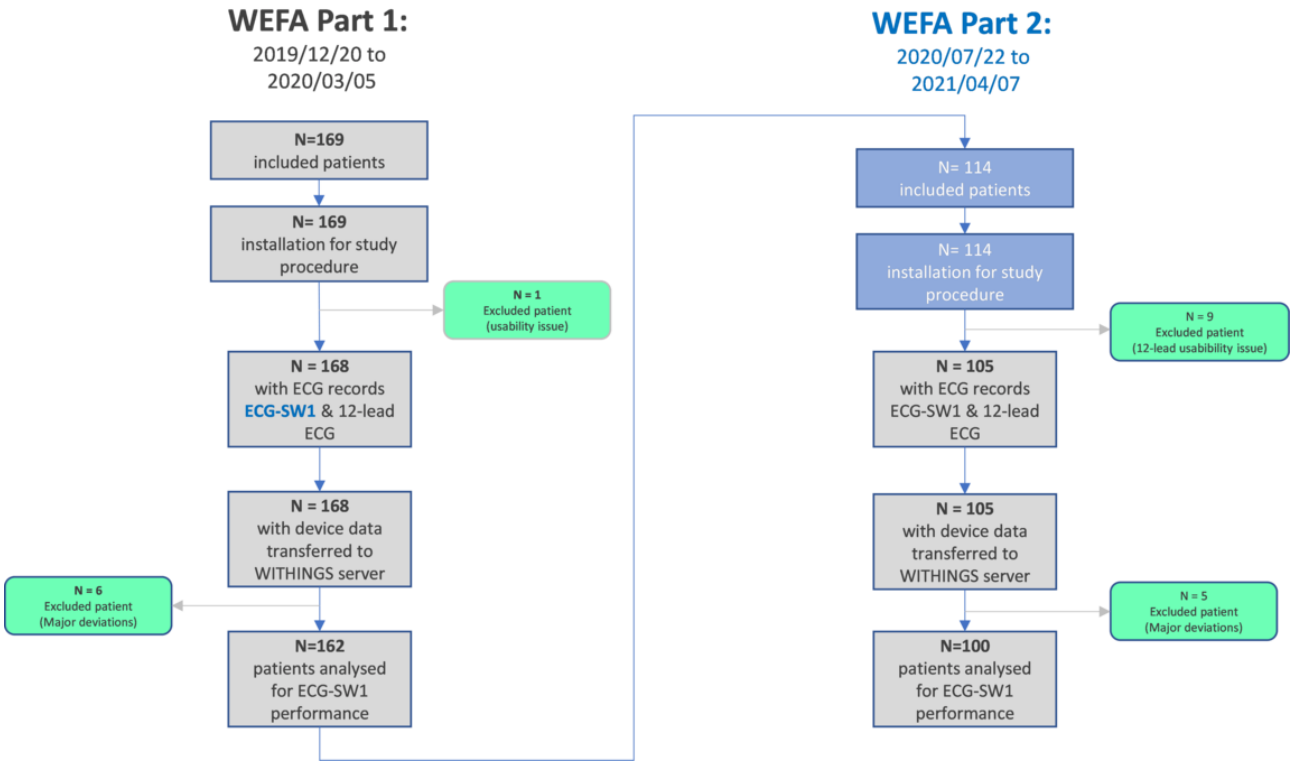


Table 1. Baseline patient characteristics.

Characteristic	All patients (n=262)	Atrial fibrillation (n=100)	Normal sinus rhythm (n=113)	Other arrhythmias (n=45)	Unreadable ECG[a] (n=4)
Age (years), mean (SD)	67.7 (14.8)	74.3 (12.3)	61.8 (14.3)	66.9 (15.2)	78.8 (12.5)
Height (cm), mean (SD)	169.2 (9.2)	168.8 (9.4)	169.7 (9.2)	169.3 (9.0)	163.8 (7.8)
Weight (kg), mean (SD)	78.9 (17.6)	79.1 (19.1)	79.1 (15.1)	78.9 (20.1)	67.8 (18.7)
BMI[b] (kg/m^2), mean (SD)	27.5 (5.7)	27.6 (6.0)	27.5 (5.1)	27.5 (6.4)	25.3 (6.5)
Sex, n (%)					
Female	102 (38.9)	42 (42.0)	39 (34.5)	18 (40.0)	3 (75.0)
Male	160 (61.1)	58 (58.0)	74 (65.5)	27 (60.0)	1 (25.0)
Patient follow-up at site during study procedure, n (%)					
In-patient	205 (78.2)	91 (91.0)	72 (63.7)	38 (84.4)	4 (100)
Out-patient	57 (21.8)	9 (9.0)	41 (36.3)	7 (15.6)	0 (0.0)
Cardiovascular medical history, n (%)					
Atrial fibrillation	122 (46.6)	91 (91.0)	11 (9.7)	19 (42.2)	1 (25.0)
Valvular heart disease with or without intervention before inclusion	46 (17.6)	25 (25.0)	12 (10.6)	8 (17.8)	1 (25.0)
Coronary artery disease with or without coronary artery bypass grafting before inclusion	72 (27.5)	27 (27.0)	30 (26.5)	14 (31.1)	1 (25.0)
Heart failure	26 (9.9)	20 (20.0)	3 (2.7)	3 (6.7)	0 (0.0)
Myocardial infarction or ischemic cardiopathy	31 (11.8)	17 (17.0)	9 (8.0)	5 (11.1)	0 (0.0)
Transient ischemic attack or stroke	12 (4.6)	9 (9.0)	2 (1.8)	1 (2.2)	0 (0.0)
Peripheral arterial obstructive disease	18 (6.9)	7 (7.0)	8 (7.1)	3 (6.7)	0 (0.0)
Abdominal aortic aneurysm	4 (1.5)	1 (1.0)	3 (2.7)	0 (0.0)	0 (0.0)
Cardiovascular risk factors, n (%)					
Hypertension	135 (51.5)	60 (60.0)	51 (45.1)	23 (51.1)	1 (25.0)
Dyslipidemia	90 (34.4)	36 (36.0)	39 (34.5)	15 (33.3)	0 (0.0)
Former or current smoker	77 (29.4)	28 (28.0)	36 (31.9)	13 (28.9)	0 (0.0)
Overweight	70 (26.7)	27 (27.0)	32 (28.3)	9 (20.0)	2 (50.0)
Obesity	42 (16.0)	14 (14.0)	22 (19.5)	6 (13.3)	0 (0.0)
Diabetes	57 (21.8)	23 (23.0)	23 (20.4)	11 (24.4)	0 (0.0)
Position during ECG recording, n (%)					
Supine	115 (43.9)	44 (44.0)	50 (44.2)	18 (40.0)	3 (75.0)
Sitting	147 (56.1)	56 (56.0)	63 (55.8)	27 (60.0)	1 (25.0)
Wrist					
Left	179 (68.3)	66 (66.0)	82 (72.6)	30 (66.7)	1 (25.0)
Right	83 (31.7)	34 (34.0)	31 (27.4)	15 (33.3)	3 (75.0)
Skin type, n (%)					
White	204 (77.9)	82 (82.0)	84 (74.3)	35 (77.8)	3 (75.0)
Mediterranean/Arabic	35 (13.4)	15 (15.0)	13 (11.5)	6 (13.3)	1 (25.0)
Black	23 (8.8)	3 (3.0)	16 (14.2)	4 (8.9)	0 (0.0)

[a]ECG: electrocardiogram.

[b]BMI: body mass index.

The atrial fibrillation group (mean 74.3 years, SD 12.3) was significantly older than the normal sinus rhythm group (mean 61.8 years, SD 14.3, $P<.001$) and other arrhythmia groups (mean 66.9 years, SD 15.2, $P=.002$). Similarly, the proportion of

in-patients was higher in the atrial fibrillation group (91/100, 91%) than those in the normal sinus rhythm (72/113, 64%) and other arrhythmia groups (38/45, 84%). Out-patients (age: mean 59.1 years) were significantly younger (P<.001) than in-patients (age: mean 70.1 years).

Among all cardiovascular risk factors, hypertension was the most represented in all subgroups (atrial fibrillation: 60/100, 60%; normal sinus rhythm: 51/113, 45%; other arrhythmia: 23/45, 51%).

Cardiovascular arterial disease was the most represented cardiovascular risk with similar prevalences across the 3 subgroups (atrial fibrillation: 27/100, 27%; normal sinus rhythm: 30/113, 27%; other arrhythmia: 14/45, 31%). Only 9 patients (9/100, 9%) in the atrial fibrillation subgroup did not have a history of atrial fibrillation before participating in the study. No adverse events occurred during the study.

Automatic Atrial Fibrillation Detection Performance

Supraventricular tachycardia was not diagnosed by any of the independent cardiologist reviewers using 12-lead ECGs. No consensus meetings were needed. The average accuracy of the 3 cardiologists was 0.92 (mean Cohen κ=0.88).

Four reference ECG were labeled as Noise, 2 of which were classified as Noise by ECG Monitor, and 2 as atrial fibrillation. ECG Monitor classified 6.9% (18/262) of the recordings performed with the watch as Noise, but none was from a patient with normal sinus rhythm (Table 2). When considering all 4 categories, the sensitivity to detect atrial fibrillation was 0.77 (95% CI lower bound 0.675), and the specificity was 0.965 (95% CI lower bound 0.912). Of the 113 normal sinus rhythm diagnoses based on the 12-lead reference ECG, only 1 (0.85%) was classified as Noise by the algorithm, and 13 (13.0%) patients diagnosed with atrial fibrillation based on the 12-lead reference ECG were classified as Noise by the algorithm. When excluding the categories *other arrhythmia* and *Noise* from the calculation, the sensitivity was 0.963 (95% CI lower bound 0.894), and the specificity was 1.000 (95% CI lower bound 0.967). Inconclusive measurements (Other and Noise) occurred more frequently in patients >65 years old (odds ratio [OR] 4.34, 95% CI 1.25-15.09; P=.02) and in patients with previously diagnosed atrial fibrillation (OR 3.75, 95% CI 1.43-9.87; P=.007). While not statistically significant, in-patients and patients with hypertension or valvular heart disease tended to have more inconclusive ECGs.

Table 2. Algorithm classification (1-lead ECG[a]) versus cardiologist diagnosis (12-lead ECG).

Algorithm classification (Smartwatch 1-lead ECG)	Cardiologist diagnosis from 12-lead ECG				
	Normal sinus rhythm, n	Atrial fibrillation, n	Other, n	Noise, n	Total, n
Normal sinus rhythm	109	3	9	0	121
Atrial fibrillation	0	77	11	2	90
Other	3	7	23	0	33
Noise	1	13	2	2	18
Total	113	100	48	4	262

[a]ECG: electrocardiogram.

ECG Signal Quality

Diagnostic Accuracy

The sensitivity for detecting atrial fibrillation was 0.89; the specificity was 0.912 (Table 3). The average accuracy between the 3 cardiologists reading single-channel recordings from the smartwatch was 0.785, and the average Cohen κ was 0.675, which reflected strong agreement between reviewers.

Table 3. Cardiologist diagnosis (1-lead ECG[a] vs 12-lead ECG).

Cardiologist diagnosis from Smartwatch 1-lead ECG	Cardiologist diagnosis from 12-lead ECG				
	Normal sinus rhythm, n	Atrial fibrillation, n	Other, n	Noise, n	Total, n
Normal sinus rhythm	103	0	1	0	104
Atrial fibrillation	2	89	17	2	110
Other	3	2	20	0	25
Noise	5	9	7	2	23
Total	113	100	45	4	262

[a]ECG: electrocardiogram.

Visibility and Polarity

P-wave visibility accuracy was 99% (99/100) in patients with atrial fibrillation and 95.7% (155/162) when excluding patients with atrial fibrillation (Table 4).

Table 4. Cardiologist review of 1-lead electrocardiogram (ECG) versus 12-lead ECG: P-wave, T-wave, and QRS-complex visibility.

	Identified (n) and accuracy (%) (n=262)
Visibility	
P wave	254 (96.9)
QRS complex	260 (99.2)
T wave	239 (99.5)
Polarity[a]	
P wave	88 (100)
QRS complex	254 (98.8)
T wave	183 (99.5)

[a]n=88, n=257, and n=184 for P wave, QRS complex, and T wave, respectively.

Interval Durations

Except for QT intervals, mean difference absolute values for difference in PR duration and QRS width were less than 3 ms and more than 97% of these differences were less than 40 ms (Table 5).

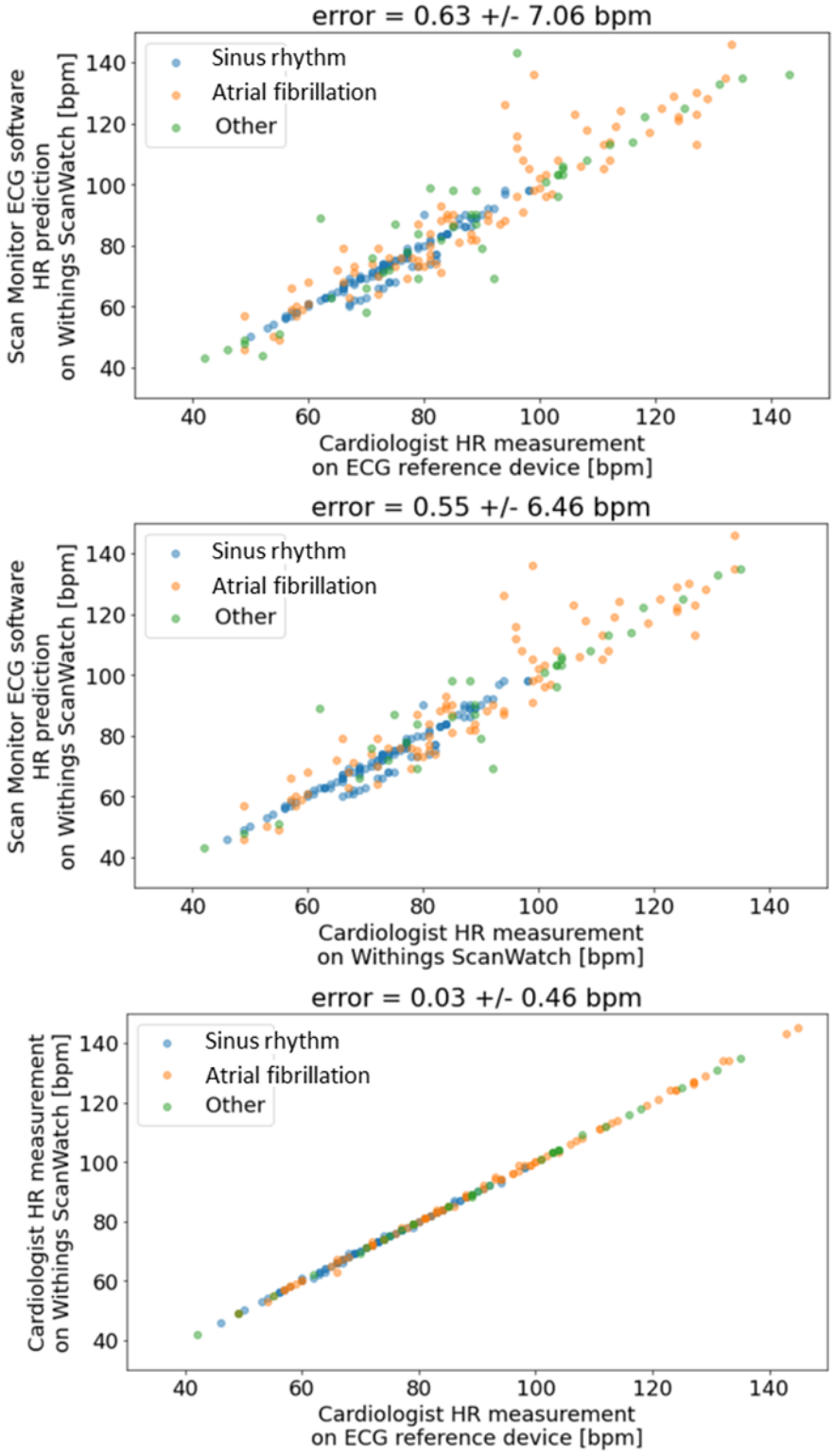
Table 5. Interval duration differences between the algorithm's and cardiologists' assessments.

	Duration difference (ms), mean (SD)	Difference <40 ms, n/N (%)
PR duration	–2.79 (17.25)	92/94 (97.9)
QRS width	–0.46 (18.91)	245/252 (97.2)
QT duration	7.13 (26.30)	161/182 (88.5)

Heart Rate

For reviewers' assessments of heart rate, 1-lead versus 12-lead ECG had the smallest difference (mean 0.03 bpm, SD 0.46). The largest difference was observed between the heart rate calculated by the algorithm compared with that by cardiologists based on 12-lead ECG (mean 0.63 bpm, SD 7.06). The mean difference between heart rates calculated by the algorithm and those by cardiologists based on 1-lead ECG was 0.55 bpm (SD 6.46) (Figure 2).

Figure 2. Heart rate calculations by the algorithm versus those by cardiologists. ECG: electrocardiogram; HR: heart rate.



Discussion

Study Strengths

Biases were limited by recruiting consecutive patients, taking the recordings simultaneously, centralizing ECG review with independent blinded reviewers, and standardizing the review process. Participants were in- and out-patients from cardiology services with multiple comorbidities. None was familiar with the device beforehand.

Principal Results and Comparison With Prior Work

Performance for Automatic Atrial Fibrillation Detection by Scan Monitor Software

Only 4 reference ECGs were deemed unreadable; this may occur for multiple reasons especially movements during the measurements.

The algorithm's ability to discriminate between atrial fibrillation and normal sinus rhythm was calculated both considering all 4 initial categories (sensitivity 0.770; specificity 0.965) and excluding Other and Noise signals (sensitivity 0.963; specificity 1.000).

Misclassifications for patients normal sinus rhythm or atrial fibrillation were very rare. In particular, the rate of false positives was 2.7% (3/113) for normal sinus rhythm, and no patients with atrial fibrillation were classified as patients with normal sinus rhythm by the algorithm. Most false negatives were misclassifications of other arrhythmias, while false positives were mostly misclassifications of patients with normal sinus rhythm or atrial fibrillation as Other. This is unfortunately associated with the algorithm development process being optimized to accurately identify atrial fibrillation and distinguish atrial fibrillation from normal sinus rhythm. Therefore, automatic classification reliability on patients with arrhythmias other than atrial fibrillation may be decreased.

These results are quite similar to those from studies on other wearable devices, such as the Apple Watch (Apple Inc) [23], the Kardia Band (AliveCor) [24], and devices by other manufacturers [25-28] (Table 6). The Apple study recruited, by far, the most patients: almost twice the number included in our study and 4 times that in [24]. All the devices had high rates of noisy or poor-quality signals (6.9% to 16.6%). For data loss, there was only 1 case (1/283, 0.4%) with ScanWatch, none (0/169, 0%) in the Kardia band study [24], and 7.6% (46/602) in Apple Watch study. Moreover, ScanWatch and Kardia band classified 11.3% (24/213) and 17.2% (29/169), respectively, of atrial fibrillation and normal sinus rhythm signals as other arrhythmias (or as unclassified or inconclusive); Apple Watch had only 2.2% (13/602) of such errors, and consequently, had higher sensitivities and specificities than ScanWatch and Kardia band. Overall, Kardia band was less accurate than ScanWatch and Apple Watch, with ScanWatch being more accurate than the Apple Watch for detecting normal sinus rhythm.

Only sparse data were found for other manufacturers. For Samsung's ECG monitor, 16.8% of ECG recordings were considered either inconclusive or of poor quality [26], similar to Kardia band and Apple Watch; sensitivity was 98.1%, and specificity was 100%. Similarly, Fitbit ECG app performances were 98.7% and 100% for correctly detecting atrial fibrillation and normal sinus rhythm, respectively [28]. Similar results were published for Amazfit [27,29].

Table 6. Performances of commercially available devices with electrocardiogram (ECG) sensors.

	Withings ScanWatch	Apple Watch [23]	Kardia Band [24]	MyDiagnostick [25]	Samsung ECG Monitor [26]	Amazfit Health Band 1S [27]	Amazfit Cardidoc app [29]	Fitbit ECG app [28]
Patients recruited, n	283	602	169	192	544	401	114	472
Patients analyzed, n	262	553	169	192	544	401	114	440
Noise or unreadable, n (%)	182 (6.9)	49 (8.9)	28 (16.6)	—[a]	— (16.8)	15 (3.74)	2 (1.75)	—
Other or unclassifiable, n (%)	33 (12.6)	19 (3.4)	29 (17.2)	—	—	—	—	—
All, n/N (%)								
Sensitivity	77/100 (77.0)	236/277 (85.2)	63/91 (69.2)	—	—/— (87.1)	—	—	—
Specificity	109/118 (92.4)	238/263 (90.5)	37/78 (47.4)	—	—/— (82.5)	—	—	—
Atrial fibrillation and normal sinus rhythm, n/N (%)								
Sensitivity	77/80 (96.3)	236/240 (98.3)	63/68 (92.6)	—/— (100)	—/— (98.1)	—/— (96.7)	—/— (88.7)	—/— (98.7)
Specificity	109/109 (100)	237/238 (99.6)	37/44 (84.1)	—/— (95.9)	—/— (100)	—/— (98.0)	—/— (100)	—/— (100)

[a]No data were available.

In addition, the ScanWatch algorithm had lower performance in heart failure (sensitivity 0.800, 95% CI lower bound 0.519), myocardial infarction or ischemic cardiopathy (sensitivity 0.769, 95% CI lower bound 0.462), and peripheral arterial obstructive disease (sensitivity 0.714, 95% CI lower bound 0.290) subgroups in detecting atrial fibrillation and normal sinus rhythm. All these patients had various comorbidities and associated medications. Moreover, 11 out of the 31 patients (35.5%) with a history of myocardial infarction or ischemic cardiopathy were also diagnosed with heart failure (ie, had multiple comorbidities).

In a registry-based study [30] conducted in 136 and 211 European cardiology centers, major ECG abnormalities were present in the majority of patients with heart failure. Out of 1460 patients with heart failure, 1222 had major ECG abnormalities with various patterns across the heart failure types [30], and the Euroheart Failure survey showed that ECG

abnormalities were present in 98% of patients with heart failure [31]. Older age and being male were associated with an increased risk of ECG abnormalities, as were history of advanced heart valve disease, chronic kidney disease, signs of heart failure decompensation, and use of diuretics and anticoagulants [31].

ECG abnormalities, such as abnormal rhythm, PR duration >250 ms, QRS interval ≥120 ms, and pathological Q waves can be found in most patients with multiple severe underlying diseases and risks factors. Similar patterns were observed in our study (Table 7). Although patients' medications were not recorded in our study, specific heart failure ECG abnormality risk factors were recorded (sex, age, valvular heart disease, kidney disease). Patients with heart failure have multiple comorbidities that affect their ECGs, and such a population can present ECG abnormalities that are too complex for our algorithm—this represents an extreme worst-case scenario for the device's algorithm—these patients alone represent 21.4% (56/262) of the study cohort.

Table 7. Patient subgroups (heart failure, peripheral arterial obstructive disease, myocardial infarction, or ischemic cardiopathy).

	All other (n=206)	Heart failure (n=26)	Peripheral arterial obstructive disease (n=18)	Myocardial infarction or ischemic cardiopathy (n=31)
Sex, n (%)				
Male	124 (60.2)	18 (69.2)	10 (55.6)	20 (64.5)
Female	82 (39.8)	8 (30.8)	8 (44.4)	11 (35.5)
Age (years), mean (SD)	66.2 (15.5)	75.3 (11.6)	74.3 (8.5)	71.9 (9.8)
Heart rate (bpm), mean (SD)	82.9 (19.9)	86.6 (23.0)	72.2 (10.9)	75.1 (18.8)
Valvular heart disease, n (%)	37 (18.0)	8 (30.8)	3 (16.7)	5 (16.1)
Diabetes, n (%)	42 (20.4)	7 (26.9)	5 (27.8)	6 (19.4)
QRS length (ms), mean (SD)	96.6 (29.4)	114.8 (35.3)	104.8 (26.4)	103.8 (26.5)
QRS >120 ms, n (%)	32 (15.5)	10 (38.5)	5 (27.8)	7 (22.6)

1-Lead ECG Quality Assessment

To the best of our knowledge, this study proposes for the first time several criteria for device signal quality assessment without bias (quantitative assessment) that are based on criteria taken from clinical practice.

In total, cardiologists declared 8.8% (23/262) of the signals as uninterpretable, and 2 patients with atrial fibrillation were classified as normal sinus rhythm (2/100). These misclassifications may be explained by artifacts. However in some cases (eg, paroxystic atrial fibrillation is more difficult to diagnose with a 30-second recording), a confirmatory ECG can improve the reliability of the diagnosis.

Accuracies in the assessment of T-wave, P-wave, and QRS-complex visibilities and polarities were high (over 96%, except for T wave: 91%). The accuracy of PR, QRS, QT interval durations was good. In particular, the standard deviation of the differences fell below 20 ms for PR and QRS time intervals, which is less than half the length of the smallest graduation (1 mm) on a standard paper ECG trace (40 ms). The standard deviation of QT interval difference was slightly higher (SD 26.3 ms).

Because Apple used a different method to assess ECG waveform quality, a comparison of ECG quality between the devices was not possible.

Comparison of cardiologist-measured heart rate on the single-lead and lead I of the 12-lead reference ECG, yielded a 0 bias and a standard deviation of the difference of 0.5 bpm.

The standard deviation between automatic heart rate calculation by ScanWatch and heart rate measurements by cardiologists on lead I of a reference ECG was 7 bpm. This difference is a consequence of the different calculation methods between reviewers (mean over 10 seconds) and the algorithm (median heart rate over 30 seconds). Given the mean heart rate (mean 82.35, SD 19.9 bpm), the mean error between the heart rate measured by the device and that measured by the reviewers on the reference ECG (mean 0.63, SD 7.09 bpm) was considered acceptable (the accuracy of the detected heart rate shall be ±10% or ±5/min, whichever is greater [32]).

Limitations

Despite good diagnostic performance in discriminating between atrial fibrillation and normal sinus rhythm, arrhythmias other than atrial fibrillation were poorly identified: only 51.1% (23/45) of these signals were correctly classified. This poor performance was expected since, by design, the algorithm was trained to specifically identify atrial fibrillation and normal sinus rhythm. Flutter signals were not negligible in our population (13/45) and were misclassified as atrial fibrillation by the algorithm: they are not easily identifiable on lead I, even for a cardiologist. Nevertheless, it is encouraging that 80.0% (36/45) of Other signals were not classified as normal sinus rhythm. Unfortunately, patients' medications at the time of the measurements were not recorded: the study lacks information about drugs that may affect cardiac rhythm.

The population enrolled in the study may not reflect the real-world use of the device: the study was conducted in cardiology services, but the device is intended for home use, some of the recordings were performed while patients were supine rather than seated, and no patients were familiar with the device before measurements, which may have increased the amount of noisy data collected.

Reduced accuracy of the machine learning classification algorithm in patients with multiple risk factors and comorbidities may be explained by the fact that data sets used to develop the algorithm did not include such specific populations. As we mentioned, these patients commonly present ECG abnormalities that may affect the algorithm's performance.

Finally, only a single measurement was recorded per patient in order to limit potential bias and test the device in the worst-case scenario. In real-use conditions, a diagnosis from a cardiologist may improve using a second recording, and recording quality may also improve with patient's practice.

Conclusions

ScanWatch, with its embedded software (Scan Monitor), was able to provide high-quality ECG traces that are adequate for clinical diagnosis of atrial fibrillation (ie, accurately discriminate between atrial fibrillation and normal sinus rhythm). Additional studies and additional machine learning work will be needed to increase the algorithm's performance in distinguishing atrial fibrillation from other arrhythmias.

Such a device offers clinicians the ability to remotely monitor patients at risk of atrial fibrillation in their daily practice with a simple, accurate, and noninvasive wearable device.

Acknowledgments

This work was supported by Withings Inc. We thank all the investigators, for their work on the clinical trial, and patients, for their participation. We thank the reviewers for comments that brought more clarity to the manuscript. We would also like to thank our cardiologists (in charge of the independent electrocardiogram [ECG] review for the study): Dr Tanios Akiki, Dr Paul Masi, Dr Ardalan Sharifzadehgan, Dr Victoria Tea, Dr Noemie Tence, Dr Victor Waldmann for their contributions in the EGC review. Finally, we would like to thank Delphine Ladarre for her contributions during the early stages of the study preparations.

Authors' Contributions

DC wrote the protocol, supervised the data analysis, and wrote part of the manuscript. VE supervised the study monitoring and wrote part of the manuscript. TdG and VE conducted the data management. TdG and PB conducted the statistical analyses. NG participated in the study design and reviewed the data analysis. JN was the coordinating investigator. DF, EP, and AP-R were lead site investigators. AF and TM-B were site investigators and wrote part of the manuscript.

Conflicts of Interest

DC, VE, TdG, PB, and NG are employees of Withings. JN received subsidies from Withings for the calibration study and this validation study. TM-B received subsidies from Withings for the calibration study. AF received consultancy fees from Withings.

References

1. Zulkifly H, Lip GYH, Lane DA. Epidemiology of atrial fibrillation. Int J Clin Pract 2018 Mar 01;72(3):e13070. [doi: 10.1111/ijcp.13070] [Medline: 29493854]
2. Colilla S, Crow A, Petkun W, Singer DE, Simon T, Liu X. Estimates of current and future incidence and prevalence of atrial fibrillation in the U.S. adult population. Am J Cardiol 2013 Oct 15;112(8):1142-1147. [doi: 10.1016/j.amjcard.2013.05.063] [Medline: 23831166]
3. Kirchhof P, Benussi S, Kotecha D, Ahlsson A, Atar D, Casadei B, et al. 2016 ESC guidelines for the management of atrial fibrillation developed in collaboration with EACTS. Eur Heart J 2016 Oct 07;37(38):2893-2962. [doi: 10.1093/eurheartj/ehw210] [Medline: 27567408]
4. Davis RC, Hobbs FDR, Kenkre JE, Roalfe AK, Iles R, Lip GYH, et al. Prevalence of atrial fibrillation in the general population and in high-risk groups: the ECHOES study. Europace 2012 Nov;14(11):1553-1559. [doi: 10.1093/europace/eus087] [Medline: 22490371]
5. Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. Stroke 1991 Aug;22(8):983-988 [FREE Full text] [Medline: 1866765]
6. Healey JS, Connolly SJ, Gold MR, Israel CW, Van Gelder IC, Capucci A, ASSERT Investigators. Subclinical atrial fibrillation and the risk of stroke. N Engl J Med 2012 Jan 12;366(2):120-129. [doi: 10.1056/NEJMoa1105575] [Medline: 22236222]
7. Boriani G, Pettorelli D. Atrial fibrillation burden and atrial fibrillation type: Clinical significance and impact on the risk of stroke and decision making for long-term anticoagulation. Vascul Pharmacol 2016 Aug;83:26-35. [doi: 10.1016/j.vph.2016.03.006] [Medline: 27196706]
8. Chen LY, Chung MK, Allen LA, Ezekowitz M, Furie KL, McCabe P, American Heart Association Council on Clinical Cardiology, Council on Cardiovascular and Stroke Nursing, Council on Quality of Care and Outcomes Research, Stroke Council. Atrial fibrillation burden: moving beyond atrial fibrillation as a binary entity: a scientific statement from the American Heart Association. Circulation 2018 May 15;137(20):e623-e644. [doi: 10.1161/CIR.0000000000000568] [Medline: 29661944]

9. Noseworthy PA, Kaufman ES, Chen LY, Chung MK, Elkind MSV, Joglar JA, American Heart Association Council on Clinical Cardiology, Electrocardiography, and Arrhythmias Committee, Council on Arteriosclerosis, Thrombosis, and Vascular Biology, Council on Cardiovascular and Stroke Nursing, Stroke Council. Subclinical and device-detected atrial fibrillation: pondering the knowledge gap: a scientific statement from the American Heart Association. Circulation 2019 Dec 17;140(25):e944-e963 [FREE Full text] [doi: 10.1161/CIR.0000000000000740] [Medline: 31694402]
10. Lopes RD, Alings M, Connolly SJ, Beresh H, Granger CB, Mazuecos JB, et al. Rationale and design of the apixaban for the reduction of thrombo-embolism in patients with device-detected sub-clinical atrial fibrillation (ARTESiA) trial. Am Heart J 2017 Jul;189:137-145. [doi: 10.1016/j.ahj.2017.04.008] [Medline: 28625370]
11. Kirchhof P, Blank BF, Calvert M, Camm AJ, Chlouverakis G, Diener H, et al. Probing oral anticoagulation in patients with atrial high rate episodes: rationale and design of the non-vitamin K antagonist oral anticoagulants in patients with atrial high rate episodes (NOAH-AFNET 6) trial. Am Heart J 2017 Aug;190:12-18 [FREE Full text] [doi: 10.1016/j.ahj.2017.04.015] [Medline: 28760205]
12. Steg PG, Alam S, Chiang C, Gamra H, Goethals M, Inoue H, RealiseAF investigators. Symptoms, functional status and quality of life in patients with controlled and uncontrolled atrial fibrillation: data from the RealiseAF cross-sectional international registry. Heart 2012 Feb;98(3):195-201. [doi: 10.1136/heartjnl-2011-300550] [Medline: 21948959]
13. Lip GYH, Laroche C, Dan G, Santini M, Kalarus Z, Rasmussen LH, et al. A prospective survey in European Society of Cardiology member countries of atrial fibrillation management: baseline results of EURObservational research programme atrial fibrillation (EORP-AF) pilot general registry. Europace 2014 Mar;16(3):308-319. [doi: 10.1093/europace/eut373] [Medline: 24351881]
14. Camm AJ, Corbucci G, Padeletti L. Usefulness of continuous electrocardiographic monitoring for atrial fibrillation. Am J Cardiol 2012 Jul 15;110(2):270-276. [doi: 10.1016/j.amjcard.2012.03.021] [Medline: 22503584]
15. Furberg CD, Psaty BM, Manolio TA, Gardin JM, Smith VE, Rautaharju PM. Prevalence of atrial fibrillation in elderly subjects (the Cardiovascular Health Study). Am J Cardiol 1994 Aug 01;74(3):236-241. [Medline: 8037127]
16. Rho RW, Page RL. Asymptomatic atrial fibrillation. Prog Cardiovasc Dis 2005;48(2):79-87. [doi: 10.1016/j.pcad.2005.06.005] [Medline: 16253649]
17. Steinhubl SR, Waalen J, Edwards AM, Ariniello LM, Mehta RR, Ebner GS, et al. Effect of a home-based wearable continuous ECG monitoring patch on detection of undiagnosed atrial fibrillation: the mSToPS randomized clinical trial. JAMA 2018 Jul 10;320(2):146-155 [FREE Full text] [doi: 10.1001/jama.2018.8102] [Medline: 29998336]
18. Halcox JPJ, Wareham K, Cardew A, Gilmore M, Barry JP, Phillips C, et al. Assessment of remote heart rhythm sampling using the AliveCor heart monitor to screen for atrial fibrillation: the REHEARSE-AF study. Circulation 2017 Nov 07;136(19):1784-1794. [doi: 10.1161/CIRCULATIONAHA.117.030583] [Medline: 28851729]
19. Isakadze N, Martin SS. How useful is the smartwatch ECG? Trends Cardiovasc Med 2020 Oct;30(7):442-448 [FREE Full text] [doi: 10.1016/j.tcm.2019.10.010] [Medline: 31706789]
20. Kalla M, Fabritz L, Kirchhof P. SMART about watches: we need technical and biological validation of atrial fibrillation screening. JACC Clin Electrophysiol 2019 Feb;5(2):209-211 [FREE Full text] [doi: 10.1016/j.jacep.2018.11.018] [Medline: 30784692]
21. Cardiovit FT-1 K170182. US Food and Drug Administration. URL: https://www.accessdata.fda.gov/cdrh_docs/pdf17/K170182.pdf [accessed 2022-05-18]
22. WMA Declaration of Helsinki – ethical principles for medical research involving human subjects. World Medical Association. URL: https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/ [accessed 2022-05-16]
23. Using Apple Watch for arrhythmia detection. Apple. 2020. URL: https://www.apple.com/healthcare/docs/site/Apple_Watch_Arrhythmia_Detection.pdf [accessed 2021-09-20]
24. Bumgarner JM, Lambert CT, Hussein AA, Cantillon DJ, Baranowski B, Wolski K, et al. Smartwatch algorithm for automated detection of atrial fibrillation. J Am Coll Cardiol 2018 May 29;71(21):2381-2388. [doi: 10.1016/j.jacc.2018.03.003] [Medline: 29535065]
25. Tieleman RG, Plantinga Y, Rinkes D, Bartels GL, Posma JL, Cator R, et al. Validation and clinical use of a novel diagnostic device for screening of atrial fibrillation. Europace 2014 Sep;16(9):1291-1295 [FREE Full text] [doi: 10.1093/europace/euu057] [Medline: 24825766]
26. Samsung health monitor application instructions for uses. Samsung. 2020 Nov 20. URL: https://images.samsung.com/is/content/samsung/p5/in/apps/samsung-health-monitor/ifu/pdf/IFU_ECG_IN_20210116.pdf [accessed 2022-05-18]
27. Chen E, Jiang J, Su R, Gao M, Zhu S, Zhou J, et al. A new smart wristband equipped with an artificial intelligence algorithm to detect atrial fibrillation. Heart Rhythm 2020 May;17(5 Pt B):847-853 [FREE Full text] [doi: 10.1016/j.hrthm.2020.01.034] [Medline: 32354449]
28. Fitbit ECG app instructions for use. Fitbit. 2021 Nov 15. URL: https://help.fitbit.com/manuals/manual_ecg_en_US.pdf [accessed 2022-05-18]
29. Fu W, Li R. Diagnostic performance of a wearing dynamic ECG recorder for atrial fibrillation screening: the HUAMI heart study. BMC Cardiovasc Disord 2021 Nov 20;21(1):558 [FREE Full text] [doi: 10.1186/s12872-021-02363-1] [Medline: 34800984]

30. Tymińska A, Ozierański K, Balsam P, Kapłon-Cieślicka A, Maciejewski C, Marchel M, et al. The prevalence and association of major ECG abnormalities with clinical characteristics and the outcomes of real-life heart failure patients - Heart Failure Registries of the European Society of Cardiology. Kardiol Pol 2021;79(9):980-987 [FREE Full text] [doi: 10.33963/KP.a2021.0053] [Medline: 34227675]
31. Khan NK, Goode KM, Cleland JGF, Rigby AS, Freemantle N, Eastaugh J, EuroHeart Failure Survey Investigators. Prevalence of ECG abnormalities in an international survey of patients with suspected or confirmed heart failure at death or discharge. Eur J Heart Fail 2007 May;9(5):491-501 [FREE Full text] [doi: 10.1016/j.ejheart.2006.11.003] [Medline: 17218150]
32. ANSI/AAMI/IEC 60601-2-27:2011 Medical electrical equipment - part 2-27: particular requirements for the basic safety and essential performance of electrocardiographic monitoring equipment. International Electrotechnical Commission. URL: https://webstore.iec.ch/publication/2638 [accessed 2022-05-18]

Abbreviations

BMI: body mass index
ECG: electrocardiogram
OR: odds ratio

Edited by T Leung; submitted 14.02.22; peer-reviewed by E Chen, JB Park; comments to author 14.03.22; revised version received 26.04.22; accepted 28.04.22; published 04.11.22.

Please cite as:
Campo D, Elie V, de Gallard T, Bartet P, Morichau-Beauchant T, Genain N, Fayol A, Fouassier D, Pasteur-Rousseau A, Puymirat E, Nahum J
Atrial Fibrillation Detection With an Analog Smartwatch: Prospective Clinical Study and Algorithm Validation
JMIR Form Res 2022;6(11):e37280
URL: https://formative.jmir.org/2022/11/e37280
doi:10.2196/37280
PMID:35481559

Original Paper

HPV Vaccine Communication Competency Scale for Medical Trainees: Interdisciplinary Development Study

Gabrielle Darville-Sanders[1*], PhD, MPH, CHES; Humberto Reinoso[2*], PhD, FNP-BC, ENP-BC; Jann MacInnes[3*], PhD; Emilie Corluyan[1*], MPH; Dominique Munroe[1*], MD; Mary W Mathis[4*], DrPH, MPH; Suzie Lamarca Madden[4*], DrPH, MPH; Johnathan Hamrick[5*], PharmD; Lisa Dickerson[6*], MD; Cheryl Gaddis[1*], DrPH, MPH, CHES

[1]Department of Public Health, Mercer University, Atlanta, GA, United States

[2]Georgia Baptist College of Nursing, Mercer University, Atlanta, GA, United States

[3]Department of Research and Evaluation Methodology, College of Education, The University of Florida, Gainesville, FL, United States

[4]Department of Public Health, Mercer University, Macon, GA, United States

[5]College of Pharmacy, Mercer University, Atlanta, GA, United States

[6]Physician Assistant Studies, Mercer University, Atlanta, GA, United States

[*]all authors contributed equally

Corresponding Author:
Gabrielle Darville-Sanders, PhD, MPH, CHES
Department of Public Health
Mercer University
3001 Mercer University Drive
Atlanta, GA, 30341
United States
Phone: 1 678 547 6492
Email: darville-sanders_gc@mercer.edu

Abstract

Background: Human papillomavirus (HPV) infection is the most common sexually transmitted infection in the United States. High-risk HPV strains are associated with cancer of the cervix, oropharynx, anus, rectum, penis, vagina, and vulva. To combat increasing HPV-related cancers, the 9-valent HPV vaccine Gardasil was developed. Recommendation of the HPV vaccine by a health care provider has been cited as the number one factor affecting vaccine uptake among adolescents and young adults. Physician assistants, nurse practitioners, and pharmacists have been enlisted to bridge the gap.

Objective: The specific aim of this research study was to develop a reliable and valid HPV vaccine communication scale that can be used to measure the competency of primary care providers when recommending the need for vaccination to parents and patients.

Methods: Using a descriptive study, we collected data via a literature review, focus groups, and an expert panel to inform the scale domains and blueprint design. Pretesting (cognitive interviews) was used to inform item revision decisions. An item analysis was also conducted for the responses provided in the cognitive interviews. Item statistics (means and SDs), interitem correlations, and reliability were examined. Data were analyzed using SPSS (IBM Corp) software.

Results: A valid and reliable 42-item HPV vaccine communication competency scale was developed. The scale included 6 domains of interest. Scale items were moderately to strongly correlated with one another, and Cronbach α indicated good internal consistency with each scale. Scale items included were related to provider introduction or rapport (α=.796), patient respect or empathy (α=.737), provider interview or intake (α=.9), patient counseling or education (α=.935), provider communication closure (α=.896), and provider knowledge (α=.824).

Conclusions: Pharmacists, nurse practitioners, and physician assistants should be trained to be competent in HPV vaccine communication and recommendation due to their expanded roles. Interdisciplinary collaboration is important to account for the trainee's individual differences and ensure the best health care outcomes for patients. A standardized HPV communication scale can be used to ensure effective and consistent recommendation by health care providers, thus affecting immunization rates.

(JMIR Form Res 2022;6(11):e38164) doi:10.2196/38164

KEYWORDS

human papillomavirus; HPV; HPV vaccine; provider communication; medical trainees; immunization; vaccine; communication; student; sexually transmitted infection; STI; United States of America; USA; young adult; teen; patient; parent; immunization; mobile phone

Introduction

Background

An increased demand for nurse practitioners and physician assistants as primary care providers has been observed in the last 10 years in the United States. This was primarily initiated because of population growth, the aging population living longer, and health insurance expansion [1]. However, primary care providers are also pivotal to immunization administration and uptake. Apart from the traditional clinical setting, we have seen an increase in the number of nurse practitioners and physician assistants working in settings such as urgent care clinics, convenient care clinics, and retail clinic service provider sites such as The Minute Clinic [1,2]. This is a great advancement in trying to bridge the gap in the health care (specifically immunizations) industry particularly as it relates to accessibility, cost, and convenience. However, the scope of practice (full, reduced, or restricted practice) is determined by the state legislative and regulatory barriers [1]. In the state of Georgia, a nurse practitioner's and physician assistant's practice and prescription authority are overseen by the supervising physician and possible written protocol enacted by that physician [3].

Apart from nurse practitioners and physician assistants, pharmacists have also been identified as critical to immunization uptake among adolescents [4]. As the COVID-19 pandemic has affected medical visits and school-based interactions because of social distancing, stay at home orders, and governmental bans, we have observed that the number of vaccinations, including the vaccination for human papillomavirus (HPV), declined dramatically over a short period especially among children and adolescents [5]. Coupled with the already lagging vaccination rates in some states before the COVID-19 pandemic, this is cause for concern [6]. It is estimated that since March 2020, ordering and billing for the HPV vaccine have dropped by almost 20% with administration rates remaining down between 20% and 40% in June 2020 [7]. Before the pandemic, Georgia had stricter laws prohibiting a pharmacist's ability to administer immunizations. However, pharmacists are now able to provide HPV vaccination in addition to other routine vaccines [8]. Therefore, pharmacists, nurse practitioners, and physician assistants must be competent concerning HPV vaccine recommendations and patient communications to help bridge the gap.

As curriculum competencies for HPV vaccination information and communication strategies vary within each program, including other programs such as family medicine, there is a need for a standardized communication scale [9]. Although there is a Medical Communication Competency Scale [10,11], this scale is primarily focused on general medical communication during medical interviews and does not tailor scale items to measure the effectiveness of communication language and skill set necessary for HPV vaccination uptake. Therefore, an HPV vaccination communication scale must be developed. This is particularly important when interacting with specific communities, community groups, and disparate population groups as they are most affected by HPV-related conditions and diseases. Research literature indicated that health care providers and parents were more accepting of females being vaccinated for HPV than males, which limited researchers' understanding of HPV-associated diseases among men [12]. According to McGhee et al [13], there is limited understanding of HPV-associated diseases among men, which may be related to low acceptance of HPV vaccinations in males. In addition, HPV incidence disproportionately affects minority racial and ethnic groups. Black women are less likely to complete the series of vaccines compared with White women [14]. Black and Hispanic individuals are more likely to be affected by HPV-related morbidity and mortality rates compared with White individuals [15]. This is a direct result of missed clinical opportunities and a lack of proper recommendations during patient visits with health care providers, which is possibly because of provider discomfort [13,16]. Due to this, the Centers for Disease Control and Prevention cancer panel encourages efforts to improve comprehensive communication strategies for primary care providers and other health care professionals [16].

Clerkship and Residential Training

Although the enrollment in medical, physician assistant, nurse practitioner, and pharmacy programs have increased, we are witnessing shortages of clinical education through clerkship training in the United States. Clerkships (clinical rotations) are immersive learning opportunities where students are provided real-world opportunities to apply their understanding of clinical and scientific concepts to patient care during their rotations. They are supervised and observed by senior faculty and take a primary role in obtaining information and developing final treatment plans [17].

Shortages are even further compounded by a decrease of preceptors, hospital mergers, or health system closures [18]. It is expected that during the clinical phase of education, a medical trainee puts into practice what they have learned. Furthermore, it requires clinical preceptors to supervise students in their performance of routine tasks for optional and fundamental learning [19]. The structure of medical and clinical education relies on the premise that supervisors possess the competencies needed to guide and assist in the training of those under their direction. It also relies on the premise that medical trainees are competent to operate as a health care provider or primary care provider. Using the newly proposed HPV communication competency scale in clinical education training before real-world experiences, medical trainees would have experiential learning opportunities that would improve long-term health for their future patients. For primary care providers already practicing in health care, the use of the HPV communication competency scale with the addition of simulation-based training (modules

providing practice examples of effective patient-provider communication) would provide safe spaces to improve upon skills needed to promote HPV vaccination. It could also highlight the need for enrollment in up-to-date continuing education training and curriculum programs, as well as effective system-wide assessments. This would ensure that from a health care system perspective, providers can receive the level of clinical education needed to meet performance, communication, and recommendation standards.

Communication Training in Medical Programs

Finally, although the demand to fill the immunization health care gaps is increasing, there is no way currently to assess the level and extent of communication training and preparation that currently exists in programs throughout the United States surrounding HPV vaccine uptake. However, there is an opportunity to develop a communication scale measure with common standards to be used as an assessment tool among medical trainees surrounding HPV vaccination. This assessment can be implemented at key points in the curriculum of students in nurse practitioner, physician assistant, and pharmacy programs to ensure that graduating students are fully competent to provide effective recommendations and are comfortable with the information that should be conveyed, thereby resulting in parents and patients who can make informed decisions related to HPV vaccine uptake.

Literature Review

Physician and medical practitioner recommendations are a key predictor of HPV vaccination. However, the literature continually cites missed opportunities by providers leading to low vaccination rates. According to the Henry J Kaiser Foundation, >1 in 10 parents of adolescent girls and 1 in 5 parents of adolescent boys said the vaccine was not recommended to them by health care providers [20]. In another research study, 23 focus group sessions (n=112) were conducted with women (aged 18-26 years), parents, community leaders, and health care providers in Ohio Appalachia. During these sessions, it was found that health care providers "were the only type of group that did not mention the importance of explaining how vaccines work or the pros and cons of HPV vaccinations a part of educational programs" [21]. In a national survey of 1400 respondents conducted by the National Cancer Institute, it was found that only 1 in 4 youths talks to a health care provider about the HPV vaccine. In the same study, it was reported that when asked about vaccine efficacy, 70% of the providers did not know how effective the vaccine was [22]. This is primarily because of (1) their lack of knowledge about HPV and the manifestations of the disease, (2) being more likely to discuss HPV with patients if they had a positive HPV diagnosis, and (3) their reluctance to talk to their young patients about sex [23-25]. In a qualitative study on physician HPV vaccination practices, it was found that physicians (1) did not feel it was their role to provide HPV vaccination and that auxiliary health care service agencies such as the health department should be charged with that responsibility, (2) stated the need for more information about the safety and efficacy of the vaccine before they could recommend and administer it, and (3) were critical of the policy recommendations for the HPV vaccines [24].

When we look at key knowledge of the HPV vaccine, only 35% of health care providers who participated in large-scale US surveys were able to correctly identify that most genital HPV infections resolve without any treatment. In addition, only 47% of the health care providers knew that the HPV strains associated with genital warts differ from strains usually associated with cervical cancer and only 63% of the health care providers believed that HPV infection increases the risk of anogenital cancer in men [26,27]. This lack of information can partly be because of the ever-changing narrative surrounding the HPV vaccine. As the vaccine was first recommended in early 2009 for use, we have seen several changes and adjustments in the last 10 years including several iterations of the vaccine (Cervarix, Gardasil 4, and Gardasil 9 at present), the shift from only female focus to include males, catch up vaccination groups, the emphasis for vaccination among the age group of 9 to 12 years for best immune response, and now new recommendations for 2 versus 3 dose series dependent on the age of initiation [28]. Due to this, many health care providers including primary care providers have been cited in the research as having limited understanding, evident knowledge gaps, uncomfortableness, and even confusion on how to recommend the vaccine to their young patients and parents [27,29,30]. Therefore, it has been recommended that patient-provider communication incorporates a strong endorsement with an emphasis on cancer prevention and same-day vaccination to improve the lagging rates [31]. Therefore, the proposed project will not only add to the research literature and thus increase the body of knowledge surrounding HPV patient-provider communication but also create a practical tool that can be used to improve health care and public health outcomes overall. Therefore, the specific aim of this research study was to develop a reliable and valid HPV vaccine communication scale that can be used to measure the competency of primary care providers when recommending the need for vaccination to parents and patients. Specific research questions included were as follows:

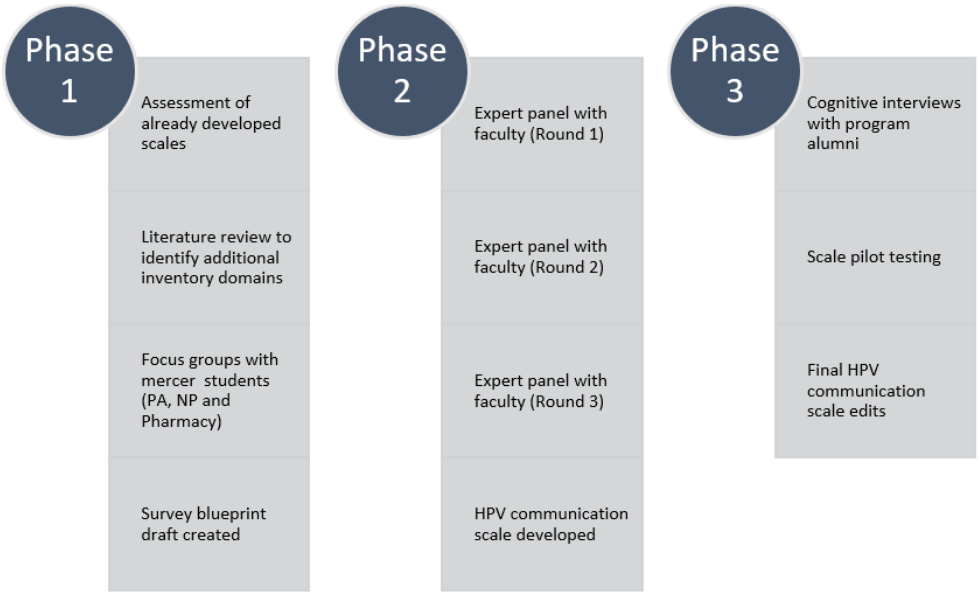
1. What factors related to provider communication comfort should be considered to increase the effectiveness of HPV vaccine recommendations when given?
2. What HPV vaccine communication competencies are relevant to the nurse practitioner, physician assistant, and pharmacy professions?

Methods

Overview

Scale domains were developed using a variety of different approaches identified below (Figure 1). Previously developed scales and HPV literature were reviewed to identify scale initial domains. Focus groups were facilitated to identify perspectives related to HPV vaccine communication that will be incorporated in the development of scale items. An expert panel was conducted to provide insight on scale content, word choice, and appropriateness, and finally, cognitive interviews were conducted with practicing professionals to assess the thought process involved with responding to the scale items and pretest scale items [32-35].

Figure 1. Methodology schema. HPV: human papillomavirus; NP: nurse practitioner; PA: physician assistant.



To mitigate the risk of contracting COVID-19, the project methodology was designed to collect data for all phases using web-based platforms or technology. This ensured that the risk of COVID-19 exposure to the investigators, staff, research participants, and extended networks was minimal to none.

Phase 1—Conceptualization and Domain Specification

We conducted a literature review to identify existing validated measures and concepts related to health communication, HPV, and HPV recommendation to begin our research. Although some scales relevant to medical communication and communication skill assessment exist, none of them were specific for HPV and HPV vaccination. Therefore, while the literature review data collection process was being conducted, focus group sessions were also held to capture data from current medical trainees in Mercer University programs. The focus group sessions were used as a complementary data collection inquiry to fill in the gaps relevant to additional scale domains. As focus groups are a frequently used qualitative approach to gain an in-depth understanding of issues, it was determined appropriate for this study [36].

Focus group participants were recruited using convenience sampling. Clinical and medical directors of all programs were asked to distribute a recruitment email to students for participation in the study. Using SuperSaaS web-based scheduling software, students were able to self-enroll in the focus group session that worked best for their schedule [37]. Following confirmation of enrollment, a web-based informed consent form was emailed to the participant encouraging completion before participation by the lead principal investigator. Qualtrics (Qualtrics XM) was used to collect informed consent and basic participant information. All focus group sessions were facilitated over the Zoom videoconferencing platform.

In February 2021, the team facilitated its first focus group session with students in the nurse practitioner or nursing, physician assistant, and pharmacy programs. On the basis of the feedback provided, the focus group protocol was updated, expanded upon, and resubmitted to the institutional review board for approval. This was done because some questions needed to be reworded or updated for additional clarity. In March and April 2021, additional focus group sessions were facilitated. The average size for the focus group sessions was 8 persons. Sample size per group was 9 for the first group, 8 for the second group, 7 for the third group, and 9 for the fourth group. Key topics explored during the focus group sessions included curriculum exposure to HPV vaccine or virus information, current HPV vaccination recommendations for adolescents, recommendation competency and comfort, additional training in HPV, health provider communication, communication tools to improve HPV vaccine recommendations, and vaccination strategy. The entire focus group protocol questionnaire can be viewed in Multimedia Appendix 1.

All focus group sessions were recorded and transcribed for further analysis by our graduate research assistant. Our graduate research assistant also served as a notetaker, so no information was missing. All the qualitative data collected from the focus group sessions were coded using NVivo (QSR International) software. A codebook was established deductively using the interview script as a baseline. Additional codes were added, defined, and expanded upon as the coding process continued using the constant comparative method based in grounded theory. To establish interrater reliability of emergent themes, the researchers developed a schedule for coding. Each week, the researchers would code 10% of the transcripts independently and then merge coding findings into NVivo to run a coding analysis. The research team would then meet to discuss the codes identified until 90% to 100% agreement was achieved for the data set. Differences were also discussed throughout the

data analysis until a consensus on the coding scheme was established.

Following the completion of the literature review and focus group sessions, a blueprint of a scale was drafted. Included components were derived from 10 articles or scales of interest [10,38-46]. Feedback and data collected from the focus group sessions were also integrated into the initial blueprint of the scale. The blueprint of the scale was then provided to the research team for individual feedback. Once all individual feedback was combined into a master document, the principal investigator and coinvestigators collectively assessed specific domains and agreed upon the final draft of the questionnaire to be used in the next phase of the research study. In total, 25 domains were derived from existing instruments or concepts from the research literature and 20 domains were newly developed or adapted from discussions from focus group participants given a total number of 45 items.

Phase 2—Expert Review and Item Development

We established a panel of expert faculty at Mercer University trained in the fields of nurse practitioner, physician assistant, and pharmaceutical studies in July 2021. The principal investigators identified 10 potential members based on their experience with training medical trainees, communication expertise, and immunization practice. Before their enrollment in the Delphi panel, the scale was entered into Qualtrics. Items were cross rated within Qualtrics by each expert panelist relevant to the domains of interest: *provider introduction or rapport, patient respect or empathy, provider interview or intake, patient counseling or education, provider communication closure, and provider knowledge*. Each expert panelist was provided 1 week to review the scale items in Qualtrics and provide feedback on each section. The assessments of the domains were done individually and free from the influence of other members. The feedback was then compiled by the research team, discussed, agreed upon, and then implemented for subsequent rounds. In total, 3 rounds of feedback were conducted, and items were evaluated within Qualtrics by the expert panel for relevance and appropriateness (1 round) and clarity (2 rounds).

A fourth round was implemented to collect preliminary data concerning the scale's reliability and validity. To facilitate this, a digital stimulus (1 minute and 40 seconds video) was provided to the expert panel, and they were asked to evaluate the provider's HPV vaccine recommendation as viewed in the video. An item analysis was conducted for the purpose of determining the quality of the items. Item analysis included item statistics (means and SDs), interitem correlations, and reliability. SPSS software was used to conduct data analysis. As this version of the scale included not applicable as a response item, response option differences emerged between the disciplines regarding their evaluation of the providers competence. This led to some scale items not being well correlated to one another. However, for all domains of interest, Cronbach α values were reported as between .7 and .9. Following the completion of the Delphi panel, the research team met again to review the scale items with particular interest in the items not well correlated. Upon the review of feedback provided by expert panelists, the data itself and recommendations from our research and evaluation methodologist items were (1) updated, (2) reorganized, or (3) removed from the scale. After the Delphi Panel, a total of 43 scale items remained.

Phase 3—Cognitive Testing

A completed list of items was then finalized to be examined during the cognitive interview phase of our research study. An additional graduate research assistant was hired who assisted with the enrollment of study participants using SuperSaaS software [37]. Convenience sampling was conducted with alumni of Mercer University's nurse practitioner, physician assistant, and pharmaceutical studies programs. However, due to competing COVID-19 pandemic response demands on practitioner enrollment and study attrition, the recruitment strategy was updated to include all available practicing practitioners that were interested in the study. Strategies implemented to recruit the latter group of practitioners included emailing national and state professional organizations relevant to the disciplines. Cognitive interviews were facilitated to garner additional feedback regarding the clarity and appropriateness of the scale items. Once scheduled, the first round of interviews was conducted by both graduate research assistants using Zoom web-based videoconferencing software. Transcripts were auto-produced and then cleaned before being analyzed using NVivo software. Notes taken during the cognitive interviews were also compiled and assessed by the research team. Final edits to the scale were made before round 2 of the cognitive interviews.

Similar to the fourth round of the Delphi panel, a second round of the cognitive interview was implemented to collect preliminary reliability measures and conduct a correlation analysis between scale items. To facilitate this, a digital stimulus (1-minute-and-40-second video) was provided to the round one cognitive interview participants, and they were asked to evaluate the provider's HPV vaccine recommendation as viewed in the video. To address issues with correlation identified previously, the not applicable (N/A) response item was completely removed from the scale, and a question asking respondents to identify their discipline was added. An item analysis was also conducted for the responses provided in the cognitive interviews. Item statistics (means and SDs), interitem correlations, and reliability were examined. Data was analyzed using SPSS software. Overall, scale items and subscales were well correlated with one another. However, there was 1 section where statements were too highly correlated with one another (redundant in nature). After further analysis and investigation, 1 statement was removed.

Ethics Approval

This study was approved by the Institutional Review Board for Human Subjects Research at Mercer University in accordance with the 2018 Federal Regulations 21 CFR 56.110(b) and 45 CFR 46.110(b) (for expedited review) and was approved under categories _6, _7 per 63 FR 60364 in November 2019 (approval number: H1911287).

Results

Focus Group Sessions

A total of 4 focus group sessions were conducted. This phase sought to seek insight from a targeted 40 students; however, only 33 of them participated. Demographic breakdown of the participants was as follows: nurse practitioner (20/33, 61%), physician assistant (9/33, 27%), and pharmacy (4/33, 12%). Emergent themes that arose from the focus groups were focused on trainees' comfort with effectively communicating HPV and HPV vaccine information to all patients (Textbox 1).

Textbox 1. Emergent themes and subthemes related to medical trainee human papillomavirus (HPV) vaccine experiences and competence (focus groups).

Themes and subthemes

- Additional insight: participants shared statements related to additional insight or ideas relating to scale development
- Additional training opportunities: statements related to ideas on ways to prepare medical trainees to be comfortable and competent in their ability to recommend the HPV vaccine
- Comfort: participants' self-rated comfort level as a health care provider in recommending the HPV vaccine to current and future patients on a Likert-like scale from 1 to 10
- Comfort scale development: statements related to the development of a scale to measure the comfort level of health care providers who recommend or communicate the need for HPV vaccination
- Competence: participants' self-rated competency as a health care provider in recommending the HPV vaccine to current and future patients on a Likert-like scale from 1 to 10
- Competence scale development: statements related to the development of a scale to measure the competence of health care providers who recommend or communicate the need for HPV vaccination
- Current HPV knowledge: the demonstration of HPV knowledge including general information about virus and vaccine. As well as including information about current vaccine recommendations
- HPV: participant knowledge of HPV
- HPV vaccine: participant's knowledge on HPV vaccine
- Human papillomavirus: participant's knowledge on the human papillomavirus
- Health care provider HPV communication: statements related to what the participant's believe health care providers should discuss and communicate about HPV vaccination
- Additional communication tools: statements related to additional tools or methods health care providers can use to communicate with patients regarding the HPV vaccine
- Knowledge exposure: statements relating to participant's exposure to information on HPV as well the vaccine
- Class or opportunity: setting in which participants were exposed to HPV information or experience
- Depth: level of exposure to HPV information
- Profession: chosen field of study or work, including reasons why the profession was chosen
- Vaccination completion: statements about methods a health care provider can use to ensure vaccination initiation and completion

For example, when discussing their comfort level, 1 participant shared that bringing up HPV and the vaccine with parents may feel taboo because "nobody wants to think that their 12-year-old is having sex." Participants indicated that their lack of competency is because of a deficiency in knowledge regarding certain aspects of HPV and HPV vaccination as seen in the following quotes:

> *Like I think I would struggle with the most of that, and then I probably would need to do a little bit more research on protocols and guidelines, as you know, like when the HPV vaccine like should really be given and the doses, and things like that.*

> *I'll be able to educate them about the purpose of the actual vaccine, the signs and symptoms. To look for what [is] HPV, how [it] is transmitted, prevention methods, but in terms of like the frequency of the vaccination I'm not quite sure and, like the age cutoff, I'm not too familiar with what I do know what the adverse outcome is if you have HPV.*

The participants also discussed the differences in competency between medical professionals concerning HPV and the HPV vaccine. Nursing and physician assistant students who saw and interacted with adolescents in primary care settings felt more competent in their ability to give a strong high-quality HPV vaccine recommendation that compared with their peers who do not work in that setting or disciplines such as pharmacy. This is exemplified in the following quotes:

> *I think the difference in like pharmacy and nursing is [that] I would be able to talk more about the disease than I probably could [on the] vaccine itself, whereas pharmacy would probably know all about the vaccine.*

> *I could definitely do like the risk and the benefits of it, but as far as like the details of the guidelines, I*

would definitely need to go and like update myself and like just educate myself a little bit more before I went in and tried to give somebody like the rundown on what that is.

I also haven't had any experience giving vaccines, but I have had a lot of experience talking with patients and deciding treatment plans and everything like that.

I've not been exposed to an opportunity to give it. I don't deal with children, you know. And I'm dealing with primarily pregnant women, anyway, so we're well beyond all that, so I just don't have any exposure to like giving lots of vaccines anyways.

The quotes further emphasize the importance of developing a scale to assess medical trainees on their ability to recommend the HPV vaccine. The focus group sessions also provided vital information on the differences in HPV and HPV vaccine education and training among medical professionals. This scale could provide the opportunity for medical trainees to be assessed and increase their comfort and competence, allowing them to appropriately bridge the gap between science and fiction as they move forward in their careers as health professionals.

Cognitive Interviews

In total, research sought to complete 33 sessions but only 15 participants completed both rounds of the cognitive interviews. Emergent themes and subthemes from this phase of our data collection are included in Textbox 2.

Textbox 2. Emergent themes and subthemes related to human papillomavirus (HPV) Vaccination Communication Scale clarity and relevance (cognitive interview sessions).

Themes and description
• Scale: a graduated range of values forming a standard system for measuring or grading something. Including statements relating to scale measurements, that is, novice-expert
• Generalizability: statements related to the applicability of the HPV Recommendation Competency Scale across multiple disciplines
• Scale item: statements about individual scale items including statements related to the necessity, appropriateness, and flow of the items
• Assessment potential: individual describes either ease or difficulty in using scale item to assess student providers, including feasibility
• Clarity: statements that relate to the participant's understanding of the scale item
• Confusion: participants indicate confusion with scale item or wording
• Rephrase: participants suggest rephrasing or changing words in items
• Positive comments: positive statements about scale items including wording, flow, and appropriateness. Can include statements of agreement regarding scale items.
• Redundancy: statements indicating either items or sections as being repetitive
• Scale section: statements related to any of the 6 scale sections, including overall flow and appropriateness

Upon the completion of this phase of the study, we were able to run reliability and validity tests. Cronbach α values for each section were as follows:

- The provider introduction or rapport subscale included 7 items (α=.796).
- The patient respect or empathy subscale included 4 items (α=.737).
- The provider interview or intake subscale included 6 items (α=.9).
- The patient counseling or education subscale included 10 items (α=.935).
- The provider communication closure subscale included 7 items (α=.896).
- The provider knowledge subscale included 8 items (α=.824).

Discussion

Principal Findings

HPV, a sexually transmitted infection, is a highly prevalent disease [47,48]. Approximately 79 million Americans are currently infected with HPV, with 14 million becoming newly infected annually [47]. To help prevent the increasing rate of cancers associated with HPV, the HPV vaccine (Gardasil 9) was developed to protect against persistent infections of associated HPV subtypes [49,50]. Despite having access to licensed HPV vaccines for more than a decade, HPV vaccination rates are still lagging among many adolescents and the coverage varies in the United States. According to Sonawane et al [51], nearly 46% of adolescents who are eligible are not up to date with the recommended HPV vaccination as of 2019. According to the National Immunization Survey 2008 to 2018 teen data, although vaccination initiation rates have increased over the last years, completion and up-to-date rates still fall below 50% in most US states [52].

Although there are many barriers to HPV vaccination, parental vaccine hesitancy has been one of the most constant factors affecting vaccination rates among the adolescents [51]. To combat this, it has been recommended that providers have strong high-quality recommendations. Current research has shown that adolescents who report receiving a strong high-quality recommendation from their provider are >8 times more likely to initiate and completed HPV vaccination than those who do not [53]. However, many providers and trainees do not feel as though they possess the competence to provide a strong high-quality recommendation. This is not only evident in the

previous research but also in the results obtained in this study. To increase communication comfort, study participants indicated that they needed to receive more information or knowledge about the HPV and vaccine guidelines especially if they were not working in a clinical setting where they would interact with them on a consistent basis. Nurse practitioner and physician assistant trainees communicated that they felt more confident in their ability, when compared with pharmacy trainees who are a new addition to the implementation of HPV immunization programs. It was evident that professional protocols concerning patient interactions, communication standards, and even training differed among the disciplines. This is further compounded by state variation concerning the scope of practice to include vaccine age—eligible adolescents (independent authority, collaborative practice agreement, or by prescription only) [54]. However, with the changes to pharmacy immunization protocols owing to the COVID-19 pandemic challenges, there is increasing support by primary care physicians and parents alike for trained pharmacist HPV vaccination administration as an independent authority [54,55].

Our study also sought to identify what specific communication competencies were needed relevant to the 3 disciplines of focus (nurse practitioner, physician assistant, and pharmacy). Although there were key differences that emerged, the 3 phases involved in this research study concluded that the communication process was an interactive and yet integral role of each clinician. Upon information collected from the literature, and feedback gathered by our focus groups and Delphi panel of experts and practicing clinicians, and validation through our cognitive interviews, it was identified that (1) provider introduction or rapport, (2) patient respect or empathy, (3) provider interview or intake, (4) patient counseling or education, (5) provider communication closure, and (6) provider knowledge were important components during a patient-provider interaction. The subscales identified and confirmed throughout our study match the ones identified in similar research studies that focused on enhanced HPV communication tools for vaccine hesitant parents. As our scale statements assess a medical trainee's competency related to presumptive communication and motivational interviewing techniques it can be more effective in alleviating parental vaccine hesitancy [56]. As a strategy to also address the diverse population subgroups served, the scale questions included provide the opportunity for tailoring recommendations and vaccine communication based on race and ethnicity, gender, sexual identity, and biology differences, to name a few. While this may not close the gap completely, as the scale was developed through a health equity lens, it can serve as a tool that moves the needle forward concerning improvements in HPV vaccination disparities.

Despite the successes of this study, there were some limitations. First, the curriculum development and training of nurse practitioners, physician assistants, and pharmacy trainees varies heavily, and as such, exposure to HPV and HPV vaccination information is not consistent and equally weighed. Therefore, although these 3 practitioners have been integral in addressing the lag in medical care provisions, specifically HPV vaccination, the differences among professions may also influence how and in which ways the scale can practically be used. Second, it requires that supplemental training be developed and tailored to each individual clinical program's curriculum to ensure standardization and effectiveness of the scale tool when used in the field. Third, although health care providers and professionals are integral to the success of vaccination initiatives and programs, they are not the only stakeholder involved. Other important adolescent reference groups for decision-making may also involve parents, teachers, or peers. However, due to the scope of this study, and the intentionality of addressing the pressing needs for high-quality provider communication as cited heavily throughout existing research, they were not included as participants in this study.

Future Implications

Although there has been much effort toward making HPV recommendation communication standardized through the publications of clinician action guides (physician, physician assistants, nurse practitioners, nurses, medical assistants, and dentist or dental care providers) by the National HPV Vaccination Roundtable, more efforts need to be made concerning the development of standardized education and training modules for each profession involved in HPV vaccination [57]. Since research has shown that strong provider recommendation is highly associated with increased vaccine initiation, completion, follow-through, and decreased parental hesitancy, it is important that training is broad enough yet specific to ensure that no matter the profession, all participating practitioners possess the communication skill sets needed to provide the strong and high-quality recommendation needed [58]. This is even more important now, as we see key disparities among the HPV-related cancers in the United States. Additional work should also explore the perspectives of key stakeholders such as parents, peers, and teachers on HPV vaccine decision-making for initiation and completion, with specific understanding toward health equity and application toward multiple priority populations.

Conclusions

Immunization against preventable cancers associated with the HPV has been a top priority for more than a decade. With the introduction of Cervarix, Gardasil 4, and now Gardasil 9 vaccines, greater importance has been placed on vaccinating adolescent girls and boys starting at the age of 9 years for achieving the best immune response and vaccine efficacy. However, there have been many barriers for HPV vaccination, with the most influential barriers prohibiting initiation and completion being parental resistance and refusal [59]. To combat this, the Centers for Disease Control and Prevention, Advisory Committee on Immunization Practices, President's Cancer Council, and American Academy of Pediatrics have all emphasized the need for strong high-quality recommendations by health care providers [60]. However, since there is no standardized training program on HPV and HPV vaccine education, and major differences exist among the nursing, physician assistant, and pharmacy professions, the development of a reliable and valid HPV vaccine communication scale is important for increasing vaccination rates. Since this study collected data using an interdisciplinary approach, we believe that the communication scale developed accounts for

communication needs and training differences among health care professions. Using the scale alongside a training module such as the *HPV Vaccine: Same Way Same Day* smartphone app will ensure that medical trainees and current practicing professionals can maintain the competence needed to provide high-quality HPV vaccine recommendations regardless of the setting, improving immunization rates and reducing future HPV-related cancer trends [61-63].

Authors' Contributions

All authors wrote, reviewed, and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Focus group facilitator protocol.
[PDF File (Adobe PDF File), 110 KB - formative_v6i11e38164_app1.pdf]

References

1. Van Vleet A, Paradise J. Tapping Nurse Practitioners to meet the rising demand for primary care. KFF. 2015 Jan 20. URL: https://www.kff.org/medicaid/issue-brief/tapping-nurse-practitioners-to-meet-rising-demand-for-primary-care/ [accessed 2020-10-02]
2. Minute clinic. CVS Health. URL: https://cvshealth.com/about/our-offerings/cvs-minuteclinic [accessed 2020-10-01]
3. Georgia scope of practice policy: state profile. NCSL Scope of Practice Policy. URL: http://scopeofpracticepolicy.org/states/ga/ [accessed 2020-10-10]
4. Shah P, Calo W, Marciniak M, Gilkey M, Brewer N. Support for pharmacist-provided HPV vaccination: national surveys of U.S. Physicians and parents. Cancer Epidemiol Biomarkers Prev 2018 Aug;27(8):970-978 [FREE Full text] [doi: 10.1158/1055-9965.EPI-18-0380] [Medline: 29871883]
5. Jenco M. AAP urges vaccination as rates drop due to COVID-19. American Academy of Pediatrics. 2020 May 8. URL: https://www.aappublications.org/news/2020/05/08/covid19vaccinations050820 [accessed 2020-10-10]
6. HPV vaccination rates remain critically low among younger adolescents in the U.S. 2019. Infectious Diseases Society of America. 2019 Jan 17. URL: https://www.sciencedaily.com/releases/2019/01/190117090506.htm [accessed 2022-03-01]
7. Gilkey M, Bednarczyk R, Gerend M, Kornides ML, Perkins RB, Saslow D, et al. Getting human papillomavirus vaccination back on track: protecting our national investment in human papillomavirus vaccination in the covid-19 era. J Adolesc Health 2020 Nov;67(5):633-634 [FREE Full text] [doi: 10.1016/j.jadohealth.2020.08.013] [Medline: 32933839]
8. Pharmacist administered vaccines. National Alliance of State Pharmacy Associations. 2020 Sep. URL: https://naspa.us/wp-content/uploads/2020/08/IZ-Authority-9_2020.pdf [accessed 2020-10-09]
9. Kasting ML, Scherr CL, Ali KN, Lake P, Malo TL, Johns T, et al. Human papillomavirus vaccination training experience among family medicine residents and faculty. Fam Med 2017 Oct;49(9):714-722 [FREE Full text] [Medline: 29045989]
10. Cegala DJ, Coleman MT, Turner JW. The development and partial assessment of the medical communication competence scale. Health Commun 1998;10(3):261-288. [doi: 10.1207/s15327027hc1003_5] [Medline: 16370986]
11. Blanch-Hartigan D. Medical Communication Competence Scale (MCCS). In: The Sourcebook of Listening Research: Methodology and Measures. Hoboken, New Jersey, United States: Wiley; 2017.
12. Newman PA, Logie CH, Doukas N, Asakura K. HPV vaccine acceptability among men: a systematic review and meta-analysis. Sex Transm Infect 2013 Nov 04;89(7):568-574 [FREE Full text] [doi: 10.1136/sextrans-2012-050980] [Medline: 23828943]
13. McGhee E, Harper H, Ume A, Baker M, Diarra C, Uyanne J, et al. Elimination of cancer health disparities through the acceleration of HPV vaccines and vaccinations: a simplified version of the president's cancer panel report on HPV vaccinations. J Vaccines Vaccin 2017 Jun;8(3):361 [FREE Full text] [doi: 10.4172/2157-7560.1000361] [Medline: 28845336]
14. Okafor C, Hu X, Cook RL. Racial/ethnic disparities in HPV vaccine uptake among a sample of college women. J Racial Ethn Health Disparities 2015 Sep 4;2(3):311-316 [FREE Full text] [doi: 10.1007/s40615-014-0074-7] [Medline: 26863461]
15. Reimer RA, Schommer JA, Houlihan AE, Gerrard M. Ethnic and gender differences in HPV knowledge, awareness, and vaccine acceptability among White and Hispanic men and women. J Community Health 2014 Apr 23;39(2):274-284. [doi: 10.1007/s10900-013-9773-y] [Medline: 24150246]
16. Rimer B, Harper H, Witte O. Accelerating HPV vaccine uptake: urgency for action to prevent cancer. National Cancer Institute. 2013. URL: https://deainfo.nci.nih.gov/advisory/pcp/annualReports/HPV/LetterToPresidentObama.htm [accessed 2019-10-15]
17. Cymet T. What is a Clinical Clerkship? American College of Osteopathic Family Physicians. 2018 Jan 15. URL: https://www.acofp.org/acofpimis/acofporg/PDFs/News_Publications/Articles/Clinical_Clerkship_1.15.2018.pdf [accessed 2022-08-08]

18. Cox WJ, Desai GJ. The crisis of clinical education for physicians in training. Mo Med 2019;116(5):389-391 [FREE Full text] [Medline: 31645790]
19. Shaterjalali M, Changiz T, Yamani N. Optimal clinical setting, tutors, and learning opportunities in medical education: a content analysis. J Educ Health Promot 2018;7(1):130 [FREE Full text] [doi: 10.4103/jehp.jehp_56_18] [Medline: 30505858]
20. The HPV Vaccine: access and use in the US. Kaiser Family Foundation. URL: http://kff.org/womens-health-policy/fact-sheet/the-hpv-vaccine-access-and-use-in/ [accessed 2019-10-25]
21. Katz ML, Reiter PL, Heaner S, Ruffin MT, Post DM, Paskett ED. Acceptance of the HPV vaccine among women, parents, community leaders, and healthcare providers in Ohio Appalachia. Vaccine 2009 Jun 19;27(30):3945-3952 [FREE Full text] [doi: 10.1016/j.vaccine.2009.04.040] [Medline: 19389447]
22. Doheny K. Many young Americans know little about cervical cancer vaccine. HealthDay. URL: http://consumer.healthday.com/cancer-information-5/cervical-cancer-news-95/many-young-americans-know-little-about-cervical-cancer-vaccine-682819.html [accessed 2019-10-25]
23. Brewer N. The 'crisis' of low HPV vaccination rates: a conversation with Dr. Noel Brewer. National Cancer Institute. 2016 Feb 2. URL: http://www.cancer.gov/news-events/cancer-currents-blog/2016/hpv-vaccination-rates [accessed 2019-10-24]
24. Berenson AB, Rahman M, Hirth JM, Rupp RE, Sarpong KO. A brief educational intervention increases providers' human papillomavirus vaccine knowledge. Hum Vaccin Immunother 2015 May 06;11(6):1331-1336 [FREE Full text] [doi: 10.1080/21645515.2015.1022691] [Medline: 25945895]
25. Koshiol J, Rutten LF, Moser RP, Hesse N. Knowledge of human papillomavirus: differences by self-reported treatment for genital warts and sociodemographic characteristics. J Health Commun 2009 Jun 28;14(4):331-345 [FREE Full text] [doi: 10.1080/10810730902873067] [Medline: 19466646]
26. Rosenthal S, Weiss T, Zimet G, Ma L, Good M, Vichnin M. Predictors of HPV vaccine uptake among women aged 19–26: importance of a physician's recommendation. Vaccine 2011 Jan 29;29(5):890-895. [doi: 10.1016/j.vaccine.2009.12.063] [Medline: 20056186]
27. Sherris J, Friedman A, Wittet S, Davies P, Steben M, Saraiya M. Chapter 25: education, training, and communication for HPV vaccines. Vaccine 2006 Aug 31;24 Suppl 3:S3/210-S3/218. [doi: 10.1016/j.vaccine.2006.05.124] [Medline: 16950009]
28. Meites E, Kempe A, Markowitz LE. Use of a 2-dose schedule for human papillomavirus vaccination - updated recommendations of the advisory committee on immunization practices. MMWR Morb Mortal Wkly Rep 2016 Dec 16;65(49):1405-1408 [FREE Full text] [doi: 10.15585/mmwr.mm6549a5] [Medline: 27977643]
29. Shay L, Baldwin A, Betts A, Marks EG, Higashi RT, Street RL, et al. Parent-provider communication of HPV vaccine hesitancy. Pediatrics 2018 Jun;141(6):e20172312 [FREE Full text] [doi: 10.1542/peds.2017-2312] [Medline: 29765009]
30. Dempsey AF, O'Leary ST. Human papillomavirus vaccination: narrative review of studies on how providers' vaccine communication affects attitudes and uptake. Acad Pediatr 2018 Mar;18(2S):S23-S27. [doi: 10.1016/j.acap.2017.09.001] [Medline: 29502633]
31. Moss JL, Gilkey MB, Rimer BK, Brewer NT. Disparities in collaborative patient-provider communication about human papillomavirus (HPV) vaccination. Hum Vaccin Immunother 2016 Jun 02;12(6):1476-1483 [FREE Full text] [doi: 10.1080/21645515.2015.1128601] [Medline: 26786888]
32. McMillan SS, King M, Tully MP. How to use the nominal group and Delphi techniques. Int J Clin Pharm 2016 Jun 5;38(3):655-662 [FREE Full text] [doi: 10.1007/s11096-016-0257-x] [Medline: 26846316]
33. Keeney S, Hasson F, McKenna H. Consulting the oracle: ten lessons from using the Delphi technique in nursing research. J Adv Nurs 2006 Jan;53(2):205-212. [doi: 10.1111/j.1365-2648.2006.03716.x] [Medline: 16422719]
34. Tilley B, LaPelle N, Goetz C, Stebbins G, MDS-UPDRS Task Force. Using cognitive pretesting in scale development for Parkinson's disease: the Movement Disorder Society Unified Parkinson's Disease Rating Scale (MDS-UPDRS) example. J Parkinsons Dis 2014;4(3):395-404 [FREE Full text] [doi: 10.3233/JPD-130310] [Medline: 24613868]
35. Lenzner T, Neuert C, Otto W. Cognitive pretesting. GESIS Survey Guidelines. 2016. URL: https://www.gesis.org/en/gesis-survey-guidelines/instruments/quality-of-survey-data/cognitive-pretesting [accessed 2020-11-10]
36. Nyumba TO, Wilson K, Derrick CJ, Mukherjee N. The use of focus group discussion methodology: insights from two decades of application in conservation. Methods Ecol Evol 2018 Jan 11;9(1):20-32 [FREE Full text] [doi: 10.1111/2041-210x.12860]
37. Appointment scheduling software. SuperSaaS. URL: https://www.supersaas.com/ [accessed 2019-10-15]
38. Lang F, McCord R, Harvill L, Anderson DS. Communication assessment using the common ground instrument: psychometric properties. Fam Med 2004 Mar;36(3):189-198 [FREE Full text] [Medline: 14999576]
39. Joyce B, Steenbergh T, Scher E. Use of the kalamazoo essential elements communication checklist (adapted) in an institutional interpersonal and communication skills curriculum. J Grad Med Educ 2010 Jun;2(2):165-169 [FREE Full text] [doi: 10.4300/JGME-D-10-00024.1] [Medline: 21975614]
40. PDF Form Directory. Internal Resources A&M Commerce. URL: https://inside.tamuc.edu/academics/colleges/educationHumanServices/departments/socialWork/form-pdf-directory.aspx [accessed 2021-07-21]
41. Low intensity cognitive behavioural competency scale manual. PEARLS. URL: https://pearlsresearchlab.group.shef.ac.uk/public/1565950652Manual-Assessment.pdf [accessed 2021-06-21]

42. Iramaneerat C, Myford CM, Yudkowsky R, Lowenstein T. Evaluating the effectiveness of rating instruments for a communication skills assessment of medical residents. Adv Health Sci Educ Theory Pract 2009 Oct 5;14(4):575-594. [doi: 10.1007/s10459-008-9142-2] [Medline: 18985427]
43. Shah P, Calo W, Gilkey M, Boynton MH, Alton Dailey S, Todd KG, et al. Questions and concerns about HPV vaccine: a communication experiment. Pediatrics 2019 Feb;143(2):e20181872 [FREE Full text] [doi: 10.1542/peds.2018-1872] [Medline: 30670584]
44. Kornides ML, Fontenot HB, McRee A, Panozzo CA, Gilkey MB. Associations between parents' satisfaction with provider communication and HPV vaccination behaviors. Vaccine 2018 May 03;36(19):2637-2642 [FREE Full text] [doi: 10.1016/j.vaccine.2018.03.060] [Medline: 29627236]
45. Clark NM, Gong M, Schork MA, Maiman LA, Evans D, Hurwitz ME, et al. A scale for assessing health care providers' teaching and communication behavior regarding asthma. Health Educ Behav 1997 Apr 01;24(2):245-256. [doi: 10.1177/109019819702400211] [Medline: 9079582]
46. Recommended curriculum guidelines for family medicine residents: women's health and gynecologic care. American Academy of Family Physicians. 2018. URL: https://www.aafp.org/dam/AAFP/documents/medical_education_residency/program_directors/Reprint282_Women.pdf [accessed 2019-10-10]
47. Genital HPV Infection Fact Sheet. Centers for Disease Control and Prevention. 2017. URL: http://www.cdc.gov/std/HPV/STDFact-HPV.htm#a4 [accessed 2019-10-10]
48. Nan X. Communicating to young adults about HPV vaccination: consideration of message framing, motivation, and gender. Health Commun 2012 Jan;27(1):10-18. [doi: 10.1080/10410236.2011.567447] [Medline: 22276999]
49. HPV vaccine information for young women. Centers for Disease Control and Prevention. 2017. URL: https://www.cdc.gov/std/hpv/stdfact-hpv-vaccine-young-women.htm [accessed 2019-10-15]
50. Rothman SM, Rothman DJ. Marketing HPV vaccine: implications for adolescent health and medical professionalism. JAMA 2009 Aug 19;302(7):781-786. [doi: 10.1001/jama.2009.1179] [Medline: 19690311]
51. Sonawane K, Lin Y, Damgacioglu H, Zhu Y, Fernandez ME, Montealegre JR, et al. Trends in human papillomavirus vaccine safety concerns and adverse event reporting in the United States. JAMA Netw Open 2021 Sep 01;4(9):e2124502 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.24502] [Medline: 34533574]
52. Chido-Amajuoyi O, Talluri R, Wonodi C, Shete S. Trends in HPV vaccination initiation and completion within ages 9-12 years: 2008-2018. Pediatrics 2021 Jun;147(6):2008 [FREE Full text] [doi: 10.1542/peds.2020-012765] [Medline: 33941585]
53. Landis K, Bednarczyk RA, Gaydos LM. Correlates of HPV vaccine initiation and provider recommendation among male adolescents, 2014 NIS-Teen. Vaccine 2018 Jun 07;36(24):3498-3504 [FREE Full text] [doi: 10.1016/j.vaccine.2018.04.075] [Medline: 29752025]
54. Calo WA, Shah PD, Gilkey MB, Vanderpool RC, Barden S, Doucette WR, et al. Implementing pharmacy-located HPV vaccination: findings from pilot projects in five U.S. states. Hum Vaccin Immunother 2019 May 07;15(7-8):1831-1838 [FREE Full text] [doi: 10.1080/21645515.2019.1602433] [Medline: 30945968]
55. Third amendment to declaration under the public readiness and emergency preparedness act for medical countermeasures against COVID–19. Department of Health and Human Services. 2020 Mar 17. URL: https://www.hhs.gov/sites/default/files/third-amendment-declaration.pdf [accessed 2022-02-22]
56. Lockhart S, Dempsey AF, Pyrzanowski J, O'Leary ST, Barnard JG. Provider and parent perspectives on enhanced communication tools for human papillomavirus vaccine-hesitant parents. Acad Pediatr 2018 Sep;18(7):776-782. [doi: 10.1016/j.acap.2018.05.012] [Medline: 30177028]
57. Clinician and health systems action guides. The National HPV Vaccination Roundtable. URL: https://hpvroundtable.org/action-guides/ [accessed 2022-02-22]
58. Oh NL, Biddell CB, Rhodes BE, Brewer NT. Provider communication and HPV vaccine uptake: a meta-analysis and systematic review. Prev Med 2021 Jul;148:106554. [doi: 10.1016/j.ypmed.2021.106554] [Medline: 33857561]
59. Gilkey MB, Calo WA, Marciniak MW, Brewer NT. Parents who refuse or delay HPV vaccine: differences in vaccination behavior, beliefs, and clinical communication preferences. Hum Vaccin Immunother 2017 Mar 04;13(3):680-686 [FREE Full text] [doi: 10.1080/21645515.2016.1247134] [Medline: 27763818]
60. Gilkey MB, Calo WA, Moss JL, Shah PD, Marciniak MW, Brewer NT. Provider communication and HPV vaccination: the impact of recommendation quality. Vaccine 2016 Feb 24;34(9):1187-1192 [FREE Full text] [doi: 10.1016/j.vaccine.2016.01.023] [Medline: 26812078]
61. HPV Vaccine: same way same day. American Pediatric Association and American Academy of Pediatrics. 2018. URL: https://apps.apple.com/us/app/hpv-vaccine-same-way-same-day/id1356847181 [accessed 2019-10-23]
62. Real FJ, Rosen BL, Bishop JM, McDonald S, DeBlasio D, Kreps GL, et al. Usability evaluation of the novel smartphone application, HPV vaccine: same way, same day, among pediatric residents. Acad Pediatr 2021 May;21(4):742-749. [doi: 10.1016/j.acap.2020.11.023] [Medline: 33279739]
63. Bishop JM, Real FJ, McDonald SL, Klein M, DeBlasio D, Kahn JA, et al. Evaluation of HPV vaccine: same way, same day: a pilot study. J Health Commun 2021 Dec 02;26(12):839-845. [doi: 10.1080/10810730.2021.2021459] [Medline: 34985403]

Abbreviations

HPV: human papillomavirus

Edited by A Mavragani; submitted 28.03.22; peer-reviewed by M Nitsch, M Katapodi; comments to author 29.07.22; revised version received 09.08.22; accepted 17.08.22; published 04.11.22.

Please cite as:
Darville-Sanders G, Reinoso H, MacInnes J, Corluyan E, Munroe D, Mathis MW, Madden SL, Hamrick J, Dickerson L, Gaddis C
HPV Vaccine Communication Competency Scale for Medical Trainees: Interdisciplinary Development Study
JMIR Form Res 2022;6(11):e38164
URL: https://formative.jmir.org/2022/11/e38164
doi:10.2196/38164
PMID:36331545

Original Paper

A Hybrid Ecological Momentary Compassion–Focused Intervention for Enhancing Resilience in Help-Seeking Young People: Prospective Study of Baseline Characteristics in the EMIcompass Trial

Isabell Paetzold[1], MSc; Anita Schick[1], PhD; Christian Rauschenberg[1], PhD; Dusan Hirjak[2], MD; Tobias Banaschewski[3], MD, PhD; Andreas Meyer-Lindenberg[2], MD, PhD; Sebastian Butz[1], PhD; Chiara Floesser[1], BSc; Leonie Schueltke[1], MSc; Jan Rasmus Boehnke[4], PhD; Benjamin Boecking[5], PhD, DClinPsy; Ulrich Reininghaus[1,6,7], PhD

[1]Department of Public Mental Health, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany

[2]Department of Psychiatry and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany

[3]Department of Child and Adolescent Psychiatry and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany

[4]School of Health Sciences, University of Dundee, Dundee, United Kingdom

[5]Tinnitus Center, Charité Universitätsmedizin Berlin, Berlin, Germany

[6]ESRC Centre for Society and Mental Health and Social Epidemiology Research Group, King's College London, London, United Kingdom

[7]Health Service and Population Research Department, Centre for Epidemiology and Public Health, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom

Corresponding Author:
Ulrich Reininghaus, PhD
Department of Public Mental Health
Central Institute of Mental Health
Medical Faculty Mannheim, Heidelberg University
J 5 1
Mannheim, 68159
Germany
Phone: 49 62117031930
Email: ulrich.reininghaus@zi-mannheim.de

Abstract

Background: Young people are a target population for mental health–related early intervention and prevention. Although evidence for early intervention is promising, availability of and access to youth mental health services remain limited. Therefore, the development of an evidence-based hybrid intervention is urgently needed.

Objective: This study aimed to present a manual for a hybrid intervention, combining an ecological momentary intervention and face-to-face sessions aimed for enhancing resilience in help-seeking young people based on compassion-focused interventions, and explore whether participants' baseline characteristics are associated with putative mechanisms and outcomes of the EMIcompass intervention. Specifically, we aimed to explore initial signals as to whether participants' sociodemographic, clinical, and functional characteristics at baseline are associated with putative mechanisms (ie, change in self-compassion, change in emotion regulation, working alliance, training frequency); and whether participants' sociodemographic, clinical, and functional characteristics, self-compassion, and emotion regulation at baseline are associated with clinical outcomes (ie, psychological distress and general psychopathology at postintervention and 4-week follow-ups) in the experimental condition and obtain first parameter estimates.

Methods: We recruited young people aged 14 to 25 years, with psychological distress, Clinical High At-Risk Mental State, or first episodes of severe mental disorder for an exploratory randomized controlled trial with assessments at baseline and postintervention and 4-week follow-ups. A structured manual was developed and optimized based on a pilot study's manual, a scoping review of existing literature and manuals, exchange with experts, the team's clinical experience of working with compassion-focused interventions, and the principles of ecological momentary interventions. This analysis focuses on the experimental condition receiving the EMIcompass intervention.

Results: A total of 46 young individuals were randomized to the experimental condition. There was evidence for initial signals of effects of age (B=0.11, 95% CI 0.00-0.22), general psychopathology (B=0.08, 95% CI −0.01 to 0.16), and clinical stage (B=1.50, 95% CI 0.06-2.93) on change in momentary self-compassion and change in emotion regulation from baseline to postintervention assessments. There was no evidence for associations of other baseline characteristics (eg, gender, minority status, and level of functioning) and putative mechanisms (eg, overall self-compassion, working alliance, and training frequency). In addition, except for an initial signal for an association of momentary self-compassion at baseline and psychological distress (B=−2.83, 95% CI −5.66 to 0.00), we found no evidence that baseline characteristics related to clinical outcomes.

Conclusions: The findings indicated the reach of participants by the intervention largely independent of sociodemographic, clinical, and functional baseline characteristics. The findings need to be confirmed in a definitive trial.

Trial Registration: German Clinical Trials Register NDRKS00017265; https://www.drks.de/drks_web/navigate.do?navigationId=trial.HTML&TRIAL_ID=DRKS00017265

International Registered Report Identifier (IRRID): RR2-10.2196/27462

(JMIR Form Res 2022;6(11):e39511) doi:10.2196/39511

KEYWORDS

mobile health intervention; mHealth intervention; digital intervention; just-in-time adaptive intervention; JITAI; blended care; public mental health; inclusiveness; transdiagnostic; clinical staging; intervention manual; mobile phone

Introduction

Background

Young people constitute a priority target population for mental health–related prevention and early intervention, as they are particularly affected by mental health problems. Mental disorders primarily emerge in adolescence and young adulthood, and >60% of all lifetime cases have their onset before the age of 25 years [1]. With a worldwide pooled prevalence of 21% of mental disorders in adolescents aged 12 to 18 years [2], mental health problems contribute substantially to the disease burden [3,4]. Addressing the co-occurrence and overlap of subclinical and clinical experiences and symptoms [5-8], especially in the early stages of psychopathology, dimensional classification frameworks [9,10] cutting across traditional diagnostic boundaries, including the Hierarchical Taxonomy of Psychopathology (HiTOP) [11], have been proposed. Clinical staging models take early, overlapping, and nonspecific psychopathological symptoms and transitional staging processes into account [12,13].

There is convincing evidence on risk factors that are modifiable, on mental health problems that can be changed, and on protective factors that can be strengthened to enhance resilience [14-16]. Traditional psychotherapeutic interventions, including standard cognitive behavioral therapy, as well as third-wave approaches, show moderate to high effect sizes in randomized controlled trials (RCTs) and meta-analyses [17-20]. However, there is considerable room for improvement, as—even after successful treatment—many service users show significant residual symptoms or relapse [21]. In addition, the availability of and access to youth mental health services remain limited [22,23]. More downstream, this may result in a longer duration of untreated illness, an important marker of poor prognosis and complex course and outcome [24,25].

Some of these problems of standard care might be caused by difficulties transferring preventive and therapeutic strategies developed in face-to face sessions to service users' daily life. Mobile health (mHealth) may be a promising approach to address these challenges by improving access to mental health care for young people by using mobile devices for the delivery of prevention and intervention [26-30]. With ecological momentary assessment (EMA), often also referred to as experience sampling methodology (ESM) [26,31], a structured diary method, momentary fluctuations in experience and behavior can be assessed in real time and real life. Ecological momentary interventions (EMIs) [26,29,32-34] offer the opportunity to deliver adaptive and personalized intervention components in daily life. The digital approach may help to lower the threshold for young people to access interventions meeting their needs and preferences and facilitates the ecological translation of techniques learned into service users' everyday lives [29]. A recent nationally representative survey indicated that young people do frequently use mHealth apps and are even more likely to do so when feeling distressed [30].

However, digital approaches are also confronted with challenges: most apps currently available in major app stores are not evidence based, and some even include potentially harmful content [28,35]. In addition, the reach of digital interventions has been subject to controversial debate, as concerns have been expressed that barriers to treatment may be created rather than removed [36,37]. A review indicated that studies of the effectiveness of mHealth apps mostly include samples of predominantly female, White participants with an average age of 30 to 45 years [38], and the degree of generalizability of findings to service users with other characteristics remains largely unexplored. Therefore, the development of evidence-based, low-threshold interventions that specifically target established candidate mechanisms that have been linked to the development and persistence of mental health conditions across various groups and settings is urgently needed. In addition, it is crucial to explore the association of participants' baseline characteristics with putative mechanisms and outcomes to examine the reach of the intervention.

Extensive research identified stress reactivity as a putative transdiagnostic mechanism in the development of psychopathology and a promising target for prevention and

early intervention [26]. Stress reactivity (ie, increases in negative affect in response to minor daily stressors) is thought to be a behavioral marker of stress sensitization, positing that frequent or chronic experiences of adversity may gradually increase individuals' stress response to subsequent adversities and minor stressors in everyday life [26,39,40].

Compassion-focused interventions (CFIs) may be a promising approach to target stress reactivity in daily life. Building on a combination of evolutionary psychology, attachment theory, and social mentality theory, the compassion-focused approach claims that various psychological problems are caused by unhelpful loops among distressing emotions, defensive behaviors, and cognitive processes such as rumination, worry, and self-criticism [41]. A model with 3 interrelated major emotional systems is suggested [41-43]: threat, drive, and soothing. Many people experience an overactive threat system, an overactive or somehow blocked drive system, and an underactive soothing system [41]. Therefore, CFIs focus on strengthening the soothing system, as it is thought to be an antagonist to an overactive threat system and a good basis for a well-functioning drive system. CFIs are not symptom specific, and previous studies demonstrated that they are an effective treatment for various mental health problems [19,44-46]. Positive imagery, a key component of CFIs, has been shown to effectively reduce a wide range of mental health problems and increases positive affect, optimism, and behavioral activation [46-50]. In laboratory studies, the application of compassion-focused techniques has been shown to reduce state negative affect and paranoia in moments of high stress [49,51].

Combining digital approaches and CFIs in a hybrid intervention using imagery-based techniques may be particularly well-suited to target stress reactivity in the daily life of young people. Previous research indicated higher acceptability and larger effect sizes for hybrid interventions in comparison with stand-alone internet- and mobile-based interventions [52,53]. Therefore, EMIcompass was developed as a hybrid intervention combining an EMI with guided face-to-face sessions. A pilot study provided initial evidence for feasibility, safety, and beneficial effects of a compassion-focused EMI for enhancing resilience in help-seeking young people [54]. Feasibility and initial signals of efficacy of the intervention have been investigated in a registered exploratory RCT in Germany [55], comparing treatment as usual (TAU) with TAU+EMIcompass in young people with early mental health problems.

Objectives

This paper aims to (1) present the intervention manual for EMIcompass, a hybrid intervention combining an EMI and face-to-face sessions aiming at enhancing resilience in help-seeking young people based on compassion-focused principles [41-43] and (2) explore whether participants' baseline characteristics are associated with putative mechanisms and outcomes of the EMIcompass intervention. To this end, we aimed to obtain first parameter estimates and explore initial signals as to whether sociodemographic, clinical, and functional characteristics at baseline (ie, clinical stage, psychological distress, general psychopathology, level of functioning, age, gender, and minority status) are associated with putative mechanisms (ie, change in self-compassion, change in emotion regulation, working alliance, and training frequency); and sociodemographic, clinical, and functional characteristics (ie, clinical stage, psychological distress, general psychopathology, level of functioning, age, gender, and minority status) as well as self-compassion and emotion regulation at baseline are associated with clinical outcomes (ie, psychological distress and general psychopathology at postintervention and 4-week follow-ups) in the experimental condition and obtain 95% CIs.

Methods

Study Design

In our exploratory RCT, participants were randomly allocated to a control condition of TAU or an experimental condition of TAU+EMIcompass in a 50:50 ratio. For this analysis, data from the experimental condition were used to examine the impact of participants' baseline characteristics on the putative mechanisms and outcomes of the intervention. In the RCT, candidate mechanisms (primary: stress reactivity; secondary: resilience, interpersonal sensitivity, threat anticipation, and negative affective appraisals) and outcomes (primary: psychological distress; secondary: primary psychiatric symptoms, general psychopathology, and quality of life) were assessed at baseline (ie, before randomization), at the end of the intervention, and at the 4-week follow-up. Observer ratings were performed by blinded assessors. The sample size was based on a power simulation for the primary outcome of the trial [56]. The RCT was conducted between August 2019 and September 2021. Appointments were held in person or via video calls (owing to the COVID-19 pandemic). Further details on study procedures are described in the study protocol [56].

Ethics Approval

The trial has received ethical approval from the local ethics committee of the Medical Faculty Mannheim, Heidelberg University (2017-602N-MA). All participants and, in case of minors, parents or legal guardians, provided written informed consent before inclusion in the study.

Manual for the EMIcompass Intervention

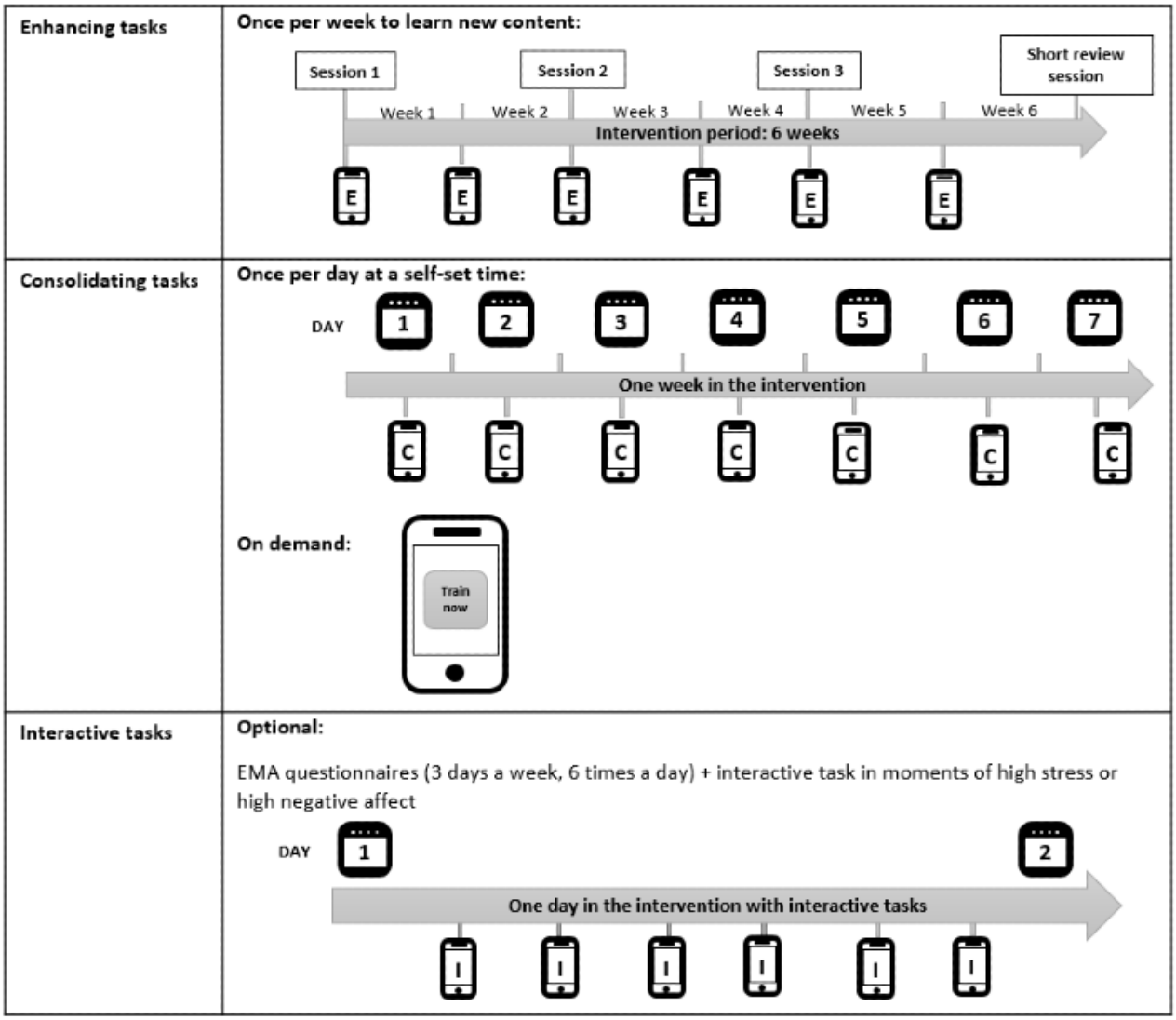
To ensure consistent delivery of the intervention, a structured manual was developed and refined building on the manual from the pilot study (Multimedia Appendix 1 provides changes to the pilot version) [54]. The development and optimization process comprised a scoping review of available literature and existing manuals. In addition, local CFI experts were consulted, and the team's clinical experience of working with these approaches was considered. The intervention was designed based on principles of EMIs [26,27,29,34].

The development and optimization process resulted in a structured manual for a 6-week intervention combining 4 individual sessions with daily training via a dedicated smartphone app. The manual is reported in the Multimedia Appendix 2 in line with state-of-the art guidelines such as World Health Organization guidelines for reporting health interventions using mobile phones [57] as well as the Template for Intervention Description and Replication Checklist [58]. An overview of the intervention structure and the types of tasks is

provided in Figure 1. Figure 2 displays a summary of the intervention content. The intervention can be aligned to participants' personal needs; for example, sessions or training weeks can be repeated if necessary. Moreover, the intervention provides 2 different study tracks with varying foci and demand levels. On the basis of the trained psychologists' impression and the participants' experiences in the first 2 weeks of the intervention, participants were allocated to the basic or the elaborate track of the intervention. The basic study track focused on creating feelings of safeness and calmness by introducing breathing techniques and soothing imagery. The elaborate track extended breathing exercises and soothing imagery by introducing self-compassionate imagery and writing.

The intervention comprised 3 guided sessions to introduce compassion-focused principles and practical tasks to activate participants' soothing system and to provide feedback on their current progress and a short review session. The content was presented on the smartphone and was discussed with the trained psychologist. All sessions could be delivered in person or via video calls. The in-person sessions were delivered in dedicated treatment and assessment rooms. For sessions delivered via video call, participants attended the sessions at home. Psychologists were trained in delivering the EMIcompass intervention and supervised by an expert in CFIs (BB) to ensure intervention quality.

Figure 1. Overview of the intervention structure and the types of tasks. EMA: ecological momentary assessment.

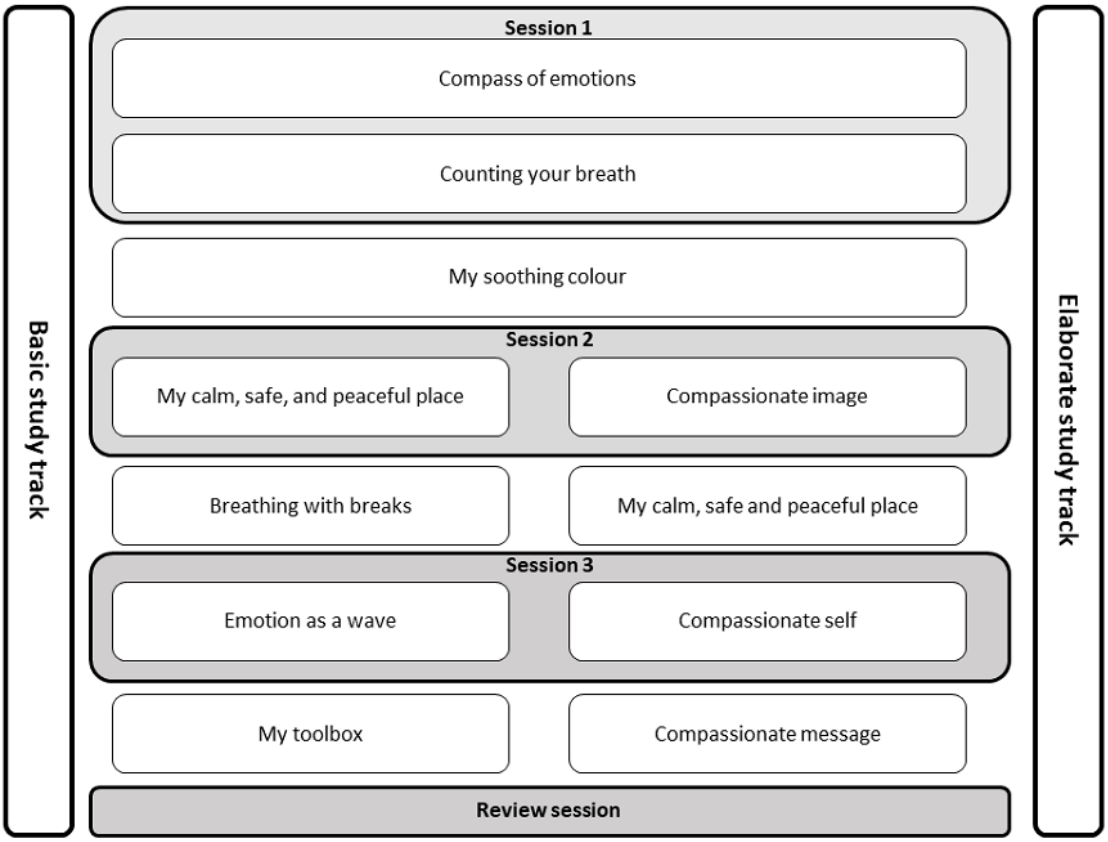


To facilitate real-time and real-world translation of techniques into participants' daily lives, an EMI was administered through an mHealth app (movisensXS) on a study smartphone that they received in the first guided session. To learn new techniques, participants were asked to complete one *enhancing* task per week, which was subsequently extended over the intervention. In the weeks with sessions, the new task was introduced in contact with the trained psychologists; in the weeks without session, participants familiarized with the new enhancing task autonomously. Short *consolidating* tasks were offered to practice the techniques previously introduced in enhancing tasks. Once a day, at a time set by the participants, a signal was prompted to offer participants a consolidating task. In addition, on-demand consolidating tasks were available at any time during the intervention. Furthermore, participants could decide whether they also wanted to allow for *interactive* tasks. To present interactive tasks, the smartphone prompted a signal 6 times per day on 3 consecutive days per week at random within set blocks

of time. At each signal, participants were asked to complete a short EMA questionnaire on momentary stress and affect. If participants indicated high stress or negative affect in the EMA, they were offered an interactive task. Thereby, the interactive tasks guided participants to use previously learned compassion-focused techniques in moments of distress, which is an essential element of CFIs [42]. A gamification element was used to provide feedback on the progress made. If appropriate, participants could choose between reading the instructions on the smartphone's screen and a guided audio version of the tasks.

Between sessions, participants received weekly feedback on their progress and were offered email and phone contact to discuss questions and technical problems. At the beginning of weeks without scheduled session (ie, weeks 2, 4, and 6), participants were contacted to notify them about a new enhancing task becoming available for them to try out autonomously. To proceed with the subsequent study week, participants had to complete at least one consolidating task per week. If this was not the case, the intervention week was repeated.

Figure 2. Summary of the intervention content.



Participants

In line with a modified version of the clinical staging model [12,56], the EMIcompass study recruited young individuals aged 14 to 25 with current psychological distress, Clinical High At-Risk Mental State (CHARMS), or a first treated episode of severe mental disorder (for a detailed description of the modified criteria, see Multimedia Appendix 3 [12,56,59-64]; age range based on suggestions of the youth mental health reform and local regulations [65]). Participants were recruited from mental health services at the Central Institute of Mental Health, Mannheim, Germany, via local registries and advertisements on the institute's webpage and social media. Self-reported and observer-rated measures were used to assess eligibility to participate. All participants (including caregivers for minors) provided informed consent and were reimbursed for their time and travel expenses. Further details on inclusion and exclusion criteria are provided in the study protocol [56].

Measures

Multimedia Appendix 4 [12,59,60,66-78] provides an overview of the measures used and the time points of administration. We used self-reports and, in the case of ethnicity, family assessments to collect data on sociodemographic characteristics. Clinical characteristics (ie, clinical stage, psychological distress, general psychopathology, and level of functioning) were assessed using self-report questionnaires, observer ratings, and standardized interviews. Self-report questionnaires were used to assess overall self-compassion, emotion regulation, and working alliance. Momentary self-compassion was assessed using EMA. The total number of training tasks completed in the EMI was used as an indicator of training frequency. Multimedia Appendix 5 displays a correlation table of the measures used.

Statistical Analysis

The study was registered on the open science framework prior to accessing the data [79]. To obtain parameter estimates for the effect of sociodemographic, clinical, and functional

characteristics on putative mechanisms and processes, we fitted linear regression models with change in self-compassion ($\delta_{postintervention-baseline}$), change in adaptive and maladaptive emotion regulation ($\delta_{postintervention-baseline}$), working alliance (patient and therapist ratings and total scores), and training frequency (total score) as dependent variables. Independent variables in the models were clinical stage (stage 1a, stage 1b, and stage 2), psychological distress, general psychopathology, level of functioning, age, gender (female and male), and ethnic minority status (minority and majority). Parameter estimates (95% CIs) were obtained for the main effects of baseline characteristics on change in self-compassion, change in adaptive and maladaptive emotion regulation, working alliance, and training frequency. We computed partial η^2 as estimators of effect size for the predictors.

To obtain parameter estimates for the effect of sociodemographic, clinical, and functional characteristics and baseline level of self-compassion, adaptive, and maladaptive emotion regulation on clinical outcomes, we fitted mixed effects regression models with psychological distress and general psychopathology at postintervention or at follow-up as the dependent variables. Independent variables in these models were time (postintervention and follow-up), clinical stage (stage 1a, stage 1b, and stage 2), level of functioning at baseline, age, gender (female and male), ethnic minority status (minority and majority), momentary and overall self-rated self-compassion at baseline, adaptive and maladaptive emotion regulation at baseline, psychological distress at baseline (as independent variable in the model with general psychopathology at postintervention or follow-up as outcome and as control variable with psychological distress at postintervention or follow-up as outcome), and general psychopathology at baseline (as independent variable in the model with psychological distress at postintervention or follow-up as outcome and as control variable with general psychopathology at postintervention or follow-up as outcome). We took into account the within-subject clustering of repeated measures by adding a level-2 random intercept. The model was fitted using restricted maximum likelihood estimation. Parameter estimates (95% CIs) were obtained for the main effects of baseline characteristics on outcomes across the 2 follow-up (ie, postintervention and 4-week follow-ups). In the next step, given the exploratory nature of this trial, 95% CIs for the 2 time-specific contrasts were obtained. For this, the aforementioned model was extended by time×characteristic interactions (time×clinical stage, time×psychological distress, time×general psychopathology, time×level of functioning, time×age, time×gender, time×self-compassion, time×adaptive emotion regulation, and time×maladaptive emotion regulation). The "margins" command was used for each interaction to obtain predicted means for both time points and all manifestations of categorical variables (eg, "margins time point #clinical stage"). For continuous variables, the "margins" command was used with z-standardized continuous variables to obtain predicted means for both time points and low (mean−1 SD), mean, and high (mean+1 SD) levels of the given continuous variable (eg, "margins, at [z_age = (−1 0 1)] over [time]").

To transform the results into an effect size, the model was run including only a random intercept for participants, the estimated target relationship, and the baseline control to obtain the conditional and pooled variance across both assessment time points [78,80,81]. The resulting estimate of variance therefore approximates the variation in the dependent variable at any cross-section in postintervention and follow-up. The resulting estimate is on a similar scale as other typical d-type effect sizes (at "0" of any random slopes, if included), and if additional random effects were strong, these variances are underestimations, and the effect sizes in the following likely are at the upper possible limit.

The analysis was conducted according to intention-to-treat principles, with data from all participants entered into the analysis, including those who have low adherence to or who dropped out of the intervention. To screen for potential collinearity problems, we computed variance inflation factors and tolerance values (Multimedia Appendix 6).

Results

Basic Sample and Clinical Characteristics

An overview of basic sample and clinical characteristics is displayed in Table 1. The sample of those randomized to the experimental condition comprised 46 individuals (50% of the total sample in the exploratory RCT of N=92), with a mean age of 21.30 (SD 2.84; range 14-25) years. Most participants (35/46, 76%) identified as girls or women, 24% (11/46) of the participants identified as boys or men, and no participant identified as nonbinary. We identified 70% (32/46) of participants as White majority (German), 9% (4/46) as White other, and 22% (10/46) as other or mixed ethnicity. Most participants were classified as stage 1a (psychological distress, 26/46, 57%), 28% (13/46) of the participants met criteria for stage 1b (CHARMS), and 15% (7/46) of the participants were classified as stage 2 (first episode of severe mental disorder). The mean level of psychological distress at baseline was 28.20 (SD 5.08), and the mean level of general psychopathology at baseline was 24.55 (SD 9.94). The average level of functioning was 71.83 (SD 9.89). Participants showed comparable levels of overall self-rated self-compassion (P=.33) and adaptive (P=.57) and maladaptive emotion regulation (P=.21) at baseline and postintervention. We observed increases in momentary self-compassion at postintervention (P=.02).

Table 1. Basic sample and clinical characteristics.

	Baseline (n_{max}=46)[a]	Postintervention (n_{max}=45)	Follow-up (n_{max}=45)	Baseline v postintervention	
				t test (*df*)	*P* value
Age at baseline (years), mean (SD)	21.30 (2.84)	—[b]	—	—	—
Gender, n (%)		—	—	—	—
Female	35 (76)				
Male	11 (24)				
Nonbinary	0 (0)				
Ethnicity, n (%)		—	—	—	—
White majority	32 (70)				
Minority					
Mixed White majority or White other	3 (7)				
White other	4 (9)				
Turkish	3 (6)				
Mixed other	2 (4)				
Middle East	1 (2)				
Asian	1 (2)				
Level of education, n (%)		—	—	—	—
School: General Certificate of Secondary Education	7 (15)				
Further: A levels	14 (30)				
Higher: university	25 (54)				
Employment status, n (%)		—	—	—	—
Student	39 (85)				
School	4 (9)				
Vocational training or university	35 (76)				
Employed	4 (9)				
Unemployed	3 (6)				
Clinical stage at baseline, n (%)		—	—	—	—
1a	26 (57)				
1b	13 (28)				
2	7 (15)				
Level of functioning at baseline, mean (SD)	71.83 (9.89)	—	—	—	—
Psychological distress, mean (SD)	28.20 (5.08)	24.11 (6.55)	22.73 (7.16)	—	—
General psychopathology, mean (SD)	24.55 (9.94)	18.0 (12.03)	16.20 (10.68)	—	—
Self-compassion, mean (SD)					
Overall self-rating	18.34 (2.77)	18.70 (2.06)	—	−0.99 (42)	.33
Momentary rating	3.89 (0.87)	4.30 (1.06)	—	−2.35 (44)	.02
Emotion regulation, mean (SD)					
Adaptive	5.51 (1.45)	5.61 (1.57)	—	−0.57 (42)	.57
Maladaptive	5.97 (1.46)	5.64 (1.45)	—	1.27 (42)	.21
Training frequency, mean (SD)	—	75.84 (85.09)	—	—	—

	Baseline (n_{max}=46)[a]	Postintervention (n_{max}=45)	Follow-up (n_{max}=45)	Baseline v postintervention	
				t test (*df*)	*P* value
Working alliance, mean (SD)					
Patient rating	—	48.07 (8.37)	—	—	—
Therapist rating	—	46.74 (6.47)	—	—	—

[a]Sample sizes varied owing to missing values at baseline (n_{max}=46; n_{min}=45).

[b]Not available.

Sociodemographic, Clinical, and Functional Characteristics at Baseline Associated With Putative Mechanisms and Processes of Change

Table 2 presents the associations of sociodemographic, clinical, and functional characteristics at baseline with change in self-compassion and emotion regulation. There was no evidence for initial signals that participants' characteristics at baseline were associated with change in overall self-rated self-compassion ($\delta_{postintervention-baseline}$). For change in momentary self-compassion, we observed a tendency for an association with age (B=0.11, 95% CI 0.00-0.22): older participants tended to show more pronounced change in momentary self-compassion from baseline to postintervention. Clinical stage was associated with change in adaptive emotion regulation such that participants in stage 2 showed more pronounced positive changes in adaptive emotion regulation compared with participants in stage 1a (B=1.50, 95% CI 0.06-2.93). For change in maladaptive emotion regulation, we found a tendency for an association with general psychopathology such that participants with lower levels of psychopathology at baseline tended to show more pronounced reductions in maladaptive emotion regulation (B=0.08, 95% CI −0.01 to 0.16).

Table 3 presents the associations of sociodemographic, clinical, and functional characteristics at baseline with working alliance and training frequency. We found no evidence for initial signals of associations of working alliance and training frequency with baseline characteristics.

Table 2. Associations of sociodemographic, clinical, and functional characteristics at baseline with change in self-compassion and emotion regulation.

	Putative mechanisms of change							
	Change in overall self-rated self-compassion (n=43)		Change in momentary self-compassion (n=45)		Change in adaptive emotion regulation (n=43)		Change in maladaptive emotion regulation (n=43)	
	B (95% CI)	Effect size[a]	B (95% CI)	Effect size	B (95% CI)	Effect size	B (95% CI)	Effect size
Age	−0.05 (−0.40 to 0.29)	0.00	0.11 (0.00 to 0.22)	0.10	−0.10 (−0.23 to 0.04)	0.05	−0.09 (−0.26 to 0.09)	0.03
Gender	0.81 (−1.68 to 3.29)	0.01	−0.07 (−0.85 to 0.71)	0.00	−0.70 (−1.70 to 0.30)	0.06	0.45 (−0.81 to 1.72)	0.02
Ethnic minority status	0.91 (−1.28 to 3.10)	0.02	−0.20 (−0.89 to 0.49)	0.01	0.07 (−0.81 to 0.95)	0.00	0.51 (−0.60 to 1.62)	0.02
Clinical stage[b]		0.01		0.04		0.15		0.00
Stage 1b	0.57 (−1.66 to 2.81)		−0.31 (−1.03 to 0.40)		0.70 (−0.20 to 1.60)		0.18 (−0.95 to 1.32)	
Stage 2	−0.34 (−3.91 to 3.23)		0.27 (−0.81 to 1.34)		1.50 (0.06 to 2.93)		0.13 (−1.68 to 1.94)	
Psychological distress	−0.01 (−0.38 to 0.35)	0.00	0.03 (−0.09 to 0.14)	0.01	−0.02 (−0.16 to 0.13)	0.00	−0.15 (−0.34 to 0.03)	0.08
General psychopathology	−0.01 (−0.18 to 0.16)	0.00	0.03 (−0.02 to 0.09)	0.04	−0.04 (−0.11 to 0.03)	0.03	0.08 (−0.01 to 0.16)	0.09
Level of functioning	−0.01 (−0.11 to 0.10)	0.00	−0.02 (−0.06 to 0.01)	0.04	0.00 (−0.04 to 0.05)	0.00	0.01 (−0.05 to 0.07)	0.00

[a]Effect size partial η^2.

[b]Stage 1a (individuals with psychological distress) as reference category.

Table 3. Associations of sociodemographic, clinical, and functional characteristics at baseline with working alliance and training frequency.

	Putative mechanisms of change					
	Working alliance—patient rating (n=44)		Working alliance—therapist rating (n=43)		Training frequency (n=45)	
	B (95% CI)	Effect size[a]	B (95% CI)	Effect size	B (95% CI)	Effect size
Age	0.57 (−0.33 to 1.46)	0.05	0.17 (−0.59 to 0.94)	0.01	2.69 (−7.38 to 12.77)	0.01
Gender	2.55 (−3.93 to 9.03)	0.02	2.72 (−2.93 to 8.37)	0.03	12.80 (−85.02 to 9.43)	0.00
Ethnic minority status	1.89 (−3.65 to 7.44)	0.01	1.54 (−3.18 to 6.26)	0.01	−26.31 (88.48 to 35.86)	0.02
Clinical stage[b]		0.09		0.04		0.05
Stage 1b	5.03 (−0.68 to 10.74)		−1.22 (−5.93 to 3.49)		26.07 (−38.45 to 90.60)	
Stage 2	−0.41 (−9.47 to 8.65)		3.51 (−4.00 to 11.02)		−41.07 (−141.05 to 58.91)	
Psychological distress	0.70 (−0.24 to 1.65)	0.06	0.33 (−0.46 to 1.12)	0.02	4.83 (−5.80 to 15.47)	0.02
General psychopathology	0.00 (−0.44 to 0.44)	0.00	−0.13 (−0.49 to 0.23)	0.02	−1.82 (−6.74 to 3.11)	0.02
Level of functioning	0.01 (−0.28 To 0.29)	0.00	0.11 (−0.14 To 0.36)	0.02	−0.31 (−3.50 to 2.87)	0.00

[a]Effect size partial η^2.

[b]Stage 1a (individuals with psychological distress) as reference category.

Sociodemographic, Clinical, and Functional Characteristics; Self-compassion; and Emotion Regulation at Baseline Associated With Clinical Outcomes

Table 4 presents findings on associations of psychological distress with participants' characteristics and level of putative mechanisms at baseline and predicted marginal means. There was some evidence for a main effect of momentary self-compassion such that higher momentary self-compassion at baseline tended to be associated with, on average, lower levels of psychological distress across postintervention and follow-up assessments (B=−2.83, 95% CI −5.66 to 0.00). There was no evidence for main effects of sociodemographic or clinical characteristics, overall self-rated self-compassion, and emotion regulation on psychological distress.

Table 5 presents findings on associations of general psychopathology with participants' characteristics and level of putative mechanisms at baseline and predicted marginal means. There was no evidence for initial signals of main effects of sociodemographic, clinical, and functional characteristics on general psychopathology.

Cross-differences between high and low levels of baseline characteristics at the time points are presented in Multimedia Appendix 7 [82].

Table 4. Associations of psychological distress with participants' characteristics and level of putative mechanisms and processes at baseline and predicted marginal means[a].

	Postintervention		Follow-up		Adjusted B (95% CI)	Effect size[b]
	Predicted marginal mean (95% CI)	SE	Predicted marginal mean (95% CI)	SE		
Time	N/A[c]	N/A	N/A	N/A	−7.46 (−37.20 to 22.29)	−1.16
Age					−0.31 (−1.04 to 0.42)	−0.05
Low[d]	24.92 (22.26 to 27.57)	1.36	24.99 (22.33 to 27.65)	1.36		
Mean	24.03 (22.33 to 25.73)	0.87	22.62 (20.93 to 24.32)	0.87		
High[e]	23.15 (20.44 to 25.86)	1.38	20.26 (17.55 to 22.96)	1.38		
Gender					−0.51 (−5.26 to 4.23)	−0.08
Female	24.16 (22.15 to 26.17)	1.03	23.90 (21.89 to 25.91)	1.03		
Male	23.65 (19.61 to 27.69)	2.06	18.45 (14.41 to 22.49)	2.06		
Ethnic minority status					2.39 (−4.40 to 9.18)	0.37
White majority	23.83 (22.02 to 25.64)	0.92	22.71 (20.91 to 24.52)	0.92		
Minority	26.22 (19.81 to 32.62)	3.27	22.11 (15.70 to 28.52)	3.27		
Clinical stage[f]					−0.19 (−3.62 to 3.24)	−0.03
Stage 1a	25.42 (23.12 to 27.73)	1.18	23.75 (21.44 to 26.05)	1.18		
Stage 1b	21.84 (18.58 to 25.10)	1.66	20.21 (16.95 to 23.46)	1.66		
Stage 2	22.62 (16.44 to 28.79)	3.15	23.37 (17.19 to 29.55)	3.15		
General psychopathology at baseline					0.04 (−0.29 to 0.38)	0.01
Low	23.62 (19.88 to 27.37)	1.91	18.79 (15.05 to 22.54)	1.91		
Mean	24.03 (22.33 to 25.74)	0.87	22.56 (20.86 to 24.26)	0.87		
High	24.45 (20.85 to 28.04)	1.83	26.33 (22.73 to 29.92)	1.83		
Level of functioning at baseline					−0.10 (−0.30 to 0.11)	-0.02
Low	25.02 (22.34 to 27.70)	1.37	22.27 (19.59 to 24.95)	1.37		
Mean	24.07 (22.37 to 25.77)	0.87	22.65 (20.95 to 24.35)	0.87		
High	23.12 (20.52 to 25.72)	1.33	23.03 (20.43 to 25.63)	1.33		
Overall self-rated self-compassion at baseline					0.06 (−0.83 to 0.94)	0.01
Low	23.81 (19.66 to 27.95)	2.11	21.28 (17.14 to 25.43)	2.11		
Mean	24.02 (22.29 to 25.76)	0.89	22.53 (20.79 to 24.26)	0.89		
High	24.24 (20.74 to 27.74)	1.79	23.77 (20.27 to 27.28)	1.79		
Momentary self-compassion at baseline					−2.83 (−5.66 to 0.00)	-0.37
Low	26.60 (23.53 to 29.68)	1.57	23.05 (19.97 to 26.12)	1.57		
Mean	24.17 (22.47 to 25.88)	0.87	22.68 (20.98 to 24.38)	0.87		
High	21.74 (18.88 to 24.60)	1.46	22.31 (19.45 to 25.17)	1.46		
Adaptive emotion regulation at baseline					−0.36 (−1.96 to 1.24)	−0.06
Low	24.57 (21.68 to 27.45)	1.47	24.01 (21.13 to 26.90)	1.47		
Mean	24.01 (22.35 to 25.74)	0.87	22.66 (20.96 to 24.63)	0.87		
High	23.52 (20.64 to 26.41)	1.47	21.30 (18.42 to 24.19)	1.47		
Maladaptive emotion regulation at baseline					0.03 (−1.74 to 1.80)	0.00
Low	24.00 (20.92 to 27.08)	1.57	20.53 (17.45 to 23.62)	1.57		
Mean	24.05 (22.35 to 25.74)	0.87	22.66 (20.96 to 24.36)	0.87		

	Postintervention		Follow-up		Adjusted B (95% CI)	Effect size[b]
	Predicted marginal mean (95% CI)	SE	Predicted marginal mean (95% CI)	SE		
High	24.09 (21.01 to 27.17)	1.57	24.78 (21.70 to 27.86)	1.57		

[a]Adjusted for baseline levels of psychological distress.

[b]d-type effect size.

[c]N/A: not applicable.

[d]Low = mean − 1 SD.

[e]High = mean + 1 SD.

[f]Stage 1a (individuals with psychological distress) is used as the reference category.

Table 5. Associations of general psychopathology with participants' characteristics and level of putative mechanisms and processes at baseline and predicted marginal means[a].

	Postintervention		Follow-up		Adjusted B (95% CI)	Effect size[b]
	Predicted marginal mean (95% CI)	SE	Predicted marginal mean (95% CI)	SE		
Time	—[c]	—	—	—	−25.10 (−56.83 to 6.63)	−2.38
Age					−0.79 (−2.05 to 0.46)	−0.08
Low[d]	19.75 (15.20 to 24.29)	2.32	18.27 (13.72 to 22.81)	2.32		
Mean	17.49 (14.58 to 20.39)	1.48	16.06 (13.15 to 18.96)	1.48		
High[e]	15.23 (10.61 to 19.86)	2.36	13.85 (9.22 to 18.48)	2.36		
Gender					−2.14 (−10.25 to 5.97)	−0.20
Female	18.01 (14.57 to 21.45)	1.76	17.44 (14.00 to 20.88)	1.76		
Male	15.87 (8.96 to 22.78)	3.53	11.49 (4.58 to 8.40)	3.53		
Ethnic minority status					1.40 (−10.21 to 13.02)	0.13
White majority	17.40 (14.30 to 20.49)	1.58	15.69 (12.60 to 18.78)	1.58		
Minority	18.80 (7.84 to 29.75)	5.59	20.10 (9.15 to 31.06)	5.59		
Clinical stage[f]					−0.97 (−6.84 to 4.89)	−0.09
Stage 1a	19.07 (15.13 to 23.01)	2.01	18.29 (14.35 to 22.23)	2.01		
Stage 1b	14.32 (8.75 to 19.88)	2.84	11.60 (6.03 to 17.17)	2.84		
Stage 2	17.82 (7.26 to 28.39)	5.39	16.34 (5.77 to 26.91)	5.39		
Psychological distress at baseline					−0.27 (1.55 to 1.01)	−0.03
Low	18.83 (12.03 to 25.63)	3.47	19.44 (12.64 to 26.25)	3.47		
Mean	17.52 (14.61 to 20.42)	1.48	16.07 (13.17 to 18.98)	1.48		
High	16.20 (9.35 to 23.06)	3.50	12.71 (5.85 to 19.56)	3.50		
Level of functioning at baseline					−0.13 (−0.48 to 0.22)	−0.01
Low	18.84 (14.26 to 23.43)	2.34	15.01 (10.43 to 19.60)	2.34		
Mean	17.56 (14.65 to 20.46)	1.48	16.06 (13.16 to 18.97)	1.48		
High	16.27 (11.82 to 20.72)	2.27	17.11 (12.66 to 21.57)	2.27		
Overall self-rated self-compassion at baseline					0.31 (−1.20 to 1.83)	0.03
Low	16.18 (9.09 to 23.27)	3.62	10.90 (3.81 to 17.99)	3.62		
Mean	17.39 (14.43 to 20.36)	1.51	15.60 (12.63 to 18.57)	1.51		
High	18.61 (12.62 to 24.60)	3.05	20.29 (14.31 to 26.28)	3.05		
Momentary self-compassion at baseline					−3.24 (−8.08 to 1.61)	−0.31
Low	20.45 (15.19 to 25.70)	2.68	15.26 (10.01 to 20.52)	2.68		
Mean	17.67 (14.76 to 20.58)	1.49	16.05 (13.14 to 18.96)	1.49		
High	14.89 (9.99 to 19.79)	2.50	16.84 (11.94 to 21.73)	2.50		
Adaptive emotion regulation at baseline					0.32 (−2.42 to 3.06)	0.03
Low	17.06 (12.13 to 22.00)	2.52	19.47 (14.53 to 24.41)	2.52		
Mean	17.52 (14.62 to 20.43)	1.48	16.09 (13.19 to 19.00)	1.48		
High	17.98 (13.05 to 22.92)	2.52	12.71 (7.78 to 17.65)	2.52		
Maladaptive emotion regulation at baseline					0.97 (−2.42 to 3.06)	0.09
Low	16.11 (10.84 to 21.38)	2.69	14.42 (9.15 to 19.69)	2.69		
Mean	17.52 (14.62 to 20.43)	1.48	16.09 (13.19 to 19.00)	1.48		

	Postintervention		Follow-up		Adjusted B (95% CI)	Effect size[b]
	Predicted marginal mean (95% CI)	SE	Predicted marginal mean (95% CI)	SE		
High	18.93 (13.67 to 24.20)	2.69	17.76 (12.49 to 23.03)	2.69		

[a]Adjusted for baseline levels of general psychopathology.

[b]d-type effect size.

[c]Not available.

[d]Low = mean – 1 SD.

[e]High = mean + 1 SD.

[f]Stage 1a (individuals with psychological distress) is used as the reference category.

Discussion

Principal Findings

First, we developed a hybrid 6-week CFI comprising 2 intervention tracks with varying foci and demand levels. Second, we observed initial signals of effects of age, general psychopathology, and clinical stage on change in momentary self-compassion and change in emotion regulation. Older participants tended to show greater differences in momentary self-compassion comparing baseline and postintervention assessments. Participants classified as stage 2 were found to show greater differences in adaptive emotion regulation comparing baseline and postintervention assessments. In addition, participants with lower levels of psychopathology at baseline showed more pronounced reductions in maladaptive emotion regulation from baseline to postintervention assessments. There was no evidence for associations of other baseline characteristics (eg, gender, minority status, and level of functioning) and putative mechanisms (ie, overall self-rated self-compassion, working alliance, and training frequency). Third, there was some evidence that higher momentary self-compassion at baseline tended to be associated with, on average, lower levels of psychological distress across postintervention and follow-up assessments. We observed no other initial signals that clinical or functional characteristics at baseline impacted clinical outcomes.

Methodological Considerations

The reported results should be interpreted in light of several methodological considerations and limitations. First, sample size and selection as well as the exploratory nature of the analyses need to be critically appraised. Although the analyses were prospectively registered, they reflect secondary analyses with an increased risk of type 1 error. As noted, our findings reflect initial signals of associations of participants' baseline characteristics with putative mechanisms, processes, and outcomes. Moreover, it should be taken into account that boys or men, individuals identifying as nonbinary, and participants from stage 2 (first episode of severe mental disorder) were considerably underrepresented in the sample. However, the gender difference in recruitment may partly be explained by higher prevalence of depressive and anxiety disorders in women and adolescent girls [83,84] and the exclusion of mental health problems that are especially prevalent in men and adolescent boys (eg, primary substance abuse disorder [66]). Randomization in a future definitive trial may therefore need to stratify by gender to rule out potential confounding by this factor. In addition, we assessed ethnicity by considering participants' self-report of citizenship, country of birth, first language, and information provided in participants' family assessment. Grouping participants into broad categories of ethnicity inevitably implies that some participants may have been assigned to a category that they do not consider belonging to and, hence, misclassification. In general, the concept of using categories, for example, with regard to ethnicity or gender, may be criticized as—of course—there is considerable heterogeneity within groups, which needs to be further explored in qualitative analyses [85,86]. These limitations can be tolerated at the exploratory stage of developing a complex intervention but should be addressed in future, definitive trials.

Second, operationalizations of putative mechanisms were not measured at multiple time points during the intervention, and difference scores were used as proxies for change in self-compassion and emotion regulation. While proxies are acceptable in this exploratory study, a future definitive trial may use multiple assessments during the intervention to yield more fine-grained data on potential changes in mechanisms.

Third, the assessment of self-compassion needs to be critically appraised: in our analyses, overall self-rated self-compassion and momentary self-compassion were not correlated, indexing low convergent validity (Multimedia Appendix 5). Similar phenomena have been observed before, for example, for negative symptoms measured with EMA and interviewer-rated measures, which may tap distinct but related constructs [87]. This may be viewed as underscoring the relevance of assessment under real-time and real-world conditions, which is supported by moderate to large correlations of momentary self-compassion with clinical characteristics (ie, clinical stage, psychological distress, general psychopathology, and level of functioning), indicating high concurrent validity. However, as the items for assessing momentary self-compassion were used for the first time in this study, they may also not fully capture the construct of self-compassion as operationalized by the subscales in the Self-Compassion Scale (ie, they are more similar in content to items from the self-kindness than mindfulness subscale) [73]. In addition, we aggregated EMA data on momentary self-compassion at the person level, which led to a loss of information in comparison with the level of EMA observations, given the repeated measurement and temporal variability EMA captures as an intensive longitudinal data collection method (Schick A, unpublished data, 2022). Nonetheless, aggregated

experience sampling measures may still capture the target constructs with less noise and greater sensitivity than recall measures [88], so this may not reduce this study's informative value substantially.

Fourth, potential influences of the COVID-19 pandemic have not been statistically accounted for in current analyses and should be considered when interpreting the findings. Owing to local regulations (eg, lockdowns and contact restrictions), the intervention sessions were shifted from face-to-face contact to video calls. Recent systematic reviews and meta-analyses indicated no differences in telehealth and in-person psychotherapy [89], but generalizability to settings in which both in-person sessions and a video call format are used flexibly remains unclear, and the impact cannot be determined with certainty without further research.

Comparison With Previous Research

To our knowledge, the EMIcompass intervention is the first hybrid CFI blending an EMI and face-to-face sessions designed to enhance self-compassion and resilience in young people with nonspecific psychological distress, CHARMS, and first episode of severe mental disorder. Building on principles of EMIs [26,27,29,34], EMIcompass combined different intervention elements: enhancing tasks provided participants with new CFI strategies. Consolidating tasks facilitated training in different contexts and translation into daily life increasing the chances of generalization. Elements of experience sampling were used to increase reflective processing improving insight and awareness of own cognitive and emotional processes [90]. This may be further improved by incorporating elements of feedback into future versions of the intervention [91]. In addition, assessing stress and affect in daily life allows the EMI to offer useful techniques in moments of high distress (ie, interactive tasks), providing participants with support in challenging life situations.

For the EMIcompass intervention, the results from an uncontrolled pilot study [54] indicated a reduction of stress reactivity at postintervention and follow-up and reduced clinical symptoms at follow-up when compared with baseline. A recent exploratory RCT [56] indicated that all feasibility criteria were met and a reduction of stress reactivity in the experimental condition as the primary candidate mechanism in comparison with a control condition of TAU. In addition, it suggests initial signals that the EMIcompass intervention may have beneficial effects on resilience in daily life and quality of life. Detailed findings on feasibility and initial signals of efficacy are described elsewhere [56].

Apart from an association of age and change in momentary self-compassion, participants' sociodemographic characteristics were not associated with putative processes, mechanisms, and outcomes of the EMIcompass intervention. This is at variance with findings in traditional psychotherapy for depression and psychosis, where reviews indicate differential treatment effects for various sociodemographic characteristics (eg, age, gender, and marital status) [92,93]. In an Acceptance and Commitment Therapy–based EMI in individuals at ultra–high risk for psychosis and with a first episode of psychosis, ethnic minority status was associated with lower compliance and higher app usefulness, whereas being female predicted lower usefulness of the app's metaphor images (van Aubel E, unpublished data, August 2022).

When examining the impact of clinical and functional characteristics, we observed associations of clinical stage and general psychopathology with putative mechanisms and processes (ie, change in momentary self-compassion and change in emotion regulation). Interestingly, *later* clinical stage was associated with a more pronounced increase in adaptive emotion regulation, whereas *lower* levels of general psychopathology tended to be associated with a more pronounced reduction of maladaptive emotion regulation. However, the findings on clinical stage must be interpreted with caution, given the small number of participants from stage 2 included in the study. The possibility of ceiling effects for a particular clinical stage could be ruled out, as the mean levels of adaptive emotion regulation were in the middle range of the scale for all clinical stages. An RCT of cognitive behavioral therapy in patients with psychotic disorders investigating predictors of improvement and dropout indicated that higher symptom severity and poor level of functioning do not pose a barrier to improvement [94]. The findings from an Acceptance and Commitment Therapy–based EMI in individuals at ultra–high risk for psychosis and with a first episode of psychosis show a differentiated perspective on symptom severity: the severity of affective symptoms was associated with higher perceived usefulness and that of negative symptoms was associated with lower perceived usefulness of the intervention (van Aubel E, unpublished data, August 2022). Besides sociodemographic, clinical, and functional characteristics at baseline, we moved beyond these previous studies and examined potential associations of baseline levels of self-compassion and emotion regulation with outcomes of the intervention. We found some evidence that higher levels of momentary self-compassion at baseline were associated with, on average, lower levels of psychological distress across assessment time points. By showing this in a longitudinal intervention study, the current findings extend evidence from a meta-analysis indicating associations of self-compassion and psychological distress in general [95]. However, in this study, this did not hold true for overall self-compassion. Apart from the effects delineated earlier, there were no initial signals of associations, tentatively suggesting that participants' sociodemographic, clinical, and functional characteristics had little influence on their response to the EMIcompass intervention. This may indicate—within the limits of the variables assessed—that the EMIcompass intervention is relatively inclusive and reach of participants is largely independent from their sociodemographic, clinical, and functional baseline characteristics.

The role of digital approaches in improving the reach of those in need within broader conceptualizations has been subject to controversial debate: qualitative studies with health professionals and service users indicate that digital approaches were viewed as having the potential to improve inclusion but also as having the risk of digital exclusion [36,37,96]. Concerns have been raised that digital approaches and the digital divide may further reinforce health inequalities (ie, systematic, avoidable, and unfair differences in health outcomes [97]) in marginalized and

underserved populations; for example, in racial and ethnic minorities [98]. Digital inequalities are suggested to comprise multiple continuous dimensions; for example, socioeconomic and educational background, migrant and ethnic minority status, and health literacy [99-101]. To further improve our understanding of the consequences of digital inequalities for individuals' response to the EMIcompass intervention, future studies may broaden their perspective by including further aspects of marginalized and underserved populations (eg, sexual minority status and socioeconomic background) and examining other criteria (eg, level of functioning, satisfaction with the intervention, goal attainment, and quality of life) in addition to those considered so far.

To address digital exclusion of marginalized and underserved populations, demands for evidence-based digital inclusion strategies have been articulated [102], and potential pathways for improving inclusion in digital approaches have been discussed. On the one hand, adaptations of interventions have been suggested; for example, feasibility and beneficial effects of cultural adaptation of interventions have already been demonstrated [103]. In addition to adapting interventions for specific groups, the needs and perspectives of individual participants should be taken into account in process evaluations combining quantitative and qualitative data [104]. In line with this, we conducted a qualitative study incorporating realist methodology [105] examining what works for whom under which circumstances in the EMIcompass study, the findings of which are reported elsewhere (Paetzold I, unpublished data, 2022). An emerging research field targets the adaptation of digital interventions on an individual level aiming at personalizing assessment and intervention [27,29,34]. On the other hand, the creation of interventions for diverse populations has been suggested, for example, in the REACT recommendations [98]. In line with this approach, a recent review of digital mental health interventions specifically designed for marginalized populations indicated promising results on feasibility and acceptability in pilot studies but also a lack of larger-scale examinations [106].

Conclusions

We developed the first hybrid CFI combining an EMI and face-to-face sessions with 2 intervention tracks and varying foci and demand levels to enhance resilience in young people with early mental health problems. We aimed at exploring whether participants' characteristics at baseline were associated with putative mechanisms and outcomes of the EMIcompass intervention. The findings indicated reach of participants by the intervention largely independent of sociodemographic, clinical, and functional baseline characteristics. The findings need to be confirmed in a definitive trial.

Acknowledgments

The authors are indebted to all individuals who participated in the EMIcompass study. The authors are grateful to Professor Stefan Priebe, Professor Stefan Wellek, Professor Nicolas Ruesch, and Dr Thomas Vaessen as well as Professor Maria Blettner, Professor Sebastian von Peter, and Professor Georg Schomerus for their being a member of the trial steering committee and data monitory and ethics committee, respectively. The authors also thank all trained psychologists and research assistants. This work was funded by a Deutsche Forschungsgemeinschaft / German Research Fundation (DFG) project grant (no. 389626655) and German research foundation Heisenberg professorship (number 389624707) to UR.

Conflicts of Interest

TB served in an advisory or consultancy role for ADHS digital, Infectopharm, Lundbeck, Medice, Neurim Pharmaceuticals, Oberberg GmbH, Roche, and Takeda. He received conference support or speaker's fee from Medice and Takeda. He received royalties from Hogrefe, Kohlhammer, CIP Medien, and Oxford University Press; this work is unrelated to these relationships. The other authors declare that they have no competing interests.

Multimedia Appendix 1
Changes to the pilot version.
[DOCX File , 28 KB - formative_v6i11e39511_app1.docx]

Multimedia Appendix 2
EMIcompass intervention manual.
[DOCX File , 16 KB - formative_v6i11e39511_app2.docx]

Multimedia Appendix 3
Modified criteria of the clinical staging model.
[DOCX File , 699 KB - formative_v6i11e39511_app3.docx]

Multimedia Appendix 4
Measures.
[DOCX File , 29 KB - formative_v6i11e39511_app4.docx]

Multimedia Appendix 5
Correlation table.
[DOCX File , 18 KB - formative_v6i11e39511_app5.docx]

Multimedia Appendix 6
Variance inflation factors and tolerance.
[DOCX File , 19 KB - formative_v6i11e39511_app6.docx]

Multimedia Appendix 7
Cross-differences.
[DOCX File , 19 KB - formative_v6i11e39511_app7.docx]

References

1. Solmi M, Radua J, Olivola M, Croce E, Soardo L, Salazar de Pablo G, et al. Age at onset of mental disorders worldwide: large-scale meta-analysis of 192 epidemiological studies. Mol Psychiatry 2022 Jan;27(1):281-295 [FREE Full text] [doi: 10.1038/s41380-021-01161-7] [Medline: 34079068]
2. Polanczyk GV, Salum GA, Sugaya LS, Caye A, Rohde LA. Annual research review: a meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. J Child Psychol Psychiatry 2015 Mar;56(3):345-365. [doi: 10.1111/jcpp.12381] [Medline: 25649325]
3. Erskine HE, Moffitt TE, Copeland WE, Costello EJ, Ferrari AJ, Patton G, et al. A heavy burden on young minds: the global burden of mental and substance use disorders in children and youth. Psychol Med 2015 May;45(7):1551-1563 [FREE Full text] [doi: 10.1017/S0033291714002888] [Medline: 25534496]
4. Global Burden of Disease Child and Adolescent Health Collaboration, Kassebaum N, Kyu HH, Zoeckler L, Olsen HE, Thomas K, et al. Child and adolescent health from 1990 to 2015: findings from the global burden of diseases, injuries, and risk factors 2015 study. JAMA Pediatr 2017 Jun 01;171(6):573-592 [FREE Full text] [doi: 10.1001/jamapediatrics.2017.0250] [Medline: 28384795]
5. Reininghaus U, Priebe S, Bentall RP. Testing the psychopathology of psychosis: evidence for a general psychosis dimension. Schizophr Bull 2013 Jul;39(4):884-895 [FREE Full text] [doi: 10.1093/schbul/sbr182] [Medline: 22258881]
6. Shevlin M, McElroy E, Bentall RP, Reininghaus U, Murphy J. The psychosis continuum: testing a bifactor model of psychosis in a general population sample. Schizophr Bull 2017 Jan;43(1):133-141 [FREE Full text] [doi: 10.1093/schbul/sbw067] [Medline: 27220965]
7. van Os J, Reininghaus U. Psychosis as a transdiagnostic and extended phenotype in the general population. World Psychiatry 2016 Jun;15(2):118-124 [FREE Full text] [doi: 10.1002/wps.20310] [Medline: 27265696]
8. Reininghaus U, Böhnke JR, Hosang G, Farmer A, Burns T, McGuffin P, et al. Evaluation of the validity and utility of a transdiagnostic psychosis dimension encompassing schizophrenia and bipolar disorder. Br J Psychiatry 2016 Aug;209(2):107-113. [doi: 10.1192/bjp.bp.115.167882] [Medline: 26989099]
9. Forbes MK, Tackett JL, Markon KE, Krueger RF. Beyond comorbidity: toward a dimensional and hierarchical approach to understanding psychopathology across the life span. Dev Psychopathol 2016 Nov;28(4pt1):971-986 [FREE Full text] [doi: 10.1017/S0954579416000651] [Medline: 27739384]
10. Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, et al. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. Am J Psychiatry 2010 Jul;167(7):748-751. [doi: 10.1176/appi.ajp.2010.09091379] [Medline: 20595427]
11. Kotov R, Krueger RF, Watson D, Achenbach TM, Althoff RR, Bagby RM, et al. The Hierarchical Taxonomy of Psychopathology (HiTOP): a dimensional alternative to traditional nosologies. J Abnorm Psychol 2017 May;126(4):454-477. [doi: 10.1037/abn0000258] [Medline: 28333488]
12. Hartmann JA, Nelson B, Spooner R, Paul Amminger G, Chanen A, Davey CG, et al. Broad clinical high-risk mental state (CHARMS): methodology of a cohort study validating criteria for pluripotent risk. Early Interv Psychiatry 2019 Jun;13(3):379-386. [doi: 10.1111/eip.12483] [Medline: 28984077]
13. McGorry PD, Purcell R, Hickie IB, Yung AR, Pantelis C, Jackson HJ. Clinical staging: a heuristic model for psychiatry and youth mental health. Med J Aust 2007 Oct 01;187(S7):S40-S42. [doi: 10.5694/j.1326-5377.2007.tb01335.x] [Medline: 17908024]
14. Forbes MK, Rapee RM, Krueger RF. Opportunities for the prevention of mental disorders by reducing general psychopathology in early childhood. Behav Res Ther 2019 Aug;119:103411 [FREE Full text] [doi: 10.1016/j.brat.2019.103411] [Medline: 31202004]
15. Bayer J, Hiscock H, Scalzo K, Mathers M, McDonald M, Morris A, et al. Systematic review of preventive interventions for children's mental health: what would work in Australian contexts? Aust N Z J Psychiatry 2009 Aug;43(8):695-710. [doi: 10.1080/00048670903001893] [Medline: 19629791]

16. Patel V, Flisher AJ, Hetrick S, McGorry P. Mental health of young people: a global public-health challenge. Lancet 2007 Apr 14;369(9569):1302-1313. [doi: 10.1016/S0140-6736(07)60368-7] [Medline: 17434406]
17. Cuijpers P, Berking M, Andersson G, Quigley L, Kleiboer A, Dobson KS. A meta-analysis of cognitive-behavioural therapy for adult depression, alone and in comparison with other treatments. Can J Psychiatry 2013 Jul;58(7):376-385. [doi: 10.1177/070674371305800702] [Medline: 23870719]
18. Hofmann SG, Smits JA. Cognitive-behavioral therapy for adult anxiety disorders: a meta-analysis of randomized placebo-controlled trials. J Clin Psychiatry 2008 Apr;69(4):621-632 [FREE Full text] [doi: 10.4088/jcp.v69n0415] [Medline: 18363421]
19. Kirby JN, Tellegen CL, Steindl SR. A meta-analysis of compassion-based interventions: current state of knowledge and future directions. Behav Ther 2017 Nov;48(6):778-792. [doi: 10.1016/j.beth.2017.06.003] [Medline: 29029675]
20. Tai S, Turkington D. The evolution of cognitive behavior therapy for schizophrenia: current practice and recent developments. Schizophr Bull 2009 Sep;35(5):865-873 [FREE Full text] [doi: 10.1093/schbul/sbp080] [Medline: 19661198]
21. Fava GA, Ruini C, Belaise C. The concept of recovery in major depression. Psychol Med 2007 Mar;37(3):307-317. [doi: 10.1017/S0033291706008981] [Medline: 17311684]
22. McGorry PD, Mei C. Early intervention in youth mental health: progress and future directions. Evid Based Ment Health 2018 Nov;21(4):182-184. [doi: 10.1136/ebmental-2018-300060] [Medline: 30352884]
23. Malla A, Shah J, Iyer S, Boksa P, Joober R, Andersson N, et al. Youth mental health should be a top priority for health care in Canada. Can J Psychiatry 2018 Apr;63(4):216-222 [FREE Full text] [doi: 10.1177/0706743718758968] [Medline: 29528719]
24. Ghio L, Gotelli S, Marcenaro M, Amore M, Natta W. Duration of untreated illness and outcomes in unipolar depression: a systematic review and meta-analysis. J Affect Disord 2014 Jan;152-154:45-51. [doi: 10.1016/j.jad.2013.10.002] [Medline: 24183486]
25. Marshall M, Lewis S, Lockwood A, Drake R, Jones P, Croudace T. Association between duration of untreated psychosis and outcome in cohorts of first-episode patients: a systematic review. Arch Gen Psychiatry 2005 Sep;62(9):975-983. [doi: 10.1001/archpsyc.62.9.975] [Medline: 16143729]
26. Myin-Germeys I, Kasanova Z, Vaessen T, Vachon H, Kirtley O, Viechtbauer W, et al. Experience sampling methodology in mental health research: new insights and technical developments. World Psychiatry 2018 Jun;17(2):123-132 [FREE Full text] [doi: 10.1002/wps.20513] [Medline: 29856567]
27. Myin-Germeys I, Klippel A, Steinhart H, Reininghaus U. Ecological momentary interventions in psychiatry. Curr Opin Psychiatry 2016 Jul;29(4):258-263. [doi: 10.1097/YCO.0000000000000255] [Medline: 27153125]
28. Rauschenberg C, Schick A, Hirjak D, Seidler A, Paetzold I, Apfelbacher C, et al. Evidence synthesis of digital interventions to mitigate the negative impact of the COVID-19 pandemic on public mental health: rapid meta-review. J Med Internet Res 2021 Mar 10;23(3):e23365 [FREE Full text] [doi: 10.2196/23365] [Medline: 33606657]
29. Reininghaus U. Ambulatorische Interventionen in der Psychiatrie: das Momentum für Veränderung im alltäglichen sozialen. Psychiatr Prax 2018 Mar;45(2):59-61. [doi: 10.1055/s-0044-101986] [Medline: 29495051]
30. Rauschenberg C, Schick A, Goetzl C, Roehr S, Riedel-Heller SG, Koppe G, et al. Social isolation, mental health, and use of digital interventions in youth during the COVID-19 pandemic: a nationally representative survey. Eur Psychiatry 2021 Mar 09;64(1):e20 [FREE Full text] [doi: 10.1192/j.eurpsy.2021.17] [Medline: 33686930]
31. Csikszentmihalyi M, Larson R. Validity and reliability of the experience-sampling method. J Nerv Ment Dis 1987 Sep;175(9):526-536. [doi: 10.1097/00005053-198709000-00004] [Medline: 3655778]
32. Myin-Germeys I, Birchwood M, Kwapil T. From environment to therapy in psychosis: a real-world momentary assessment approach. Schizophr Bull 2011 Mar;37(2):244-247 [FREE Full text] [doi: 10.1093/schbul/sbq164] [Medline: 21224492]
33. Heron KE, Smyth JM. Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. Br J Health Psychol 2010 Feb;15(Pt 1):1-39 [FREE Full text] [doi: 10.1348/135910709X466063] [Medline: 19646331]
34. Reininghaus U, Depp CA, Myin-Germeys I. Ecological interventionist causal models in psychosis: targeting psychological mechanisms in daily life. Schizophr Bull 2016 Mar;42(2):264-269 [FREE Full text] [doi: 10.1093/schbul/sbv193] [Medline: 26707864]
35. Larsen ME, Huckvale K, Nicholas J, Torous J, Birrell L, Li E, et al. Using science to sell apps: evaluation of mental health app store quality claims. NPJ Digit Med 2019 Mar 22;2:18 [FREE Full text] [doi: 10.1038/s41746-019-0093-1] [Medline: 31304366]
36. Greer B, Robotham D, Simblett S, Curtis H, Griffiths H, Wykes T. Digital exclusion among mental health service users: qualitative investigation. J Med Internet Res 2019 Jan 09;21(1):e11696 [FREE Full text] [doi: 10.2196/11696] [Medline: 30626564]
37. Bucci S, Berry N, Morris R, Berry K, Haddock G, Lewis S, et al. "They are not hard-to-reach clients. We have just got hard-to-reach services." Staff views of digital health tools in specialist mental health services. Front Psychiatry 2019 May 10;10:344 [FREE Full text] [doi: 10.3389/fpsyt.2019.00344] [Medline: 31133906]
38. 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]

39. Collip D, Myin-Germeys I, Van Os J. Does the concept of "sensitization" provide a plausible mechanism for the putative link between the environment and schizophrenia? Schizophr Bull 2008 Mar;34(2):220-225 [FREE Full text] [doi: 10.1093/schbul/sbm163] [Medline: 18203757]
40. Wichers M, Schrijvers D, Geschwind N, Jacobs N, Myin-Germeys I, Thiery E, et al. Mechanisms of gene-environment interactions in depression: evidence that genes potentiate multiple sources of adversity. Psychol Med 2009 Jul;39(7):1077-1086. [doi: 10.1017/S0033291708004388] [Medline: 18834553]
41. Gilbert P. The origins and nature of compassion focused therapy. Br J Clin Psychol 2014 Mar;53(1):6-41. [doi: 10.1111/bjc.12043] [Medline: 24588760]
42. Gilbert P. Introducing compassion-focused therapy. Adv Psychiatr Treat 2009;15(3):199-208. [doi: 10.1192/apt.bp.107.005264]
43. Gilbert P. Compassion Focused Therapy: Distinctive Features. London, UK: Routledge; 2010.
44. Cuppage J, Baird K, Gibson J, Booth R, Hevey D. Compassion focused therapy: exploring the effectiveness with a transdiagnostic group and potential processes of change. Br J Clin Psychol 2018 Jun;57(2):240-254. [doi: 10.1111/bjc.12162] [Medline: 29044607]
45. Heriot-Maitland C, McCarthy-Jones S, Longden E, Gilbert P. Compassion focused approaches to working with distressing voices. Front Psychol 2019 Feb 1;10:152 [FREE Full text] [doi: 10.3389/fpsyg.2019.00152] [Medline: 30774614]
46. Leaviss J, Uttley L. Psychotherapeutic benefits of compassion-focused therapy: an early systematic review. Psychol Med 2015 Apr;45(5):927-945 [FREE Full text] [doi: 10.1017/S0033291714002141] [Medline: 25215860]
47. Holmes EA, Blackwell SE, Burnett Heyes S, Renner F, Raes F. Mental imagery in depression: phenomenology, potential mechanisms, and treatment implications. Annu Rev Clin Psychol 2016;12:249-280. [doi: 10.1146/annurev-clinpsy-021815-092925] [Medline: 26772205]
48. Holmes EA, Mathews A. Mental imagery in emotion and emotional disorders. Clin Psychol Rev 2010 Apr;30(3):349-362. [doi: 10.1016/j.cpr.2010.01.001] [Medline: 20116915]
49. Pearson J, Naselaris T, Holmes EA, Kosslyn SM. Mental imagery: functional mechanisms and clinical applications. Trends Cogn Sci 2015 Oct;19(10):590-602 [FREE Full text] [doi: 10.1016/j.tics.2015.08.003] [Medline: 26412097]
50. Renner F, Ji JL, Pictet A, Holmes EA, Blackwell SE. Effects of engaging in repeated mental imagery of future positive events on behavioural activation in individuals with major depressive disorder. Cognit Ther Res 2017;41(3):369-380 [FREE Full text] [doi: 10.1007/s10608-016-9776-y] [Medline: 28515538]
51. Lincoln TM, Hohenhaus F, Hartmann M. Can paranoid thoughts be reduced by targeting negative emotions and self-esteem? An experimental investigation of a brief compassion-focused intervention. Cogn Ther Res 2012 Aug 14;37(2):390-402. [doi: 10.1007/s10608-012-9470-7]
52. Topooco N, Riper H, Araya R, Berking M, Brunn M, Chevreul K, E-COMPARED consortium. Attitudes towards digital treatment for depression: a European stakeholder survey. Internet Interv 2017 Jun;8:1-9 [FREE Full text] [doi: 10.1016/j.invent.2017.01.001] [Medline: 30135823]
53. Baumeister H, Reichler L, Munzinger M, Lin J. The impact of guidance on Internet-based mental health interventions — a systematic review. Internet Interv 2014 Oct;1(4):205-215. [doi: 10.1016/j.invent.2014.08.003]
54. Rauschenberg C, Boecking B, Paetzold I, Schruers K, Schick A, van Amelsvoort T, et al. A compassion-focused ecological momentary intervention for enhancing resilience in help-seeking youth: uncontrolled pilot study. JMIR Ment Health 2021 Aug 05;8(8):e25650 [FREE Full text] [doi: 10.2196/25650] [Medline: 34383687]
55. Efficacy of a novel, accessible, transdiagnostic, compassion-focused ecological momentary intervention for enhancing resilience in youth. Deutsches Register Klinischer Studien. 2019. URL: https://www.drks.de/drks_web/navigate.do?navigationId=trial.HTML&TRIAL_ID=DRKS00017265 [accessed 2022-09-20]
56. Schick A, Paetzold I, Rauschenberg C, Hirjak D, Banaschewski T, Meyer-Lindenberg A, et al. Effects of a novel, transdiagnostic, hybrid ecological momentary intervention for improving resilience in youth (EMIcompass): protocol for an exploratory randomized controlled trial. JMIR Res Protoc 2021 Dec 03;10(12):e27462 [FREE Full text] [doi: 10.2196/27462] [Medline: 34870613]
57. Agarwal S, LeFevre AE, Lee J, L'Engle K, Mehl G, Sinha C, WHO mHealth Technical Evidence Review Group. Guidelines for reporting of health interventions using mobile phones: mobile health (mHealth) evidence reporting and assessment (mERA) checklist. BMJ 2016 Mar 17;352:i1174. [doi: 10.1136/bmj.i1174] [Medline: 26988021]
58. 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 Mar 07;348:g1687. [doi: 10.1136/bmj.g1687] [Medline: 24609605]
59. Kessler RC, Andrews G, Colpe LJ, Hiripi E, Mroczek DK, Normand SL, 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]
60. Goldman HH, Skodol AE, Lave TR. Revising axis V for DSM-IV: a review of measures of social functioning. Am J Psychiatry 1992 Sep;149(9):1148-1156. [doi: 10.1176/ajp.149.9.1148] [Medline: 1386964]
61. First MB, Williams JB, Karg RS, Spitzer RL. Structured clinical interview for DSM-5—Research version (SCID-5 for DSM-5, research version; SCID-5-RV). Arlington, VA, USA: American Psychiatric Association; 2015.

62. Hamilton M. The assessment of anxiety states by rating. Br J Med Psychol 1959;32(1):50-55. [doi: 10.1111/j.2044-8341.1959.tb00467.x] [Medline: 13638508]
63. Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry 1960 Feb;23(1):56-62. [doi: 10.1136/jnnp.23.1.56] [Medline: 14399272]
64. Yung AR, Yuen HP, McGorry PD, Phillips LJ, Kelly D, Dell'Olio M, et al. Mapping the onset of psychosis: the Comprehensive Assessment of At-Risk Mental States. Aust N Z J Psychiatry 2005;39(11-12):964-971. [doi: 10.1080/j.1440-1614.2005.01714.x] [Medline: 16343296]
65. Malla A, Iyer S, McGorry P, Cannon M, Coughlan H, Singh S, et al. From early intervention in psychosis to youth mental health reform: a review of the evolution and transformation of mental health services for young people. Soc Psychiatry Psychiatr Epidemiol 2016 Mar 19;51(3):319-326. [doi: 10.1007/s00127-015-1165-4] [Medline: 26687237]
66. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry 2005 Jun;62(6):593-602. [doi: 10.1001/archpsyc.62.6.593] [Medline: 15939837]
67. Derogatis LR. BSI Brief Symptom Inventory. Administration, Scoring, and Procedures Manual. 4th edition. Minneapolis, MN, USA: National Computer Systems; 1993.
68. Derogatis L, Fitzpatrick M. The SCL-90-R, the Brief Symptom Inventory (BSI), and the BSI-18. In: Maruish ME, editor. Instruments for adults: the use of psychological testing for treatment planning and outcomes assessment. Mahwah, NJ, USA: Lawrence Erlbaum Associates; 2004:1-41.
69. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979 Mar;86(2):420-428. [doi: 10.1037//0033-2909.86.2.420] [Medline: 18839484]
70. JASP Team. JASP (Version 0.16.1). JASP. 2022. URL: https://jasp-stats.org/download/ [accessed 2022-10-03]
71. Hilsenroth MJ, Ackerman SJ, Blagys MD, Baumann BD, Baity MR, Smith SR, et al. Reliability and validity of DSM-IV axis V. Am J Psychiatry 2000 Nov;157(11):1858-1863. [doi: 10.1176/appi.ajp.157.11.1858] [Medline: 11058486]
72. Hupfeld J, Ruffieux N. Validierung einer deutschen Version der Self-Compassion Scale (SCS-D). Z Klin Psychol Psychother 2011 Apr;40(2):115-123. [doi: 10.1026/1616-3443/a000088]
73. Neff KD. The development and validation of a scale to measure self-compassion. Self Identity 2003 Jul;2(3):223-250. [doi: 10.1080/15298860309027]
74. Garnefski N, Kraaij V. Cognitive emotion regulation questionnaire – development of a short 18-item version (CERQ-short). Pers Individ Dif 2006 Oct;41(6):1045-1053. [doi: 10.1016/j.paid.2006.04.010]
75. Garnefski N, Kraaij V, Spinhoven P. CERQ - Manual for the use of the Cognitive Emotion Regulation Questionnaire. Leiderdorp, The Netherlands: DATEC; 2002.
76. Martins EC, Freire M, Ferreira-Santos F. Examination of adaptive and maladaptive cognitive emotion regulation strategies as transdiagnostic processes: associations with diverse psychological symptoms in college students. Stud Psychol 2016;58(1):59-73.
77. Horvath AO, Greenberg LS. Development and validation of the Working Alliance Inventory. J Couns Psychol 1989;36(2):223-233. [doi: 10.1037/0022-0167.36.2.223]
78. Bijleveld CC, van der Kamp LJ, Mooijaart A, Van Der Van Der Kloot WA, Van Der Leeden R, Van Der Burg E. Longitudinal Data Analysis: Designs, Models and Methods. Thousand Oaks, CA, USA: Sage Publications; 1998.
79. Paetzold I, Schick A, Rauschenberg C, Hirjak D, Banaschewski T, Meyer-Lindenberg A, et al. A hybrid ecological momentary compassion-focused intervention for enhancing resilience in help-seeking youths: findings on baseline characteristics from the EMIcompass trial. Open Science Framework. 2021 Nov 18. URL: https://osf.io/yxadg [accessed 2022-09-20]
80. Xiao Z, Kasim A, Higgins S. Same difference? Understanding variation in the estimation of effect sizes from educational trials. Int J Educ Res 2016;77:1-14.
81. Hoffman L, Stawski RS. Persons as contexts: evaluating between-person and within-person effects in longitudinal analysis. Res Hum Dev 2009;6(2-3):97-120. [doi: 10.1080/15427600902911189]
82. Puhani PA. The treatment effect, the cross difference, and the interaction term in nonlinear “difference-in-differences” models. Econ Lett 2012 Apr;115(1):85-87 [FREE Full text] [doi: 10.1016/j.econlet.2011.11.025]
83. Parker G, Roy K. Adolescent depression: a review. Aust N Z J Psychiatry 2001 Oct;35(5):572-580. [doi: 10.1080/0004867010060504] [Medline: 11551271]
84. Afifi M. Gender differences in mental health. Singapore Med J 2007 May;48(5):385-391 [FREE Full text] [Medline: 17453094]
85. Morgan C, Kirkbride J, Leff J, Craig T, Hutchinson G, McKenzie K, et al. Parental separation, loss and psychosis in different ethnic groups: a case-control study. Psychol Med 2007 Apr;37(4):495-503. [doi: 10.1017/S0033291706009330] [Medline: 17094816]
86. Bhopal R. Is research into ethnicity and health racist, unsound, or important science? BMJ 1997 Jun 14;314(7096):1751-1756 [FREE Full text] [doi: 10.1136/bmj.314.7096.1751] [Medline: 9202509]
87. Paetzold I, Hermans KS, Schick A, Nelson B, Velthorst E, Schirmbeck F, EU-GEI High Risk Study, et al. Momentary manifestations of negative symptoms as predictors of clinical outcomes in people at high risk for psychosis: experience sampling study. JMIR Ment Health 2021 Nov 19;8(11):e30309 [FREE Full text] [doi: 10.2196/30309] [Medline: 34807831]

88. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4:1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]
89. Carlbring P, Andersson G, Cuijpers P, Riper H, Hedman-Lagerlöf E. Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: an updated systematic review and meta-analysis. Cogn Behav Ther 2018 Jan;47(1):1-18. [doi: 10.1080/16506073.2017.1401115] [Medline: 29215315]
90. Telford C, McCarthy-Jones S, Corcoran R, Rowse G. Experience Sampling Methodology studies of depression: the state of the art. Psychol Med 2012 Jun;42(6):1119-1129. [doi: 10.1017/S0033291711002200] [Medline: 22008511]
91. Kramer I, Simons CJ, Hartmann JA, Menne-Lothmann C, Viechtbauer W, Peeters F, et al. A therapeutic application of the experience sampling method in the treatment of depression: a randomized controlled trial. World Psychiatry 2014 Feb 04;13(1):68-77. [doi: 10.1002/wps.20090] [Medline: 24497255]
92. O'Keeffe J, Conway R, McGuire B. A systematic review examining factors predicting favourable outcome in cognitive behavioural interventions for psychosis. Schizophr Res 2017 May;183:22-30. [doi: 10.1016/j.schres.2016.11.021] [Medline: 27889383]
93. Hamilton KE, Dobson KS. Cognitive therapy of depression: pretreatment patient predictors of outcome. Clin Psychol Rev 2002 Jul;22(6):875-893. [doi: 10.1016/s0272-7358(02)00106-x] [Medline: 12214329]
94. Lincoln TM, Rief W, Westermann S, Ziegler M, Kesting ML, Heibach E, et al. Who stays, who benefits? Predicting dropout and change in cognitive behaviour therapy for psychosis. Psychiatry Res 2014 May 15;216(2):198-205. [doi: 10.1016/j.psychres.2014.02.012] [Medline: 24602992]
95. Marsh IC, Chan SW, MacBeth A. Self-compassion and psychological distress in adolescents—a meta-analysis. Mindfulness 2018;9(4):1011-1027. [doi: 10.1007/s12671-017-0850-7] [Medline: 30100930]
96. Berry N, Bucci S, Lobban F. Use of the internet and mobile phones for self-management of severe mental health problems: qualitative study of staff views. JMIR mental health 2017 Nov 1;4(4):e52. [doi: 10.2196/mental.8311] [Medline: 29092809]
97. McCartney G, Popham F, McMaster R, Cumbers A. Defining health and health inequalities. Public Health 2019 Jul;172:22-30 [FREE Full text] [doi: 10.1016/j.puhe.2019.03.023] [Medline: 31154234]
98. Friis-Healy EA, Nagy GA, Kollins SH. It is time to REACT: opportunities for digital mental health apps to reduce mental health disparities in racially and ethnically minoritized groups. JMIR mental health. 2021 2021 Jan 26;8(1):e25456. [doi: 10.2196/25456] [Medline: 33406050]
99. Cruz-Jesus F, Vicente M, Bacao F, Oliveira T. The education-related digital divide: an analysis for the EU-28. Comput Human Behav 2016 Mar;56:72-82. [doi: 10.1016/j.chb.2015.11.027]
100. Haight M, Quan-Haase A, Corbett B. Revisiting the digital divide in Canada: the impact of demographic factors on access to the Internet, level of online activity, and social networking site usage. Inf Commun Soc 2014;17(4):503-519. [doi: 10.1080/1369118X.2014.891633]
101. Bailey SC, O'Conor R, Bojarski EA, Mullen R, Patzer RE, Vicencio D, et al. Literacy disparities in patient access and health-related use of Internet and mobile technologies. Health Expect 2015 Dec;18(6):3079-3087 [FREE Full text] [doi: 10.1111/hex.12294] [Medline: 25363660]
102. Robotham D, Satkunanathan S, Doughty L, Wykes T. Do we still have a digital divide in mental health? A five-year survey follow-up. J Med Internet Res 2016 Nov 22;18(11):e309. [doi: 10.2196/jmir.6511] [Medline: 27876684]
103. Rathod S, Phiri P, Harris S, Underwood C, Thagadur M, Padmanabi U. Cognitive behaviour therapy for psychosis can be adapted for minority ethnic groups: a randomised controlled trial. Schizophr Res. (2-3) 2013 Feb;143(2-3):319-326. [doi: 10.1016/j.schres.2012.11.007] [Medline: 23231878]
104. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. BMJ 2015 Mar 19;350:h1258. [doi: 10.1136/bmj.h1258] [Medline: 25791983]
105. Wong G, Westhorp G, Manzano A, Greenhalgh J, Jagosh J, Greenhalgh T. RAMESES II reporting standards for realist evaluations. BMC Med 2016 Jun 24;14(1):96 [FREE Full text] [doi: 10.1186/s12916-016-0643-1] [Medline: 27342217]
106. Schueller SM, Hunter JF, Figueroa C, Aguilera A. Use of digital mental health for marginalized and underserved populations. Curr Treat Options Psych 2019 Jul 5;6(3):243-255. [doi: 10.1007/s40501-019-00181-z]

Abbreviations

CFI: compassion-focused intervention
CHARMS: Clinical High At-Risk Mental State
EMA: ecological momentary assessment
EMI: ecological momentary intervention
ESM: experience sampling methodology
HiTOP: Hierarchical Taxonomy of Psychopathology
mHealth: mobile health
RCT: randomized controlled trial
TAU: treatment as usual

Edited by A Mavragani; submitted 13.05.22; peer-reviewed by S Mukherjee, K Abela , J Huber; comments to author 27.06.22; revised version received 19.07.22; accepted 09.08.22; published 04.11.22.

Please cite as:
Paetzold I, Schick A, Rauschenberg C, Hirjak D, Banaschewski T, Meyer-Lindenberg A, Butz S, Floesser C, Schueltke L, Boehnke JR, Boecking B, Reininghaus U
A Hybrid Ecological Momentary Compassion–Focused Intervention for Enhancing Resilience in Help-Seeking Young People: Prospective Study of Baseline Characteristics in the EMIcompass Trial
JMIR Form Res 2022;6(11):e39511
URL: https://formative.jmir.org/2022/11/e39511
doi:10.2196/39511
PMID:36331526

Original Paper

The Receptivity to Safety-Related Mobile Apps Among Commercial Fishing Captains: Descriptive Exploratory Study

Maria T Bulzacchelli[1], PhD; Jenna M Bellantoni[1], MHS; Leigh McCue[2], MSE, PhD; Jerry Dzugan[3], MSED

[1]Krieger School of Arts and Sciences, Johns Hopkins University, Baltimore, MD, United States

[2]Department of Mechanical Engineering, George Mason University, Fairfax, VA, United States

[3]Alaska Marine Safety Education Association, Sitka, AK, United States

Corresponding Author:
Maria T Bulzacchelli, PhD
Krieger School of Arts and Sciences
Johns Hopkins University
3505 North Charles Street
Room 202
Baltimore, MD, 21218
United States
Phone: 1 410 516 8340
Email: mbulzac1@jhu.edu

Abstract

Background: Mobile apps addressing a variety of workplace safety issues have proliferated over the last decade as mobile technology has advanced and smartphone ownership has increased. Workplace safety interventions are often designed for a specific work site. However, some of the most dangerous jobs are ones in which workers frequently change field locations, such as commercial fishing. Mobile apps may be particularly suitable for delivering safety interventions to these workers.

Objective: We sought to gauge the potential for using mobile apps to deliver safety interventions to commercial fishing workers. The purpose of this paper is to describe how fishermen use their mobile devices during fishing operations and identify any mobile apps they already use for safety.

Methods: Participants comprised commercial fishing captains who already owned an iOS or Android smartphone or tablet. They completed a questionnaire that asked about their current mobile device use and their use of safety-related mobile apps, in addition to questions about their fishing operations. We performed descriptive analyses of the data.

Results: A total of 61 participants completed the questionnaire. The most common types of mobile devices participants reported owning were iPhones (n=36, 59%) and Android phones (n=24, 39%). Most participants (n=53, 87%) reported using their mobile device for both work and personal purposes, including while out at sea (n=52, 85%). Over half of the participants reported that they had either safety-related apps (n=17, 28%) or apps that help them with their work (n=35, 57%). The types of apps most frequently mentioned were apps for weather, wind, tides, and navigation.

Conclusions: The results of this study indicate that some commercial fishing captains who own a mobile device are receptive to using safety-related apps for work. Apps that help avoid hazards by monitoring environmental conditions and apps optimized for use on smartphones may be most likely to be adopted and used. Overall, these results suggest that mobile apps are a promising avenue for improving safety among workers in commercial fishing and similar occupations.

(JMIR Form Res 2022;6(11):e33638) doi:10.2196/33638

KEYWORDS

mobile app; mobile device; mobile phone; smartphone; safety; workplace safety; occupational safety; mobile health; mHealth; commercial fishing; cross-sectional study

Introduction

In the past 15 years, smartphone ownership has grown dramatically. According to the Pew Research Center, 85% of adults in the United States owned a smartphone in 2021, compared to 70% in 2016 and 35% in 2011 [1]. With the rapid growth of mobile technology and the resulting ubiquity of smartphones, health and safety interventions delivered through

mobile apps have proliferated [2]. There are now tens of thousands of mobile health (mHealth) apps providing reference information, training and education, monitoring, clinical score calculators, reminders, and communication functions [3,4].

One area with potential for greatly expanding the reach of mHealth interventions is workplace safety. A total of 91% of full-time workers use a smartphone at work [5]. Technology that integrates interventions into a small device that almost all workers already carry with them opens up numerous possibilities for promoting safety and preventing injuries. Industries and occupations in which workers frequently change field locations, such as construction, tree care, commercial fishing, as well as certain installation, maintenance, and repair occupations, may be especially promising areas for mHealth interventions that can be accessed wherever workers are. The National Institute for Occupational Safety and Health has developed mobile apps to address a variety of occupational hazards including a ladder safety app, a heat safety tool, a sound level meter, a personal protective equipment (PPE) inventory tracker, and multiple apps addressing aspects of ergonomics [6]. Dozens of other occupational safety and health apps are available on the app market as well [7].

Little is known about how workers use occupational safety apps or even how they use their mobile devices in the workplace. Studies that have examined workers' use of mobile devices or apps relevant to various work settings have produced mixed findings. In a scoping review that looked at nurses' use of personal smartphones in clinical settings, de Jong et al [8] found that nurses frequently used their personal smartphones to look up medical information and communicate with others on the patient-care team. In a mixed methods study, Patel et al [9] gave 16 physicians access to a mobile clinical decision support tool; 9 of the 16 doctors used the app during the study period while 7 did not. They found that 5 of the doctors accounted for 90% of the interactions with the app, and just 2 accounted for 56% of the interactions with the app over the study period. Snipes et al [10] pilot tested a mobile app to promote PPE use among farmworkers. Some participants reported technical problems with the phone such as low battery or the screen freezing, but 73% encountered no barriers to using the phone. For consistency, participants were given a particular type of phone to use for the pilot testing and they were also given specific types of PPE, but, notably, 75% of participants already had a personal mobile phone at baseline and 39% of those workers used it always or sometimes. Studies of apps that monitor UV conditions and provide tailored sun protection advice and warnings have found apps to be effective in improving some sun safety behaviors among participants from the general population, but not all participants used the apps [11,12]. These studies suggest that many workers already use their smartphones for work purposes, but even when asked to try apps designed for use in specific work settings, only a portion of workers may use the apps.

Commercial fishing consistently ranks as one of the most hazardous occupations in the United States. From 2000 to 2015, fishers and related fishing workers had a fatal injury rate of 117 per 100,000 full-time equivalent workers, approximately 29 times the all-worker rate [13]. Mobile apps should be explored as low- or no-cost interventions to protect this worker population from injury and death. A first step is determining how receptive commercial fishing workers would be to such apps. Technology adoption theories suggest that experienced mobile device users are more likely to adopt new mobile apps than individuals with little experience using mobile devices [14]. Little is known about how commercial fishing workers use mobile devices, especially when out at sea. The purpose of this paper is to describe how fishermen use their mobile devices during fishing operations and identify any mobile apps they already use for safety.

Methods

Study Population

Participants were a convenience sample of commercial fishing vessel captains recruited for a study field-testing 2 mobile apps that aim to address certain safety issues related to fishing operations [15]. To be eligible for our study, participants had to be commercial fishing captains who already owned an iOS or Android smartphone or tablet. Participant recruitment and data collection began in July 2018 and were completed in October 2020. A detailed description of our recruitment and data collection procedures is published elsewhere [15].

Data Collection and Key Variables

Participants completed a baseline questionnaire that was either administered by a researcher in person or self-administered online. The questionnaire asked about participants' commercial fishing operations and experience, some of their safety practices and concerns about safety, their current mobile device use, and their use of safety-related mobile apps (Textbox 1). Prior to data collection, a member of the study team with extensive commercial fishing experience (JD) reviewed all questionnaire items to ensure their suitability for this population. The items about mobile device use and safety-related app use were exploratory in nature and left open to each participant's interpretation.

Textbox 1. Questionnaire items about participants' mobile device and app use.

What kinds of mobile devices do you have, specifically?

- Do you have an iPhone? (Yes/no)
- Do you have an iPad? (Yes/no)
- Do you have an iPod Touch? (Yes/no)
- Do you have an Android phone? (Yes/no)
- Do you have an Android tablet? (Yes/no)

Thinking about your [device]… Do you use your [device] for work, personal use, or both? (Work only/personal use only/both work and personal use)

Do you ever use your [device] when you are out at sea? (Yes/no)

On your [device], do you have any apps that help you with your commercial fishing work? (Yes/no; if yes: which ones?)

On your [device], do you have any safety apps or apps that you use for the purpose of making you safer, either at work or on your own time? (Yes/no; if yes: which ones?)

Data Analysis

We performed descriptive analyses of key variables, computing mean values for numeric variables and frequencies for categorical variables. We excluded 2 participants from the calculations of mean and median years of experience because both of these participants gave implausible values, reporting greater years of experience working as a commercial fishing captain than years of experience working on commercial fishing vessels generally. We summarized the free-text responses into broad categories by manually performing a content analysis.

Ethics Approval

This study was approved by the Johns Hopkins University Homewood Institutional Review Board (PR00015355 HIRB00011417).

Results

A total of 61 commercial fishing captains participated in our study and completed the baseline questionnaire. The mean age of our sample was 47.3 (SD 14.4) years, with a median of 47 (IQR 34-56) years. Participants had an average of 27.1 years of experience working on commercial fishing vessels and an average of 19.4 years of experience as a commercial fishing captain (Table 1). A total of 57 (93%) participants identified as male.

Per the study's eligibility criteria, all participants owned at least one mobile device. When asked what kinds of mobile devices they owned, the largest proportion reported owning an iPhone (n=36, 59%), followed by an Android phone, iPad, Android tablet, and iPod Touch (Table 2). Most participants (n=53, 87%) reported using their mobile device for both work and personal use. A similar proportion indicated that they take at least one of their devices with them out at sea. Just over half of the participants (n=35, 57%) indicated they have apps that help them with commercial fishing work, and 28% (n=17) indicated they have safety-related apps. Participants who reported having apps for commercial fishing work or safety most frequently named apps for weather, wind, tides, and navigation, including apps that combine these functions.

Table 1. Participants' age and years of experience (N=61).

Characteristic	Mean (SD)	Median (IQR)
Age	47.3 (14.4)	47 (34-56)
Years of experience working on commercial fishing vessels[a]	27.1 (13.3)	30 (15-37.5)
Years of experience working as a commercial fishing vessel captain[a]	19.4 (13.9)	18 (6.5-29.5)

[a]Two participants were excluded: n=59.

Table 2. Mobile device use among participants (N=61).

Variable	Participants, n (%)
Mobile device ownership[a]	
iPhone	36 (59)
Android phone	24 (39)
iPad	14 (23)
Android tablet	2 (3)
iPod Touch	2 (3)
Mobile device use	
Use for both work and personal use	53 (87)
Use when out at sea	52 (85)
Mobile app use	
Use apps for commercial fishing work	35 (57)
Use safety apps	17 (28)

[a]Categories are not mutually exclusive; participants could select all that apply.

Discussion

Principal Results and Implications

This study demonstrates that approximately 85% of commercial fishing captains who own an iOS or Android device use it for both work and personal use, including while at sea. Additionally, most captains already use mobile apps for commercial fishing work or safety purposes, most commonly apps for weather, wind, tides, and navigation. All but 1 participant reported owning a smartphone, whereas only about one-quarter reported owning a tablet computer.

These results suggest that commercial fishing captains may be receptive to additional safety-related mobile apps, especially apps that provide real-time information about vessel location and environmental conditions. Apps that monitor changing conditions and provide actionable information to help avoid hazards may be particularly useful to this population. Expanding some popular existing apps to incorporate additional safety functions could be a relatively easy way to increase the availability and uptake of mobile safety interventions. In addition, mobile apps aimed at protecting workers in commercial fishing and similar occupations may be more readily adopted if they are optimized for use on smartphones.

Limitations

An important limitation of this study is that our sample only included captains who already owned a smartphone or tablet and, therefore, are probably more inclined to adopt mobile technologies than those who do not already have a mobile device. On the one hand, this limitation may lead us to overestimate the potential reach of mobile interventions for commercial fishing safety. On the other hand, our findings are useful for informing the development of mobile apps targeting the specific population most likely to adopt and benefit from them.

While we do not know what proportion of all commercial fishing captains own a smartphone or tablet, our sample appears to be representative of this worker population in at least one important way. The median age of participants in this study (47 years) is very similar to the median age of workers in the agriculture, forestry, fishing, and hunting industry sectors nationally (47.8 years in 2019) [16]. This is relevant for assessing this population's receptivity to mobile interventions because workers in this industry tend to be older than workers in other industries, and smartphone ownership is related to age. The median age of all employed persons in the United States was 42.3 years in 2019 [16]. In 2021, smartphone ownership in the country was 95% among those aged 30 to 49 years but only 83% among those who are 50 to 64 years old [1].

Another limitation of this study is that we asked participants if they have safety apps on their mobile device, but we did not ask participants how frequently they use the apps mentioned and why. It is possible to have apps on a device that are rarely used, which would reduce the amount of protection afforded by the apps. Additional research is needed to determine factors that influence the actual use of safety apps among this worker population and any barriers to using the apps they might experience.

Conclusion

The results of this study indicate that some commercial fishing captains who own a mobile device are receptive to using safety-related apps for work. Apps that help avoid hazards by monitoring environmental conditions may be especially appealing to this population, and apps that are optimized for use on smartphones may be most likely to be adopted and used. Overall, these results suggest that mobile apps are a promising avenue for improving safety among commercial fishing workers.

Acknowledgments

Funding for this study was provided by the National Institute for Occupational Safety and Health (NIOSH) (2U54OH007542) through the Northeast Center for Occupational Health and Safety: Agriculture, Forestry and Fishing. This paper is solely the responsibility of the authors and does not necessarily represent the official views of the NIOSH.

Data Availability

The data sets generated and analyzed in this study are not publicly available due to participant confidentiality.

Conflicts of Interest

None declared.

References

1. Demographics of mobile device ownership and adoption in the United States. Pew Research Center. 2021 Apr 07. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2021-06-10]
2. Mosa ASM, Yoo I, Sheets L. A systematic review of healthcare applications for smartphones. BMC Med Inform Decis Mak 2012 Jul 10;12:67 [FREE Full text] [doi: 10.1186/1472-6947-12-67] [Medline: 22781312]
3. Silva BM, Rodrigues JJ, de la Torre Díez I, López-Coronado M, Saleem K. Mobile-health: a review of current state in 2015. J Biomed Inform 2015 Aug;56:265-272 [FREE Full text] [doi: 10.1016/j.jbi.2015.06.003] [Medline: 26071682]
4. Xu W, Liu Y. mHealthApps: a repository and database of mobile health apps. JMIR Mhealth Uhealth 2015 Mar 18;3(1):e28 [FREE Full text] [doi: 10.2196/mhealth.4026] [Medline: 25786060]
5. Smith A. U.S. smartphone use in 2015. Pew Research Center. 2015 Apr 01. URL: https://www.pewresearch.org/internet/2015/04/01/us-smartphone-use-in-2015/ [accessed 2021-01-23]
6. NIOSH mobile applications. Centers for Disease Control and Prevention. 2017. URL: https://www.cdc.gov/niosh/pubs/apps/default.html [accessed 2021-01-23]
7. Mobile apps. Safety+Health. 2021. URL: https://www.safetyandhealthmagazine.com/topics/924-mobile-apps [accessed 2022-03-14]
8. de Jong A, Donelle L, Kerr M. Nurses' use of personal smartphone technology in the workplace: scoping review. JMIR Mhealth Uhealth 2020 Nov 26;8(11):e18774 [FREE Full text] [doi: 10.2196/18774] [Medline: 33242012]
9. Patel R, Green W, Shahzad MW, Larkin C. Use of mobile clinical decision support software by junior doctors at a UK teaching hospital: identification and evaluation of barriers to engagement. JMIR Mhealth Uhealth 2015 Aug 13;3(3):e80 [FREE Full text] [doi: 10.2196/mhealth.4388] [Medline: 26272411]
10. Snipes SA, Montiel-Ishino FA, Smyth JM, Murphy DJ, Miranda PY, Davis LA. User perceptions of ¡Protéjase!: an intervention designed to increase protective equipment use among Mexican immigrant and Mexican American farmworkers. JMIR Mhealth Uhealth 2016 Apr 11;4(2):e28 [FREE Full text] [doi: 10.2196/mhealth.4455] [Medline: 27066727]
11. Buller DB, Berwick M, Lantz K, Buller MK, Shane J, Kane I, et al. Smartphone mobile application delivering personalized, real-time sun protection advice: a randomized clinical trial. JAMA Dermatol 2015 May;151(5):497-504 [FREE Full text] [doi: 10.1001/jamadermatol.2014.3889] [Medline: 25629710]
12. Robinson JK, Patel S, Heo SY, Gray E, Lim J, Kwon K, et al. Real-time UV measurement with a sun protection system for warning young adults about sunburn: prospective cohort study. JMIR Mhealth Uhealth 2021 May 06;9(5):e25895 [FREE Full text] [doi: 10.2196/25895] [Medline: 33955844]
13. Commercial fishing safety: national overview. Centers for Disease Control and Prevention. 2019. URL: https://www.cdc.gov/niosh/topics/fishing/nationaloverview.html#ref1 [accessed 2021-04-25]
14. Venkatesh, Thong, Xu. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Quarterly 2012;36(1):157. [doi: 10.2307/41410412]
15. Bulzacchelli M, Bellantoni J, McCue-Weil L, Dzugan J. Field test of two mobile apps for commercial fishing safety: feedback from fishing vessel captains. Work (forthcoming) 2022.
16. Labor force statistics from the current population survey: demographics. US Bureau of Labor Statistics. 2021 Apr 14. URL: https://www.bls.gov/cps/demographics.htm [accessed 2021-06-10]

Abbreviations

mHealth: mobile health
PPE: personal protective equipment

Edited by T Leung; submitted 16.09.21; peer-reviewed by J Walsh, SM Ayyoubzadeh, S Rostam Niakan Kalhori ; comments to author 09.01.22; revised version received 18.03.22; accepted 18.08.22; published 08.11.22.

Please cite as:
Bulzacchelli MT, Bellantoni JM, McCue L, Dzugan J
The Receptivity to Safety-Related Mobile Apps Among Commercial Fishing Captains: Descriptive Exploratory Study
JMIR Form Res 2022;6(11):e33638
URL: https://formative.jmir.org/2022/11/e33638
doi:10.2196/33638
PMID:36346649

Original Paper

Development of an Electronic Screening and Brief Intervention to Address Perinatal Substance Use in Home Visiting: Qualitative User-Centered Approach

Sarah Dauber[1], PhD; Cori Hammond[1], MPH; Aaron Hogue[1], PhD; Craig Henderson[2], PhD; Jessica Nugent[3], MPH; Veronica Ford[4], PhD; Jill Brown[5], MA; Lenore Scott[5], MSW; Steven Ondersma[6], PhD

[1]Partnership to End Addiction, New York, NY, United States
[2]Sam Houston State University, Huntsville, TX, United States
[3]Burke Foundation, Princeton, NJ, United States
[4]Prevent Child Abuse New Jersey, New Brunswick, NJ, United States
[5]New Jersey Department of Children and Families, Trenton, NJ, United States
[6]Michigan State University, East Lansing, MI, United States

Corresponding Author:
Sarah Dauber, PhD
Partnership to End Addiction
711 Third Avenue
New York, NY, 10017
United States
Phone: 1 212 841 5270
Email: sdauber@toendaddiction.org

Abstract

Background: Perinatal substance use (SU) is prevalent during pregnancy and the postpartum period and may increase the risks to maternal and child health. Many pregnant and postpartum women do not seek treatment for SU because of fear of child removal. Home visiting (HV), a voluntary supportive program for high-risk families during the perinatal period, is a promising avenue for addressing unmet SU needs. Confidential delivery of screening and brief intervention (BI) for SU via computers has demonstrated high user satisfaction among pregnant and postpartum women as well as efficacy in reducing perinatal SU. This study describes the development of the electronic screening and BI for HV (e–SBI-HV), a digital screening and BI program that is adapted from an existing electronic screening and BI (e-SBI) for perinatal SU and tailored to the HV context.

Objective: This study aimed to describe the user-centered intervention development process that informed the adaptation of the original e-SBI into the e–SBI-HV, present specific themes extracted from the user-centered design process that directly informed the e–SBI-HV prototype and describe the e–SBI-HV prototype.

Methods: Adaptation of the original e-SBI into the e–SBI-HV followed a user-centered design process that included 2 phases of interviews with home visitors and clients. The first phase focused on adaptation and the second phase focused on refinement. Themes were extracted from the interviews using inductive coding methods and systematically used to inform e–SBI-HV adaptations. Participants included 17 home visitors and 7 clients across 3 Healthy Families America programs in New Jersey.

Results: The e–SBI-HV is based on an existing e-SBI for perinatal SU that includes screening participants for SU followed by a brief motivational intervention. On the basis of the themes extracted from the user-centered design process, the original e-SBI was adapted to address population-specific motivating factors, address co-occurring problems, address concerns about confidentiality, acknowledge fear of child protective services, capitalize on the home visitor–client relationship, and provide information about SU treatment while acknowledging that many clients prefer not to access the formal treatment system. The full e–SBI-HV prototype included 2 digital intervention sessions and home visitor facilitation protocols.

Conclusions: This study describes a user-centered approach for adapting an existing e-SBI for SU for use in the HV context. Despite the described challenges, home visitors and clients generally reacted favorably to the e–SBI-HV, noting that it has the potential to fill a significant gap in HV services. If proven effective, the e–SBI-HV could provide a way for clients to receive help with SU within HV, while maintaining their privacy and avoiding the overburdening of home visitors. The next step in this study would be to test the feasibility and preliminary efficacy of the e–SBI-HV.

(JMIR Form Res 2022;6(11):e37865) doi:10.2196/37865

KEYWORDS

pregnant women; postpartum women; home visiting; substance use; computerized intervention; mobile health; mobile phone

Introduction

Background and Rationale

Substance use (SU) in the perinatal period is a critical public health challenge that is associated with negative birth outcomes, poor maternal and child health, and increased risk for child welfare system involvement [1-3]. Recent national data indicate that 5.4% of pregnant women reported using illicit drugs, 9.4% reported alcohol use, and 2.3% reported binge drinking [4-6]. Although many women who used substances before pregnancy decrease their use during pregnancy, 25% to 50% relapse in the first 3 months post partum [7,8]. The risk of relapse increases because of hormonal changes, the stress of caring for a newborn, sleep deprivation, and social isolation, making the early postpartum weeks a critical period for intervention [8,9]. Despite the existence of effective treatments for substance misuse, <10% of women who need SU treatment receive it, a gap that is highest among low-income and underrepresented minorities [10]. Pregnant and postpartum women experience many barriers to accessing SU treatment and many conceal their SU and do not seek treatment because of fear of child removal [11].

Home visiting (HV), a strategy for delivering voluntary preventive services aimed at optimizing parent and child outcomes across the life course, is the primary supportive intervention offered to at-risk families during the perinatal period in the United States [12,13]. HV represents a promising avenue for addressing unmet SU needs in pregnant and postpartum women, as evidence-based HV programs currently operate in all 50 states and serve the nation's highest risk families [14]. HV is typically offered to families who are identified as having specific risk factors associated with child maltreatment, such as inadequate income, unstable housing, history of SU, no prenatal care, or history of mental health concerns [15]. SU is prevalent among women served by HV programs, with nearly 40% of HV clients reporting binge drinking or using other drugs in the 3 months before HV enrollment nationally [16]. The immediate postpartum period is a particularly important time for preventing SU relapse [7,8] and HV often represents vulnerable families' only contact with the formal service system during this time. Finally, new mothers may be especially motivated to change behaviors that may negatively impact their baby, such as SU [17-19]. As a voluntary, strengths-based program, HV provides a natural framework for capitalizing on this motivation to change.

Despite this promise, the most widely implemented HV models, such as Healthy Families America (HFA) and Parents as Teachers, do not have systematic protocols for identifying and addressing SU. This leaves many HV clients with undetected and unmet SU needs that increase maternal and child risk. Screening, brief intervention, and referral to treatment (SBIRT), originally designed to reduce gaps in the service continuum from primary care to SU treatment [20,21], has been widely recommended as a public health model for addressing SU [22-24]. Recommendations for SBIRT in the perinatal period include universal SU screening followed by brief intervention (BI) for women at low-to-moderate risk (defined by high past use or low current use) and referral to specialty SU treatment for those screening as being at the highest risk for SU [25]. This approach may be a good fit for the HV context given its nonjudgmental nature, the prevalence of low to moderate SU risk among HV clients, and the small proportion of HV clients who actively seek SU treatment [26,27].

However, 2 key challenges may preclude the successful integration of traditional SBIRT procedures into HV. First, HV clients are often reluctant to disclose SU to professionals, including home visitors, because of shame, stigma, denial, and most notably, fear of child removal [28-31]. These fears are not always unfounded, as maternal SU often triggers involvement of the child protective system (CPS), and many states have laws mandating reports to CPS if SU is discovered [32,33]. Second, most home visitors are lay professionals who lack the advanced clinical training and skills needed to effectively deliver evidence-based BIs [34]. Home visitors have repeatedly demonstrated low rates of risk identification and referral to treatment for SU [35-38], have reported feeling ill-equipped to effectively address client SU and other behavioral health concerns, and desire more training and supervision related to addressing client behavioral health [39-42]. Given the combination of client reluctance to disclose SU and lay professional home visitors with minimal clinical training and skill, a traditional SBIRT model that requires disclosure of current SU followed by delivery of an evidence-based BI is unlikely to be successful in HV. Our own prior work supports this contention: training home visitors in the implementation of SU screening and brief motivational interventions yielded very few positive screens and low rates of implementation and referral to treatment [43].

Screening and BI (SBI) that is delivered digitally via a computer or smartphone has great potential to overcome both of these challenges and bolster HV capacity to address maternal SU. Digital screening can be conducted without home visitor involvement, preserving the confidentiality of clients' responses and allowing for greater comfort and honesty [44,45]. Screening via computer has been shown to promote honest responses among women in health care settings [46-48] and may be preferred by clients with concerns about confidentiality [49]. Computer-delivered BI has the advantage of greater standardization of delivery and alleviates the need for home visitors to implement complex BI techniques that are incompatible with their training and skill level [50]. Systematic reviews of studies of non–treatment-seeking adults have found positive impacts of both brief and long-term digital interventions on SU outcomes, with moderate effect sizes for self-report and biological data when compared with controls [50,51]. Digital SBI has demonstrated significant reductions in alcohol and drug use among pregnant and postpartum women when delivered in

health care settings [52,53] and Women, Infants, and Children Supplemental Nutrition Program offices [54], with outcomes comparable with provider-delivered interventions [52]. Moreover, client satisfaction ratings for computerized SU interventions are high across multiple client types, including pregnant and postpartum women [50,55-57]. The electronic delivery of SBI is also highly compatible with the virtual delivery of HV services during the COVID-19 pandemic [58].

Early evidence from health care–based trials with postpartum women has supported the efficacy of electronic delivery of SBIRT in this population. In 2 randomized controlled trials, women reporting marijuana or other drug use before pregnancy were recruited after delivery and randomly assigned to receive electronic SBI (e-SBI) or an attention control. Women who received the e-SBI significantly reduced their SU frequency [59] and had significantly higher abstinence rates at 3-month follow-up based on self-report and biological measures, with a moderate effect size [53]. Intervention effects were maintained at 6 months at a moderate effect size but were no longer significant. A study comparing the e-SBI to clinician-delivered SBIRT in obstetrics and gynecology clinics found significant declines in SU in both groups compared with controls, with no difference between computerized and clinician-delivered SBIRT, although impacts attenuated by 6 months [52]. Only 1 other study has tested this approach in HV [60] and found that a series of 8 e-SBI sessions targeting multiple child maltreatment risk factors including SU was feasible to deliver within HV. Participant satisfaction ratings were high; however, no impacts on SU or other child maltreatment risk factors were detected. Two key limitations of this study that may have contributed to the results are (1) failure to adequately consider the perspectives of end users in the development of the e-SBI sessions and (2) a lack of a structured implementation process for home visitors. In this study, we applied user-centered design methods to adapt the e-SBI described earlier into the e-SBI for HV (e–SBI-HV). The e–SBI-HV is tailored to the unique HV context and is designed to overcome challenges related to SU disclosure and home visitor delivery of evidence-based BI.

Objectives

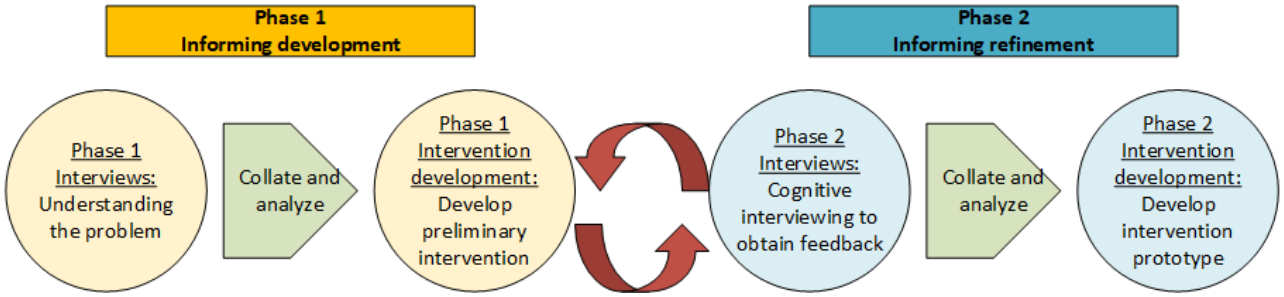
The overall goal of this study was to develop a full package of tools for implementing digital SBI for SU in HV (e–SBI-HV), which includes a digital intervention that is adapted from the existing e-SBI for perinatal SU to be fully tailored to the unique HV context, and facilitation protocols for home visitors to support the successful integration of the digital intervention into routine HV services. We used a user-centered design process that involved iterative cycles of qualitative data collection and intervention design to align the elements of the intervention with themes extracted from the qualitative data. This approach has been used in other studies [61] and yields a final product that is reflective of users' experiences and therefore more relevant and acceptable to users. The following are the specific objectives of this paper: (1) to describe the 2-phase user-centered intervention development process that informed the adaptation of the original e-SBI into the e–SBI-HV, (2) to present specific themes extracted from the user-centered design process that directly informed the e–SBI-HV prototype, and (3) to describe the e–SBI-HV prototype.

Methods

Overview of the User-Centered Design Process

Our 2-phase user-centered design process for the adaptation of the e-SBI into the e–SBI-HV combined agile methods [62] and design thinking principles [63]. Agile methods allow for an iterative, nonlinear development process in which the intervention is quickly adapted in response to user feedback. Design thinking and user-centered design frameworks focus on deep empathic engagement with end users regarding their needs, goals, and preferences to inform intervention design [64]. The focus on empathy makes design thinking a particularly appropriate approach for intervention design in clinical contexts such as HV, as it prioritizes the needs of home visitors and clients and is more likely to result in a product that can be easily integrated into an existing system of care [63]. Combining these 2 approaches led to our 2-phase design process, which is depicted in Figure 1, and included 2 phases of data collection via structured interviews of home visitors and clients. In the first phase, we gathered information to inform the initial *development* of the e–SBI-HV, which included adaptations to the original e-SBI and the development of preliminary versions of home visitor facilitation protocols. In the second phase, we presented the initial version of all the e–SBI-HV components and gathered information to inform the *refinement* of the prototype. We extracted themes from the 2 rounds of interviews using inductive coding methods and used the themes systematically to inform e–SBI-HV adaptations.

Figure 1. Schematic of the user-centered intervention development process.



Ethics Approval

Ethics approval for the study was granted in January 2019 by the Solutions Institutional Review Board (approval number #2020/08/17), the institutional review board that is used by the Partnership to End Addiction. Informed consent was obtained from home visitors and clients before participation in the first interview or in the first focus group.

Participant Eligibility and Recruitment

The user-centered intervention development process took place from November 2018 to March 2020 in 3 counties in New Jersey that were implementing the HFA HV program and volunteered to participate in the study. Eligible participants included home visitors delivering HFA HV services at a participating county and their HFA clients who were aged ≥18 years and either pregnant or within 1 year post partum. Home visitors were recruited directly by the research team, and clients were recruited by their home visitors. The home visitors introduced the study and offered the option for participation to clients who met the study eligibility criteria.

Procedures

As shown in Figure 1, we conducted 2 phases of interviews with the home visitors and clients. The home visitors were interviewed in focus groups by site and clients were interviewed individually. We conducted 4 home visitor focus groups in phase 1 and 7 in phase 2. Phase 2 focus groups were smaller, to facilitate the cognitive interviewing process. We conducted 7 client interviews in phase 1 and 2 client interviews in phase 2. All home visitor focus groups were conducted in person, while 3 client interviews were conducted in person and 6 were conducted via phone. All interviews were guided by semistructured interview guides, and all interviews that were conducted in person were audio-recorded and transcribed. Interview guides were developed by the study investigative team based on published guidelines for assessing the feasibility and acceptability of new interventions in field settings [65-68]. The focus groups were conducted by 2 interviewers who were trained and supervised by the study principal investigator (PI). One moderator facilitated the group discussion and another was responsible for logistics, including audio recording, distributing incentives, and note-taking when necessary. Individual client interviews were conducted one-on-one by either the PI or a research assistant who was trained and supervised by the PI. Interviews conducted by phone could not be recorded; however, the interviewers took detailed notes. Each interview lasted for approximately 1 hour and participants were compensated US $25 in gift cards for each interview they completed.

Phase 1 interviews focused on preliminary information gathering to inform the development of the content and structure of the e–SBI-HV components, which included adapting the original e-SBI and developing the home visitor facilitation protocols. The topics covered with home visitors included current HV activities and challenges related to SU screening and intervention, comfort discussing SU with clients, openness to integrating technology into their usual practice, past experience working with SU clients, and potential facilitators of and barriers to e–SBI-HV implementation. Client interviews conducted in phase 1 helped gather information on the challenges faced by pregnant people and new mothers, the relationship with their home visitors, their feelings about discussing sensitive information such as SU with their home visitors**,** and their comfort level with technology.

In the phase 2 interviews, we presented the first draft of the digital intervention to home visitors and clients, engaging them in a process of cognitive interviewing. Cognitive interviewing is widely used in the development of measurement tools and interventions targeting system-level changes [69,70] and asks users of a system to *think aloud* as they test the system components, responding to tailored questions to assess comprehension, usability, meaning of responses, and the need for additional content. This process allowed us to obtain detailed feedback on all aspects of the digital intervention.

Data Analysis

All focus groups and interview recordings were transcribed verbatim by research assistants. Interview transcripts and notes were analyzed using thematic content analysis, a widely used approach in qualitative research that applies inductive coding to identify themes within the data [71,72]. All transcripts and notes were coded by 2 independent raters, and the final themes and subthemes were determined by consensus. Owing to the small number of client interviews with full transcripts, we identified themes based on the home visitor interviews and then reviewed the client interviews and noted any new themes that arose. Interview coding and analysis were conducted separately for phase 1 and phase 2 interviews (the findings are grouped together for ease of presentation). The goal of the coding process was to extract themes that could be incorporated into the intervention design. A similar analytical approach has been used in other studies in which the purpose of the interviews was to inform intervention development and refinement [61].

Description of the Original e-SBI

The original e-SBI was developed using the Computerized Intervention Authoring System (CIAS), an authoring tool for developing mobile health interventions. Interventions built using the CIAS are compatible with all mobile platforms and feature synchronous interactivity; natural language reflections; branching logic; a clean user interface; and the ability to easily incorporate images, graphics, figures, text, and videos. The program is fully automated and can be completed on a touch screen tablet, smartphone, or computer with headphones for privacy. A 3D cartoon character capable of a range of animated actions narrates the e-SBI, reads all content aloud so that no reading or typing is required, and reflects back participant responses. Digital interventions developed using the CIAS have been used with thousands of participants to date, many of whom had a low socioeconomic status and have consistently received extremely high user satisfaction ratings [53,57,73], including in a study of HV clients [60].

The original e-SBI is a single 20-minute session focused on alcohol and drug use, variations of which have been tested in pregnant and postpartum women in delivery hospitals and obstetrics and gynecology clinics [57,59,74,75]. The session begins with *screening* for SU using the Alcohol, Smoking, and

Substance Involvement Screening Test [76-78], modified to ask about SU in the 3 months before pregnancy [59], which has been shown to yield more accurate responses with greater sensitivity for identifying active substance users during pregnancy and post partum than asking about current use [31,79].

After screening, participants are branched to a *BI* that is tailored to pregnancy status (pregnant vs postpartum) and primary substance of concern. The BI applies motivational interviewing (MI) principles and the Feedback, Responsibility, Advice, Menu Options, Empathy, and Self-Efficacy BI framework [80]. The BI content is tailored to the participant's reported level of readiness to change and includes the following components: (1) *personalized feedback* on participant-reported negative consequences of SU, readiness to change, and how their SU compares with that of other women; (2) *pros and cons* of SU and behavior change; (3) *menu of options* for strategies that have helped other women change their SU behavior; and (4) *optional goal setting* regarding changing SU behavior for those who report a desire to make a change. In accordance with MI theory that suggests that MI strategies should be matched to participants' level of motivation to change [81], participants who express limited interest in change receive motivational enhancement interventions and those desiring change proceed directly to goal-setting interventions. Participants who report having already quit receive motivational content aimed at maintaining their successful change in SU. The BI concludes with a video testimonial from a mother describing her struggles with SU during pregnancy and her success in overcoming them.

Results

Sample Characteristics

The participants included 17 home visitors and 7 clients across the 3 sites. All home visitors participated in both phases of the interviews, and 2 clients participated in both phases. Across sites, home visitors (n=17) were female, with an average age of 33 (SD 10.3) years, Latina (14/17, 82%), White (4/17, 24%), Black or African American (2/17, 12%), and from other racial or ethnic backgrounds (1/17, 6%). Most home visitors (12/17, 70%) had a bachelor's degree or less, with 6% (1/17) having some postcollege education, and 24% (4/17) having a graduate degree. Home visitors had an average of 5 (SD 7.2) years of experience in HV. Clients (n=7) were female, having an average age of 26 (SD 5.15) years, Latina (3/7, 43%), Black or African American (4/7, 57%), and from other racial or ethnic backgrounds (2/7, 28%). None of the clients were pregnant at the time of enrollment. In total, 43% (3/7) of clients had a high school education, 29% (2/7) had some college education, and 29% (2/7) had graduated from college. Less than half (3/7, 43%) of the participants were employed. Clients had an average of 1.6 children aged <5 (SD 0.79) years and had been in the HFA program for an average of 7.5 (SD 6.68) months.

Interview Themes and e–SBI-HV Adaptations

On the basis of themes extracted from the 2 phases of focus groups and interviews, the e–SBI-HV was adapted to (1) address population-specific motivating factors; (2) address co-occurring problems; (3) address concerns about confidentiality; (4) acknowledge the fear of CPS involvement; (5) capitalize on the home visitor–client relationship, while avoiding interfering with the relationship; and (6) provide information about how to access treatment, while understanding that many clients prefer not to access the formal treatment system. The subsequent sections describe the feedback that informed specific adaptations to the e–SBI-HV in each of these 6 areas as well as the adaptations that were made in response to the feedback. We have also presented several recommendations from home visitors that could not be incorporated into this version of the e–SBI-HV but that will be incorporated into a future version. Representative quotes for each domain are included in the sections that follow as well as in Multimedia Appendix 1.

Address Population-Specific Motivating Factors

Home visitors described a variety of factors that motivate their clients to change their SU behaviors. The primary motivator for reducing SU was the health and safety of the baby. For example, 1 home visitor stated:

> *My client says, 'he saved my life. The catalyst to help me get clean was the pregnancy'.*

Fear of CPS involvement was also described as a strong motivator for change. The client interviews described shame and embarrassment as a primary reason for not seeking help with SU and mental health problems.

The adapted digital intervention includes an assessment of the factors that motivate participants to want to quit or cut down their SU. The adapted digital intervention provides users with a list of potential motivating factors and allows them to select those that are most meaningful to them. The health of the baby and the desire to have a healthy pregnancy is featured as a primary motivator, and much of the content of the BI attempts to capitalize on this motivation. To address shame and embarrassment, the digital intervention includes reflections on the options that participants select in response to questions about SU that attempt to normalize their feelings. For example, the narrator might say the following:

> *You said you sometimes used marijuana when you were pregnant because it helped relieve stress. Many other pregnant women feel the same way you do.*

Address Co-occurring Problems

Home visitors described clients who use substances as complex, often with co-occurring mental health conditions and other unmet basic needs. As 1 home visitor stated:

> *I guess I would say that a significant number of clients that have substance abuse issues have mental health issues with that.*

Home visitors and clients recommended that the digital intervention be expanded to include additional topics beyond SU that are often of concern to pregnant and postpartum women, including mental health and intimate partner violence, which are also difficult topics for clients to discuss with home visitors. On the basis of this feedback, we expanded the digital intervention to include a second session devoted to some of the issues that often co-occur with SU, including intimate partner violence and mental health. We also included smoking and

vaping as a topic in the second session, as it was not included in the first session on SU but is prevalent among HV clients according to home visitors.

Address Concerns About Confidentiality

Home visitors described clients' reluctance to disclose SU to their home visitor as the primary challenge to successfully addressing SU in HV. While some clients did disclose SU, home visitors reported that most were not comfortable answering questions about SU and that asking about SU could interfere with the home visitor–client relationship. Home visitors were not certain that clients would automatically trust the confidentiality of the e–SBI-HV or that they would be willing to enter information about SU into a web-based program. One home visitor explained:

> *During our enrollment, one of the things we have to say is that we are a mandated reporter. So I feel like once we say that that sort of sticks out of everything we've said during the enrollment. Then to actually trust that we don't know what their answers are... some might believe it, some might not.*

The home visitor facilitation protocol scripts reinforce the confidentiality of the program. The scripts emphasize that it is the client's choice whether to disclose SU to the home visitor, and the home visitor will not know what they enter into the digital intervention. The confidentiality of the program is also emphasized in the introductory sections of both digital intervention sessions, as well as at several other points throughout the sessions.

Acknowledge Fear of CPS Involvement

Home visitors reported that fear of being reported to CPS was the primary reason for clients being reluctant to disclose SU to home visitors. Home visitors described the fear of losing their children as paramount for their clients, leading them to conceal their SU and avoid seeking help. For example, 1 home visitor said the following:

> *I think a big concern is [CPS] involvement for a lot of families. Obviously if you have children or about to have children while battling addiction, there's always that factor there. If they relapse, I feel like a lot of them feel like they're being watched and told what to do.*

The home visitors made suggestions regarding the best ways to address clients' fears of CPS reporting, including describing what happens when a call is made, describing the support provided by CPS, and encouraging discussions with the home visitor. The digital intervention includes a section on CPS reporting that acknowledges the fear that many clients have and provides information about when home visitors are required to report SU to CPS and when they are not. The program encourages clients to talk with their home visitor about their concerns and provides examples of ways they might do so without directly disclosing their own SU.

Capitalize on the Home Visitor–Client Relationship While Avoiding Interfering With the Relationship

Home visitors emphasized that establishing a trusting relationship with the client is critical. They described their role as being a supportive, nonjudgmental listener for their clients. As 1 home visitor described:

> *To be there supporting them is very good for me and very good for them. I try to help them in everything I can.*

Once that trust is established, clients may be more willing to disclose SU. Similarly, clients described the home visitor as a trusted and nonjudgmental source of support and the person they can talk to about sensitive topics, including SU. However, they noted that it takes time for that relationship to develop and that not every client will achieve that level of comfort with their home visitor.

Home visitors raised the concern that the digital intervention may interfere with the home visitor–client relationship, as they would not know what information the client was entering into the program and would therefore not be able to take appropriate action if needed. Some home visitors were concerned that clients would enter information into the digital intervention that should prompt a CPS call, but if they did not know what was being entered, they would not be able to make the call. In addition, they were concerned that if clients reacted negatively to the program, they would blame the home visitor and leave the HV program. One home visitor stated:

> *You may lose that relationship with them if it seems like you're forcing.*

However, despite these concerns, most home visitors reacted favorably to the e–SBI-HV approach and believed that it would ultimately increase the clients' comfort level and might help open the door to a conversation about SU with clients, leading to a positive change.

The home visitor facilitation component of the e–SBI-HV was designed specifically to capitalize on the trusting relationship between the home visitor and client that is at the core of HV. This component includes scripted protocols and an accompanying training for home visitors to enable them to introduce each digital intervention session and debrief after the client's completion of each session. In addition, both digital intervention sessions highlight the ways in which home visitors can help clients with concerns around SU and encourage participants to use their home visitor as a resource. However, the program emphasizes that it is ultimately the client's choice whether to disclose SU to the home visitor.

Provide Information About Treatment While Acknowledging That Many Clients Prefer Not to Access the Formal Treatment System

Home visitors reported that many of their clients did not want to attend treatment, despite a clear need. Reasons for this, as reported by home visitors, typically included denial of the problem, fear of losing their children if they entered treatment, and having other needs that were more pressing to address, such

as food and housing insecurity. For example, 1 home visitor said the following:

> *It just depends on the client. Some of them are motivated and really want to change, but some of them are still in that denial stage and they don't want to seek out services or help.*

Accessing treatment is particularly challenging for undocumented families, whose fear of deportation often prevents them from seeking the necessary treatment. Home visitors also described a lack of available culturally-sensitive treatment options as well as stigma that is more pervasive in certain cultures, preventing mothers from seeking treatment.

The digital intervention sessions are based on MI principles, and the primary aim is to motivate participants to achieve the goals around SU that they set for themselves. The BI provides information about treatment as one option out of several that are reasonable approaches for participants to take to make progress on their goals. Participants who are interested in treatment can select the option to learn more about available treatment providers in their county. Those who are not interested in treatment are able to bypass that option in the digital intervention. Participants' choice is emphasized throughout the digital intervention sessions.

Feedback to Be Addressed in Future Iterations of the e–SBI-HV

Home visitors recommended several additional domains to be addressed in the future, including prevention-oriented content for clients who do not endorse SU and content directed at other family members who are using substances. Partner SU was emphasized as a particular concern for many HV clients. In addition, the need for a Spanish version of the program was highlighted. While the current version of the e–SBI-HV does not address all areas of need related to SU, home visitors agreed that the program would help fill an important gap in HV services, which currently does not include a standardized protocol for identifying and addressing SU.

Final e–SBI-HV Prototype

Overview

The final e–SBI-HV prototype incorporated the feedback into 2 digital intervention sessions and their accompanying home visitor facilitation protocols. Digital intervention sessions may be completed either during home visits or on clients' own devices in between home visits. Home visitor facilitation may be done during either virtual or in-person home visits.

Digital Intervention

The digital intervention, adapted from the original e-SBI, includes 2 sessions, each approximately 20 minutes in duration, with content tailored to pregnancy status (pregnant vs post partum). Session 1 is focused on alcohol and drug use and follows the basic structure of the original e-SBI, with the adaptations based on the feedback gathered in the user-centered design process. Most notably, HV is featured throughout the session as an important resource, with emphasis on the different ways in which the home visitor can be helpful to participants in addressing SU concerns. Concerns about CPS reporting are acknowledged, and participants are encouraged to ask their home visitor to explain the criteria for making a CPS call. Information about local and web-based SU treatment and support resources for their county of residence are also included in the session, so that participants can access that information without the involvement of their home visitor. Participants are also provided with the option to enter their email address and receive a list of local resources via email that they could easily access after completing the session. Confidentiality and participant choices are emphasized throughout the session.

Session 2 is structured similar to session 1 and focuses on behavioral health concerns that often co-occur with SU in pregnant and postpartum women: smoking and vaping, depression, and intimate partner violence. Participants can choose which of the 3 topics they are interested in learning more about, with the option at the end of each topic to explore the remaining 2 topics. For each of the 3 topics, the program is based on MI principles and contains the following components: (1) a short psychoeducational video, (2) a brief question about their own experience followed by a reflection, (3) a menu of options that have been helpful to other women (eg, talking to home visitors), (4) links to resources, (5) an opportunity to select something from the menu of options that they would like to try, and (6) reflection on choice. The session ends with a brief recap.

Home Visitor Facilitation

The home visitor facilitation protocols were designed in the spirit of the strengths-based perspective that underpins HV and aimed to leverage the trusting, supportive relationship that is key to effective HV [82,83]. The goal of the facilitation protocols is to support the successful integration of the digital intervention sessions within the HV context; they are not intended to be therapeutic for the client. The protocols are brief and scripted to facilitate delivery in the context of virtual HV services. The facilitation protocols do not require clients to disclose SU to the home visitor, although they will not be prevented from doing so and may do so if they wish. The protocols include an *Introduction* to be delivered in the home visit before each digital intervention session and a *Debriefing* to be delivered in the home visit following completion of each digital intervention session.

In the *Introduction,* the home visitors provide the clients with information on what to expect in the digital intervention session, reinforce the confidentiality of the session, and answer any questions the clients have. In the *Debriefing*, the home visitors ask the client if they would like to discuss any aspect of the session or share their reactions, while emphasizing the confidentiality of the session and that it is the client's decision whether to discuss SU with the home visitor.

Discussion

Principal Findings

This study describes a user-centered approach for adapting an existing e-SBI for SU for use in the HV context. Although e-SBI for SU has shown promise in health care settings such as prenatal care clinics and delivery hospitals [52,84,85], its potential for impact in social service settings such as HV is

understudied. The user-centered design process used in this study yielded a deeper understanding of the complexity of addressing SU in the HV context and directly informed the content and structure of the e–SBI-HV.

Although the interviews described here were intended solely to inform the development of the intervention, their apparent themes resonate with prior studies on pregnant and postpartum women who use substances [86]. The theme of children as a primary motivator for mothers to access SU treatment and reduce their use has been documented in other studies [86,87]. However, despite being highly motivated, the pervasive stigma around SU in pregnant people and mothers and fear of child removal often prevent mothers from accessing the necessary help for SU [86]. Distrust of formal systems of care has been documented among pregnant and postpartum women who use substances [87]. This distrust can be due to prior interactions in which they were stigmatized [88] and is most prevalent among women of color because of histories of racial discrimination within these systems. Despite the voluntary nature of HV and the trust built between home visitors and clients, fear of CPS reporting often prevents families from fully engaging in HV services [89], and this is particularly likely for mothers who use substances. Home visitors recognize this distrust and may respond by avoiding the topic of SU in an attempt to retain families in HV services. The goal of the e–SBI-HV is to enable mothers to obtain information and support for SU confidentially, capitalizing on their motivation to reduce SU without requiring disclosure to a professional.

Although home visitors noted challenges to this approach within HV, most indicated that the e–SBI-HV has the potential to fill a significant gap in HV services. A recent national survey of HV programs on service coordination activities for addressing maternal mental health, SU, and intimate partner violence found that the most commonly used approach for addressing SU in HV was offering a referral to treatment [90]. However, a recent study found that only 21% of referrals from HV resulted in the receipt of services, suggesting that a referral alone may be insufficient for many HV clients [91]. If proven effective, the intervention developed in this study could provide a way for clients to receive help with SU within the HV context while maintaining their privacy and without overburdening home visitors.

Strengths and Limitations

The strengths of the e–SBI-HV include the 2-session digital intervention, the focus on a broad range of substances, and the user-centered design approach to intervention development applied in the adaptation of the original e-SBI into the e–SBI-HV. Existing e-SBIs for pregnant and postpartum women have typically focused on a single substance [54] and have consisted of only a single brief session [52,57]. The e–SBI-HV includes modules covering a range of substances, allowing the program content to be tailored to the primary substance reported by the client, while acknowledging that many people use multiple substances at the same time. In addition, the e–SBI-HV includes 2 separate sessions, consistent with the larger SBI literature, indicating that more than one BI session may increase efficacy [92]. The addition of a second session to the e–SBI-HV also provides an opportunity to address other concerns that often co-occur with SU in pregnant and postpartum women, including mental health and intimate partner violence [18,93-95].

This study also has several limitations. The study was conducted in a single state in the context of a single HV model, limiting its generalizability to other states and HV models. The generalizability is further limited by the small sample of home visitors and, in particular, by the very small client sample. Moreover, none of the participating clients were pregnant at the time of the interview. Unfortunately, our ability to interview clients was curtailed by the pandemic. However, previous studies on pregnant and postpartum women suggest that many of the concerns raised in this study, such as fears of CPS involvement and concerns about confidentiality, would be shared by pregnant people [86,87]. The use of convenience sampling may have limited the sample to those more predisposed to be supportive of the e–SBI-HV. However, other studies surveying home visitors generally found that the need for tools to address SU in HV is high [96,97]. Despite the low representation of clients in the sample, the perspectives of home visitors on the e–SBI-HV are informative, as home visitors are the ultimate purveyors of the program to clients; thus, their buy-in and support of the tool are critical to its successful implementation. In addition, the data were collected before the COVID-19 pandemic, so some of the information may be outdated. Although HV shifted to providing all services virtually in March 2020, it has since shifted back to some extent. A national survey of HV programs conducted in July 2021 found that 83% of the programs surveyed had resumed in-person visits and that nearly half of all visits on average were being conducted in person [98]. This survey found that over 90% of the programs planned to offer both in-person and virtual visits going forward. Finally, the e-SBI approach for addressing SU is most appropriate for those at the lower end of the SU risk continuum [25]. Although most HV clients would fit into that category, for those in need of more than a BI or those with an SU disorder, the e–SBI-HV will likely be insufficient.

Conclusions and Future Directions

The home visitor interviews conducted in this study provided several suggestions for future refinement of the e–SBI-HV. First, the development of a culturally tailored version of the e–SBI-HV for Latinx HV clients whose preferred language is Spanish is of high priority. With 1 exception [54], existing research on e-SBI for perinatal SU has been limited to those who are able to complete the program in English. In 2021, nearly 30% of families served by evidence-based HV programs across the United States were Hispanic or Latinx and 15% indicated Spanish as their primary language [99], supporting the need for cultural tailoring. In addition, one of the primary barriers to e–SBI-HV implementation reported by home visitors in this study was client concern about who would have access to their SU information once it was entered into the digital intervention program, despite reassurances of confidentiality. Security of the information entered into mobile health intervention apps is a critical concern that needs to be addressed to facilitate the widespread use of the e–SBI. Future research should assess the specific security features needed to assure users of their privacy. Finally, the expansion of the e–SBI-HV to address SU in

partners and other family members is also a critical area of need. The engagement of fathers is a priority in HV research and practice [100] and enhancements to HV to address mental health in fathers have recently been developed [101].

The next step in this research is to test the feasibility and preliminary efficacy of the e–SBI-HV. If proven feasible and effective, the e–SBI-HV has the potential for widespread dissemination throughout HV networks to improve reach among perinatal women with unmet needs for help with SU, particularly those who are reluctant to engage in face-to-face services. Current HV practices for addressing SU vary widely by model and state and are mostly decided upon and implemented at the level of local implementing agencies [102]. The practices include varying approaches to screening and referral and range from the provision of education in the form of pamphlets and other materials that home visitors may review with clients to the use of MI techniques to encourage clients to attend SU treatment [102]. These practices are not standardized across models or across programs within a specific model, and home visitors vary in their level of skill and training regarding addressing SU. If successful, the e–SBI-HV could provide a standardized approach for addressing SU in HV. The program allows for the standardization of evidence-based components of the intervention while also enabling its tailoring to a specific local site by including links to local resources. The e–SBI-HV also provides an excellent fit with the current move toward precision HV [103,104], which aims to deliver HV models with fidelity while tailoring the program to individual families' needs [105,106]. Enhancements to HV services such as the e–SBI-HV will allow for better tailoring of HV services to meet families' needs in different areas that are not directly part of the HV curriculum and that home visitors may not be well-equipped to address on their own.

Acknowledgments

This study was funded by the National Institute on Drug Abuse (grant R34DA045831). The authors acknowledge their partners at the New Jersey Department of Children and Families, Department of Health, and Division of Family Development for their contributions to this study. They are grateful to the home visitors and clients who provided their valuable insights to help develop the electronic screening and brief intervention for home visiting.

The authors gratefully acknowledge Anne Duggan for her thoughtful insights that helped shape the study design and procedures.

Data Availability

Access to an updated, noncommercial, and open-source version of the Computerized Intervention Authoring System authoring tool (funded by grant EB028990 from the National Institute of Biomedical Imaging and Bioengineering) [107].

Conflicts of Interest

SO is part owner of a company that licenses the intervention authoring software that was used to develop the electronic screening and brief intervention for this study. As of January 2022, the company is no longer providing new licenses.

Multimedia Appendix 1
Electronic screening and brief intervention for home visiting feedback domains and representative quotes from home visitor interviews.
[DOCX File , 25 KB - formative_v6i11e37865_app1.docx]

References

1. Dubowitz H, Kim J, Black MM, Weisbart C, Semiatin J, Magder LS. Identifying children at high risk for a child maltreatment report. Child Abuse Negl 2011 Feb;35(2):96-104. [doi: 10.1016/j.chiabu.2010.09.003] [Medline: 21376396]
2. New Directions in Child Abuse and Neglect Research. Washington, DC: National Academies Press; Mar 25, 2014.
3. Alhusen JL, Lucea MB, Bullock L, Sharps P. Intimate partner violence, substance use, and adverse neonatal outcomes among urban women. J Pediatr 2013 Aug;163(2):471-476 [FREE Full text] [doi: 10.1016/j.jpeds.2013.01.036] [Medline: 23485028]
4. Results from the 2013 National Survey on Drug Use and Health: Summary of National Findings, NSDUH Series H-48, HHS Publication No. (SMA) 14-4863. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2014.
5. Martin CE, Longinaker N, Terplan M. Recent trends in treatment admissions for prescription opioid abuse during pregnancy. J Subst Abuse Treat 2015 Jan;48(1):37-42 [FREE Full text] [doi: 10.1016/j.jsat.2014.07.007] [Medline: 25151440]
6. Patrick SW, Schumacher RE, Benneyworth BD, Krans EE, McAllister JM, Davis MM. Neonatal abstinence syndrome and associated health care expenditures: United States, 2000-2009. JAMA 2012 May 09;307(18):1934-1940. [doi: 10.1001/jama.2012.3951] [Medline: 22546608]
7. Forray A, Foster D. Substance use in the perinatal period. Curr Psychiatry Rep 2015 Nov;17(11):91 [FREE Full text] [doi: 10.1007/s11920-015-0626-5] [Medline: 26386836]

8. Forray A, Merry B, Lin H, Ruger JP, Yonkers KA. Perinatal substance use: a prospective evaluation of abstinence and relapse. Drug Alcohol Depend 2015 May 01;150:147-155 [FREE Full text] [doi: 10.1016/j.drugalcdep.2015.02.027] [Medline: 25772437]
9. Gopman S. Prenatal and postpartum care of women with substance use disorders. Obstet Gynecol Clin North Am 2014 Jun;41(2):213-228. [doi: 10.1016/j.ogc.2014.02.004] [Medline: 24845486]
10. Martin CE, Scialli A, Terplan M. Unmet substance use disorder treatment need among reproductive age women. Drug Alcohol Depend 2020 Jan 01;206:107679. [doi: 10.1016/j.drugalcdep.2019.107679] [Medline: 31740208]
11. Stone R. Pregnant women and substance use: fear, stigma, and barriers to care. Health Justice 2015 Feb 12;3(1):2. [doi: 10.1186/S40352-015-0015-5]
12. Minkovitz CS, O'Neill KM, Duggan AK. Home visiting: a service strategy to reduce poverty and mitigate its consequences. Acad Pediatr 2016 Apr;16(3 Suppl):S105-S111. [doi: 10.1016/j.acap.2016.01.005] [Medline: 27044687]
13. Lanier P, Maguire-Jack K, Welch H. A nationally representative study of early childhood home visiting service use in the United States. Matern Child Health J 2015 Oct;19(10):2147-2158. [doi: 10.1007/s10995-015-1727-9] [Medline: 25680701]
14. 2019 Home visiting yearbook an overview. National Home Visiting Resource Center. 2019. URL: https://nhvrc.org/wp-content/uploads/NHVRC_Yearbook_Summary_2019_FINAL.pdf [accessed 2022-01-10]
15. HFA Best Practice Standards. Prevent Child Abuse America. 2017 Jan 7. URL: https://www.dhs.state.il.us/OneNetLibrary/27896/documents/GATA_2020Grants/FCS/OtherDocuments/2018_2021HFABestPracticeStandardsJuly2017_.pdf [accessed 2022-01-10]
16. Michalopoulos C, Lee H, Duggan A, Lundquist E, Tso A, Crowne S, et al. The mother and infant home visiting program evaluation: early findings on the maternal, infant, and early childhood home visiting program – a report to congress. U.S. Department of Health & Human Services. 2015 Feb 6. URL: https://www.acf.hhs.gov/opre/report/mother-and-infant-home-visiting-program-evaluation-early-findings-maternal-infant-and
17. Kuo C, Schonbrun YC, Zlotnick C, Bates N, Todorova R, Kao JC, et al. A qualitative study of treatment needs among pregnant and postpartum women with substance use and depression. Subst Use Misuse 2013 Dec;48(14):1498-1508 [FREE Full text] [doi: 10.3109/10826084.2013.800116] [Medline: 23819737]
18. Lee King PA, Duan L, Amaro H. Clinical needs of in-treatment pregnant women with co-occurring disorders: implications for primary care. Matern Child Health J 2015 Jan;19(1):180-187. [doi: 10.1007/s10995-014-1508-x] [Medline: 24770992]
19. Goodman D, Saunders EC, Wolff KB. In their own words: a qualitative study of factors promoting resilience and recovery among postpartum women with opioid use disorders. BMC Pregnancy Childbirth 2020 Mar 18;20(1):178 [FREE Full text] [doi: 10.1186/s12884-020-02872-5] [Medline: 32188411]
20. Systems-Level Implementation of Screening, Brief Intervention, and Referral to Treatment (SBIRT). Substance Abuse and Mental Health Services Administration. 2013 May. URL: https://store.samhsa.gov/product/TAP-33-Systems-Level-Implementation-of-Screening-Brief-Intervention-and-Referral-to-Treatment-SBIRT/SMA13-4741 [accessed 2022-01-10]
21. Babor TF, McRee BG, Kassebaum PA, Grimaldi PL, Ahmed K, Bray J. Screening, Brief Intervention, and Referral to Treatment (SBIRT): toward a public health approach to the management of substance abuse. Subst Abus 2007;28(3):7-30. [doi: 10.1300/J465v28n03_03] [Medline: 18077300]
22. Alcohol Screening and Brief Intervention (SBI) for trauma patients. American College of Surgeons. 2007. URL: https://www.facs.org/media/bshd3vaq/sbirtguide.pdf [accessed 2022-01-10]
23. Unhealthy alcohol use in adolescents and adults: screening and behavioral counseling interventions. U.S. Preventive Services Task Force. 2018 Nov 13. URL: https://www.uspreventiveservicestaskforce.org/uspstf/recommendation/unhealthy-alcohol-use-in-adolescents-and-adults-screening-and-behavioral-counseling-interventions [accessed 2022-01-10]
24. Babor TF, Del Boca F, Bray JW. Screening, Brief Intervention and Referral to Treatment: implications of SAMHSA's SBIRT initiative for substance abuse policy and practice. Addiction 2017 Feb;112 Suppl 2:110-117. [doi: 10.1111/add.13675] [Medline: 28074569]
25. Wright TE, Terplan M, Ondersma SJ, Boyce C, Yonkers K, Chang G, et al. The role of screening, brief intervention, and referral to treatment in the perinatal period. Am J Obstet Gynecol 2016 Nov;215(5):539-547. [doi: 10.1016/j.ajog.2016.06.038] [Medline: 27373599]
26. Farr SL, Hutchings YL, Ondersma SJ, Creanga AA. Brief interventions for illicit drug use among peripartum women. Am J Obstet Gynecol 2014 Oct;211(4):336-343 [FREE Full text] [doi: 10.1016/j.ajog.2014.04.005] [Medline: 24721261]
27. Ondersma SJ, Martino S, Svikis DS, Yonkers KA. Commentary on Kim et al. (2017): staying focused on non-treatment seekers. Addiction 2017 May;112(5):828-829 [FREE Full text] [doi: 10.1111/add.13736] [Medline: 28378329]
28. Gerbert B, Johnston K, Caspers N, Bleecker T, Woods A, Rosenbaum A. Experiences of battered women in health care settings: a qualitative study. Women Health 1996;24(3):1-17. [doi: 10.1300/j013v24n03_01] [Medline: 9046550]
29. Rodriguez MA, Quiroga SS, Bauer HM. Breaking the silence. Battered women's perspectives on medical care. Arch Fam Med 1996 Mar;5(3):153-158. [doi: 10.1001/archfami.5.3.153] [Medline: 8620256]
30. Fugate M, Landis L, Riordan K, Naureckas S, Engel B. Barriers to domestic violence help seeking: implications for intervention. Violence Against Women 2005 Mar;11(3):290-310. [doi: 10.1177/1077801204271959] [Medline: 16043551]

31. Grekin ER, Svikis DS, Lam P, Connors V, Lebreton JM, Streiner DL, et al. Drug use during pregnancy: validating the Drug Abuse Screening Test against physiological measures. Psychol Addict Behav 2010 Dec;24(4):719-723 [FREE Full text] [doi: 10.1037/a0021741] [Medline: 21198230]
32. Berger LM, Slack KS, Waldfogel J, Bruch SK. Caseworker-perceived caregiver substance abuse and child protective services outcomes. Child Maltreat 2010 Aug;15(3):199-210 [FREE Full text] [doi: 10.1177/1077559510368305] [Medline: 20460304]
33. Parrish JW, Lanier P, Newby-Kew A, Arvidson J, Shanahan M. Maternal intimate partner violence victimization before and during pregnancy and postbirth child welfare contact: a population-based assessment. Child Maltreat 2016 Feb;21(1):26-36. [doi: 10.1177/1077559515616704] [Medline: 26627838]
34. Paulsell D, Del Grosso P, Supplee L. Supporting replication and scale-up of evidence-based home visiting programs: assessing the implementation knowledge base. Am J Public Health 2014 Sep;104(9):1624-1632. [doi: 10.2105/AJPH.2014.301962] [Medline: 25033117]
35. Duggan A, Caldera D, Rodriguez K, Burrell L, Rohde C, Crowne SS. Impact of a statewide home visiting program to prevent child abuse. Child Abuse Negl 2007 Aug;31(8):801-827. [doi: 10.1016/j.chiabu.2006.06.011] [Medline: 17822764]
36. Duggan A, Fuddy L, Burrell L, Higman SM, McFarlane E, Windham A, et al. Randomized trial of a statewide home visiting program to prevent child abuse: impact in reducing parental risk factors. Child Abuse Negl 2004 Jun;28(6):623-643. [doi: 10.1016/j.chiabu.2003.08.008] [Medline: 15193852]
37. Tandon SD, Parillo KM, Jenkins C, Duggan AK. Formative evaluation of home visitors' role in addressing poor mental health, domestic violence, and substance abuse among low-income pregnant and parenting women. Matern Child Health J 2005 Sep;9(3):273-283. [doi: 10.1007/s10995-005-0012-8] [Medline: 16240078]
38. Jones Harden B, Denmark N, Saul D. Understanding the needs of staff in head start programs: the characteristics, perceptions, and experiences of home visitors. Children Youth Serv Rev 2010 Mar;32(3):371-379. [doi: 10.1016/j.childyouth.2009.10.008]
39. Eddy T, Kilburn E, Chang C, Bullock L, Sharps P. Facilitators and barriers for implementing home visit interventions to address intimate partner violence: town and gown partnerships. Nurs Clin North Am 2008 Sep;43(3):419-35, ix [FREE Full text] [doi: 10.1016/j.cnur.2008.04.005] [Medline: 18674673]
40. Lecroy CW, Whitaker K. Improving the quality of home visitation: an exploratory study of difficult situations. Child Abuse Negl 2005 Sep;29(9):1003-1013. [doi: 10.1016/j.chiabu.2005.04.003] [Medline: 16159664]
41. Tandon SD, Mercer CD, Saylor EL, Duggan AK. Paraprofessional home visitors' perspectives on addressing poor mental health, substance abuse, and domestic violence: a qualitative study. Early Childhood Res Q 2008 Jul;23(3):419-428. [doi: 10.1016/j.ecresq.2008.02.002]
42. Zeanah PD, Larrieu JA, Boris NW, Nagle GA. Nurse home visiting: perspectives from nurses. Infant Ment Health J 2006 Jan;27(1):41-54. [doi: 10.1002/imhj.20079] [Medline: 28640422]
43. Dauber S, Hogue A, Henderson CE, Nugent J, Hernandez G. Addressing maternal depression, substance use, and intimate partner violence in home visiting: a quasi-experimental pilot test of a screen-and-refer approach. Prev Sci 2019 Nov;20(8):1233-1243 [FREE Full text] [doi: 10.1007/s11121-019-01045-x] [Medline: 31432378]
44. Choo E, Ranney M, Wetle T, Morrow K, Mello M, Squires D, et al. Attitudes toward computer interventions for partner abuse and drug use among women in the emergency department. Addict Disord Their Treat 2015 Jun;14(2):95-104 [FREE Full text] [doi: 10.1097/ADT.0000000000000057] [Medline: 26167133]
45. Tzilos GK, Sokol RJ, Ondersma SJ. A randomized phase I trial of a brief computer-delivered intervention for alcohol use during pregnancy. J Womens Health (Larchmt) 2011 Oct;20(10):1517-1524 [FREE Full text] [doi: 10.1089/jwh.2011.2732] [Medline: 21823917]
46. Nelson HD, Bougatsos C, Blazina I. Screening women for intimate partner violence: a systematic review to update the U.S. Preventive Services Task Force recommendation. Ann Intern Med 2012 Jun 05;156(11):796-808, W [FREE Full text] [doi: 10.7326/0003-4819-156-11-201206050-00447] [Medline: 22565034]
47. Ahmad F, Hogg-Johnson S, Stewart DE, Skinner HA, Glazier RH, Levinson W. Computer-assisted screening for intimate partner violence and control: a randomized trial. Ann Intern Med 2009 Jul 21;151(2):93-102 [FREE Full text] [doi: 10.7326/0003-4819-151-2-200907210-00124] [Medline: 19487706]
48. Rhodes KV, Drum M, Anliker E, Frankel RM, Howes DS, Levinson W. Lowering the threshold for discussions of domestic violence: a randomized controlled trial of computer screening. Arch Intern Med 2006 May 22;166(10):1107-1114. [doi: 10.1001/archinte.166.10.1107] [Medline: 16717173]
49. Plichta SB. Interactions between victims of intimate partner violence against women and the health care system: policy and practice implications. Trauma Violence Abuse 2007 Apr;8(2):226-239. [doi: 10.1177/1524838007301220] [Medline: 17545576]
50. Moore BA, Fazzino T, Garnet B, Cutter CJ, Barry DT. Computer-based interventions for drug use disorders: a systematic review. J Subst Abuse Treat 2011 Apr;40(3):215-223 [FREE Full text] [doi: 10.1016/j.jsat.2010.11.002] [Medline: 21185683]
51. Donoghue K, Patton R, Phillips T, Deluca P, Drummond C. The effectiveness of electronic screening and brief intervention for reducing levels of alcohol consumption: a systematic review and meta-analysis. J Med Internet Res 2014 Jun 02;16(6):e142 [FREE Full text] [doi: 10.2196/jmir.3193] [Medline: 24892426]

52. Martino S, Ondersma SJ, Forray A, Olmstead TA, Gilstad-Hayden K, Howell HB, et al. A randomized controlled trial of screening and brief interventions for substance misuse in reproductive health. Am J Obstet Gynecol 2018 Mar;218(3):322.e1-32.e12 [FREE Full text] [doi: 10.1016/j.ajog.2017.12.005] [Medline: 29247636]
53. Ondersma SJ, Svikis DS, Thacker LR, Beatty JR, Lockhart N. Computer-delivered screening and brief intervention (e-SBI) for postpartum drug use: a randomized trial. J Subst Abuse Treat 2014 Jan;46(1):52-59 [FREE Full text] [doi: 10.1016/j.jsat.2013.07.013] [Medline: 24051077]
54. Nayak MB, Kaskutas LA, Mericle AA. Randomized trial of an innovative electronic screening and brief intervention for reducing drinking among women of childbearing age. J Addict Med 2019;13(6):450-459 [FREE Full text] [doi: 10.1097/ADM.0000000000000518] [Medline: 30882553]
55. Chang G, McNamara TK, Orav EJ, Koby D, Lavigne A, Ludman B, et al. Brief intervention for prenatal alcohol use: a randomized trial. Obstet Gynecol 2005 May;105(5 Pt 1):991-998 [FREE Full text] [doi: 10.1097/01.AOG.0000157109.05453.84] [Medline: 15863535]
56. Ondersma SJ, Chase SK, Svikis DS, Schuster CR. Computer-based brief motivational intervention for perinatal drug use. J Subst Abuse Treat 2005 Jun;28(4):305-312 [FREE Full text] [doi: 10.1016/j.jsat.2005.02.004] [Medline: 15925264]
57. Ondersma SJ, Svikis DS, Thacker LR, Beatty JR, Lockhart N. A randomised trial of a computer-delivered screening and brief intervention for postpartum alcohol use. Drug Alcohol Rev 2016 Nov;35(6):710-718 [FREE Full text] [doi: 10.1111/dar.12389] [Medline: 27004474]
58. Marshall J, Kihlström L, Buro A, Chandran V, Prieto C, Stein-Elger R, et al. Statewide implementation of virtual perinatal home visiting during COVID-19. Matern Child Health J 2020 Oct;24(10):1224-1230 [FREE Full text] [doi: 10.1007/s10995-020-02982-8] [Medline: 32691359]
59. Ondersma SJ, Svikis DS, Schuster CR. Computer-based brief intervention a randomized trial with postpartum women. Am J Prev Med 2007 Mar;32(3):231-238 [FREE Full text] [doi: 10.1016/j.amepre.2006.11.003] [Medline: 17236741]
60. Ondersma SJ, Martin J, Fortson B, Whitaker DJ, Self-Brown S, Beatty J, et al. Technology to augment early home visitation for child maltreatment prevention: a pragmatic randomized trial. Child Maltreat 2017 Nov;22(4):334-343 [FREE Full text] [doi: 10.1177/1077559517729890] [Medline: 28954530]
61. Eaves ER, Doerry E, Lanzetta SA, Kruithoff KM, Negron K, Dykman K, et al. Applying user-centered design in the development of a supportive mHealth app for women in substance use recovery. Am J Health Promot 2022 Jul 10:8901171221113834. [doi: 10.1177/08901171221113834] [Medline: 35815770]
62. Wilson K, Bell C, Wilson L, Witteman H. Agile research to complement agile development: a proposal for an mHealth research lifecycle. NPJ Digit Med 2018;1:46 [FREE Full text] [doi: 10.1038/s41746-018-0053-1] [Medline: 31304326]
63. Farao J, Malila B, Conrad N, Mutsvangwa T, Rangaka MX, Douglas TS. A user-centred design framework for mHealth. PLoS One 2020 Aug 19;15(8):e0237910 [FREE Full text] [doi: 10.1371/journal.pone.0237910] [Medline: 32813711]
64. Graham AK, Munson SA, Reddy M, Neubert SW, Green EA, Chang A, et al. Integrating user-centered design and behavioral science to design a mobile intervention for obesity and binge eating: mixed methods analysis. JMIR Form Res 2021 May 10;5(5):e23809 [FREE Full text] [doi: 10.2196/23809] [Medline: 33970114]
65. Aarons GA, Palinkas LA. Implementation of evidence-based practice in child welfare: service provider perspectives. Adm Policy Ment Health 2007 Jul;34(4):411-419. [doi: 10.1007/s10488-007-0121-3] [Medline: 17410420]
66. Bowen DJ, Kreuter M, Spring B, Cofta-Woerpel L, Linnan L, Weiner D, et al. How we design feasibility studies. Am J Prev Med 2009 May;36(5):452-457 [FREE Full text] [doi: 10.1016/j.amepre.2009.02.002] [Medline: 19362699]
67. Palinkas LA, Schoenwald SK, Hoagwood K, Landsverk J, Chorpita BF, Weisz JR, Research Network on Youth Mental Health. An ethnographic study of implementation of evidence-based treatments in child mental health: first steps. Psychiatr Serv 2008 Jul;59(7):738-746. [doi: 10.1176/ps.2008.59.7.738] [Medline: 18586990]
68. Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health 2011 Mar;38(2):65-76 [FREE Full text] [doi: 10.1007/s10488-010-0319-7] [Medline: 20957426]
69. Bergh A, Johansson I, Persson E, Karlsson J, Friberg F. Nurses' Patient Education Questionnaire – development and validation process. J Res Nursing 2014 May 21;20(3):181-200. [doi: 10.1177/1744987114531583]
70. Young HN, Hwang MJ, Dilworth TJ, Mott D, Cox ED, Moreno MA. Development and evaluation of an instrument to measure community pharmacists' self-efficacy beliefs about communicating with Spanish-speaking patients. Res Social Adm Pharm 2011 Dec;7(4):330-337 [FREE Full text] [doi: 10.1016/j.sapharm.2010.08.003] [Medline: 21272530]
71. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
72. Thomas DR. A general inductive approach for analyzing qualitative evaluation data. Am J Eval 2016 Jun 30;27(2):237-246. [doi: 10.1177/1098214005283748]
73. Gray J, Beatty JR, Svikis DS, Puder KS, Resnicow K, Konkel J, et al. Electronic brief intervention and text messaging for marijuana use during pregnancy: initial acceptability of patients and providers. JMIR Mhealth Uhealth 2017 Nov 08;5(11):e172 [FREE Full text] [doi: 10.2196/mhealth.7927] [Medline: 29117931]

74. Ondersma SJ, Beatty JR, Svikis DS, Strickler RC, Tzilos GK, Chang G, et al. Computer-delivered screening and brief intervention for alcohol use in pregnancy: a pilot randomized trial. Alcohol Clin Exp Res 2015 Jul;39(7):1219-1226 [FREE Full text] [doi: 10.1111/acer.12747] [Medline: 26010235]
75. Ondersma SJ, Svikis DS, Thacker C, Resnicow K, Beatty JR, Janisse J, et al. Computer-delivered indirect screening and brief intervention for drug use in the perinatal period: a randomized trial. Drug Alcohol Depend 2018 Apr 01;185:271-277 [FREE Full text] [doi: 10.1016/j.drugalcdep.2017.12.022] [Medline: 29482051]
76. The Alcohol, Smoking and Substance Involvement Screening Test (ASSIST). Geneva: World Health Organization; Jan 1, 2010.
77. Newcombe DA, Humeniuk RE, Ali R. Validation of the world health organization alcohol, smoking and substance involvement screening test (assist): report of results from the Australian site. Drug Alcohol Rev 2005 May;24(3):217-226. [doi: 10.1080/09595230500170266] [Medline: 16096125]
78. WHO ASSIST Working Group. The Alcohol, Smoking and Substance Involvement Screening Test (ASSIST): development, reliability and feasibility. Addiction 2002 Sep;97(9):1183-1194. [doi: 10.1046/j.1360-0443.2002.00185.x] [Medline: 12199834]
79. Ondersma SJ, Svikis DS, LeBreton JM, Streiner DL, Grekin ER, Lam PK, et al. Development and preliminary validation of an indirect screener for drug use in the perinatal period. Addiction 2012 Dec;107(12):2099-2106 [FREE Full text] [doi: 10.1111/j.1360-0443.2012.03982.x] [Medline: 22882721]
80. Miller W, Rollnick S. Motivational Interviewing: Helping People Change. New York, United States: Guilford Publications; 2013.
81. Miller WR, Rose GS. Toward a theory of motivational interviewing. Am Psychol 2009 Sep;64(6):527-537 [FREE Full text] [doi: 10.1037/a0016830] [Medline: 19739882]
82. Schaefer JK. Personal characteristics of effective home visitors. J Social Serv Res 2015 Nov 16;42(1):84-95. [doi: 10.1080/01488376.2015.1078868]
83. Saïas T, Lerner E, Greacen T, Emer A, Guédeney A, Dugravier R, et al. Parent-provider relationship in home visiting interventions. Children Youth Serv Rev 2016 Oct;69:106-115. [doi: 10.1016/j.childyouth.2016.08.004]
84. Wouldes TA, Crawford A, Stevens S, Stasiak K. Evidence for the effectiveness and acceptability of e-SBI or e-SBIRT in the management of alcohol and illicit substance use in pregnant and post-partum women. Front Psychiatry 2021;12:634805 [FREE Full text] [doi: 10.3389/fpsyt.2021.634805] [Medline: 34025470]
85. Toquinto SM, Berglas NF, McLemore MR, Delgado A, Roberts SC. Pregnant women's acceptability of alcohol, tobacco, and drug use screening and willingness to disclose use in prenatal care. Womens Health Issues 2020;30(5):345-352. [doi: 10.1016/j.whi.2020.05.004] [Medline: 32622582]
86. Adams ZM, Ginapp CM, Price CR, Qin Y, Madden LM, Yonkers K, et al. "A good mother": impact of motherhood identity on women's substance use and engagement in treatment across the lifespan. J Subst Abuse Treat 2021 Nov;130:108474. [doi: 10.1016/j.jsat.2021.108474] [Medline: 34118710]
87. Wright ME, Temples HS, Shores E, Chafe O, Lannamann R, Lautenschlager C. Pregnant and parenting women's experiences with substance use disorder. MCN Am J Matern Child Nurs 2021;46(5):271-276. [doi: 10.1097/NMC.0000000000000741] [Medline: 34398828]
88. Cockroft JD, Adams SM, Bonnet K, Matlock D, McMillan J, Schlundt D. "A scarlet letter": stigma and other factors affecting trust in the health care system for women seeking substance abuse treatment in a community setting. Subst Abus 2019;40(2):170-177. [doi: 10.1080/08897077.2018.1544184] [Medline: 30759047]
89. Beasley LO, Ridings LE, Smith TJ, Shields JD, Silovsky JF, Beasley W, et al. A qualitative evaluation of engagement and attrition in a nurse home visiting program: from the participant and provider perspective. Prev Sci 2018 May;19(4):528-537. [doi: 10.1007/s11121-017-0846-5] [Medline: 29022144]
90. West A, Duggan A, Gruss K, Minkovitz C. Service coordination to address maternal mental health, partner violence, and substance use: findings from a national survey of home visiting programs. Prev Sci 2021 Jul;22(5):633-644. [doi: 10.1007/s11121-021-01232-9] [Medline: 33835378]
91. Goldberg J, Greenstone Winestone J, Fauth R, Colón M, Mingo MV. Getting to the warm hand-off: a study of home visitor referral activities. Matern Child Health J 2018 Oct;22(Suppl 1):22-32 [FREE Full text] [doi: 10.1007/s10995-018-2529-7] [Medline: 29858963]
92. Sundström C, Blankers M, Khadjesari Z. Computer-based interventions for problematic alcohol use: a review of systematic reviews. Int J Behav Med 2017 Oct;24(5):646-658 [FREE Full text] [doi: 10.1007/s12529-016-9601-8] [Medline: 27757844]
93. Hyer J, Ulrickson C, Yerelian E, Metz TD, Allshouse AA, Hoffman MC. Self-reported alcohol, tobacco, and marijuana use in pregnant women with depressive symptomatology. Am J Perinatol 2020 Oct;37(12):1223-1227. [doi: 10.1055/s-0039-1692685] [Medline: 31238345]
94. Chapman SLC, Wu L. Postpartum substance use and depressive symptoms: a review. Women Health 2013;53(5):479-503 [FREE Full text] [doi: 10.1080/03630242.2013.804025] [Medline: 23879459]
95. Metz VE, Brown QL, Martins SS, Palamar JJ. Characteristics of drug use among pregnant women in the United States: opioid and non-opioid illegal drug use. Drug Alcohol Depend 2018 Feb 01;183:261-266 [FREE Full text] [doi: 10.1016/j.drugalcdep.2017.11.010] [Medline: 29310077]

96. Dauber S, Ferayorni F, Henderson C, Hogue A, Nugent J, Alcantara J. Substance use and depression in home visiting clients: home visitor perspectives on addressing clients' needs. J Community Psychol 2017 Apr;45(3):396-412 [FREE Full text] [doi: 10.1002/jcop.21855] [Medline: 28408768]
97. Tandon D, Perry DF, Edwards K, Mendelson T, SCRIPT Advisory Panel. Developing a model to address mental health, substance use, and intimate partner violence among home visiting clients. Health Promot Pract 2020 Mar;21(2):156-159. [doi: 10.1177/1524839919886293] [Medline: 31718313]
98. O'Neill K, Burrell L, Chute D, Korfmacher J, Duggan A. COVID-19's continued impact on home visiting. Report of Results from a National HARC-Beat Follow-up Survey of Local Home Visiting Programs. 2021 Nov. URL: http://www.hvresearch.org/wp-content/uploads/2022/03/COVID-II-report-_FINAL.pdf [accessed 2021-10-10]
99. Who is being served? National Home Visiting Resource Center. 2021 Dec 26. URL: https://nhvrc.org/yearbook/2021-yearbook/who-is-being-served/ [accessed 2022-02-10]
100. Father engagement in home visiting: benefits, challenges, and promising strategies. National Home Visiting Resource Center. 2019 Apr. URL: https://www.urban.org/sites/default/files/publication/100229/father_engagement_in_home_visiting.pdf [accessed 2022-01-10]
101. Hamil J, Gier E, Garfield CF, Tandon D. The development and pilot of a technology-based intervention in the United States for father's mental health in the perinatal period. Am J Mens Health 2021;15(5):15579883211044306 [FREE Full text] [doi: 10.1177/15579883211044306] [Medline: 34587839]
102. Hossain M, Lauren A, Del Grosso P, Shenk M, Cavanaugh M, Azur M. Touchpoints for addressing substance use issues in home visiting: phase 1 final report. Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services. 2020. URL: https://www.acf.hhs.gov/opre/report/touchpoints-addressing-substance-use-issues-home-visiting-phase-1-final-report [accessed 2021-10-10]
103. Supplee LH, Parekh J, Johnson M. Principles of precision prevention science for improving recruitment and retention of participants. Prev Sci 2018 Jul;19(5):689-694. [doi: 10.1007/s11121-018-0884-7] [Medline: 29532364]
104. The Home Visiting Applied Research Collaborative. Introduction to precision home visiting. Child Trends and James Bell Associates. 2018. URL: https://www.hvresearch.org/precision-home-visiting/introduction-to-precision-home-visiting/ [accessed 2022-01-10]
105. Haroz EE, Ingalls A, Wadlin J, Kee C, Begay M, Neault N, et al. Utilizing broad-based partnerships to design a precision approach to implementing evidence-based home visiting. J Community Psychol 2020 May;48(4):1100-1113 [FREE Full text] [doi: 10.1002/jcop.22281] [Medline: 31970805]
106. O'Brien RA, Moritz P, Luckey DW, McClatchey MW, Ingoldsby EM, Olds DL. Mixed methods analysis of participant attrition in the nurse-family partnership. Prev Sci 2012 Jun;13(3):219-228 [FREE Full text] [doi: 10.1007/s11121-012-0287-0] [Medline: 22562646]
107. CIAS. URL: https://www.cias.app/ [accessed 2022-10-22]

Abbreviations

BI: brief intervention
CIAS: Computerized Intervention Authoring System
CPS: child protective system
e-SBI: electronic screening and brief intervention
e–SBI-HV: electronic screening and brief intervention for home visiting
HFA: Healthy Families America
HV: home visiting
MI: motivational interviewing
PI: principal investigator
SBI: screening and brief intervention
SBIRT: screening, brief intervention, and referral to treatment
SU: substance use

Edited by A Mavragani; submitted 09.03.22; peer-reviewed by S Diestelkamp, M Jenkins; comments to author 16.06.22; revised version received 06.09.22; accepted 29.09.22; published 08.11.22.

Please cite as:
Dauber S, Hammond C, Hogue A, Henderson C, Nugent J, Ford V, Brown J, Scott L, Ondersma S
Development of an Electronic Screening and Brief Intervention to Address Perinatal Substance Use in Home Visiting: Qualitative User-Centered Approach
JMIR Form Res 2022;6(11):e37865
URL: https://formative.jmir.org/2022/11/e37865
doi:10.2196/37865
PMID:36346648

Original Paper

Shared Autonomy to Reduce Sedentary Behavior Among Sit-Stand Desk Users in the United States and India: Web-Based Study

Lawrence H Kim[1,2], PhD; Gourab Saha[3], MS; Annel Amelia Leon[4]; Abby C King[5], PhD; Matthew Louis Mauriello[6], PhD; Pablo E Paredes[2,7], PhD

[1]School of Computing Science, Simon Fraser University, Burnaby, BC, Canada
[2]Department of Psychiatry and Behavioral Sciences, School of Medicine, Stanford University, Stanford, CA, United States
[3]Department of Mechanical Engineering, Stanford University, Stanford, CA, United States
[4]Department of Computer Science, Stanford University, Stanford, CA, United States
[5]Department of Epidemiology and Population Health, School of Medicine, Stanford University, Stanford, CA, United States
[6]Department of Computer & Information Sciences, University of Delaware, Newark, DE, United States
[7]Department of Computer Science, University of Maryland, College Park, MD, United States

Corresponding Author:
Pablo E Paredes, PhD
Department of Computer Science
University of Maryland
8125 Paint Branch Drive
College Park, MD, 20742
United States
Phone: 1 650 723 4000
Email: paredes@umd.edu

Abstract

Background: Fitness technologies such as wearables and sit-stand desks are increasingly being used to fight sedentary lifestyles by encouraging physical activity. However, adherence to such technologies decreases over time because of apathy and increased dismissal of behavioral nudges.

Objective: To address this problem, we introduced shared autonomy in the context of sit-stand desks, where user input is integrated with robot autonomy to control the desk and reduce sedentary behavior and investigated user reactions and preferences for levels of automation with a sit-stand desk. As demographics affect user acceptance of robotic technology, we also studied how perceptions of nonvolitional behavior change differ across cultures (United States and India), sex, familiarity, dispositional factors, and health priming messages.

Methods: We conducted a web-based vignette study in the United States and India where a total of 279 participants watched video vignettes of a person interacting with sit-stand desks of various levels of automation and answered questions about their perceptions of the desks such as ranking of the different levels of automation.

Results: Participants generally preferred either manual or semiautonomous desks over the fully autonomous option (P<.001). However, participants in India were generally more amenable to the idea of nonvolitional interventions from the desk than participants in the United States (P<.001). Male participants had a stronger desire for having control over the desk than female participants (P=.01). Participants who were more familiar with sit-stand desks were more likely to adopt autonomous sit-stand desks (P=.001). No effects of health priming messages were observed. We estimated the projected health outcome by combining ranking data and hazard ratios from previous work and found that the semiautonomous desk led to the highest projected health outcome.

Conclusions: These results suggest that the shared autonomy desk is the optimal level of automation in terms of both user preferences and estimated projected health outcomes. Demographics such as culture and sex had significant effects on how receptive users were to autonomous intervention. As familiarity improves the likelihood of adoption, we propose a gradual behavior change intervention to increase acceptance and adherence, especially for populations with a high desire for control.

(JMIR Form Res 2022;6(11):e35447) doi:10.2196/35447

KEYWORDS

shared autonomy; automation; sedentary behavior; sit-stand desk; nonvolitional behavior change; culture

Introduction

Background

Humans are engaging in increasingly more sedentary lifestyles that are correlated with increasing levels of technology use [1,2]. Various studies have been conducted to assess the damage such sedentary behavior can induce to overall health, including increased risks for stress, anxiety, depression, premature mortality, and decreased telomere length, a biological measure associated with longevity [3-6].

As the number of sedentary workers has increased in the modern work environment, technologies that promote physical activities have garnered heightened attention, including the use of prompts and similar behavioral strategies via text messages, websites, wearables, and mobile apps [7]. An additional technology of particular interest is the sit-stand desk, which interrupts periods of sedentary behavior when a person moves to use the sit-stand functionality. Breaking and reducing sedentary time with frequent light-intensity movements (eg, moving to stand or sit) has been found to improve health outcomes [3,8]. Such movements every 30 minutes may help people live longer and healthier lives [8,9]. In addition, sit-stand desks combine the comfort of sitting and the health benefits of standing by allowing the user to switch between the 2 positions easily while maintaining a functional workspace. In theory, consistent use of a sit-stand desk can reduce sedentary time, which may mitigate factors related to metabolic risk and other health outcomes [3].

However, adherence to the consistent use of a sit-stand desk is typically low because of apathy and low motivation [10,11]. Recent research suggests that only about one-third of sit-stand desk owners use their sit-stand functionality after the first few months of ownership, and they typically do so less than once a month [10]. Most workers simply do not use the feature despite being aware of the health implications of sitting down for too long [11] and a desire for a healthier lifestyle [12]. These results align with prior research indicating apathy toward the more active use of sit-stand desks [11].

Given the behavioral, cognitive, and motivational demands often accompanying such behavioral choices and decision-making throughout the day [13,14], reducing the cognitive load accompanying volitional sit-stand desk use through automation may be particularly valuable. For example, even simple interventions such as setting the default desk height to the standing position at the beginning of the workday increase the standing work rates for employees [15], as users have a strong tendency to go along with default options [16]. However, there has been little work expanding upon the idea of moving from default options to full automation in sit-stand desks, which makes for an interesting environment to explore behavior change in occupational settings and examine how much control can be given to such systems without negative implications for worker adoption.

In this study, we proposed the integration of shared autonomy in sit-stand desks. Contrary to the typical use of shared autonomy, where task performance such as accuracy, speed, and robustness is of the highest interest [17,18], our goal is to improve users' physical well-being through the consistent use of sit-stand desks. Regarding health behavior change such as regular physical exercise, eating behavior, and alcohol consumption, conscious intentions are typically insufficient and generally have limited effects [19]. Instead, nonconscious and nonintentional processes can, for many people, be more instrumental in self-regulation.

Although most research on behavior change tends to focus on nudges or reminders that can easily be ignored, the concept of nonvolitional physical behavior change using robotic furniture was introduced and defined in previous work as an infrastructure-mediated intervention that enforces a change in behavior, such as activity or posture [20]. The *Haunted Desk* is an instance of nonvolitional behavior change to proactively promote healthy movements in users by automating the transitions between sitting and standing, thus alleviating users from the burden of decision-making. However, even the participants who preferred the autonomous desk desired some sense of control. On the basis of this work, we investigated how user perceptions and preferences change when desks with shared autonomy, in addition to binary extremes (ie, manual vs autonomous), are presented in the context of sit-stand desks.

As described in various technology acceptance models, such as the Unified Theory of Acceptance and Use of Technology model [21] and the 3-layered trust model [22], demographics (eg, culture and sex) and prior experience play an important role in moderating user attitude and behavior toward automation and new technology; for instance, culture has an influence on technology acceptance, as per the observations by Im et al [23] that the effect of effort expectancy (ie, how easy the technology is to use) on behavioral intention and the impact of behavioral intention on actual use were both greater for US users compared with Korean users. Sex and gender are also significant factors. For example, in health care robotics, Kuo et al [24] found that men have a more positive attitude toward robots. The genders of both the human and the robot are also important, as participants of both genders tend to rate the robot of the opposite gender as more credible, trustworthy, and engaging [25]. Familiarity with recommendation agents was also found to improve the intention to adopt through cognitive and emotional trust [26]. Given these findings, we aim to understand how culture (in particular, the United States and India for this project), sex, and familiarity affect user acceptance and shared autonomy preferences for sit-stand desks.

In addition to demographics, dispositional factors such as self-regulation and desirability of control (DC) are important to consider for adoption of technology that enforces health-promoting behaviors. From a health perspective, self-regulation has been recognized as an important factor in the uptake of and adherence to health-promoting behaviors,

such as physical activity [27-30]. From a technology adoption perspective, the DC can influence how users respond to automated technology [31,32].

Recent research efforts have focused on designing nonconscious interventions, such as goal priming, a cueing intervention tool to activate health goals, and encourage healthier behavior [19,33,34]. For instance, Chen et al [34] demonstrated that participants who were primed to view active video games as exercise used the system significantly longer than those primed to view them as gameplay. In our study, we investigated the effects of priming by emphasizing the health benefits of autonomous sit-stand desks.

Objectives

The central questions for this research were as follows: (Q1) How do users perceive and react to sit-stand desks with varying levels of automation? (Q2) How do demographics, such as culture and sex, along with familiarity with sit-stand desks, affect these perceptions? (Q3) How do dispositional factors such as the DC and self-regulation affect user perception? (Q4) Can goal priming alter preferences? (Q5) Can we estimate the approximate projected health outcomes with these different levels of automation based on the adoption likelihood and hazard ratio?

As a first step toward answering these questions, we designed a formative web-based video vignette study, a technique commonly used in the field of psychology, human-robot interaction (HRI), and human–computer interaction to better understand user reactions to technologies [35-38]. Participants were given hypothetical situations to which they responded, thereby revealing their perceptions, values, social norms, and impressions of the events. On the basis of participants' feedback from prior work [20], the levels of automation were expanded from binary (ie, manual vs autonomous) to include 2 intermediate levels (ie, notification and set-and-forget desks), as shown in Figure 1. For both intermediate levels, control over the height changes was distributed to the user to varying degrees. The notification desk provided regular notifications (every 30 minutes) to the user to either sit or stand but ultimately left the decision to the user. In contrast, the set-and-forget desk asked the user daily for the desired height switch frequency but executed the height changes autonomously for the rest of the day.

We gathered and analyzed responses from 279 adult participants from the United States and India, as shown in Table 1, to investigate perceptions and preferences regarding levels of automation, explore differences across people of different demographics (ie, culture and sex) and familiarity with the technology, and study the effects of goal priming on these perceptions. In addition to culture, we also gathered users' individual traits, such as the DC and relative autonomous motivation (RAM) index, as measured by the Treatment Self-Regulation Questionnaire (TSRQ) to understand its impact on user preference.

Figure 1. Four levels of automation were studied in the context of sit-stand desks. (A) For the manual desk, the user is the sole controller over the desk. (B) For the notification desk, the user is notified every 30 minutes to change the height of the desk which the user can accept or reject. (C) For the set-and-forget desk, the user decides in the morning how often the desk will change heights throughout the day. (D) For the autonomous desk, the desk decides when to change heights.

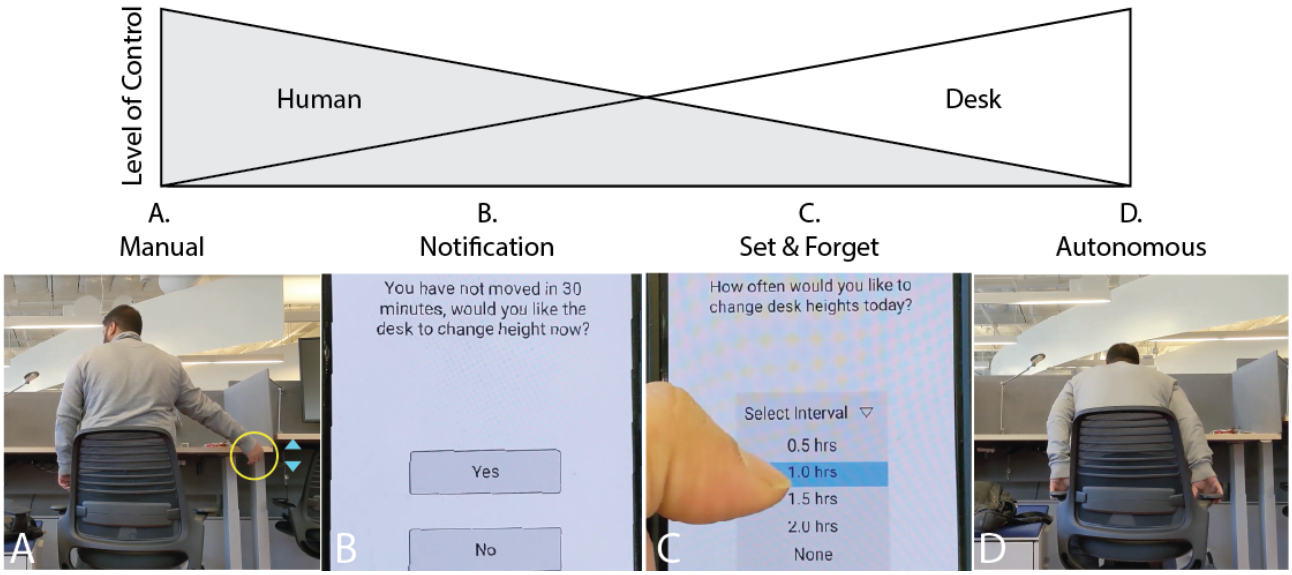


Table 1. Participant demographics: equal numbers of participants were recruited from 2 countries (India and United States) and sex. After removing responses that did not pass attention checks, we had slightly fewer participants from India, mainly female.

	Sex, n		Race or ethnicity, n				Age (years), n				Familiarity, n		
	F[a]	M[b]	W[c]	A[d]	B[e]	O[f]	19-29	30-39	40-49	>50	Read[g]	Seen[h]	Used[i]
India	59	70	2	125	0	2	75	46	6	2	40	29	60
United States	75	75	117	17	14	6	47	61	21	21	65	35	50
Total	134	145	119	142	14	8	122	107	27	23	105	64	110

[a]F: female.

[b]M: male.

[c]W: White.

[d]A: Asian.

[e]B: Black or African American.

[f]O: other race or ethnicity.

[g]Read about sitting or seeing a video of a sit-stand desk.

[h]Seen someone using a sit-stand desk in real life.

[i]Used a sit-stand desk before.

Methods

Overview

Previous work has shown that people are divided in their reactions and acceptance of automated sit-stand desks [20]. To improve the adoption likelihood of such an intervention and thus the projected user health outcome, we introduced sit-stand desks with shared autonomy in addition to binary levels (ie, manual and autonomous). We explored user reactions and preferences for these levels of shared autonomy (Figure 1) in the context of sit-stand desks through a web-based video vignette study.

Ethical Considerations

This research was approved by the Committee on the Use of Human Subjects, University's Institutional Review Board (IRB #45825), and complies with all relevant ethical regulations. Individuals provided informed web-based digital consent before participation.

Independent Variables

For this study, we used 4 independent variables: level of shared autonomy, country of the participants, sex of the participants, and priming.

Level of Automation

In prior studies, only binary levels of automation were tested: manual and autonomous [20]. In that preliminary study, most participants voiced their desire to have features, such as being notified before the height adjustment intervention and having the option to set the height adjustment frequency for each day. Thus, for this work, we added 2 new shared autonomy options where different levels of actuation are provided to the user: a *notification* desk that would send reminders to the user's phone at preset intervals and change heights if the user consents, and a *set-and-forget* (semiautonomous) desk that would ask once at the beginning of the day and automatically change heights at preset intervals for the remaining of the day. More details on the video vignettes of these conditions can be found in the Video Vignettes section.

Culture

As discussed in the Introduction section, cross-cultural studies are becoming increasingly common and important with respect to evaluating technology adoption. We aimed to improve the health of the global population through future interventions. Thus, it is necessary to understand the differences between different countries and cultures to tailor such interventions to user populations to increase the likelihood of adoption. We recruited participants from the United States and India because they represent 2 well-identified cultures (individualistic vs collectivistic [39]) with vastly opposing levels of penetration regarding automation [40]. In addition, people in these 2 countries are sufficiently fluent in the same language (ie, English), reducing the chance of miscommunication because of language [30].

Sex

Similar to culture, sex and gender have also been extensively investigated for technology adoption and HRI [24,25,41-44]. To study the differences between sexes, we explicitly recruited approximately equal numbers of male and female participants from both India and the United States.

Familiarity

Familiarity or prior experience with technology can often lead to a more positive attitude toward technology adoption [21,22,26]. In this study, we only recruited participants who had some degree of exposure to sit-stand desks to minimize the novelty effect; we are still interested in whether the level of familiarity with sit-stand desks impacts participants' perception and reaction. Thus, we asked participants to report whether they had previously "read about or seen a video of a sit-stand desk," "seen someone use a sit-stand desk in real life," or "used a sit-stand desk before."

Self-regulation

To operationalize self-regulation (ie, RAM), we used the TSRQ [45,46]. The TSRQ measures the likelihood that someone would engage in healthy behavior (eg, exercise, diet, and quit smoking) and the degree to which their motivation for engagement is autonomous (ie, simply for the pleasure, interest, and satisfaction derived from the engagement) or controlled (ie, engaged to obtain a reward or to avoid negative consequences) [47]. We were interested in the participant's motivation to alternate between sitting and standing and thus used the TSRQ-exercise questionnaire. It consists of 15 questions that measure controlled and autonomous motivation [45]. We used the difference between these 2 as our RAM index [48].

Desirability of Control

The questionnaire for DC [49] contains 20 questions that measure the likelihood that control over events is a major motivational force in decision-making. In theory, higher scores indicate an unwillingness to give up control over a sit-stand desk and an aversion to autonomous behavior change technologies.

Priming

To study the effects of health-focused priming, the participants were randomly distributed into 3 conditions: no priming, loss-framed priming, and gain-framed priming. Before answering questions compared with the 4 sit-stand desk prototypes, the loss-framed group of participants read an informational description that stated that autonomous sit-stand desks help users alternate between sitting and standing positions throughout the day because of the *harmful* effects of sedentary lifestyles, such as poor blood circulation, muscle stiffness, and back pain. In the gain-framed group, this message was altered to state that sit-stand desks were used to *alleviate* the effects of sedentary lifestyles.

Measure

The dependent variables for the study were perception and preferences regarding the levels of automation in sit-stand desks. Participants were asked to rate the following aspects for each level of automation on a 7-point Likert scale: likeability, ease of use, safety, improvement in productivity, reduction in stress, and health improvement. They also ranked the 4 levels of automation based on which level they preferred to use regularly, and provided an open-text explanation. Afterward, they rated on a 7-point Likert scale their likelihood of adopting an autonomous sit-stand desk at either work or home and provided an open-text explanation. Finally, we asked them to choose their preferred alternating frequency among (1) 30 to 45 minutes, (2) 45 to 60 minutes, (3) 60 to 75 minutes, (4) 75 to 90 minutes, and (5) ≥2 hours.

Procedure

After answering the demographics questions and passing the inclusion criteria described in the Participants section, participants provided informed web-based digital consent. Then, they were presented with videos of the 4 different sit-stand desks described in the Video Vignettes section in a randomized order. After each video, participants were asked about their perceptions in terms of likeability, ease of use, safety, productivity, stress, and health benefits of the desk. Participants were then randomly assigned to 1 of 3 priming conditions: no priming, loss-framed priming, and gain-framed priming. After reading the priming text, participants were asked to rank the desks from their most preferred desk to their least preferred desk and rate their likelihood of adopting an autonomous desk (ie, the set-and-forget desk and the autonomous desk) both at work and at home. We then collected three open-response questions in which respondents were asked to explain their reasoning for their (1) most preferred desk, (2) least preferred desk, and (3) adoption likelihood rating of an autonomous sit-stand desk for regular use at work and at home. Finally, participants filled out the DC questionnaire and the TSRQ, which is used to measure the RAM index. The average completion time was 31 (SD 12) minutes.

Participants

We recruited 397 participants from India and the United States using Amazon Mechanical Turk. To ensure comprehension of the survey, we created inclusion criteria, such as fluency in English, normal or corrected-to-normal vision, and hearing. We also recruited individuals who spent more than 2 hours at a desk on a typical weekday and had at least read about or viewed a video of a sit-stand desk before the survey to ensure that their impressions would not be affected by the novelty of the sit-stand desk itself. Of the 386 completed responses, we removed 107 because they either did not satisfy the inclusion criteria or did not correctly answer 2 test questions designed to differentiate those who are properly following the instructions from those who are not. The demographics of the remaining 279 participants included in the analysis are described in Table 1.

Video Vignettes

Overview

In the literature, HRI researchers have compared video-based studies to live in-laboratory studies. When studying how a robot should approach users, there was high agreement between live in-laboratory and video-based studies [50]. In contrast, physically present robots were found to yield greater emotional and social user feedback than robots through video or text [51], while people trusted and provided more personal space to physically present robots compared with robots that were video displayed [52]. However, both physical and video-displayed robots were still effective in conveying contextual information and eliciting feedback on general attitudes [51] and were greeted by and cooperated with participants equally [52]. On the basis of these studies, we believe that video-based studies will serve as useful design probes for understanding user reactions and perceptions of shared autonomy in sit-stand desks.

For this study, we created 4 video vignettes to serve as design probes, as shown in Figure 1 in the main text. Each video was silent, used subtitles to convey the features of the desk, and was approximately 1.5 minutes long. These videos were embedded in Qualtrics, and a brief text note below them described the key features. Below are brief descriptions of each scenario and the sit-stand desk portrayed in these videos.

Manual Desk

Raj is working at his office while sitting down. After a while, Raj decided to stand and uses the buttons on the desk to raise it to the desired height. He repeats this procedure several times throughout the day.

Notification Desk

Raj begins his work sitting down. Every 30 minutes, he receives a notification on his phone asking whether he would like to change desk positions. Raj can select either "Yes" to have the desk change to precomputed optimal heights or "No" to skip. He can make small adjustments using the buttons. Notes below the video shows that various antipinch features and mechanical "click" sounds are used for safety purposes.

Set-and-Forget Desk

As Raj begins his work, he receives a phone notification asking him how often he would like to change the desk's position. After Raj selects the desired height change frequency, the desk automatically changes position at that frequency, and Raj alternates between sitting and standing.

Autonomous Desk

Raj begins work at his office. After every 30 minutes, the desks automatically change heights, and Raj alternates between sitting and standing accordingly.

Analysis

To examine the effects of the categorical independent variables (eg, shared autonomy, country, sex, familiarity, and priming) on nonnormal data, such as the ranking of automation level and preferred switch frequency, a nonparametric Friedman test (for repeated measures) and Kruskal-Wallis test (for nonrepeated measures) were conducted. Bonferroni-corrected post hoc tests were used to determine the pairs that were statistically significantly different.

To examine the effects of the categorical independent variables (eg, shared autonomy, country, sex, familiarity, and priming), including interaction on normal data such as the Likert scale responses and the DC or RAM scores, a Mauchly Test of Sphericity, and an N-way repeated measures ANOVA (or N-way ANOVA) were performed for each dependent variable. If the Mauchly Test of Sphericity was violated, we used a Greenhouse-Geisser correction for F and P values from ANOVA, indicated by F* and P*. If any independent variable or combination had statistically significant effects (P<.05), Bonferroni-corrected post hoc tests were used to determine which pairs were significantly different. If the effect is statistically significant, the effect size (η^2p) is also reported. For reference, η^2p=0.01, 0.059, and 0.138 corresponds to small, medium, and large particles, respectively [53,54].

For ordinal or continuous independent variables (eg, age, DC, and RAM index), Spearman correlation was used to evaluate the correlation with ordinal dependent variables such as Likert scale ratings, ranking, and switch frequency data. The correlation coefficients and their significance levels are presented.

For each of the 3 open-response questions, 3 of the authors coded the first 50 responses together and developed a codebook that applied to all 3 questions, as shown in Table 2. We then proceeded to divide the remaining responses among 2 of the authors. Each author individually coded the rest of their assigned responses and tagged any responses for which they were unsure of the code to be assigned. Three authors then reconvened to discuss the tagged responses and decided on the code. In total, 5.6% (47/837) of the responses were not clear (eg, typos) in their meaning and were excluded from further analysis.

Table 2. Aggregated list of reasons for the participant's automation preference and adoption likelihood was used in our thematic analysis of the open-response survey questions.

Code	Definition	Example
Safety	Concern about the risk of injury	Automatic switching can lead to accidents.
Ease	Absence of difficulty; comfort	This desk is comfortable and easy to use.
Generic	No concrete reasoning	Very nice desk.
Annoying	Disturbing or obtrusive	It can irritate me if it forces me to stand.
Sense of control	Favorable for direct control over desk	I can decide the height of the desk at any time.
Health benefits	Positive health outcomes	The desk will improve my blood circulation when working.
Productivity	Work-related efficiency	The desk will allow me to focus on my work.
Automation	Favorable to the automated aspect of the desk	I forget to stand so an automatic desk would be nice.
External barriers	Concern about cost or space	It is too expensive.

Ultimately, our objective was to improve user health outcomes. To understand the potential effects of each level of automation in sit-stand desks, we combined the standardized ranking data, which is correlated with adoption likelihood, and the standardized hazard ratio associated with each desk, which is estimated using the data from Diaz et al [9]. For hazard ratio, we used the average projected hazard ratio for each desk's corresponding mean sedentary bout duration. For the autonomous desk, we assumed a 30-minute sedentary bout duration, and this frequency was strictly enforced. For the set-and-forget desk, we used the mean frequency (55 minutes) that users reported to prefer in our survey, while we assumed a mean bout duration of 2 hours for the manual desk as people changed their desk height once every 4 hours [55]. For the notification desk, we assumed that users would alternate between sitting and standing at a rate in between that for the

set-and-forget desk and the manual desk; hence, a mean bout duration of 90 minutes.

Results

Overview

We used hypothetical video vignettes that describe interaction with a sit-stand desk of varying levels of automation to answer our questions on (1) user preference in the level of automation, (2) effects of user demographics (ie, culture and sex), and familiarity on user preference, (3) effects of dispositional factors, (4) effects of priming on user preference, and (5) estimation of project health outcomes for different levels of automation based on adoption likelihood and hazard ratio.

Q1: Perception and Preference in Levels of Automation

To answer Q1, we first analyzed the data from all participants.

Level of Automation

As shown in Figure 2, there were statistically significant differences between the levels of automation in terms of the participants' rankings (χ^2_3=87.2; P<.001). There were statistically significant differences between the autonomous desk and manual desk (P<.001), notification desk (P<.001), set-and-forget desk (P<.001), and between the manual desk and notification desk (P=.047) after Bonferroni adjustments.

Shared autonomy had statistically significant effects on perceived likeability ($F_{2.7,760.1}$=19.0; P<.001; η2p=0.064), ease of use ($F_{2.8,781.1}$=5.9; P<.001; η^2p=0.021), safety ($F_{2.7,753.8}$=19.9; P<.001; η^2p=0.067), productivity ($F_{2.8,764.6}$=8.8; P<.001; η^2p=0.031), and stress reduction ($F_{2.8,780.6}$=14.6; P<.001; η^2p=0.05; Greenhouse-Geisser correction) but not on rated health, as shown in Figure 3A. The manual desk consistently had the highest ratings in both countries, whereas the fully autonomous desk had the lowest scores.

Figure 2. Shared autonomy preferences for sit-stand desk users across culture and sex with means and SEs reported via the black line and error bars. (∗.01≤P<.05, ∗∗.001≤P<.01, ∗∗∗P<.001) The clustered data points in the background of the image provide a visual encoding of the number of participants who provided that ranking; thus, the size of the clusters in each subplot offers a visual of the composition of the data in the top-left corner.

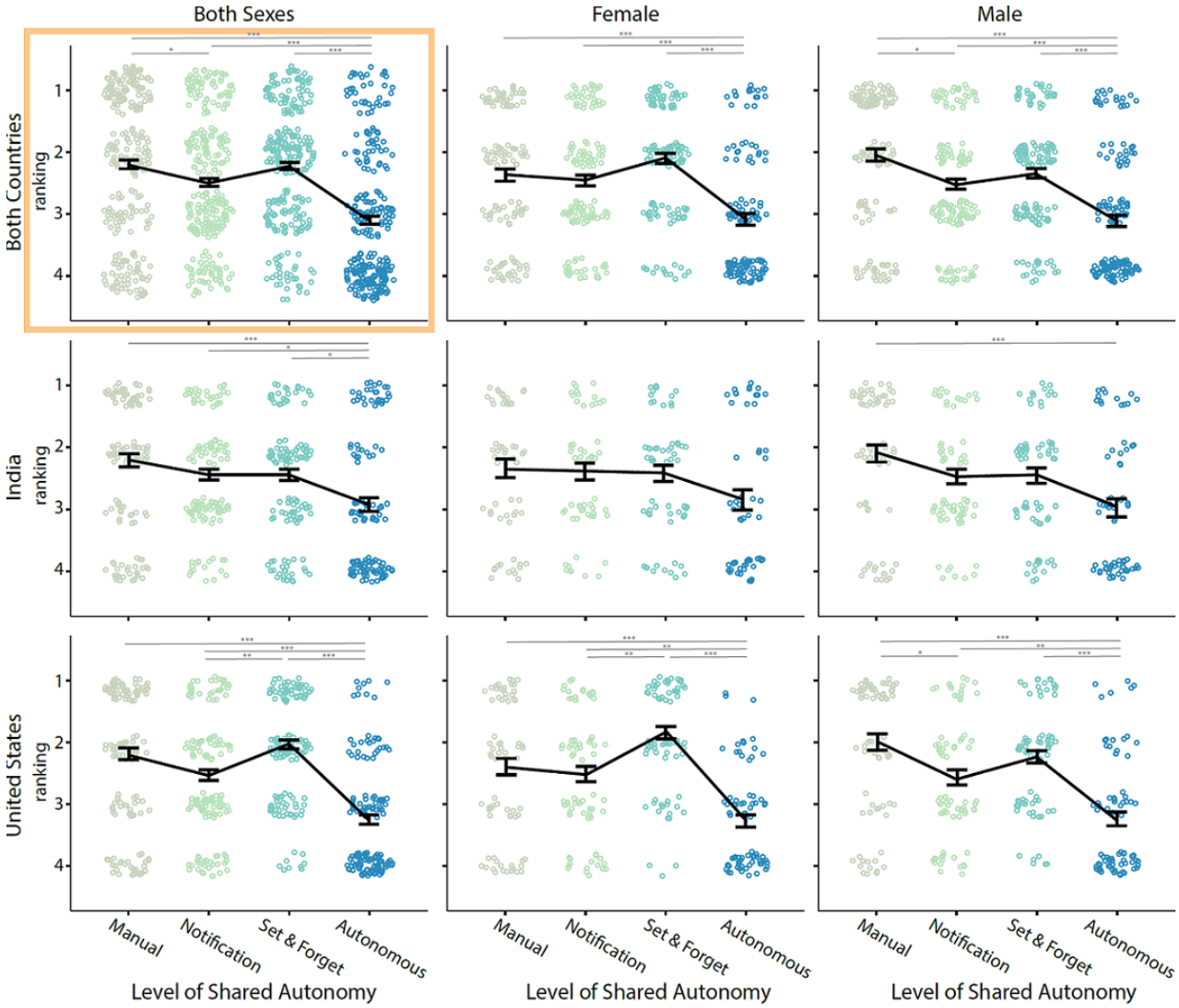
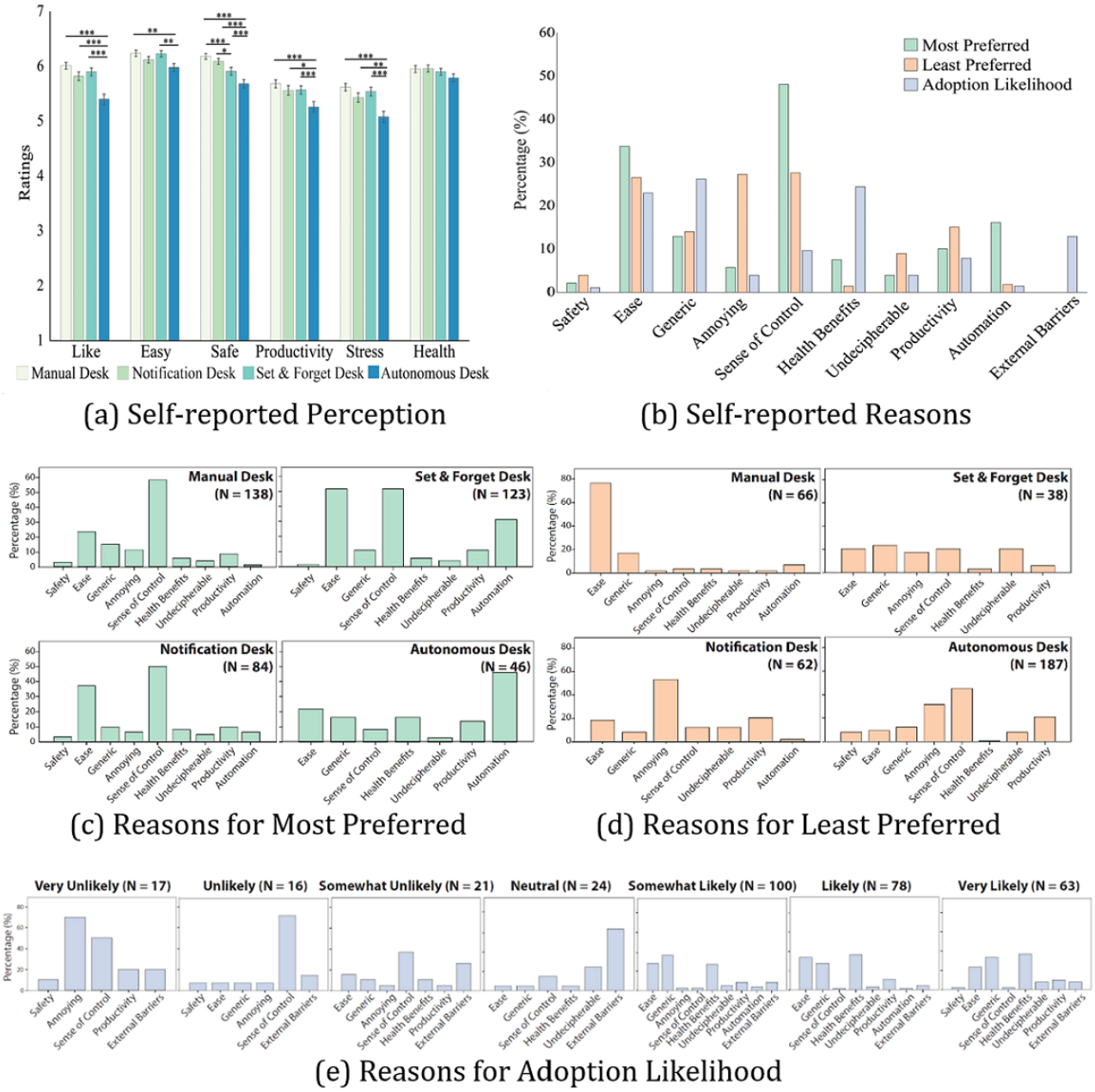


Figure 3. (A) Perception of different levels of automation in terms of likeability, ease of use, safety, productivity, stress, and health. There are statistically significant differences among the 4 desks for all measures except health. The means and SEs are reported. (*.01≤ *P*<.05, **.001≤ *P*<.01, ***P<.001) (B) Overall Breakdown of self-reported reasons for the participants' shared autonomy preferences and the likelihood of adopting autonomous sit- stand desks. Ease, annoyance, sense of control, and health benefits were the most frequent reasons. Detailed breakdown of self-reported reasons for (C) the most preferred level of automation, (D) the least preferred level of automation, and (E) adoption likelihood.



Reasons for Shared Autonomy Preference and Adoption Likelihood

Figure 3B plots the breakdown of the reasons for participants' preferences and adoption likelihood. Sense of control and ease were the 2 most frequent factors for the most preferred level of shared autonomy, whereas annoyance was also a frequent factor for the least preferred level of automation. For the adoption likelihood, most of the participants selected the health benefits and ease of the autonomous desk as their primary reasons.

Q2: Effects of Demographics and Familiarity

To answer Q2, we investigated the effects of countries, sex, and familiarity on user-shared autonomy preferences.

Country

As shown on the left side of Figure 2, participants from India ranked the autonomous desk higher (χ^2_1=11.0; *P*<.001) and the set-and-forget desk lower (χ^2_1=4.1; *P*=.04) than participants from the United States. Participants from India also preferred more frequent height changes than those from the United States (χ^2_1=7.7; *P*=.005).

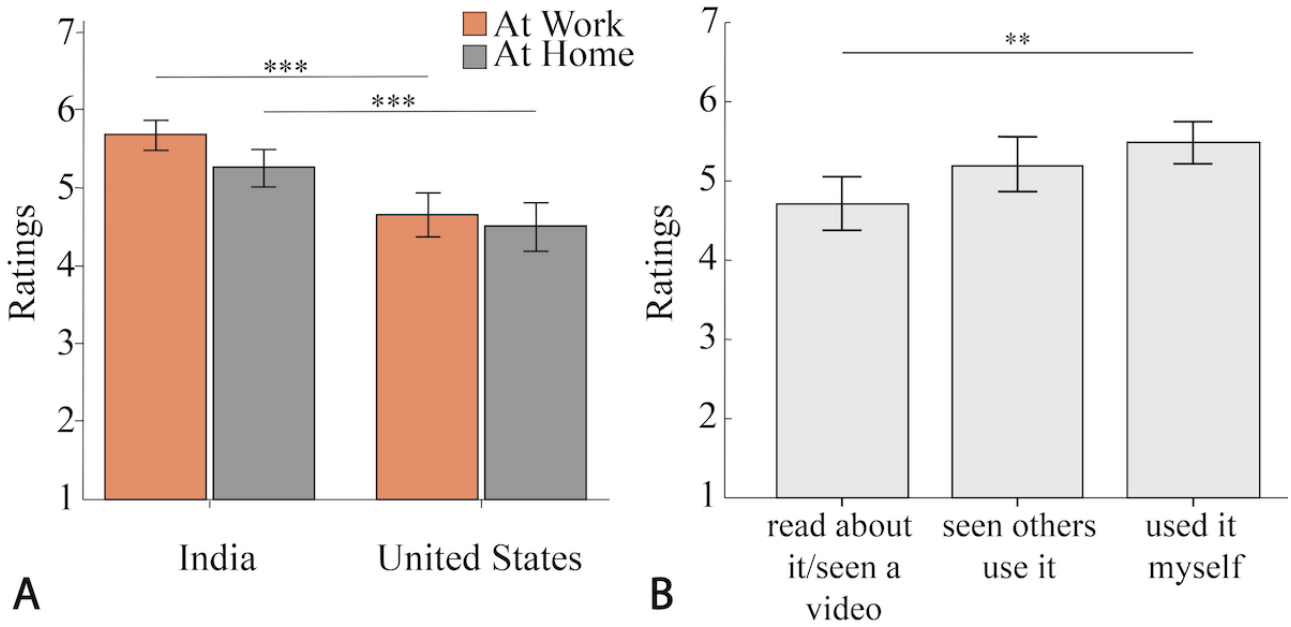
Participants from India were younger than those from the United States, as shown in Table 1. Thus, to understand how age affected user ranking of levels of automation, we conducted multinomial logistic regressions with country and age group as the independent variables and the most and least preferred level

of automation as the dependent variable. For both the most and least preferred levels of automation, only country was the statistically significant independent variable with $P<.001$ and $P=.048$, respectively.

As shown in Figure 4A, people in India were more likely to adopt an autonomous desk both at work ($P<.001$) and at home ($P<.001$) than people in the United States. In terms of DC, participants from the United States had statistically higher scores (mean 102.0, SE 1.1) than participants from India (mean 99.0, SE 0.8). For RAM, participants from the United States also had statistically higher indices (mean 2.26, SE 0.15) than participants from India (mean 1.00, SE 0.12).

Country had a statistically significant effect on perceived likeability ($F_{1,277}=12.6$; $P<.001$; $\eta^2p=0.043$), productivity ($F_{1,277}=31.0$; $P<.001$; $\eta^2p=0.101$), and stress reduction ($F_{1,277}=32.5$; $P<.001$; $\eta^2p=0.105$). Participants from India rated desks as more likable, productive, and useful in lowering stress than participants from the United States. Statistically significant interaction effects were also observed between automation and country. Statistically significant interaction effects were observed on perceived likeability ($F_{2.7,754.5}=3.4$; $P=.02$; $\eta^2p=0.012$) and productivity ($F_{2.8,762.4}=4.5$; $P=.004$; $\eta 2p=0.016$). Although the autonomous desk was rated the lowest in terms of likeability for both countries, participants from India rated the set-and-forget desks closer to the autonomous desk, whereas participants from the United States rated the notification desk closer to the autonomous desk. Participants from India reported that all desks were comparable in terms of productivity, whereas participants from the United States found the autonomous desk to be worse than the manual desk and set-and-forget desk.

Figure 4. Self-reported adoption ratings by (A) country and (B) familiarity with a sit-stand desk. Participants from India report that they are more likely to adopt autonomous (the set-and-forget desk or the autonomous desk) both at work and home than participants from the United States. Participants who have used a sit-stand desk were more likely to adopt an autonomous desk at work than participants who have only read about it or seen a video of it. The means and SEs are reported. (*.01≤ $P<.05$, **.001≤ $P<.01$, ***$P<.001$).



Sex

As shown at the top of Figure 2, male participants ranked the manual desk higher ($\chi^2_1=6.1$; $P=.01$) and the set-and-forget desk lower ($\chi^2_1=4.6$; $P=.03$) compared with female participants. However, there were no statistically significant main or interaction effects of sex on the overall impression of desks. An independent samples 2-tailed *t* test on desk adoption at home or work, DC score, and RAM index found no statistically significant differences in adoption likelihood, DC score, and RAM index.

Familiarity With Sit-Stand Desks

There were statistically significant differences among the 3 levels of familiarity with the sit-stand desk on the autonomous desk adoption Likert scale ratings ($F_{2,276}=7.0$; $P=.001$). There was a statistically significant difference between people who had read or seen a video and those who had used it before ($P=.001$), as shown in Figure 4B.

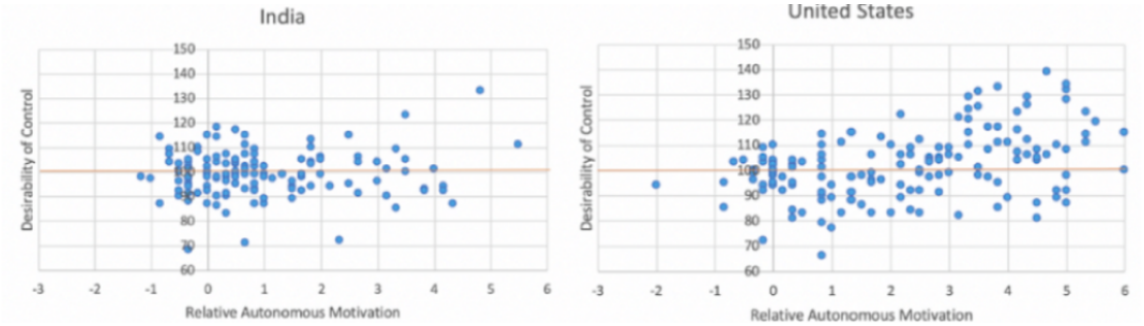
Q3: Effects of Dispositional Factors

Overview

Figure 5 shows the distributions of DC and RAM for India and the United States. There were statistically significant differences in both the DC and RAM. For DC, United States (mean 102.0, SE 0.94) had higher mean than India (mean 98.8, SE 1.0) with $P=.02$. For RAM, the US (mean 2.26, SE 0.14) also had a higher mean than India (mean 0.99, SE 0.15) with $P<.001$.

In terms of the effects of dispositional factors on gender, a statistically significant effect was observed for DC ($P=.04$), where male participants (mean 101.8, SE 0.96) had higher DC scores than female participants (mean 99.0, SE 1.0).

Figure 5. Distribution of relative autonomous motivation and desirability of control for India and the United States.



Desirability of Control

As shown in Table 3, DC had positive correlations with all impressions of the sit-stand desks, except for stress reduction in the manual desk. DC also had positive correlations with adoption likelihood at work (r_s=0.118; P=.049) and home (r_s=0.218; P<.001). In terms of ranking, DC had a positive correlation with the autonomous desk (r_s=0.119; P=.047) and a negative correlation with the notification desk (r_s=−0.140; P=.02).

Table 3. Correlation coefficients between desirability of control and impressions of the sit-stand desks.

	Manual desk	Notification desk	Set-and-forget desk	Autonomous desk
Like	0.288[a]	0.273[a]	0.188[a]	0.213[a]
Easy	0.28[a]	0.285[a]	0.279[a]	0.314[a]
Safe	0.222[a]	0.374[a]	0.291[a]	0.248[a]
Productivity	0.178[a]	0.23[a]	0.165[a]	0.131[b]
Stress	0.097	0.21[a]	0.231[a]	0.129[b]
Health	0.229[a]	0.303[a]	0.204[a]	0.217[a]
Ranking	−0.035	−0.14[b]	0.081	0.119[b]

[a]P<.01.
[b].01≤P<.05.

Self-regulation

As shown in Table 4, the RAM index had positive correlations with ease and safety ratings for all levels of automation but had negative correlations with productivity for notification and autonomous desks and with stress for notification, set-and-forget, and autonomous desks. For adoption likelihood, the RAM index had negative correlations at work (r_s=−0.242; P<.001) and home (r_s=−0.204; P<.001) but had a positive correlation with the preferred height change frequency (r_s=0.152; P=.01). In terms of ranking, the RAM index had a positive correlation with the set-and-forget desk (r_s=−0.138; P=.02).

Table 4. Correlation coefficients between relative autonomous motivation index and impressions of the sit-stand desks.

	Manual desk	Notification desk	Set-and-forget desk	Autonomous desk
Like	0.096	−0.022	0.084	−0.043
Easy	0.213[a]	0.152[b]	0.169[a]	0.158[a]
Safe	0.257[a]	0.275[a]	0.186[a]	0.15b
Productivity	−0.115	−0.155[a]	−0.106	−0.256[a]
Stress	−0.086	−0.126[b]	−0.164[a]	−0.186[a]
Health	0.017	0.123[b]	0.144[b]	0.036
Ranking	0.052	0.016	−0.138[b]	0.063

[a]P<.01.
[b].01≤P<.05.

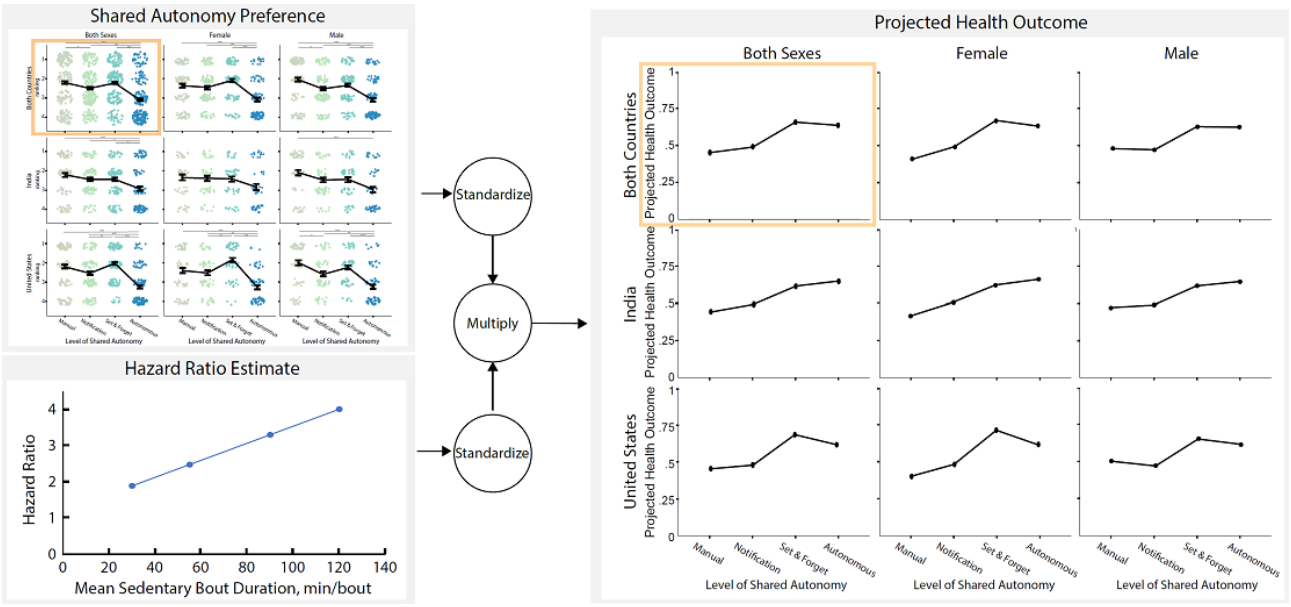
Q4: Effects of Goal Priming

There were no statistically significant effects of priming on autonomous desk adoption Likert scale ratings both at home and at work (P=.14 and P=.83, respectively), the rankings of desks (manual P=.39, notification P=.43, set and forget P=.12, and autonomous P=.14), and the preferred frequency (P=.86) for height.

Q5: Projected Health Outcome

As shown in Figure 6, the set-and-forget desks and autonomous desks have the highest projected overall health outcomes based on the potential use that is estimated using the ranking data of each level of automation. For participants from India, the autonomous desk has the highest projected health outcome, whereas the set-and-forget desk has the highest projected health outcome for participants from the United States.

Figure 6. Projected health outcomes were estimated using our user preference data and the hazard ratios that were adapted using the data from Diaz et al [9].



Discussion

Summary of Findings

The aim of Q1 of this study is to explore shared autonomy in sit-stand desks and study user preferences. From the overall results, finding a simple answer to Q1 is not straightforward. Similar to a previous study [20], we found that users desire some degree of control over height changes despite being aware of the health benefits of using autonomous sit-stand desks. Most participants preferred either the manual desk or the set-and-forget desk, which asked users once a day for their desired height switch frequency. Sense of control was cited as the most frequent reason for both desks, and ease of use was another frequently mentioned factor for the semiautonomous condition. These findings align well with the results from Wunderlich et al [31], where control was found to be an important aspect for smart interactive services. Overall, a fully autonomous intervention, where users have no control, may not be accepted or adopted, as demonstrated with the autonomous sit-stand desk. However, even a minor sense of control (ie, daily selection of the sit or stand frequency) can provide sufficient motivation to adhere to an intervention that remains reasonably autonomous for the duration of the entire workday.

Demographics had a significant impact on user perceptions and reactions to sit-stand desks. Culture is an important factor in this process. The study results provide a convincing answer to Q2 that participants from India are much more receptive to higher levels of automation for sit-stand desks than are participants from the United States. Participants from India also rated the likeability of the notification desk and the autonomous desk higher than participants from the United States. They also rated all desks comparable in terms of productivity, whereas participants from the United States rated the autonomous desk as less conducive for productivity than the manual desk or the set-and-forget desk. Participants from India were more likely to report being willing to adopt semiautonomous or autonomous desks than participants from the United States. Participants from India also ranked the autonomous desk higher, but the set-and-forget desk lower than participants from the United States. The DC scores and RAM index also shed light on the reason for this difference. As shown in prior surveys [56], the United States has a lower power distance index than India, suggesting a limited dependence of subordinates on their bosses. Similarly, we see that participants from the United States have a higher desirability for control and higher RAM than participants from India. Similar to prior work [57,58], these findings reinforce the importance of tailoring to each culture to maximize the likelihood of adoption.

Sex also significantly affected user impression of the autonomous interventions. We observed differences between male and female participants' preferences for the level of automation in that male participants in both countries preferred the manual one significantly more than female participants, whereas female participants preferred the semiautonomous set-and-forget desk more, providing insights on Q2. This trend contradicts the findings of Kotze et al [41], who found that

female participants are less optimistic than male participants and exhibit higher levels of risk aversion toward high-technology products [42] but aligns well with previous results with digital media, where female participants value its perceived ease of use more than male participants [43,44]. The discrepancy between male and female participants' preferences could be because of male participants being more wary of losing control over the desk than female participants. Further studies are needed to confirm this hypothesis.

Familiarity with a sit-stand desk also influences how likely users are to adopt autonomous desks at work and at home. Participants who had personal experience using a sit-stand desk were more likely to rate the adoption likelihood higher than participants who had only read or seen one. This finding is in line with previous work, where familiarity has improved the adoption of wellness games among older adults [59] and the adoption of recommendation agents [26]. This suggests that incremental increases in autonomous intervention using a sit-stand desk could build trust and increase adoption likelihood.

Overall, for Q3, we observed mild effects of DC and TSRQ on the overall population. However, the differences observed in the distribution of TSRQ per country (Figure 5) may help explain the preference in India for the autonomous desk. When looking into the Spearman correlation analysis for these 2 metrics per country, we observe that in the case of India, DC is positively correlated with the adoption likelihood at home (r_s=0.5; P<.001) and at work (r_s=0.428; P<.001), while RAM (TSRQ) is negatively correlated with the adoption likelihood in the United States at home (r_s=−0.262; P=.001) and at work (r_s=−0.21; P=.01). Furthermore, DC is positively correlated with ranking preferences for Desk D (r_s=0.167; P<.05). Overall, it seems like a culture with more homogeneous and lower levels of DC, such as India, compared with a culture with a higher variety of RAM (TSRQ) could have a better predisposition toward adoption of nonvolitional behavior change technologies.

In terms of priming, we found no evidence to suggest that it has any influence on user-shared autonomy preferences when answering Q4. Compared with the condition without priming, conditions with loss-framed or gain-framed priming did not lead to any differences in participants' preference for shared autonomy or their adoption likelihood. This contradicts findings from prior work [34] and indicates that priming alone is not sufficient to persuade users to accept this type of autonomous intervention.

Finally, we estimated the projected health outcome (Figure 6) for different levels of automation across cultures and sexes based on hazard ratio estimates and user preferences, answering Q5. Although there are some potential limitations because of some assumptions made (eg, using a 90-minute bout duration for the notification desk), it helps us understand the potential benefits of embedding automation into sit-stand desks. Although users do not always prefer higher levels of automation, embedding automation has great potential in reducing sedentary behaviors, thus lowering the hazard ratio estimate and ultimately leading to a higher level of projected health outcomes. To verify this, we plan to follow up with a longitudinal study.

Design Insights

On the basis of these findings, we present a brief series of insights toward more effective autonomous intervention designs.

A sense of control is an essential factor in the initial impression and adoption likelihood of autonomous interventions. Given the various preferences observed, we recommend offering several different automation modes to autonomous sit-stand desks and similar automated furniture. It is possible that providing some degree of control initially may be the best strategy to help users become comfortable with the idea of a fully autonomous intervention.

To predict the amount of control users might desire or to determine what types of intervention modes to support, demographics may be useful information to obtain, given our observations of different segments of our populations.

Our findings suggest that it is important to provide a clear explanation of the benefits of behavior change. For example, to promote the adoption of an autonomous sit-stand desk, it is important to provide a clear explanation of the health benefits. Although health benefits were not the primary factor for users' preferences, they were one of the most frequently cited reasons for adoption, as shown in Figure 3E. Although priming on the health benefits per se did not impact users' preferences or adoption likelihood, participants might have already been aware of the health benefits as they had some degree of prior exposure to a sit-stand desk.

Ethical Considerations for Future Deployment

By design, this technology can only be deployed in a home setting with full disclosure by the user because of the agency people have when purchasing a desk. However, in situations where an employer may require their workers to such a desk, there are important ethical concerns about the loss of autonomy that the workers would experience. Therefore, it is paramount that when introducing technology that enforces nonvolitional behavior change, the user has complete disclosure concerning the automation of the technology and the extent to which their autonomy will be limited. Designers are wary of technologies designed to trick users, such as technologies that use dark patterns and purposeful design elements to mislead people in a certain direction [60]. In contrast, the consensual loss of autonomy and transparency, which is critical to our design, separates our approach from the technology that attempts to trick the user.

Limitations

One main limitation of this study is that we used web-based video vignettes instead of in-person interaction to indirectly convey the user experience of using sit-stand desks of different automation levels. As previous work has shown comparable results between vignette-based and in-person user evaluations [35,36,50], we believe vignettes can be sufficient to obtain the first impression of sit-stand desks and nonvolitional behavior change. However, to accurately measure the ecological validity and adherence to continuous use of our sit-stand desk prototypes, we plan to conduct a longitudinal user study based on the findings of this study.

Another limitation is that we did not incorporate the optimal timing for the height change intervention. Rather, the intervention was given at preset intervals for both the formative and main web-based studies. Previous work suggests that providing interventions during a task change is ideal for both a sit-stand desk [61] and graphical user interface [62]. Although our participants cited a lack of sense of control as the primary reason for disliking the autonomous condition, providing interventions at the ideal timing may improve user's impression of autonomous sit-stand desks, especially during longitudinal studies.

Future Research

For longitudinal studies, our findings demonstrate the necessity to consider the background of the participants, such as culture, sex, and familiarity with sit-stand desks, as it significantly impacts how they will react to automation. For instance, as we observed a linear relationship between familiarity with sit-stand desks and the adoption likelihood of autonomous sit-stand desks, it may be prudent to start with a lower level of automation for participants new to sit-stand desks, whereas participants who are already familiar could be more accepting of higher levels of automation from the beginning.

In addition to culture, sex, and familiarity, other factors should be investigated in the future. For example, it would be interesting to study the effects of income, education level, or age and whether these would be better predictors for user acceptance of autonomous interventions for physical well-being.

This study aimed to explore the possibility of autonomous behavior changes using robotic furniture. Although we began, for simplicity, with a 1 *df* sit-stand desk that can only change its height, this concept of autonomous intervention could be applied to a wide range of actuated objects. There are already instances of robotic furniture in the form of ottomans [63] and computer monitors [61]. Even with the new actuated furniture with additional *df*s, we expect that our study findings will hold regarding user perceptions and reactions to different levels of automation. In addition, we expect the demographics of the users to play a role. In the future, we plan to investigate using a wider range of actuated objects to encourage healthy behavior in users.

Conclusions

With automation becoming increasingly embedded in our environment, it is important to consider how we can best leverage it to improve the mental and physical well-being of people. In this study, we investigated user perception and preference at the level of automation in the context of sit-stand desks. The results suggest that, despite being aware of the health benefits and effectiveness of autonomous interventions, participants regarded having a sense of control over the desk as an important factor. Culture and sex significantly affect the importance of this sense of control in adopting an autonomous desk. As we observed positive effects from familiarity, we believe that a gradual approach with incremental exposure to autonomous behavior change will be suitable for populations less receptive to automation. We hope that this work will spur more research into shared autonomy-augmented health behavior changes with robotic furniture and bring us closer to a future of well-being–driven physical computing.

Acknowledgments

The authors thank all the volunteers who participated in this study. This work was supported by the Stanford Catalyst for Collaborative Solutions Award led by Scott Delp. PEP was supported by the Precision Health and Integrated Diagnostics Center and the Center for Population Health Sciences at Stanford University. Contributions from LHK, MLM, and PEP were made during the transition from Stanford University to Simon Fraser University, University of Delaware, and the University of Maryland, respectively.

Conflicts of Interest

None declared.

References

1. Straker L, Mathiassen SE. Increased physical work loads in modern work--a necessity for better health and performance? Ergonomics 2009 Oct;52(10):1215-1225. [doi: 10.1080/00140130903039101] [Medline: 19787501]
2. van Dommelen P, Coffeng JK, van der Ploeg HP, van der Beek AJ, Boot CR, Hendriksen IJ. Objectively measured total and occupational sedentary time in three work settings. PLoS One 2016 Mar 3;11(3):e0149951 [FREE Full text] [doi: 10.1371/journal.pone.0149951] [Medline: 26937959]
3. Healy GN, Dunstan DW, Salmon J, Cerin E, Shaw JE, Zimmet PZ, et al. Breaks in sedentary time: beneficial associations with metabolic risk. Diabetes Care 2008 Apr;31(4):661-666. [doi: 10.2337/dc07-2046] [Medline: 18252901]
4. Owen N, Healy GN, Matthews CE, Dunstan DW. Too much sitting: the population health science of sedentary behavior. Exerc Sport Sci Rev 2010 Jul;38(3):105-113 [FREE Full text] [doi: 10.1097/JES.0b013e3181e373a2] [Medline: 20577058]
5. Sjögren P, Fisher R, Kallings L, Svenson U, Roos G, Hellénius ML. Stand up for health--avoiding sedentary behaviour might lengthen your telomeres: secondary outcomes from a physical activity RCT in older people. Br J Sports Med 2014 Oct;48(19):1407-1409. [doi: 10.1136/bjsports-2013-093342] [Medline: 25185586]
6. Lee E, Kim Y. Effect of university students' sedentary behavior on stress, anxiety, and depression. Perspect Psychiatr Care 2019 Apr;55(2):164-169 [FREE Full text] [doi: 10.1111/ppc.12296] [Medline: 29797324]

7. Connelly J, Kirk A, Masthoff J, MacRury S. The use of technology to promote physical activity in type 2 diabetes management: a systematic review. Diabet Med 2013 Dec;30(12):1420-1432. [doi: 10.1111/dme.12289] [Medline: 23870009]
8. Diaz KM, Howard VJ, Hutto B, Colabianchi N, Vena JE, Safford MM, et al. Duration of sedentary episodes is associated with risk for death. Ann Intern Med 2017 Oct 03;167(7):1-24 [FREE Full text] [doi: 10.7326/P17-9045] [Medline: 28892830]
9. Diaz KM, Howard VJ, Hutto B, Colabianchi N, Vena JE, Safford MM, et al. Patterns of sedentary behavior and mortality in U.S. middle-aged and older adults: a national cohort study. Ann Intern Med 2017 Oct 03;167(7):465-475 [FREE Full text] [doi: 10.7326/M17-0212] [Medline: 28892811]
10. Renaud LR, Huysmans MA, van der Ploeg HP, Speklé EM, van der Beek AJ. Long-term access to sit-stand workstations in a large office population: user profiles reveal differences in sitting time and perceptions. Int J Environ Res Public Health 2018 Sep 15;15(9):2019 [FREE Full text] [doi: 10.3390/ijerph15092019] [Medline: 30223582]
11. Wilks S, Mortimer M, Nylén P. The introduction of sit-stand worktables; aspects of attitudes, compliance and satisfaction. Appl Ergon 2006 May;37(3):359-365. [doi: 10.1016/j.apergo.2005.06.007] [Medline: 16102719]
12. Chau JY, Daley M, Srinivasan A, Dunn S, Bauman AE, van der Ploeg HP. Desk-based workers' perspectives on using sit-stand workstations: a qualitative analysis of the Stand@Work study. BMC Public Health 2014 Jul 25;14:752 [FREE Full text] [doi: 10.1186/1471-2458-14-752] [Medline: 25059500]
13. Hagger MS, Wood C, Stiff C, Chatzisarantis NL. Ego depletion and the strength model of self-control: a meta-analysis. Psychol Bull 2010 Jul;136(4):495-525. [doi: 10.1037/a0019486] [Medline: 20565167]
14. Baumeister RF, Vohs KD, Tice DM. The strength model of self-control. Curr Dir Psychol Sci 2007 Dec;16(6):351-355. [doi: 10.1111/j.1467-8721.2007.00534.x]
15. Venema TA, Kroese FM, De Ridder DT. I'm still standing: a longitudinal study on the effect of a default nudge. Psychol Health 2018 May;33(5):669-681. [doi: 10.1080/08870446.2017.1385786] [Medline: 28992735]
16. Thaler RH, Sunstein CR. Nudge: Improving Decisions About Health, Wealth, and Happiness. New York, NY, USA: Penguin Publishing Group; 2009.
17. Nikolaidis S, Zhu YX, Hsui D, Srinivasa S. Human-robot mutual adaptation in shared autonomy. In: Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction. 2017 Mar Presented at: HRI '17; March 6-9, 2017; Vienna, Austria p. 294-302. [doi: 10.1145/2909824.3020252]
18. Gopinath D, Jain S, Argall BD. Human-in-the-loop optimization of shared autonomy in assistive robotics. IEEE Robot Autom Lett 2017 Jan;2(1):247-254 [FREE Full text] [doi: 10.1109/LRA.2016.2593928] [Medline: 30662953]
19. Papies EK. Health goal priming as a situated intervention tool: how to benefit from nonconscious motivational routes to health behaviour. Health Psychol Rev 2016 Dec;10(4):408-424 [FREE Full text] [doi: 10.1080/17437199.2016.1183506] [Medline: 27144729]
20. Kim LH, Leon AA, Sankararaman G, Jones BM, Saha G, Spyropolous A, et al. The haunted desk: exploring non-volitional behavior change with everyday robotics. In: Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction. 2021 Presented at: HRI '21 Companion; March 8-11, 2021; Boulder, CO, USA p. 71-75. [doi: 10.1145/3434074.3447131]
21. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003 Sep;27(3):425-478. [doi: 10.2307/30036540]
22. Hoff KA, Bashir M. Trust in automation: integrating empirical evidence on factors that influence trust. Hum Factors 2015 May;57(3):407-434. [doi: 10.1177/0018720814547570] [Medline: 25875432]
23. Im I, Hong S, Kang MS. An international comparison of technology adoption: testing the UTAUT model. Inf Manag 2011 Jan;48(1):1-8. [doi: 10.1016/j.im.2010.09.001]
24. Kuo IH, Rabindran JM, Broadbent E, Lee YI, Kerse N, Stafford RM, et al. Age and gender factors in user acceptance of healthcare robots. In: Proceedings of the RO-MAN 2009 - The 18th IEEE International Symposium on Robot and Human Interactive Communication. 2009 Presented at: ROMAN '09; September 27-October 2, 2009; Toyama, Japan p. 214-219. [doi: 10.1109/ROMAN.2009.5326292]
25. Siegel M, Breazeal C, Norton MI. Persuasive Robotics: the influence of robot gender on human behavior. In: Proceedings of the 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems. 2009 Presented at: IROS '09; October 10-15, 2009; St. Louis, MO, USA p. 2563-2568. [doi: 10.1109/IROS.2009.5354116]
26. Komiak SY, Benbasat I. The effects of personalization and familiarity on trust and adoption of recommendation agents. MIS Q 2006 Dec;30(4):941-960. [doi: 10.2307/25148760]
27. Aarts H. Health and goal-directed behavior: the nonconscious regulation and motivation of goals and their pursuit. Health Psychol Rev 2007 Mar;1(1):53-82. [doi: 10.1080/17437190701485852]
28. Hagger MS. Sleep, self-regulation, self-control and health. Stress Health 2010;26(3):181-185. [doi: 10.1002/smi.1345]
29. de Ridder DT, de Wit JB. Self-Regulation in Health Behavior. Hoboken, NJ, USA: Wiley Online Library; 2006.
30. Hall PA, Fong GT. Temporal self-regulation theory: a model for individual health behavior. Health Psychol Rev 2007 Mar;1(1):6-52. [doi: 10.1080/17437190701492437]
31. Wünderlich NV, Wangenheim FV, Bitner MJ. High tech and high touch: a framework for understanding user attitudes and behaviors related to smart interactive services. J Serv Res 2012 Jun 11;16(1):3-20. [doi: 10.1177/1094670512448413]

32. Dvash A, Mannheim B. Technological coupling, job characteristics and operators' well-being as moderated by desirability of control. Behav Inf Technol 2001;20(3):225-236. [doi: 10.1080/01449290116930]
33. Consolvo S, McDonald DW, Toscos T, Chen MY, Froehlich J, Harrison B, et al. Activity sensing in the wild: a field trial of ubifit garden. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2008 Presented at: CHI '08; April 5-10, 2008; Florence, Italy p. 1797-1806.
34. Chen FX, King AC, Hekler EB. "Healthifying" exergames: improving health outcomes through intentional priming. In: Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems,. 2014 Presented at: CHI '14; April 26-May 1, 2014; Toronto, Canada p. 1855-1864.
35. Kim LH, Follmer S. Ubiswarm: ubiquitous robotic interfaces and investigation of abstract motion as a display. Proc ACM Interact Mob Wearable Ubiquitous Technol 2017 Sep 11;1(3):1-20. [doi: 10.1145/3130931]
36. Takayama L, Dooley D, Ju W. Expressing thought: improving robot readability with animation principles. In: Proceedings of the 6th ACM/IEEE International Conference on Human-Robot Interaction. 2011 Presented at: HRI '11; March 8-11, 2011; Lausanne, Switzerland p. 69-76. [doi: 10.1145/1957656.1957674]
37. Awad E, Levine S, Kleiman-Weiner M, Dsouza S, Tenenbaum JB, Shariff A, et al. Drivers are blamed more than their automated cars when both make mistakes. Nat Hum Behav 2020 Feb;4(2):134-143. [doi: 10.1038/s41562-019-0762-8] [Medline: 31659321]
38. Kitagawa R, Liu Y, Kanda T. Human-inspired motion planning for omni-directional social robots. In: Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction. 2021 Presented at: HRI '21; March 8-11, 2021; Boulder, CO, USA p. 34-42 URL: https://doi.org/10.1145/3434073.3444679 [doi: 10.1145/3434073.3444679]
39. Robot density rises globally. IFR - International Federation of Robotics. URL: https://ifr.org/ifr-press-releases/news/robot-density-rises-globally [accessed 2021-11-11]
40. Ramesh A, Gelfand MJ. Will they stay or will they go? The role of job embeddedness in predicting turnover in individualistic and collectivistic cultures. J Appl Psychol 2010 Sep;95(5):807-823. [doi: 10.1037/a0019464] [Medline: 20836585]
41. Kotze TG, Anderson O, Summerfield K. Technophobia: gender differences in the adoption of high-technology consumer products. South Afr J Bus Manag 2016;47(1):21-28 [FREE Full text]
42. Simon SJ, Peppas SC. Attitudes towards product website design: a study of the effects of gender. J Mark Commun 2005 Jun;11(2):129-144. [doi: 10.1080/1352726042000286507]
43. Venkatesh V, Morris MG. Why don't men ever stop to ask for directions? Gender, social influence, and their role in technology acceptance and usage behavior. MIS Q 2000 Mar;24(1):115-139. [doi: 10.2307/3250981]
44. Ong CS, Lai JY. Gender differences in perceptions and relationships among dominants of e-learning acceptance. Comput Human Behav 2006 Sep;22(5):816-829. [doi: 10.1016/j.chb.2004.03.006]
45. Levesque CS, Williams GC, Elliot D, Pickering MA, Bodenhamer B, Finley PJ. Validating the theoretical structure of the Treatment Self-Regulation Questionnaire (TSRQ) across three different health behaviors. Health Educ Res 2007 Oct;22(5):691-702. [doi: 10.1093/her/cyl148] [Medline: 17138613]
46. Williams GC, Grow VM, Freedman ZR, Ryan RM, Deci EL. Motivational predictors of weight loss and weight-loss maintenance. J Pers Soc Psychol 1996 Jan;70(1):115-126. [doi: 10.1037//0022-3514.70.1.115] [Medline: 8558405]
47. Deci EL, Ryan RM. Intrinsic Motivation and Self-Determination in Human Behavior. New York, NY, USA: Springer; 1985.
48. Vaz A, Pratley P, Alkire S. Measuring women's autonomy in chad using the relative autonomy index. Fem Econ 2016;22(1):264-294. [doi: 10.1080/13545701.2015.1108991]
49. Burger JM, Cooper HM. The desirability of control. Motiv Emot 1979 Dec;3(4):381-393. [doi: 10.1007/BF00994052]
50. Woods SN, Walters ML, Koay KL, Dautenhahn K. Methodological issues in HRI: a comparison of live and video-based methods in robot to human approach direction trials. In: Proceedings of the ROMAN 2006 - The 15th IEEE International Symposium on Robot and Human Interactive Communication. 2006 Presented at: ROMAN '06; September 6-8, 2006; Hatfield, UK p. 51-58. [doi: 10.1109/ROMAN.2006.314394]
51. Xu Q, Ng J, Cheong YL, Tan O, Wong JB, Tay TC, et al. Effect of scenario media on elder adults' evaluation of human-robot interaction. In: Proceedings of the 2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication. 2012 Presented at: ROMAN '12; September 9-13, 2019; Paris, France p. 455-462. [doi: 10.1109/ROMAN.2012.6343794]
52. Bainbridge WA, Hart JW, Kim ES, Scassellati B. The benefits of interactions with physically present robots over video-displayed agents. Int J Soc Robotics 2011;3(1):41-52. [doi: 10.1007/s12369-010-0082-7]
53. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd edition. Cambridge, MA, USA: Academic Press; 1988.
54. Miles J, Shevlin M. Applying Regression and Correlation: A Guide for Students and Researchers. Thousand Oaks, CA, USA: Sage Publications; 2001.
55. Sharma PP, Mehta RK, Pickens A, Han G, Benden M. Sit-stand desk software can now monitor and prompt office workers to change health behaviors. Hum Factors 2019 Aug;61(5):816-824 [FREE Full text] [doi: 10.1177/0018720818807043] [Medline: 30296174]

56. Hofstede G, Hofstede GJ. Cultures and Organizations: Software for the Mind. New York, NY, USA: McGraw Hill Professional; 2005.
57. E JL, E IL, Landay JA, Cauchard JR. Drone and Wo: cultural influences on human-drone interaction techniques. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 2017 Presented at: CHI '17; May 6-11, 2017; Denver, CO, USA p. 6794-6799 URL: https://doi.org/10.1145/3025453.3025755 [doi: 10.1145/3025453.3025755]
58. Lee SG, Trimi S, Kim C. The impact of cultural differences on technology adoption. J World Bus 2013 Jan;48(1):20-29. [doi: 10.1016/j.jwb.2012.06.003]
59. Pan Z, Miao C, Yu H, Leung C, Chin JJ. The effects of familiarity design on the adoption of wellness games by the elderly. In: Proceedings of the 2015 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology. 2015 Presented at: WI-IAT '15; December 6-9, 2015; Singapore, Singapore p. 387-390.
60. Bösch C, Erb B, Kargl F, Kopp H, Pfattheicher S. Tales from the dark side: privacy dark strategies and privacy dark patterns. Proc Priv Enhanc Technol 2016;2016(4):237-254. [doi: 10.1515/popets-2016-0038]
61. Lee B, Wu S, Reyes MJ, Saakes D. The effects of interruption timings on autonomous height-adjustable desks that respond to task changes. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. 2019 Presented at: CHI '19; May 4-9, 2019; Glasgow, UK p. 1-10. [doi: 10.1145/3290605.3300558]
62. Adamczyk PD, Bailey BP. If not now, when? The effects of interruption at different moments within task execution. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2004 Presented at: CHI '04; April 24-29, 2004; Vienna, Austria p. 271-278. [doi: 10.1145/985692.985727]
63. Li J, Cuadra A, Mok B, Reeves B, Kaye J, Ju W. Communicating dominance in a nonanthropomorphic robot using locomotion. J Hum Robot Interact 2019 Mar 22;8(1):1-14. [doi: 10.1145/3310357]

Abbreviations

DC: desirability of control
HRI: human-robot interaction
RAM: relative autonomous motivation
TSRQ: Treatment Self-Regulation Questionnaire

Edited by A Mavragani; submitted 20.12.21; peer-reviewed by C Kirkland, L Schumacher; comments to author 10.03.22; revised version received 03.06.22; accepted 26.07.22; published 09.11.22.

Please cite as:
Kim LH, Saha G, Leon AA, King AC, Mauriello ML, Paredes PE
Shared Autonomy to Reduce Sedentary Behavior Among Sit-Stand Desk Users in the United States and India: Web-Based Study
JMIR Form Res 2022;6(11):e35447
URL: https://formative.jmir.org/2022/11/e35447
doi:10.2196/35447
PMID:36350687

Original Paper

Digital Technology in Psychiatry: Survey Study of Clinicians

William Andrew Sterling[1,2,3], MD; Michael Sobolev[4], PhD; Anna Van Meter[1,2,4], PhD; Daniel Guinart[1,2,4], MD, PhD; Michael L Birnbaum[1,2,4], MD; Jose M Rubio[1,2,4], MD; John M Kane[1,2,4], MD

[1]Institute of Behavioral Science, Department of Psychiatry, Zucker Hillside Hospital, Glen Oaks, NY, United States

[2]The Donald and Barbara Zucker School of Medicine, Hofstra University, Northwell Health, New York, NY, United States

[3]Department of Psychiatry, Grossman School of Medicine, New York University Langone Health, New York, NY, United States

[4]Center for Psychiatric Neuroscience, Feinstein Institute for Medical Research, Manhasset, NY, United States

Corresponding Author:
Daniel Guinart, MD, PhD
Institute of Behavioral Science
Department of Psychiatry
Zucker Hillside Hospital
75-59 263rd St
Glen Oaks, NY, 11004
United States
Phone: 1 718 470 8005
Email: daniguinart@gmail.com

Abstract

Background: Digital technology has the potential to transform psychiatry, but its adoption has been limited. The proliferation of telepsychiatry during the COVID-19 pandemic has increased the urgency of optimizing technology for clinical practice. Understanding clinician attitudes and preferences is crucial to effective implementation and patient benefit.

Objective: Our objective was to elicit clinician perspectives on emerging digital technology.

Methods: Clinicians in a large psychiatry department (inpatient and outpatient) were invited to complete a web-based survey about their attitudes toward digital technology in practice, focusing on implementation, clinical benefits, and expectations about patients' attitudes. The survey consisted of 23 questions that could be answered on either a 3-point or 5-point Likert scale. We report the frequencies and percentages of responses.

Results: In total, 139 clinicians completed the survey—they represent a variety of years of experience, credentials, and diagnostic subspecialties (response rate 69.5%). Overall, 83.4% (n=116) of them stated that digital data could improve their practice, and 23.0% (n=32) of responders reported that they had viewed patients' profiles on social media. Among anticipated benefits, clinicians rated symptom self-tracking (n=101, 72.7%) as well as clinical intervention support (n=90, 64.7%) as most promising. Among anticipated challenges, clinicians mostly expressed concerns over greater time demand (n=123, 88.5%) and whether digital data would be actionable (n=107, 77%). Furthermore, 95.0% (n=132) of clinicians expected their patients to share digital data.

Conclusions: Overall, clinicians reported a positive attitude toward the use of digital data to not only improve patient outcomes but also highlight significant barriers that implementation would need to overcome. Although clinicians' self-reported attitudes about digital technology may not necessarily translate into behavior, our results suggest that technologies that reduce clinician burden and are easily interpretable have the greatest likelihood of uptake.

(JMIR Form Res 2022;6(11):e33676) doi:10.2196/33676

KEYWORDS

digital psychiatry; passive monitoring technology; digital phenotype; psychiatry; mental health; clinicians; clinician perspectives; digital health; physicians; psychiatrists

Introduction

Digital technology is a central feature of modern life and is becoming an increasingly prominent component of modern medical practice. In psychiatry, there is established evidence that digital data can be used to monitor multiple health outcomes. Metrics including the frequency of web-based activity, the content and language uploaded to social media, and smartphone sensing of biometrics including sleep and

physical activity have strong predictive value for medication compliance, current mental status, or risk of relapse and can be used to guide treatment decisions effectively [1-5]. Mental health apps show broad applications across many patient demographics and diagnostic categories [6]. The integration of digital data into clinical practice represents not only a new means of interacting with patients—allowing clinicians to monitor dynamic symptoms in real time, providing patients and clinicians alike early notice of relapse and reducing latency during intervention—but also, indeed, a new way of understanding patients, thus synthesizing symptoms data into individualized profiles for each unique patient who uses the technology—a digital phenotype [7]. However, while academic research has explored patients' digital phenotypes, this work has not yet fully integrated multiple streams of digital data, nor has it been deployed systematically in actual clinical settings [8].

Clinician attitudes toward the use of digital data in psychiatry—in terms of benefits to clinical care, barriers to effective implementation, understanding of the capabilities of current technology, and willingness to change current practice—are a key factor in the development and adoption of future technological platforms, but they remain underexplored and underappreciated [9-11]. Indeed, clinician enthusiasm has been shown to be a vital factor to successful implementation of new digital platforms in clinical settings [12]. Studies of telepsychiatry suggest that successful implementation of digital technology into clinical settings depends on clinician enthusiasm and confidence in the safety and efficacy of the new platform, as well as access to proper training [13-15]. Previous studies of clinicians' attitudes toward digital data in psychiatry indicate some enthusiasm regarding administrative improvements that technology could provide—including ease of scheduling and monitoring patients between sessions—but strong concerns about privacy, and limited understanding of the capabilities of digital platforms has been demonstrated previously [16]. Elsewhere, surveys of clinicians have demonstrated a negative correlation between professional experience and clinicians' attitudes toward the use digital data in psychiatry [17]. To develop a successful platform for clinicians to use digital data in practice, researchers must determine not only clinicians' attitudes toward the technology but also their literacy about its efficacy and potential as a clinical aid.

Although digital mental health data have been researched for over a decade, their development has recently become urgent owing to the COVID-10 pandemic. Clinical practices worldwide experienced a massive shift toward remote assessment as a result of COVID-19 restrictions, and clinicians were rapidly introduced to telepsychiatry, who under, normal circumstances may never have been exposed to it. This shift in practice highlighted not only the need for increased clinical services for individuals without access but also the reluctance of many clinicians to engage with telehealth technologies owing to concerns about its efficacy and compromised privacy [18,19]. As future clinical practice in psychiatry may continue to rely on remote assessment, understanding clinician perspectives on this technology remains crucial to optimizing service and user experience.

The aim of our study was to capture clinician attitudes and expectations regarding emerging digital technology in psychiatry, in light of the unprecedented shift in clinician experience that the COVID-19 pandemic represents. Previously reluctant or unfamiliar clinicians will now have had significant experience of telehealth practice, and attitudes toward digital data in psychiatry may have shifted. Understanding clinician attitudes and preferences is necessary to overcome implementation challenges and achieve patient benefits.

Methods

Methods Overview

The survey was developed for the purpose of this study by the Digital Clinic research group, which is composed of clinicians working in outpatient and inpatient settings as well as dedicated research staff familiar with emerging digital technology in psychiatry, working on the implementation of digital technology in clinical settings across our health system. The survey was designed to study clinician attitudes toward the use of digital technology in psychiatry in a variety of dimensions using varied Likert scales. Questions were written on the basis of expert research group members' existing knowledge of digital technology in psychiatry as described in current literature, and in anticipation of developing a digital data platform for use in our campus' outpatient clinic (see attached survey in Multimedia Appendix 1).

Survey items included (1) an assessment of whether clinicians thought digital data about patients could inform their practice (using a 5-point Likert scale: 0%="No," 25%="I'm not sure, I need to know more," 50%="I think so, but would need to try it out," 75%="Yes," and 100%="Definitely, I incorporate this data already"), (2) ratings of perceived relative usefulness of different types of digital patient data (such as sleep, physical activity, location, web-based search activity, etc, using a 3-point Likert scale: 0%="Low," 50%="Medium," and 100%="High"), (3) anticipated barriers to using digital data in clinical practice (such as "Patient participation," "Increased time demands tending to flagged digital events," "Volume of data created by digital monitoring," etc, using a 3-point Likert scale: 0%="Not a barrier," 50%="Somewhat of a barrier," and 100%="A significant barrier"), (4) anticipated benefits to using digital data in clinical practice (such as "Having a consistent source of collateral data," "Helping patients feel better understood," "As an alert system when patient activities change," etc, using a 3-point Likert scale: 0%="Not a benefit," 50%="Somewhat of a benefit," and 100%="A significant benefit"), (5) clinicians' expectations of whether patients would be willing to share digital data in a clinical setting and experience using digital patient data from social media (using multiple-choice questions about how data were used and free-text responses to allow responders to elaborate), (6) rating level of agreement with statements about the incorporation of digital data into psychiatric practice (such as "Having access to information collected in a Digital Clinic will... Lead to more frequent patient encounters," "...Improve clinical outcomes," or "...Increase the amount of documentation to complete," using a 5-point Likert scale: 0%="Strongly disagree," 25%="Disagree," 50%="Neither agree nor disagree,"

75%="Agree," and 100%="Strongly agree"), (7) assessment of if or how clinician attitudes toward digital data in psychiatry had changed in response to the COVID-19 pandemic and the changes in practice that many of the surveyed participants had experienced (using free-text responses), and (8) demographic characteristics (including age, years in practice, nature of practice, and role in clinic).

The survey was administered to all clinicians on the campus of an academic psychiatric treatment facility. The facility surveyed here is in a demographically diverse section of a major urban area, with high access to mental health services. The survey was written in Survey Monkey and administered during a Grand Rounds event. It was subsequently emailed to all clinicians on campus. A total of 200 clinicians were surveyed—150 outpatients and 50 inpatients. The survey was open for 6 months from May to November 2020. Participation in the survey was voluntary. There was no compensation for completing the survey. Responders were given the option to share their contact information (deanonymizing themselves) to participate in a possible focus group, but owing to the COVID-19 pandemic, this focus group was not held.

Ethical Considerations

The survey was reviewed by the Northwell institutional review board and granted an exemption (#19-0958).

Statistical Analysis

Descriptive statistics were used to report survey results. Chi-square tests were used to compare categorical variables. First, omnibus comparisons were conducted by prescribing status (yes/no). If significant differences were detected, we then tested the individual interactions of interest post hoc. Wilcoxon tests for nonnormally distributed, continuous variables (age), as determined by a Shapiro–Wilk W test, were conducted. All analyses were conducted using JMP, (version 13; SAS Institute Inc, 1989-2019).

Results

Overview

We received a total of 139 completed survey responses (response rate 69.5%). The median age of responders was 42 (IQR 34-70) years, with a median 12 (IQR 4-33) years of clinical experience. In total, 50 (36.0%) responders were psychiatrists, 27 (27.0%) were residents, 20 (14.4%) were psychologists, 14 (10.1%) were social workers, 3 (2.2%) were nurse practitioners, and 25 (16.5%) were nurses or other clinical staff. Furthermore, 79 (56.8%) responders provided medication management and 82 (59.0%) provided individual psychotherapy (Table 1).

Table 1. Characteristics of responders (N=139).

Characteristics	Value
Age (years), median (IQR)	42 (34-70)
Years in practice, median (IQR)	12 (4-33)
Gender, n (%)	
Male	53 (38.1)
Female	74 (53.2)
Declined to answer	12 (8.7)
Role in clinic, n (%)	
Psychiatrist	50 (36.0)
Resident or intern	27 (19.4)
Psychologist	20 (14.4)
Nurse practitioner	3 (2.2)
Nurse	2 (1.4)
Social worker	14 (10.1)
Other	23 (16.6)
Type of practice (allowed to select more than one), n (%)	
Medication Management	79 (56.8)
Individual Psychotherapy	82 (59.0)
Group Therapy	30 (21.6)
Neuromodulation	5 (3.6)
Administration	26 (18.7)
Research	3 (2.2)

Attitudes Toward Digital Data in Psychiatry

Out of 139 responders, 116 (83.4%) stated that digital data could improve their clinical practice. Among different categories of digital data, responders rated data on Sleep (n=100, 72.0%), treatment adherence (n=101, 79.5%), substance use (n=105, 76.0%), self-reported symptom ratings (n=90, 70.0%), and physical activity (n=73, 53%) as having the highest potential utility in practice. Responders rated location (n=33, 26.0%), screen time (n=34, 24.0%), and criminal justice data (inmate registries, WebCrims, etc; n=25, 18.0%) as having the lowest potential use in practice (Table 2).

Among anticipated benefits, responders rated "Helping patients track their activities and symptoms" (n=101, 72.7%), "As a support for clinical intervention" (n=90, 64.7%), and "Having a consistent source of collateral data" (n=86, 62.5%) as offering the most significant benefit. Responders rated "Helping patients feel better understood" as the least potentially beneficial aspect of using digital data in psychiatry ("Not a benefit" n=14, 10.1%; Figure 1). Among anticipated barriers to the use of digital data, responders rated "Increased time demands tending to flagged digital events" (n=54, 38.8%), "Volume of data created by digital monitoring" (n=52, 37.4%), "Increased documentation" (n=57, 41%), "Patient participation" (n=46, 33.8%), and "Increased time demands tending to digital data during clinic visits" (n=54, 38.8%) as the most likely barriers. Responders rated "Lack of trust in digital data" (n=51, 36.7%), "Interference with alliance" (n=46, 33.1%), and "Uncertainty about how to integrate digital data into practice" (n=45, 32.4%) as the least likely barriers (Figure 2). Overall, 74.1% (n=103) of responders stated that they would consult a dashboard of patient digital data prior to a clinic visit, and 85.6% (n=119) of them stated that they thought it would be beneficial for patients to have access to an app-based dashboard with information about their digital data.

Table 2. Attitudes toward digital data in psychiatry.

Rating of the relative level of perceived usefulness that each category has for individuals' practice	Low, n (%)	Medium, n, (%)	High, n (%)
Sleep	3 (2.2)	27 (19.4)	100 (71.9)
Physical activity	4 (2.9)	59 (42.2)	73 (52.5)
Social media activity (frequency of posts or content of posts concerning symptoms of mental illness)	13 (9.4)	69 (49.6)	57 (41.0)
Web-based search activity (content of which concerns symptoms of mental illness)	19 (13.7)	58 (41.7)	62 (44.6)
Mobility (time spent away from home versus at home)	14 (10.1)	59 (42.4)	66 (47.5)
Location	36 (25.9)	65 (46.8)	33 (23.7)
Substance use	10 (7.2)	20 (14.4)	105 (75.5)
Screen time	34 (24.5)	66 (47.5)	38 (27.3)
Treatment adherence	7 (5.0)	22 (15.8)	110 (79.1)
Homework completion	26 (18.7)	63 (45.3)	48 (34.5)
Criminal justice data (inmate registries, WebCrims, etc)	25 (18.0)	59 (42.4)	55 (39.6)

Figure 1. Clinicians' responses to the question, "What do you anticipate would be the greatest benefits of digital data?".

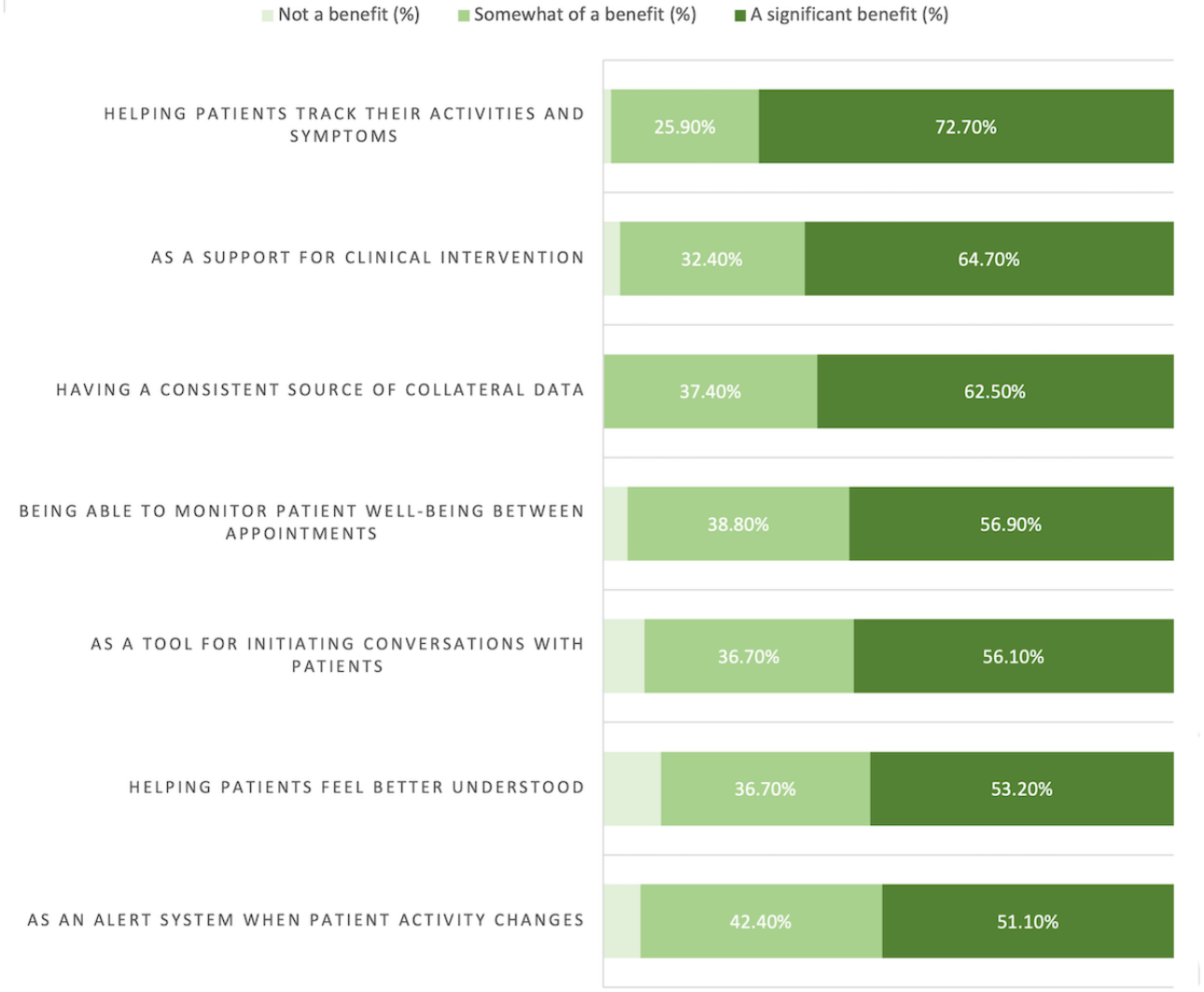
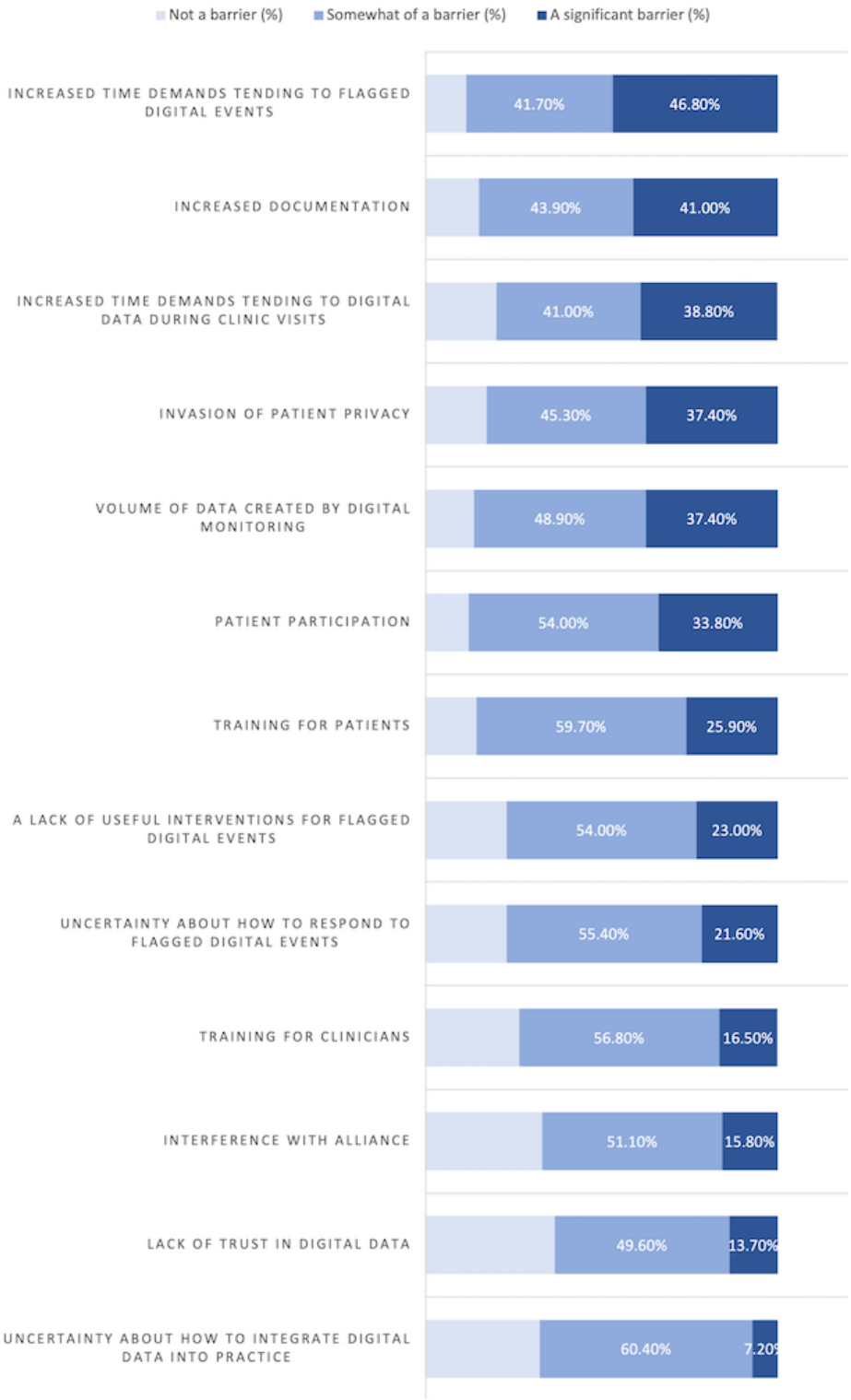


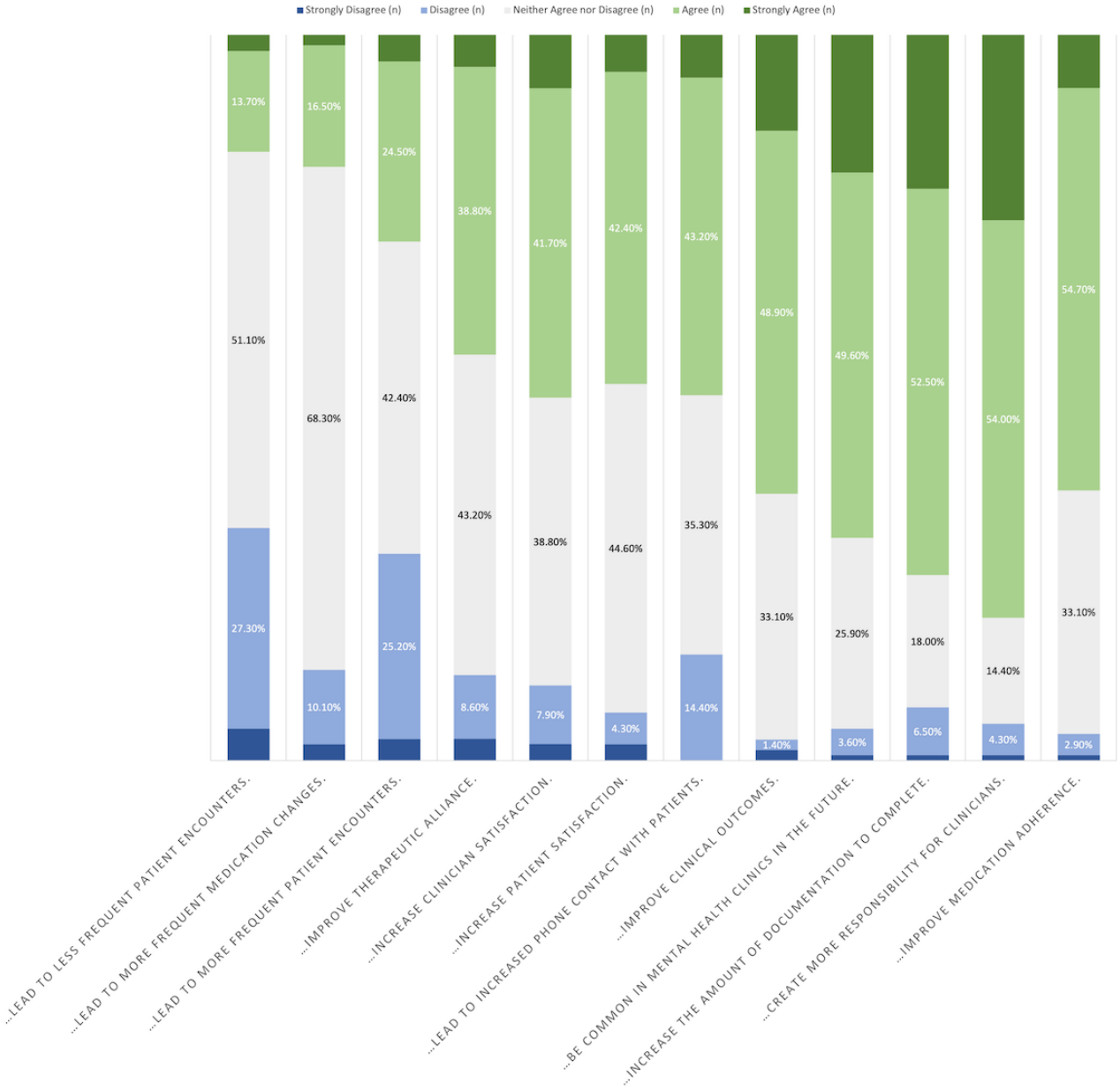
Figure 2. Clinicians' responses to the question, "What do you anticipate would be the greatest barriers against your use of digital data?".



Expectations From Digital Technology in Psychiatry

Responses skewed negatively regarding whether or not the introduction of digital technology into psychiatry would change the frequency of patient encounters ("...lead to more frequent patient encounters": 39, 28.1% respondents either disagreed or strongly disagreed; "...lead to less frequent patient encounters": 44, 31.7% either disagreed or strongly disagreed). Responders expressed agreement that digital technology could "...Increase the amount of documentation to complete" and "...Create more responsibility for clinicians" (n=102, 73.4% and n=110, 79.2%, respectively; Figure 3).

Figure 3. Clinicians' responses to the following statement: "Please rate your level of agreement with the following statements: Having access to information collected in a digital clinic will…".



Regarding past use of digital data in psychiatric practice, 32 (23.0%) participants responded affirmatively that they had viewed a patient's social media accounts (Facebook, Twitter, Instagram, and other social media accounts) while delivering care. A total of 21 (33.9%) of those who answered yes cited gaining collateral information as the rationale, while others reported viewing social media to evaluate the clinical significance of the posts, to evaluate patients' interpretations of posts, or for use in psychoeducation or therapeutic skill-building. In total, 120 (94.5%) responders stated that they thought their patients would be willing to share at least some digital data. Furthermore, 79 (63.2%) responders expressed agreement that digital technology in psychiatry could improve medication adherence.

Additional analyses were completed to assess differences in response patterns based on age and prescriber versus nonprescriber status (prescriber: MD physician, resident, or nurse practitioner; nonprescriber: nurse or social worker). There were overall few significant differences between groups, but younger responders expressed greater willingness to consult a digital dashboard (P=.01) and greater concern about increased time demands from digital data (P=.04) and expected greater patient satisfaction from a digital clinic (P=.04), and prescribers expressed greater interest in monitoring social media activity for posts concerning mental illness (P=.02) and greater uncertainty about how to respond to flagged digital events (P=.03). Younger responders were likelier to report having looked up a patient's social media accounts (P=.005).

Changes in Opinion After the COVID-19 Pandemic

When asked if the COVID-19 pandemic had affected their attitudes toward digital technology in psychiatry, 79 (56.8%) responders answered "yes." Regarding how their attitudes had changed, responders provided a variety of responses in a

free-text portion of the survey. Many were optimistic about digital technology in psychiatry after having had a positive initial experience with telepsychiatry: "More positive, more interest, greater comfort with telepsychiatry that makes me appreciate how digital data could inform my practice"; "It has normalized the use of technology in routine encounters"; "COVID is a disaster but I hope that a silver lining is that Telepsychiatry and digital psychiatry becomes more the accepted norm. Wider use of telepsychiatry is an excellent way to address the national shortage of psychiatrists available to deliver care"; and "Working from home greatly improves my work-life balance as a clinician. I find that patients feel the same way as they no longer have to travel to clinic, take time off work, or wait to see the doctor -- appointments start on time. For the stable, relatively high functioning outpatients that comprise my (very small) patient panel, telemedicine is equivalent to, and in many ways better than, in person psychiatry." Some expressed skepticism, stating that "less interaction with the patient; not the same as in person visits."

Discussion

Principal Findings

The purpose of our study was to examine clinician attitudes toward the use of digital data in mental health care via a survey of clinicians on the campus of an academic psychiatric hospital. We collected 139 survey responses. Overall, 83.4% (n=116) of responders reported positive expectations about digital data in clinical practice. Responders reported the highest enthusiasm for patient self-monitoring technologies and the strongest concern about potential increases in workload and actionability of digital data. As the use of digital data continues to gain prominence in mental health care, the results of our survey serve as a useful indicator of clinician expectations and concerns regarding this new technology in mental health. This study examined clinician attitudes toward the usefulness of different forms of digital data in clinical practice, clinicians' anticipated benefits and barriers to employing digital technology in clinical practice, expectations of the ways in which clinical practice will change with the inclusion of digital data in psychiatry, experiences using patients' social media data in practice, and the ways in which clinician perspectives on digital data in psychiatry have changed in response to the COVID-19 pandemic and the wide-scale shift to telehealth practice in medicine. Overall, the surveyed clinicians expressed enthusiasm to include digital data in their clinical practice and confidence that patients would be willing to share data; however, they expressed strong concern about increases in workload related to the inclusion of new technology in clinical practice.

Comparison to Prior Work

While much attention has been paid to research developments, expert consensus, and patient expectations on digital data in psychiatry, clinician attitudes remain under-reported [8,9,11,20,21]. Our study offers a unique contribution to the literature on this rapidly developing area of mental health care, detailing clinician expectations of forthcoming digital technology in psychiatry, attitudes regarding how digital data can both help and hinder their practice, and current clinician behaviors regarding patients' digital information (via web-based or social media search) at a unique point in time. These results will be useful in designing future digital clinics and be instructive for those disseminating digital data research as to which technologies clinicians are ready to embrace or remain reluctant.

Survey responders reported largely high expectations that digital data could improve their clinical practice, and more than half of them reported a shift in attitude after the onset of the COVID-19 pandemic, when most clinical practices had transitioned to a telehealth model. Providers also expressed high willingness to consult a digital dashboard and have patients do the same. These findings are consistent with the broader literature on the topic but higher than clinician expectation statistics reported elsewhere [20,22,23]. It may be that this enthusiasm reflected changing attitudes related to the COVID-19 pandemic or could be a reflection of the relatively young age of the group surveyed. Similarly, responders reported high confidence that patients would be willing to share digital data, consistent with other similar studies [23]. Our findings indicate that enthusiasm for the use of digital data in psychiatry and high expectations among clinicians that patients will feel the same way. Although there is a known gap between research enthusiasm and clinician enthusiasm [8,11], our data indicate that clinicians remain enthusiastic about digital data in psychiatry. This study should serve as encouragement for those seeking to implement digital technology into clinical practice.

Regarding the anticipated usefulness of different types of digital data in psychiatry, responders ranked items including sleep, substance use, patient-rated symptom scales, and physical activity as having high utility, but they rated location, internet search activity, and criminal data as having substantially less utility. One interpretation of this finding is that clinicians favor monitoring actionable metrics that could allow for discrete intervention over other types of digital measurement. This sentiment would appear echoed in the finding that clinicians rated "Helping patients track their activities and symptoms" and "As a support for clinical intervention" as the two most prominent benefits of using digital monitoring in mental health care. Previous work on digital technology in psychiatry has acknowledged a gap between what is notable in research versus what is readily adopted in clinical practice [8,11]. Our results indicate something similar and underscore the need to demonstrate clinical utility to encourage clinician adoption. Similarly, survey responders expressed concern about increases in documentation with the adoption of digital technology in clinical practice. In addition to demonstrating the actionability of digital data, designers of future clinical platforms will have to streamline the modes in which data are presented to clinicians to minimize the documentation burden.

The response to the question, "Have you ever viewed a patient's social media (Facebook page, Twitter account, Instagram account etc)?" is a notable finding in itself because this behavior is extremely underreported in the literature [24,25]. Based on available research, our results are consistent with those of other studies [26]. Although not a strict equivalent to consulting a digital dashboard in the clinical setting (for both practical and ethical reasons), looking up patients on social media can be

viewed as somewhat of a proxy for current use of digital data in clinical practice among those whom we surveyed. By that standard, despite their enthusiasm for new technology, our responders' use of social media data was low. Given the controversy regarding this practice, it is plausible that our responders underreported how often they viewed patients' social media accounts. Future research of other psychiatric clinicians in other settings could help to further elucidate how clinicians incorporate patient social media data into their assessments.

Limitations

This study had several limitations. The first and most significant limitation is that it was conducted on a campus on which psychiatric digital technology research has had high visibility for several years. Although few responders identified themselves as researchers, many had likely had exposure to emerging digital technology and may have been predisposed to a positive response regarding potential benefits and lower trepidation about potential barriers of this technology in practice. This limits the generalizability of our data, when compared to clinician attitudes in community or nonacademic settings. Second, the survey was composed largely of closed-ended questions. Though there were opportunities for free-text responses where responders could express thoughts or feelings not captured in the survey questions, this mode of surveying may have left some attitudes unaccounted for.

A third limitation of our survey is timing, and the limitation is 2-fold. First, clinicians completed the survey beginning in May 2020, several months after a wide-scale shift to telehealth on our campus. In their free-text responses, many mentioned positive experiences with telepsychiatry but less specifically any of the digital technology described in the survey. It is possible that some responders conflated their experience with telepsychiatry with the technology described as part of the hypothetical digital clinic and provided contrived, more positive responses as a result. Future research must investigate how clinicians' attitudes toward technology have continued to evolve in the postpandemic world. The second issue is that emerging technology is a dynamic, ever-shifting field, and that the generalizability of survey results, such as that of our survey, is limited to the moment in time the survey was administered. This limits the generalizability of our study or rather any study that discusses this topic. Lastly, the dimensions of our survey were not calculated, and specific validity and reliability tests were not conducted. Studies similar to this one will remain useful for understanding clinicians' perspectives on these issues, but so will studies that measure the safety and efficacy of how these technologies are implemented in clinical practice.

Conclusions

Our survey results indicate that clinicians' attitudes toward the implementation of digital data in psychiatry are largely positive. However, responders voiced some trepidation about the actionability of digital data and increased time demands from addressing or documenting data. Overall, 23.0% (n=32) of responders reported having looked up patients on social media. More than half of the responders reported a change in attitude toward digital data in psychiatry following the COVID-19–related transition to telepsychiatry service models. Our survey results underscore the need for clinician engagement and education as digital data platforms are developed for clinical use in psychiatry.

Acknowledgments

We thank the clinicians who took the time to participate in our study despite their busy schedule in such challenging times. We would also like to thank Lillian Smith for her guidance in creating the initial drafts of our survey.

Conflicts of Interest

DG has been a consultant for or has received speaker honoraria from Otsuka America Pharmaceuticals and Janssen Pharmaceuticals. JMR has received honoraria from Teva and Lundbeck, royalties from UpToDate, and grant funding from Alkermes and the National Institute of Mental Health. JMK has been a consultant or advisor for or has received honoraria from Alkermes, Allergan, LB Pharmaceuticals, H Lundbeck, Intracellular Therapies, Janssen Pharmaceuticals, Johnson and Johnson, Merck, Minerva, Neurocrine, Newron, Otsuka, Pierre Fabre, Reviva, Roche, Sumitomo Dainippon, Sunovion, Takeda, Teva, and UpToDate, and is a shareholder in LB Pharmaceuticals and Vanguard Research Group. MLB receives equity from NorthShore Therapeutics and serves as a consultant for HearMe. WAS, MS, and AVM report no conflicts of interest.

Multimedia Appendix 1
Sample of the digital clinic clinician survey.
[PDF File (Adobe PDF File), 405 KB - formative_v6i11e33676_app1.pdf]

References

1. Ben-Zeev D, Brian R, Wang R, Wang W, Campbell AT, Aung MSH, et al. CrossCheck: integrating self-report, behavioral sensing, and smartphone use to identify digital indicators of psychotic relapse. Psychiatr Rehabil J 2017 Sep;40(3):266-275 [FREE Full text] [doi: 10.1037/prj0000243] [Medline: 28368138]
2. Buck B, Scherer E, Brian R, Wang R, Wang W, Campbell A, et al. Relationships between smartphone social behavior and relapse in schizophrenia: a preliminary report. Schizophr Res 2019 Jun;208:167-172 [FREE Full text] [doi: 10.1016/j.schres.2019.03.014] [Medline: 30940400]

3. Birnbaum ML, Wen H, Van Meter A, Ernala SK, Rizvi AF, Arenare E, et al. Identifying emerging mental illness utilizing search engine activity: a feasibility study. PLoS One 2020 Oct 16;15(10):e0240820 [FREE Full text] [doi: 10.1371/journal.pone.0240820] [Medline: 33064759]
4. Lewis S, Ainsworth J, Sanders C, Stockton-Powdrell C, Machin M, Whelan P, et al. Smartphone-enhanced symptom management in psychosis: open, randomized controlled trial. J Med Internet Res 2020 Aug 13;22(8):e17019 [FREE Full text] [doi: 10.2196/17019] [Medline: 32788150]
5. Hirschtritt ME, Insel TR. Digital technologies in psychiatry: present and future. Focus (Am Psychiatr Publ) 2018 Jul;16(3):251-258 [FREE Full text] [doi: 10.1176/appi.focus.20180001] [Medline: 31975919]
6. Beard C, Silverman AL, Forgeard M, Wilmer MT, Torous J, Björgvinsson T. Smartphone, social media, and mental health app use in an acute transdiagnostic psychiatric sample. JMIR Mhealth Uhealth 2019 Jun 07;7(6):e13364 [FREE Full text] [doi: 10.2196/13364] [Medline: 31199338]
7. Insel TR. Digital phenotyping: technology for a new science of behavior. JAMA 2017 Oct 03;318(13):1215-1216. [doi: 10.1001/jama.2017.11295] [Medline: 28973224]
8. Mohr DC, Lyon AR, Lattie EG, Reddy M, Schueller SM. Accelerating digital mental health research from early design and creation to successful implementation and sustainment. J Med Internet Res 2017 May 10;19(5):e153 [FREE Full text] [doi: 10.2196/jmir.7725] [Medline: 28490417]
9. Bhugra D, Tasman A, Pathare S, Priebe S, Smith S, Torous J, et al. The WPA-Lancet Psychiatry Commission on the future of psychiatry. Lancet Psychiat 2017 Oct;4(10):775-818. [doi: 10.1016/s2215-0366(17)30333-4]
10. Tonn P, Reuter SC, Kuchler I, Reinke B, Hinkelmann L, Stöckigt S, et al. Development of a questionnaire to measure the attitudes of laypeople, physicians, and psychotherapists toward telemedicine in mental health. JMIR Ment Health 2017 Oct 03;4(4):e39 [FREE Full text] [doi: 10.2196/mental.6802] [Medline: 28974485]
11. Mohr DC, Weingardt KR, Reddy M, Schueller SM. Three problems with current digital mental health research . . . and three things we can do about them. Psychiatr Serv 2017 May 01;68(5):427-429 [FREE Full text] [doi: 10.1176/appi.ps.201600541] [Medline: 28412890]
12. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 01;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
13. Wade VA, Eliott JA, Hiller JE. Clinician acceptance is the key factor for sustainable telehealth services. Qual Health Res 2014 May 31;24(5):682-694. [doi: 10.1177/1049732314528809] [Medline: 24685708]
14. Cowan KE, McKean AJ, Gentry MT, Hilty DM. Barriers to use of telepsychiatry: clinicians as gatekeepers. Mayo Clin Proc 2019 Dec;94(12):2510-2523. [doi: 10.1016/j.mayocp.2019.04.018] [Medline: 31806104]
15. Guinart D, Marcy P, Hauser M, Dwyer M, Kane JM. Mental health care providers' attitudes toward telepsychiatry: a systemwide, multisite survey during the COVID-19 pandemic. Psychiatr Serv 2021 Jun;72(6):704-707. [doi: 10.1176/appi.ps.202000441] [Medline: 33593104]
16. Schueller SM, Washburn JJ, Price M. Exploring mental health providers' interest in using web and mobile-based tools in their practices. Internet Interv 2016 May;4(2):145-151 [FREE Full text] [doi: 10.1016/j.invent.2016.06.004] [Medline: 28090438]
17. Mayer G, Gronewold N, Alvarez S, Bruns B, Hilbel T, Schultz J. Acceptance and expectations of medical experts, students, and patients toward electronic mental health apps: cross-sectional quantitative and qualitative survey study. JMIR Ment Health 2019 Nov 25;6(11):e14018 [FREE Full text] [doi: 10.2196/14018] [Medline: 31763990]
18. Rodriguez-Villa E, Rauseo-Ricupero N, Camacho E, Wisniewski H, Keshavan M, Torous J. The digital clinic: implementing technology and augmenting care for mental health. Gen Hosp Psychiatry 2020 Sep;66:59-66 [FREE Full text] [doi: 10.1016/j.genhosppsych.2020.06.009] [Medline: 32688094]
19. Uscher-Pines L, Sousa J, Raja P, Mehrotra A, Barnett ML, Huskamp HA. Suddenly becoming a "Virtual Doctor": experiences of psychiatrists transitioning to telemedicine during the COVID-19 pandemic. Psychiatr Serv 2020 Nov 01;71(11):1143-1150 [FREE Full text] [doi: 10.1176/appi.ps.202000250] [Medline: 32933411]
20. Hatch A, Hoffman JE, Ross R, Docherty JP. Expert consensus survey on digital health tools for patients with serious mental illness: optimizing for user characteristics and user support. JMIR Ment Health 2018 Jun 12;5(2):e46 [FREE Full text] [doi: 10.2196/mental.9777] [Medline: 29895514]
21. Jacob C, Sanchez-Vazquez A, Ivory C. Social, organizational, and technological factors impacting clinicians' adoption of mobile health tools: systematic literature review. JMIR Mhealth Uhealth 2020 Feb 20;8(2):e15935. [doi: 10.2196/15935]
22. Bauer R, Glenn T, Monteith S, Whybrow PC, Bauer M. Survey of psychiatrist use of digital technology in clinical practice. Int J Bipolar Disord 2020 Oct 03;8(1):29 [FREE Full text] [doi: 10.1186/s40345-020-00194-1] [Medline: 33009954]
23. Armstrong CC, Odukoya EJ, Sundaramurthy K, Darrow SM. Youth and provider perspectives on behavior-tracking mobile apps: qualitative analysis. JMIR Ment Health 2021 Apr 22;8(4):e24482 [FREE Full text] [doi: 10.2196/24482] [Medline: 33885364]
24. Ashby GA, O'Brien A, Bowman D, Hooper C, Stevens T, Lousada E. Should psychiatrists 'Google' their patients? BJPsych Bull 2015 Dec 02;39(6):278-283 [FREE Full text] [doi: 10.1192/pb.bp.114.047555] [Medline: 26755985]

25. Clinton B, Silverman BC, Brendel DH. Patient-targeted googling: the ethics of searching online for patient information. Harv Rev Psychiatry 2010;18(2):103-112. [doi: 10.3109/10673221003683861] [Medline: 20235775]
26. Bosslet GT, Torke AM, Hickman SE, Terry CL, Helft PR. The patient-doctor relationship and online social networks: results of a national survey. J Gen Intern Med 2011 Oct 25;26(10):1168-1174 [FREE Full text] [doi: 10.1007/s11606-011-1761-2] [Medline: 21706268]

Edited by A Mavragani; submitted 18.09.21; peer-reviewed by V Guan, S Steingrimsson, N Martin-Key; comments to author 19.10.21; revised version received 14.02.22; accepted 05.08.22; published 10.11.22.

Please cite as:
Sterling WA, Sobolev M, Van Meter A, Guinart D, Birnbaum ML, Rubio JM, Kane JM
Digital Technology in Psychiatry: Survey Study of Clinicians
JMIR Form Res 2022;6(11):e33676
URL: https://formative.jmir.org/2022/11/e33676
doi:10.2196/33676
PMID:36355414

Original Paper

A Digital Video and Text Messaging Intervention to Support People With Chronic Pain During Opioid Tapering: Content Development Using Co-design

Michael R Magee[1], MA; Ali Gholamrezaei[1], MD, PhD; Amy G McNeilage[1], BA; Alison Sim[1], MA; Leah Dwyer[2], BA; Manuela L Ferreira[3], BPTh, MSc, PhD; Beth D Darnall[4], PhD; Paul Glare[1], MBBS, MM; Claire E Ashton-James[1], PhD

[1]Pain Management Research Institute, Faculty of Medicine and Health, The University of Sydney, Sydney, Australia

[2]Consumer Advisory Group, Pain Australia, Deakin, Australia

[3]Sydney Musculoskeletal Health, School of Health Sciences, Kolling Institute, Faculty of Medicine and Health, The University of Sydney, Sydney, Australia

[4]Department of Anaesthesiology, Perioperative and Pain Medicine, Stanford University School of Medicine, Stanford, CA, United States

Corresponding Author:
Michael R Magee, MA
Pain Management Research Institute
Faculty of Medicine and Health
The University of Sydney
Douglas Building, Royal North Shore Hospital
Reserve Rd
Sydney, St Leonards 2065
Australia
Phone: 61 2 9463 1526
Email: mmag9080@uni.sydney.edu.au

Abstract

Background: People living with chronic pain report that tapering prescribed opioids is challenging and more support is needed. In our formative research, consumers indicated that mobile health (mHealth) technology could be an acceptable form of support for opioid tapering and may improve tapering self-efficacy.

Objective: We aimed to evaluate and improve the content of an mHealth intervention before pilot-testing, based on consumer and clinician feedback.

Methods: Participants were 12 consumers and 12 clinicians who evaluated an initial draft of a video script and 90 SMS text messages. Consumers and clinicians rated the appropriateness and likely usefulness (consumers) or likely effectiveness (clinicians) of a video script and a random selection of 15 SMS text messages using a 5-point Likert-type scale (1=totally disagree; 5=totally agree). Each draft SMS text message was reviewed by 2 consumers and 2 clinicians. Texts were deemed acceptable for inclusion in the pilot intervention only if the summed participant ratings of text appropriateness and usefulness or effectiveness were ≥8. Participants were also invited to provide open-text feedback on the draft script and SMS text messages.

Results: Consumers generally agreed that the draft video script and text content were likely to be appropriate (video: mean 4.4, SD 0.52; text: mean 4.3, SD 0.79) and useful (video: mean 4.3, SD 0.65; text: mean 4.2, SD 0.84). Similarly, clinicians generally agreed that the draft video script and text content were likely to be appropriate (video: mean 4.5, SD 0.67; text: mean 4.4, SD 0.81) and effective (video: mean 4.0, SD 0.43; text: mean 4.3, SD 0.76). Overall, 77% (69/90) of the draft texts met the threshold rating for acceptability for inclusion in the pilot test of mHealth intervention by consumers, and 82% (74/90) met the threshold for acceptability by clinicians. Consumers' and clinicians' ratings were used to rank order the texts. The top 56 draft texts (all meeting the threshold levels of acceptability) were selected for inclusion in the pilot intervention. When consumer or clinician feedback was provided, the texts meeting the criteria for inclusion in the pilot were further revised and improved. Feedback on the video script was also used to further improve the acceptability of the video script before pilot-testing the intervention.

Conclusions: This study describes the process by which a 28-day mHealth intervention to support patients with chronic pain to taper opioid medications was evaluated and improved before pilot-testing. The mHealth intervention consisted of a 10-minute psychoeducational video about pain and opioid tapering and 56 unique SMS text messages providing information and reassurance

(texts delivered twice per day for 28 days). Having established that the content of the mHealth intervention is acceptable to both consumer and clinician groups, the mHealth intervention will be piloted in future research.

(JMIR Form Res 2022;6(11):e40507) doi:10.2196/40507

KEYWORDS

chronic pain; deprescribing; tapering; dose reduction; opioids; mHealth; mobile health; SMS; text messaging; digital health; behavior change; self-efficacy; consumer engagement; co-design; coproduction

Introduction

Chronic Pain and Opioid Tapering

The reduction or cessation of prescription opioid medications (opioid tapering) is recommended for people living with chronic nonmalignant pain in cases where the potential harms of long-term high-dose opioid therapy outweigh the benefits [1-4]. Tapering can lead to unpleasant withdrawal symptoms, increased distress, and a potential increase in pain intensity and interference, especially in the absence of alternative nonpharmacological pain management strategies [3-7]. Research suggests that consumers with chronic pain, who are tapering opioids, benefit from education regarding the negative consequences of long-term opioid therapy, nonpharmacological (ie, cognitive and behavioral) active pain management strategies, and strategies to help with the management of withdrawal symptoms [1,8]. Consistent with this, research indicates that patients with chronic pain who have access to multidisciplinary pain management programs (with care provided by a team of physicians, physiotherapists, and clinical psychologists) are more likely to reduce their opioid dose with fewer negative side effects [8,9]. However, there is a limited workforce of physicians, psychologists, and physiotherapists with expertise in chronic pain treatment, which is far outweighed by consumer demand for this specialized support [10-12]. Hence, there is a need for innovative, scalable solutions to provide specific support to people living with chronic pain who are tapering opioids [10,13].

Digital Health

Digital and mobile health (mHealth) services might help address this gap in access to support for opioid tapering in patients with chronic pain [13,14]. Accumulating research indicates that mHealth interventions may be an acceptable form of opioid tapering support for individuals living with chronic pain [15,16]. SMS text message–based interventions, in particular, may be most effective in reaching the largest number of consumers, as texts do not require access to the internet, have low technology literacy requirements, and are low cost and familiar, as there is a high current use of text messaging worldwide [16-18]. Mobile phone text interventions have demonstrated effectiveness in facilitating health-related behavior change [18,19], helping people to quit smoking [20], improving outcomes in the management of musculoskeletal conditions [21], and even supporting people to manage perioperative pain [17,22].

Co-design

Although text message–delivered mHealth interventions have the potential to be efficacious, the effectiveness of digital health interventions in practice is likely to be enhanced by consumer co-design. Research demonstrates that using a co-design methodology in the development and evaluation of mHealth interventions improves their acceptability, increases the likelihood that researchers will achieve recruitment and retention goals in clinical research trials, and ensures that interventions meet the needs of targeted populations, such as people living with health issues like chronic pain [23-27]. Our research group used a 4-step co-design methodology [28] to develop and test an effective mHealth intervention to support patients with chronic pain to taper prescription opioids. The co-design steps include (1) conducting formative research with consumers to develop insights into consumer needs and preferences, (2) developing the content of the intervention to meet consumer-stated needs and preferences, (3) coevaluating the intervention content in consultation with consumers and key stakeholders (eg, clinicians or caregivers), and (4) revising and improving the intervention in response to feedback [28].

Previously, we conducted formative research with consumers to identify whether people living with chronic pain who were tapering their opioid medication dose reported favorable attitudes toward the feasibility and acceptability of mHealth interventions to support tapering [16]. Our formative research revealed the following key insights:

1. For most consumer participants, SMS text messaging was preferred over an app for the delivery of support [16].
2. Consumers wanted the mHealth intervention to provide emotional and informational support concerning managing chronic pain without opioids; withdrawal symptom management; and strategies for improving functioning, quality of life, and mood [16].
3. Consumers suggested that text message–delivered support would be more engaging and effective if they had some familiarity with the information being provided (ie, pain self-management strategies) [16].

On the basis of this consumer feedback, and in collaboration with a consumer with lived experience (LD), as well as with clinician representatives (AS, BDD, MLF, MRM, and PG), we developed the concept of a text-delivered mHealth intervention to support patients with chronic pain to taper prescription opioids. The content of the intervention should be developed to provide consumers with emotional and informational support with respect to pain self-management and strategies for minimizing withdrawal symptoms and improving mood. In response to consumers' suggestions to familiarize users with intervention content before it is received through texts, we identified the need for a brief preintervention psychoeducational video. Supporting the benefits of including an educational video in the mHealth intervention, research indicates that after viewing a brief educational video, including patient testimonials about

the benefits of opioid tapering, patients develop more positive attitudes toward tapering [29]. Further research indicates that videos including testimonials about the effectiveness of digital health interventions can improve their acceptability [30].

Self-efficacy and Behavior Change

Consumers are more likely to engage with and respond to health interventions rooted in behavior change theory compared with interventions that are not theory-derived [31,32]. Research investigating the experiences of people who are tapering opioids suggests that low self-efficacy is a barrier to tapering and that improvements in both pain self-efficacy and opioid tapering self-efficacy are associated with improved outcomes for people living with chronic pain who are tapering [15,33]. Self-efficacy for opioid tapering has been demonstrated to be amenable to change through digitally delivered education [29,34,35]. Self-efficacy, a concept similar to confidence, refers to an individual's self-appraisal of their ability to perform a certain task or achieve a specific outcome [34,35]. Social learning theory suggests that enhancing self-efficacy for specific health-related behaviors is one of the key drivers for enduring behavioral change [36]. In 1 study that delivered a digital pain self-management program, improved pain self-efficacy was associated with reduced opioid misuse [37]. Accordingly, we theorized that self-efficacy to taper should be associated with a greater likelihood of reducing reliance on opioid medications to manage pain. As such, improving opioid tapering self-efficacy was posited as the proximal goal of our mHealth intervention [28].

This Study

Building on our formative research, this study describes the process by which mHealth intervention content was developed to meet consumer-stated needs and preferences, evaluated in consultation with consumers and key stakeholders (clinicians), and revised in response to feedback.

Methods

Setting

The study was conducted within a specialist pain clinic and research institute at a metropolitan public hospital in Sydney from December 2020 to August 2021.

Participants

Purposeful recruitment was selected as our aim was to develop an intervention for a very specific population: people with chronic pain who were tapering their opioid medications under the supervision and guidance of a pain specialist [38]. Hence, for the initial evaluation of intervention content, we recruited consumers from specialist pain clinics and clinicians with experience in delivering tertiary care to patients with chronic pain who are tapering opioids. Within this specialized consumer population, we aimed to recruit a diverse sample of participants regarding socioeconomic, cultural, gender, linguistic, and other characteristics. We drew on our extensive Australia-wide network of clinical colleagues to identify patients with chronic pain who had some experience of tapering opioids (not necessarily successful or good experiences).

Overall, 12 individuals with chronic pain who had previous experience tapering prescription opioid medications were recruited through purposeful sampling and targeted recruitment [38]. The 12 consumers varied with respect to their education, income, culture, preferred language, and geographic region (metropolitan, regional, and rural). These consumers who participated in the evaluation of video and SMS content had varied tapering experiences. Most consumer participants had experienced difficulties with tapering. Some had been successful in overcoming a reliance on opioids to reduce their dose. Other participants were still tapering or had paused their dose reduction. Hence, our consumer participants had experience of what does not work, what is hard about tapering, and what can be helpful. They were known to the researchers as consumers of pain services through previous research participation and consumer involvement, affiliation with an Australian consumer advisory group [39] or had previously expressed interest in pain research or consumer advocacy. Of these 12 consumers, 4 (33%) were born overseas, while 8 (67%) were born in Australia. Three participants lived in regional areas, and the remainder lived in metropolitan areas.

We matched our sample of consumers with 12 clinicians from varied pain management disciplines. The sample included registered multidisciplinary health professionals who currently work with patients who experience chronic pain and have also supported consumers to reduce or cease opioid medications directly (ie, the physician supervising dose reduction) or within a multidisciplinary team (eg, a psychologist or physiotherapist in an individual or group pain program supporting consumers who are tapering). The 12 clinicians included 5 specialist pain physicians (3 anesthetists, 1 neurologist, and 1 psychiatrist), 3 clinical psychologists, 3 physiotherapists, and 1 osteopath, who were practicing in private and public clinics in rural or regional and metropolitan clinical settings. Clinicians were recruited through purposeful sampling and direct invitation. The clinicians invited to participate in this study all delivered evidence-based interventions within pain management programs, supported consumers who were tapering prescription opioids, had diverse clinical experiences across different areas of Australia (ie, working in private and public clinics, working across regional and metropolitan areas, and supporting consumers with diverse cultural and linguistic backgrounds), and kept up to date with research and treatment advances in the area.

Ethical Considerations

This study was approved by the Northern Sydney Local Health District Research and Ethics Committee (2020ETH03288). All participants were provided with study information before completing the surveys. As indicated in the study, participants provided informed consent. To ensure privacy and confidentiality, only minimal demographic information was collected from participants. No compensation was provided for participating in the study.

Initial Design Considerations for the Content of the mHealth Program

Overview

In response to consumer feedback, the mHealth content design and aims are to support users to successfully reduce or discontinue their prescription opioid medication when engaging in a voluntary taper. On the basis of self-efficacy theory [34,36], as well as clinical resources and research on pain management [40-46] and opioid tapering [6,33,37,47-50], members of the research group (MRM, AG, AGM, CEAJ, and PG) drafted content for the preliminary video script and texts. This content was drawn from digital and traditional written resources and included information about pain self-management education, cognitive behavioral therapy concepts, and informational and supportive content from previous chronic pain and tapering research [7,9,15,16,40,42-45,47-51]. The researchers who developed the draft script and texts screened and adapted the information and resources to develop content that matched consumer-identified themes from formative research [29]. The other authors (AS, LD, BDD, and MLF) provided iterative feedback on the draft content and supported further editing of the script or texts.

Video Design

The video was designed to familiarize participants with information about chronic pain, provide a rationale for opioid tapering (eg, emphasizing long-term side effects and enhancing motivation to taper [29,30]), strengthen opioid tapering concepts (benefits of tapering on quality of life and functioning), and address common concerns about tapering (withdrawal management and information on pain intensity) [6,33,37,47-50], while introducing various pain self-management strategies (including flare-up management, relaxation, sleep hygiene, thought management, and pacing) [40-46]. These psychological concepts were also linked to improving mood, and the association of tapering, mood, and pain was also included in the content [15,42,43,46,52]. To reinforce informational content, 3 patient testimonials were solicited and included in the video design. Previous research has demonstrated that patients who watch testimonials express more favorable attitudes toward opioid tapering than those who receive information without testimonials [29]. On the basis of the informational content, a series of slides with accompanying images were developed by members of the research team (MRM, AG, AGM, CEAJ, and AS). The video script, including testimonials, was transcribed for review. First, by the rest of the authors, then following further revisions, the script was prepared for review by consumers and clinicians for this study.

Text Message Design

The texts aimed to provide daily informational and emotional support and reinforce key concepts about pain management and opioid tapering, particularly those concepts introduced in the video. The research groups (MRM, AG, AGM, CEAJ, and PG) generated 200 texts in the first instance. The team (all authors) made initial revisions. Duplicates were removed and the texts were edited for an appropriate length; 160 characters or less was optimal, but texts of up to 320 characters (2 standard texts) were permitted. The language of the texts was also simplified throughout revisions. Following these edits, the list was reduced to 90 texts. The texts included a range of content based on the communication objectives, information, and resources cited earlier. Content addressed pain education and self-management strategies and information about opioids and opioid tapering. Texts were designed to enhance motivation to taper opioids, provide emotional support (ie, messages were validating and encouraging), minimize nocebo, and optimize placebo. Content was also designed to be engaging, principally because engagement with mHealth content may improve the likelihood that users achieve their behavioral change goal [53-56]. Specific strategies to increase engagement included simplifying the reading load and including validating [57], reassuring [58], encouraging, motivating [59], and supportive content [60,61]. Texts were also designed to be personalized, with intermittent use of the users' preferred name [18,61,62]. Content was tailored for specific times of the day [19]. For example, sleep hygiene information was included in evening texts and motivational content was included in morning texts [61].

Sample Size

It was determined that 12 consumers would be needed to ensure that all messages could be evaluated by 2 participants from each group. The sample size was estimated based on similar previous designs, considering that to obtain sufficient qualitative and survey data to evaluate a message, each total SMS text message should be reviewed by at least 2 clinicians and 2 consumers [41]. However, to reduce the burden and in the hope that more content would be deeply reviewed, we wanted only 15 messages reviewed per participant. Overall, 12 participants per group (consumers and clinicians) provided 2 opinions from each group for every 15 messages.

Stakeholder Evaluation of Intervention Content

Consistent with previous research [28,40], the proposed content of the mHealth intervention was next evaluated by patients with pain who had experienced tapering opioids (consumers) and clinicians with experience supporting patients with pain to taper opioid medications.

Procedure

Consumers and clinicians who consented to participate were sent a web-based survey using REDCap (Research Electronic Data Capture; Vanderbilt University) [63]. The participants were asked to provide feedback on the video script and 15 texts from the text library. Each text was reviewed by 2 consumers and 2 clinicians.

Measures

Consumers were asked to rate the appropriateness ("easy to understand, sensitive to individual circumstances, and accurate") and usefulness ("supporting and motivating patients to persist with reducing their dose of opioids") of the script and SMS text messages on a 5-point Likert-type scale (1=totally disagree, 2=disagree, 3=neutral, 4=agree, and 5=totally agree). Clinicians rated the appropriateness and effectiveness ("supporting and motivating patients to persist with opioid tapering") of the script and SMS text messages on the same scale. All participants were invited to provide open-text feedback about the script and SMS

text messages. This methodology was adopted from a similar study that co-designed texts for an mHealth intervention [41].

Analytic Strategy

For the video script, the agreed appropriateness, usefulness, and effectiveness were evaluated by reviewing all participant responses to calculate the mean ratings of the appropriateness (both groups), likely usefulness (consumers only), and likely effectiveness (clinicians only) items. It was predetermined that the average appropriateness, usefulness, and effectiveness scores of "agree" (mean 4) or "totally agree" (mean 5) would indicate the acceptability of the script. Open-text feedback was also reviewed by the research group to improve the script and texts.

For the texts, as with the script, the average appropriateness, usefulness, and effectiveness scores of "agree" (mean 4) or "totally agree" (mean 5) would indicate the overall acceptability of the content. Unlike the script, in which all participants reviewed the same content, each text was evaluated by only 2 consumers and 2 clinicians. This was to make it more feasible for the study participants to complete the survey. The mean sum score (MSS) was calculated and used to rank each text. For each text, the consumer MSS is the sum of the appropriateness and usefulness ratings for the 2 consumers divided by 4. For each text, the clinician MSS is the sum of the appropriateness and perceived effectiveness ratings for the 2 clinicians divided by 4. The MSS has a range of 0 to 10. The objective of this analysis was to identify the most suitable messages for inclusion in mHealth support. Suitability was determined as an MSS that would suggest most responses on the usefulness and effectiveness scores were "agree" (mean 4) or "totally agree" (mean 5). An MSS <8 for clinicians and consumers was set as a cut-off for inclusion in the final mHealth content [41]. Open-text feedback was also reviewed for each text message and used to guide the revision of text messages. MSS was used to rank each text message for revision and inclusion decisions, which are described in the Results section.

Results

Overview

The Multimedia Appendix 1 contains a detailed breakdown of descriptive statistics for survey items (scripts and texts) and a summary of the wide range of open-text feedback toward the content provided by consumers and clinicians. Key results are also described in subsequent sections, generally separated by script or texts and consumers or clinicians. Descriptive statistics for the script and text survey results (both groups) were not normally distributed, so median and IQR values were reported.

Video Script

Table 1 reports the survey results for consumer and clinician ratings of the video script. On average, consumers endorsed responses indicating they "agreed" the video script was appropriate and likely to be useful. On average, clinicians endorsed responses indicating they "agreed" the video script was appropriate and likely to be effective. In total, 83% (10/12) of consumers provided feedback on the script. The video script was praised for the broad range of pain self-management strategies provided ("[It] gives a full snapshot") and for the inclusion of testimonials ("Seeing someone that has gone through something similar will be beneficial"). Other suggestions were used to revise the video script (described here), such as providing more "coping mechanisms," emphasizing the "real and tangible" benefits of tapering and clarifying the difference between acute and chronic pain. Overall, 58% (7/12) of clinicians provided qualitative feedback on the script. The video script was praised for the use of simplified language ("Obviously a lot of thought has gone into the lay examples and terms"), the repetition of "values-based activities," and the use of real patient testimonies ("The cases can be very helpful"). Other suggestions were used to revise the video script (described here), such as changing the metaphor used to describe chronic pain ("I find people relate more to a faulty smoke alarm rather than a car alarm") and providing more validation ("I wonder if less information and more recognition of potential emotional distress might be helpful").

Table 1. Descriptive statistics for script.

Question	Value, mean (SD)	Value, median (IQR)	Range (minimum-maximum)
Appropriate[a]			
Consumer	4.42 (0.515)	4 (1)	1 (4-5)
Clinician	4.50 (0.674)	5 (1)	2 (3-5)
Likely useful[b]			
Consumer	4.33 (0.651)	4 (1)	2 (3-5)
Likely effective[c]			
Clinician	4.00 (0.426)	4 (1)	2 (3-5)

[a]Appropriate: easy to understand, sensitive to individual circumstances, and accurate.

[b]Likely to be useful: supporting and motivating patients to persist by reducing their dose of opioids.

[c]Likely to be effective: supporting and motivating patients to persist with opioid tapering.

Text Messages

Table 2 presents the descriptive statistics of the text message survey. On average, consumers indicated they "agreed" the texts were appropriate and likely to be useful. On average, clinicians indicated they "agreed" texts were appropriate and likely to be effective. The mean of all text sum scores for both consumers and clinicians was >8 (acceptability cut-off for texts). A total of 77% (69/90) of texts had a consumer MSS ≥8, and 82% (74/90) of texts had a clinician MSS ≥8. Of the 90 total texts, 41 (46%) had consumer comments and 28 (31%) were commented on by clinicians. Overall, 10 consumers and 8 clinicians provided general open-text feedback on the text content. They made suggestions about the frequency (eg, "Every day is great," consumer; "I think two per day," clinician), timing (eg, "First thing in the morning for motivation and in the afternoon when we start lagging," consumer; "One to start the day that has some strategies e.g. relaxation and one at the end of the day/late arvo that is more reassuring," clinician), and tone (eg, "Change the language to a more active voice rather than passive," consumer; "Less bossy," clinician) of the text messages. Over 80% of clinician and consumer responses to individual texts endorsed participants "agreed" (mean 4) or "totally agreed" (mean 5) the texts were appropriate, likely to be useful, and likely to be effective. There was considerable agreement between groups. Overall, 66% (60/90) of the total texts were rated MSS <8 by both groups. In total, 10% (9/90) of texts were rated MSS <8 by consumers but MSS <8 by clinicians; 16% (14/90) of texts were rated MSS <8 by clinicians, but not by consumers. Overall, 8% (7/90) of texts were rated MSS <8 by both groups.

Table 2. Descriptive statistics for text message survey.

Question	Value, mean (SD)	Value, median (IQR)	Range (minimum-maximum)
Appropriate[a]			
Consumer	4.26 (0.785)	4 (1)	3 (2-5)
Clinician	4.38 (0.807)	5 (1)	4 (1-5)
Likely useful[b]			
Consumer	4.23 (0.840)	4 (1)	3 (2-5)
Likely effective[c]			
Clinician	4.26 (0.758)	4 (1)	4 (1-5)
Sum score[d]			
Consumer	8.49 (1.537)	8 (2)	6 (4-10)
Clinician	8.55 (1.291)	8 (2)	6 (4-10)

[a]Appropriate: easy to understand, sensitive to individual circumstances, and accurate.

[b]Useful: supporting and motivating patients to persist with reducing their dose of opioids.

[c]Likely to be effective: supporting and motivating patients to persist with opioid tapering.

[d]Sum score (both survey items for texts combined for each participant).

Revise mHealth Intervention

Video Script

On the basis of consumer and clinician feedback (see Multimedia Appendix 1 for a full summary), the following revisions were made to the draft video script. The overall length was shortened because of feedback that "the script is long. It takes many paragraphs to get to the point that the study is about supporting opioid reduction because opioids don't 'work' for chronic pain." The script initially described both chronic and acute pain in detail, with comprehensive definitions of both. The content in the revised script introduction was narrowed in length and focus to provide mostly an introduction to, and education about, chronic pain. A metaphor used to explain chronic pain was changed as suggested by a clinician. Emphasis on pain self-management strategies was prioritized. Information about the mHealth intervention was reduced, with a focus on increasing positive expectation bias, nocebo effects, and acceptability, rather than explaining the procedure of the study. On the basis of feedback, acknowledgments and normalization of tapering concerns were added to the introduction, and reasons for tapering opioids (ie, the benefits of reducing dose and the harms of continued use) were emphasized. The language used in the video was further simplified and changes were made (broadly and based on specific feedback). For example, clinicians suggested that the term "tapering" could be simplified to "reducing your dose," and this change was made in both the script and most texts. Finally, we added to the script to emphasize the importance of working collaboratively with the GP or prescribing physician.

Suggestions were made to change the content of consumer testimonials. As these testimonials had already taken place and the content could not be changed, the research group conducted interviews with 2 additional consumers (including culturally and linguistically diverse participants). The additional content included more discussions of strategies that were found helpful during tapering (eg, pain flare management). The interviewees also discussed their perspectives on managing chronic pain and the benefits of tapering opioid medications. Interview videos were edited using Adobe Premier Pro [64] video editing

software, and informational content was recorded using Microsoft PowerPoint [65]. The slideshow and consumer videos were then iteratively edited, improving the quality of the audio recording and making changes to images used in the slide deck. Finally, the video was uploaded to a web-based video streaming platform so that it could easily be made available for viewing within an internet browser. A transcription was prepared for closed captions for those with hearing or comprehension difficulties. As the video contains recordings of consumers, it is not publicly shared; however, a full transcript of the final script is included in Multimedia Appendix 1.

Text Messages

The texts were edited based on survey responses and open-text feedback. All texts were placed in rank order using their MSS score. First, the texts were ranked by consumer MSS. If the consumer MSS was the same, those scores were ranked by the clinician MSS. All individual text feedback was reviewed by members of the research team, even those with an MSS ≤8 in either group (ie, those that were not considered for inclusion in the final list). Where there was a large divergence between group MSS, or there was conflict in the open-text feedback, revisions attempted to address the concerns but not in a way that interfered with the original intention of the text (ie, the communication strategy). Texts were not included in the final list if the MSS for that item (across both clinicians and consumers) was ≤8. For each text, based on the evaluation score and feedback, it was decided to keep, modify, or remove the messages. Of the 90 messages, 57 (63%) were commented on, 36 (40%) modifications were made, and 3 (3%) texts were removed from the draft message bank. These comments also resulted in the development of 2 additional messages. The 2 examples are listed in Table 3. The final message list and further examples of revisions are provided in the Multimedia Appendix 1.

Table 3. Examples of text revisions.

Original message	Feedback	Decision and revised message
Message 7, pain education and self-management: "Relaxation strategies can help to reduce muscle tension and pain. Try deep breathing exercise for a few minutes several times per day."	• Consumer MSS[a] 9; consumer feedback: "Link with doing relaxation with somebody if really tense or use a tape that is specific to you." • Clinician MSS 9.5; no clinician feedback.	• Revised message, included in the final list: "Relaxation strategies can help to reduce muscle tension and pain. Try deep breathing exercise for a few minutes several times per day. Finding a recording online that works for you can be a great help too! USYD"
Message 35, opioid tapering support: "Opioid medications can be useful for pain in the short term (a few days or weeks). However, research shows that opioids offer only limited relief in the long term."	• Consumer MSS 7.5; consumer feedback: "It is important to remember that if you have had experience with long-term opioid use, even short-term use will need to be closely monitored as it can open you to thinking you may need them long term again." • Clinician MSS 6; clinician feedback: "It seems like a backward step to receive a message promoting opioids for acute pain and irrelevant."	• Removed message from the library. • Accept feedback from the clinician reviewer. • Unnecessary message.

[a]MSS: mean sum score.

In response to general open-text feedback on the texts, the decision was made to include 2 texts per day, with 1 text in the morning and 1 in the afternoon for 28 days. The morning texts were sent between 9 AM and 11 AM, and the afternoon texts were sent between 2 PM and 5 PM. The texts were reviewed to ensure that the communication strategy was effective throughout the day. Texts were also reviewed to ensure that the language was simple, that the key concepts were concisely delivered, that language was persuasive without being "bossy" (a specific piece of feedback), and that an active voice was used for more of the texts. The sign-off "USYD" (for the University of Sydney) was added to the end of each text to ensure consistency and clarity for consumers [40].

In total, 54 of the draft texts were included in the final text list. In total, 59% (32/54) of the texts in the final list received qualitative feedback from a consumer or clinician, and 19% (10/54) were modified based on feedback during the review. Of the 54 final texts, the MSS for consumers was 8.98 (SD 0.693), and the MSS for clinicians was 8.981 (SD 0.725). Of 32 "pain education and self-management" texts, 17 (53%) were in the final list. Of 32 "motivation to taper opioids" texts, 19 (59%) were in the final list. Of 26 "opioid tapering support" texts, 18 (69%) were in the final list. The remaining 2 texts included an introductory text on commencing the study (day 1, text 1), which asked participants to "Save this phone number as TXTSupport@USYD" and another text that included links to web-based pain management resources (day 1, text 2). One "motivation to taper opioids" message was modified to include a final message for participants (day 28, text 2). To ensure fidelity for the evaluation of the content, texts were standardized so that the content is delivered in the same order and at the same time across the intervention, and no additional texts would be sent from the content pool [66]. Therefore, based on the design features of the platform selected for the pilot study (see the Discussion section), it was decided that only 1-way text communication (mHealth to user) could be used for this iteration of the intervention. An automated feature of the sending technology is to cease the sending of any texts if a user sends a reply saying "Stop."

Discussion

Principal Findings

This study describes the development of an mHealth intervention to support people with chronic pain who are tapering prescription opioid medications. Using a co-design methodology, we sought input from 12 consumers with chronic pain who had experience tapering prescription opioids and 12 clinicians with experience supporting opioid tapering in chronic pain treatment services to develop an mHealth program consisting of a 10-minute video and 56 mobile phone texts. On the basis of consumer perspectives (formative research [16]) and research and resources from the areas of digital health, chronic pain, and opioid tapering, content was developed in which consumers and clinicians agreed was, on average, acceptable, likely to be useful (consumer perspective), and likely to be effective (clinician perspective) in supporting consumers to taper opioids. Most study participants provided optional open-text feedback on the proposed script or the texts. Participant feedback on the proposed content enabled further revisions and improvements.

Strengths of the Study

Consumers were involved in the conception and co-design of the intervention and intervention content. Consumers and experienced clinicians ("stakeholders") were represented within the research team, and an independent group of consumers and clinicians participated in the evaluation of the initial content development and design concept. Stakeholder participation in co-design should improve consumer and clinician engagement with digital health interventions and may also help improve the retention of participants in subsequent pilot studies and trials of mHealth support [23-27].

The conceptualization of this mHealth support occurred before the COVID-19 pandemic having disruptive effects on human movement and increasing the need for and use of technology-assisted health care. The current content development research was conducted during a period where Australia, similar to many other countries, was engaged in a variety of social distancing measures and, at times, strict isolation and quarantine. Whether acceptability or preferred content delivery preferences changed within this timeline of disruption because of the COVID-19 pandemic remains to be determined. However, it could be predicted that as digital health technologies have become more normalized since the onset of the COVID-19 pandemic, the acceptability of mHealth interventions may have increased further since this study was conducted.

Various reviews have indicated the need for digital health interventions to support chronic pain [21,67,68]. Most of the interventions described in these reviews are web-based interventions and often include some degree of clinician support. A variety of digital interventions have been designed to support people with chronic pain, including texting interventions [69], the use of smartphone-based chatbots [70,71], and more sophisticated technologies such as virtual reality [72,73]. However, to our knowledge, no published studies have described mHealth interventions developed specifically to support people with chronic pain to reduce their reliance on opioids. Only 1 other planned research project has been identified, where the purpose is specifically to support individuals with chronic pain to reduce their reliance on opioids [74]. A pilot study protocol describes the proposed intervention, which uses a mobile app [74]. This study also used a co-design process for content development. The content developed in our study represents an advancement in the support available to individuals with chronic pain who are tapering, particularly end users who prefer to text over a more sophisticated app or other digital support [16,74].

Limitations of the Study

Video Script

Notably, the participants did not evaluate the final video; they only read a script. This may have influenced the evaluation of the content, as video testimonials may be more engaging and persuasive than written testimonials [62]. Furthermore, despite the survey ratings indicating agreement on the appropriateness and likely effectiveness of the script to support consumers, only 7 clinicians provided text feedback for the script. To overcome these difficulties, the research group (including clinicians and a consumer representative) iteratively edited the video multiple times to ensure that it was of a professional standard.

SMS Text Messages

This study aimed to balance the desire for a rigorous co-design methodology with a need for efficiency and to reduce the demands on study participants. To address the latter issue, participants evaluated only 15 of the 90 texts from the draft list. We believe that the use of MSS and a cut-off of an MSS 8 to evaluate the content, which has been used in a similar study, was appropriate for our objectives [41]. However, because the data only included 2 participants' ratings from each group per text to calculate an MSS, it is acknowledged that this is less representative than all participants rating all texts that we would have done if we believed it feasible for participants. In our study, 1 participant's perspective significantly influenced the MSS. To mitigate this, open-text feedback was carefully considered, rather than relying solely on survey responses. Given the reasonably small sample groups (n=12 per group), it should be acknowledged that this study largely aimed to gain perspectives from individuals with experience in the domains of tapering and chronic pain using purposeful sampling and targeted recruitment of participants known to have chronic pain and experience tapering, or as a clinician with experience supporting this group. Therefore, the demographic information collected from participants, as reported in the paper, was minimal. The design could not be considered a completely representative sample of potential end users of the proposed mHealth support (Australians with chronic pain who are tapering prescription opioids). In retrospect, with the aim of describing the co-design methods in the development of this intervention and to facilitate evaluation and discussion in the academic community, demographic details should have been collected with more rigor. Despite this limitation, we believe that a good balance was struck to meet the aims of developing, pretesting, and revising the mHealth content minimizing the participant burden [28].

Next Steps

An important step is to confirm and validate our mHealth content in a diverse group of patients (eg, age, race or ethnicity, and pain condition). A multisite pilot randomized controlled trial will be conducted to further evaluate the intervention using the content developed in this study [75,76]. The pilot randomized controlled trial will recruit consumers from pain clinics who are tapering opioids to evaluate the acceptability, feasibility, and preliminary effectiveness (improving tapering self-efficacy) of the mHealth content (video and texts) and usual care, compared with usual care alone [75,76]. Further revisions and refinement of the intervention (content and design) can be made based on feedback from the trial participants [28,76]. Of particular interest is the acceptability and feasibility of the videos. The video helps address consumer-identified needs for education about tapering and chronic pain before receiving texts [16]. However, the inclusion of internet-based videos (accessed through a website or downloaded) requires higher technological literacy and more sophisticated technology (eg, smart phones, tablets, or laptops with access to the internet) than receiving texts alone. Furthermore, while a range of strategies are used in this mHealth support, those included are not exhaustive, and various other strategies could be integrated into the text message format [46,77]. Feedback from trial participants may also identify potential modifications to the text. These could include manipulating delivery characteristics (eg, 2-way texting or changing the frequency of texts) [16] or content (eg, different information or communication strategies). Our chosen content development model encourages an ongoing process of piloting and reviewing content, and our future research will aim to continuously improve the content developed in this study [28]. A range of different consumer and stakeholder engagement strategies can be used in the codevelopment of interventions [78,79], and many commercially available chronic pain apps (mHealth interventions) do not feature co-design [80]. The model selected for the co-design in this study was chosen because it is specifically for text intervention development [28]. However, any researchers engaged in future research into mHealth co-design should explore various models and methods to ensure that their design methodology meets the needs of their end users and the systems that they will develop (eg, texting, app, or browser-based).

Conclusions

Using a co-design methodology with 12 consumers and 12 clinicians, this study describes the development of a 28-day mHealth intervention consisting of a 10-minute video and 56 text messages. The intervention aimed to support people with chronic pain to taper prescription opioids. Overall, the content of the video and text messages were considered generally acceptable and likely to be useful and effective by both consumers and clinicians. Feedback provided by study participants informed improvements to both the script and texts, which should improve the acceptability of the intervention in a pilot randomized controlled trial. A randomized controlled trial will investigate the acceptability and feasibility of the mHealth intervention among a group of patients with chronic pain recruited from tertiary pain clinics who are tapering prescription opioids under the supervision of a pain specialist [75,76].

Acknowledgments

This research was supported by a philanthropic gift to the University of Sydney from the Ernest Heine Family Foundation and by an Australian Government Research Training Program Scholarship (for MRM). MLF holds a National Health and Medical Research Council of the Australia Investigator Leader Fellowship. The supporting sources had no role in the conceptualization, design, interpretation, or decision to publish the results.

Authors' Contributions

PG and CEAJ conceptualized the interventions. PG, CEAJ, MRM, AG, and MLF designed and evaluated the procedures. The roles of the initial content development are described in the study methods. Recruitment and data collection were supported by PG, CEAJ, MRM, AG, AGM, AS, MLF, and LD. AG was responsible for data capture and reporting. MRM, AG, AGM, PG, and CEAJ conducted the initial data review and analysis. All authors have contributed to the final content (at various stages, as described in the manuscript). MRM drafted the original manuscript. All authors critically reviewed and edited the manuscript and approved its final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Summary of open-text feedback for script and text messages, final script, examples of the review process for text messages, and text message content.
[DOCX File , 86 KB - formative_v6i11e40507_app1.docx]

References

1. Dowell D, Compton WM, Giroir BP. Patient-centered reduction or discontinuation of long-term opioid analgesics: the HHS guide for clinicians. JAMA 2019 Nov 19;322(19):1855-1856 [FREE Full text] [doi: 10.1001/jama.2019.16409] [Medline: 31600366]
2. Dowell D, Haegerich TM, Chou R. CDC guideline for prescribing opioids for chronic pain--United States, 2016. JAMA 2016 Apr 19;315(15):1624-1645 [FREE Full text] [doi: 10.1001/jama.2016.1464] [Medline: 26977696]
3. Baldini A, Von Korff M, Lin EH. A review of potential adverse effects of long-term opioid therapy: a practitioner's guide. Prim Care Companion CNS Disord 2012;14(3):PCC.11m01326 [FREE Full text] [doi: 10.4088/PCC.11m01326] [Medline: 23106029]
4. Ashburn MA, Fleisher LA. Increasing evidence for the limited role of opioids to treat chronic noncancer pain. JAMA 2018 Dec 18;320(23):2427-2428. [doi: 10.1001/jama.2018.19327] [Medline: 30561463]
5. Agnoli A, Xing G, Tancredi DJ, Magnan E, Jerant A, Fenton JJ. Association of dose tapering with overdose or mental health crisis among patients prescribed long-term opioids. JAMA 2021 Aug 03;326(5):411-419 [FREE Full text] [doi: 10.1001/jama.2021.11013] [Medline: 34342618]
6. Goesling J, DeJonckheere M, Pierce J, Williams DA, Brummett CM, Hassett AL, et al. Opioid cessation and chronic pain: perspectives of former opioid users. Pain 2019 May;160(5):1131-1145 [FREE Full text] [doi: 10.1097/j.pain.0000000000001493] [Medline: 30889052]
7. Frank JW, Levy C, Matlock DD, Calcaterra SL, Mueller SR, Koester S, et al. Patients' perspectives on tapering of chronic opioid therapy: a qualitative study. Pain Med 2016 Oct;17(10):1838-1847 [FREE Full text] [doi: 10.1093/pm/pnw078] [Medline: 27207301]
8. Avery N, McNeilage AG, Stanaway F, Ashton-James CE, Blyth FM, Martin R, et al. Efficacy of interventions to reduce long term opioid treatment for chronic non-cancer pain: systematic review and meta-analysis. BMJ 2022 Apr 04;377:e066375 [FREE Full text] [doi: 10.1136/bmj-2021-066375] [Medline: 35379650]
9. Nicholas MK, Asghari A, Sharpe L, Beeston L, Brooker C, Glare P, et al. Reducing the use of opioids by patients with chronic pain: an effectiveness study with long-term follow-up. Pain 2020 Mar;161(3):509-519. [doi: 10.1097/j.pain.0000000000001763] [Medline: 31764391]
10. Pain Australia. The National Strategic Action Plan for Pain Management. Department of Health and Aged Care, Australian Government. Deakin, Australia; 2021 May 17. URL: https://www.health.gov.au/resources/publications/the-national-strategic-action-plan-for-pain-management [accessed 2022-05-04]
11. Choinière M, Peng P, Gilron I, Buckley N, Williamson O, Janelle-Montcalm A, et al. Accessing care in multidisciplinary pain treatment facilities continues to be a challenge in Canada. Reg Anesth Pain Med 2020 Dec;45(12):943-948. [doi: 10.1136/rapm-2020-101935] [Medline: 33024007]
12. Impact of opioid regulatory reforms on people living with chronic pain. Pain Australia. 2022 Apr. URL: https://www.painaustralia.org.au/static/uploads/files/painaustralia-impact-of-opioid-regulatory-reforms-on-people-living-with-chronic-pai-wfopkycfkmnq.pdf [accessed 2022-05-04]
13. Ashton-James CE, Glare P, Darnall BD. Out of office hours: scalable, on-demand, digital support for patients tapering prescription opioids. Pain 2020 Oct;161(10):2252-2254. [doi: 10.1097/j.pain.0000000000001947] [Medline: 32501860]
14. Watts G. COVID-19 and the digital divide in the UK. Lancet Digit Health 2020 Aug;2(8):e395-e396 [FREE Full text] [doi: 10.1016/S2589-7500(20)30169-2] [Medline: 32835198]
15. McNeilage AG, Avery NS, Holliday S, Glare PA, Ashton-James CE. A qualitative trajectory analysis of patients' experiences tapering opioids for chronic pain. Pain 2022 Feb 01;163(2):e246-e260. [doi: 10.1097/j.pain.0000000000002336] [Medline: 33990111]
16. Magee MR, McNeilage AG, Avery N, Glare P, Ashton-James CE. mHealth interventions to support prescription opioid tapering in patients with chronic pain: qualitative study of patients' perspectives. JMIR Form Res 2021 May 18;5(5):e25969 [FREE Full text] [doi: 10.2196/25969] [Medline: 34003133]
17. Buck C, Keweloh C, Bouras A, Simoes EJ. Efficacy of short message service text messaging interventions for postoperative pain management: systematic review. JMIR Mhealth Uhealth 2021 Jun 16;9(6):e20199 [FREE Full text] [doi: 10.2196/20199] [Medline: 34132646]
18. Head KJ, Noar SM, Iannarino NT, Grant Harrington N. 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]
19. Orr JA, King RJ. Mobile phone SMS messages can enhance healthy behaviour: a meta-analysis of randomised controlled trials. Health Psychol Rev 2015;9(4):397-416. [doi: 10.1080/17437199.2015.1022847] [Medline: 25739668]
20. Scott-Sheldon LA, Lantini R, Jennings EG, Thind H, Rosen RK, Salmoirago-Blotcher E, et al. Text messaging-based interventions for smoking cessation: a systematic review and meta-analysis. JMIR Mhealth Uhealth 2016 May 20;4(2):e49 [FREE Full text] [doi: 10.2196/mhealth.5436] [Medline: 27207211]
21. Hewitt S, Sephton R, Yeowell G. The effectiveness of digital health interventions in the management of musculoskeletal conditions: systematic literature review. J Med Internet Res 2020 Jun 05;22(6):e15617 [FREE Full text] [doi: 10.2196/15617] [Medline: 32501277]

22. Anthony CA, Rojas E, Glass N, Keffala V, Noiseux N, Elkins J, et al. A psycholgical intervention delivered by automated mobile phone messaging stabilized hip and knee function during the COVID-19 pandemic: a randomized controlled trial. J Arthroplasty 2022 Mar;37(3):431-7.e3 [FREE Full text] [doi: 10.1016/j.arth.2021.12.006] [Medline: 34906660]
23. Institute of Medicine, Board on Health Sciences Policy, Forum on Drug Discovery, Development, and Translation. In: English RA, Lebovitz Y, Giffin RB, editors. Transforming Clinical Research in the United States: Challenges and Opportunities: Workshop Summary. Washington, DC, USA: National Academies Press (US); 2010.
24. Caldwell PH, Hamilton S, Tan A, Craig JC. Strategies for increasing recruitment to randomised controlled trials: systematic review. PLoS Med 2010 Nov 09;7(11):e1000368 [FREE Full text] [doi: 10.1371/journal.pmed.1000368] [Medline: 21085696]
25. Bagley HJ, Short H, Harman NL, Hickey HR, Gamble CL, Woolfall K, et al. A patient and public involvement (PPI) toolkit for meaningful and flexible involvement in clinical trials - a work in progress. Res Involv Engagem 2016 Apr 27;2:15 [FREE Full text] [doi: 10.1186/s40900-016-0029-8] [Medline: 29062516]
26. Symons T, McKenzie A, Straiton N. Consumer Involvement and Engagement Toolkit. ACTA (Austraian Clinical Trials Alliance). 2019. URL: https://involvementtoolkit.clinicaltrialsalliance.org.au/ [accessed 2022-01-02]
27. Crocker JC, Ricci-Cabello I, Parker A, Hirst JA, Chant A, Petit-Zeman S, et al. Impact of patient and public involvement on enrolment and retention in clinical trials: systematic review and meta-analysis. BMJ 2018 Nov 28;363:k4738 [FREE Full text] [doi: 10.1136/bmj.k4738] [Medline: 30487232]
28. Abroms LC, Whittaker R, Free C, Mendel Van Alstyne J, Schindler-Ruwisch JM. Developing and pretesting a text messaging program for health behavior change: recommended steps. JMIR Mhealth Uhealth 2015 Dec 21;3(4):e107 [FREE Full text] [doi: 10.2196/mhealth.4917] [Medline: 26690917]
29. Feng B, Malloch YZ, Kravitz RL, Verba S, Iosif AM, Slavik G, et al. Assessing the effectiveness of a narrative-based patient education video for promoting opioid tapering. Patient Educ Couns 2021 Feb;104(2):329-336 [FREE Full text] [doi: 10.1016/j.pec.2020.08.019] [Medline: 32900605]
30. Baumeister H, Seifferth H, Lin J, Nowoczin L, Lüking M, Ebert D. Impact of an acceptance facilitating intervention on patients' acceptance of Internet-based pain interventions: a randomized controlled trial. Clin J Pain 2015 Jun;31(6):528-535. [doi: 10.1097/AJP.0000000000000118] [Medline: 24866854]
31. Morrison LG. Theory-based strategies for enhancing the impact and usage of digital health behaviour change interventions: a review. Digit Health 2015 Jul 17;1:2055207615595335 [FREE Full text] [doi: 10.1177/2055207615595335] [Medline: 29942544]
32. 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]
33. Quinlan J, Willson H, Grange K. Hopes and fears before opioid tapering: a quantitative and qualitative study of patients with chronic pain and long-term opioids. Br J Pain 2021 May;15(2):120-128 [FREE Full text] [doi: 10.1177/2049463720974053] [Medline: 34055333]
34. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977 Mar;84(2):191-215. [doi: 10.1037//0033-295x.84.2.191] [Medline: 847061]
35. Bandura A. Self-Efficacy: The Exercise of Control. New York, NY, USA: W.H. Freeman; 1997.
36. Bandura A. The explanatory and predictive scope of self-efficacy theory. J Soc Clin Psychol 1986 Sep;4(3):359-373. [doi: 10.1521/jscp.1986.4.3.359]
37. Wilson M, Finlay M, Orr M, Barbosa-Leiker C, Sherazi N, Roberts ML, et al. Engagement in online pain self-management improves pain in adults on medication-assisted behavioral treatment for opioid use disorders. Addict Behav 2018 Nov;86:130-137. [doi: 10.1016/j.addbeh.2018.04.019] [Medline: 29731244]
38. Palinkas LA, Horwitz SM, Green CA, Wisdom JP, Duan N, Hoagwood K. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. Adm Policy Ment Health 2015 Sep;42(5):533-544 [FREE Full text] [doi: 10.1007/s10488-013-0528-y] [Medline: 24193818]
39. Consumer Advisory Group. PainAustralia. 2022. URL: https://www.painaustralia.org.au/about-us/consumer-advisory-group [accessed 2021-10-12]
40. Fritsch CG, Ferreira PH, Prior JL, Clavisi O, Chow CK, Redfern J, et al. TEXT4myBACK: a text message intervention to improve function in people with low back pain-protocol of a randomized controlled trial. Phys Ther 2021 Jul 01;101(7):pzab100. [doi: 10.1093/ptj/pzab100] [Medline: 33764461]
41. Gassen Fritsch C, Prior J, Ferreira P, McLachlan A, Chow C, Redfern J, et al. 'TEXT4MYBACK' - the development process of a self-management intervention delivered via text message for low back pain. Osteoarthritis Cartilage 2019 Apr 1;27(2):S458 [FREE Full text] [doi: 10.1016/j.joca.2019.02.499]
42. Butler DS, Moseley GL. Explain Pain. 2nd Edition. Adelaide, Australia: Noigroup publications; 2013.
43. Nicholas M, Molloy A, Beeston L, Tonkin L. Manage Your Pain: Practical and Positive Ways of Adapting to Chronic Pain. Chicago, IL, USA: Souvenir Press; 2012.
44. Moseley GL. Painful Yarns: Metaphors & Stories to Help Understand the Biology of Pain. Minneapolis, MN, USA: Dancing Giraffe Press; 2007.

45. Darnall BD, Roy A, Chen AL, Ziadni MS, Keane RT, You DS, et al. Comparison of a single-session pain management skills intervention with a single-session health education intervention and 8 sessions of cognitive behavioral therapy in adults with chronic low back pain: a randomized clinical trial. JAMA Netw Open 2021 Aug 02;4(8):e2113401 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.13401] [Medline: 34398206]
46. Sharpe L, Jones E, Ashton-James CE, Nicholas MK, Refshauge K. Necessary components of psychological treatment in pain management programs: a Delphi study. Eur J Pain 2020 Jul;24(6):1160-1168. [doi: 10.1002/ejp.1561] [Medline: 32187442]
47. Henry SG, Paterniti DA, Feng B, Iosif AM, Kravitz RL, Weinberg G, et al. Patients' experience with opioid tapering: a conceptual model with recommendations for clinicians. J Pain 2019 Feb;20(2):181-191 [FREE Full text] [doi: 10.1016/j.jpain.2018.09.001] [Medline: 30243859]
48. Darnall BD, Ziadni MS, Krishnamurthy P, Flood P, Heathcote LC, Mackey IG, et al. "My surgical success": effect of a digital behavioral pain medicine intervention on time to opioid cessation after breast cancer surgery-a pilot randomized controlled clinical trial. Pain Med 2019 Nov 01;20(11):2228-2237 [FREE Full text] [doi: 10.1093/pm/pnz094] [Medline: 31087093]
49. Darnall BD, Ziadni MS, Stieg RL, Mackey IG, Kao MC, Flood P. Patient-centered prescription opioid tapering in community outpatients with chronic pain. JAMA Intern Med 2018 May 01;178(5):707-708 [FREE Full text] [doi: 10.1001/jamainternmed.2017.8709] [Medline: 29459978]
50. Henry SG, Feng B, Verba S, Kravitz RL, Iosif AM. The story vs the storyteller: factors associated with the effectiveness of brief video-recorded patient stories for promoting opioid tapering. Health Expect 2021 Jun;24(3):991-999 [FREE Full text] [doi: 10.1111/hex.13243] [Medline: 33835644]
51. Nicholas M. Manage Your Pain. 3rd edition. Sydney, Australia: HarperCollins; 2011.
52. Bohnert AS, Ilgen MA. Understanding links among opioid use, overdose, and suicide. N Engl J Med 2019 Jan 03;380(1):71-79. [doi: 10.1056/NEJMra1802148] [Medline: 30601750]
53. Szinay D, Jones A, Chadborn T, Brown J, Naughton F. Influences on the uptake of and engagement with health and well-being smartphone apps: systematic review. J Med Internet Res 2020 May 29;22(5):e17572 [FREE Full text] [doi: 10.2196/17572] [Medline: 32348255]
54. Wei Y, Zheng P, Deng H, Wang X, Li X, Fu H. Design features for improving mobile health intervention user engagement: systematic review and thematic analysis. J Med Internet Res 2020 Dec 09;22(12):e21687 [FREE Full text] [doi: 10.2196/21687] [Medline: 33295292]
55. Short CE, Rebar AL, Plotnikoff RC, Vandelanotte C. Designing engaging online behaviour change interventions: a proposed model of user engagement. Eur Health Psychol 2015;17(1):32-38.
56. Cole-Lewis H, Ezeanochie N, Turgiss J. Understanding health behavior technology engagement: pathway to measuring digital behavior change interventions. JMIR Form Res 2019 Oct 10;3(4):e14052 [FREE Full text] [doi: 10.2196/14052] [Medline: 31603427]
57. Linton SJ, Boersma K, Vangronsveld K, Fruzzetti A. Painfully reassuring? The effects of validation on emotions and adherence in a pain test. Eur J Pain 2012 Apr;16(4):592-599. [doi: 10.1016/j.ejpain.2011.07.011] [Medline: 22396087]
58. Linton SJ, McCracken LM, Vlaeyen JW. Reassurance: help or hinder in the treatment of pain. Pain 2008 Jan;134(1-2):5-8. [doi: 10.1016/j.pain.2007.10.002] [Medline: 18035496]
59. Hors-Fraile S, Schneider F, Fernandez-Luque L, Luna-Perejon F, Civit A, Spachos D, et al. Tailoring motivational health messages for smoking cessation using an mHealth recommender system integrated with an electronic health record: a study protocol. BMC Public Health 2018 Jun 05;18(1):698 [FREE Full text] [doi: 10.1186/s12889-018-5612-5] [Medline: 29871595]
60. Graham AL, Papandonatos GD, Jacobs MA, Amato MS, Cha S, Cohn AM, et al. Optimizing text messages to promote engagement with internet smoking cessation treatment: results from a factorial screening experiment. J Med Internet Res 2020 Apr 02;22(4):e17734 [FREE Full text] [doi: 10.2196/17734] [Medline: 32238338]
61. Redfern J, Santo K, Coorey G, Thakkar J, Hackett M, Thiagalingam A, et al. Factors influencing engagement, perceived usefulness and behavioral mechanisms associated with a text message support program. PLoS One 2016 Oct 14;11(10):e0163929 [FREE Full text] [doi: 10.1371/journal.pone.0163929] [Medline: 27741244]
62. Jensen JD, King AJ, Carcioppolo N, Davis L. Why are tailored messages more effective? A multiple mediation analysis of a breast cancer screening intervention. J Commun 2012 Oct;62(5):851-868 [FREE Full text] [doi: 10.1111/j.1460-2466.2012.01668.x] [Medline: 26405350]
63. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
64. Adobe Premier Pro. Adobe Systems. 2017. URL: https://www.adobe.com/products/premiere.html [accessed 2021-04-15]
65. Microsoft PowerPoint. Microsoft Corporation. 2022. URL: https://www.microsoft.com/en-in/microsoft-365/powerpoint [accessed 2021-05-17]

66. Murray E, Hekler EB, Andersson G, Collins LM, Doherty A, Hollis C, et al. Evaluating digital health interventions: key questions and approaches. Am J Prev Med 2016 Nov;51(5):843-851 [FREE Full text] [doi: 10.1016/j.amepre.2016.06.008] [Medline: 27745684]
67. Gandy M, Pang ST, Scott AJ, Heriseanu AI, Bisby MA, Dudeney J, et al. Internet-delivered cognitive and behavioural based interventions for adults with chronic pain: a systematic review and meta-analysis of randomized controlled trials. Pain 2022 Oct 01;163(10):e1041-e1053. [doi: 10.1097/j.pain.0000000000002606] [Medline: 35121696]
68. Mariano TY, Wan L, Edwards RR, Jamison RN. Online teletherapy for chronic pain: a systematic review. J Telemed Telecare 2021 May;27(4):195-208. [doi: 10.1177/1357633X19871746] [Medline: 31488004]
69. Sherman KA, Pehlivan MJ, Singleton A, Hawkey A, Redfern J, Armour M, et al. Co-design and development of ENDOTEXT, a supportive text message intervention for individuals living with endometriosis: mixed methods study. JMIR Form Res 2022. [doi: 10.2196/preprints.40837]
70. Hauser-Ulrich S, Künzli H, Meier-Peterhans D, Kowatsch T. A smartphone-based health care chatbot to promote self-management of chronic pain (SELMA): pilot randomized controlled trial. JMIR Mhealth Uhealth 2020 Apr 03;8(4):e15806 [FREE Full text] [doi: 10.2196/15806] [Medline: 32242820]
71. Gupta M, Malik T, Sinha C. Delivery of a mental health intervention for chronic pain through an artificial intelligence-enabled app (Wysa): protocol for a prospective pilot study. JMIR Res Protoc 2022 Mar 31;11(3):e36910 [FREE Full text] [doi: 10.2196/36910] [Medline: 35314423]
72. Darnall BD, Krishnamurthy P, Tsuei J, Minor JD. Self-administered skills-based virtual reality intervention for chronic pain: randomized controlled pilot study. JMIR Form Res 2020 Jul 07;4(7):e17293 [FREE Full text] [doi: 10.2196/17293] [Medline: 32374272]
73. Giravi HY, Biskupiak Z, Tyler LS, Bulaj G. Adjunct digital interventions improve opioid-based pain management: impact of virtual reality and mobile applications on patient-centered pharmacy care. Front Digit Health 2022 Jun 13;4:884047 [FREE Full text] [doi: 10.3389/fdgth.2022.884047] [Medline: 35770137]
74. Edmond SN, Wesolowicz DM, Moore BA, Ibarra J, Chhabra M, Fraenkel L, et al. Opioid tapering support using a web-based app: development and protocol for a pilot randomized controlled trial. Contemp Clin Trials 2022 Aug;119:106857. [doi: 10.1016/j.cct.2022.106857] [Medline: 35863697]
75. A Pilot Randomised Controlled Trial of SMS Text Messages to Support Tapering Opioids for Chronic Pain. Australian New Zealand Clinical Trials Registry. 2021. URL: https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=381757&isReview=true [accessed 2021-07-07]
76. Magee M, Gholamrezaei A, McNeilage AG, Dwyer L, Sim A, Ferreira M, et al. Evaluating acceptability and feasibility of a mobile health intervention to improve self-efficacy in prescription opioid tapering in patients with chronic pain: protocol for a pilot randomised, single-blind, controlled trial. BMJ Open 2022 Apr 26;12(4):e057174 [FREE Full text] [doi: 10.1136/bmjopen-2021-057174] [Medline: 35473742]
77. Anthony CA, Rojas EO, Keffala V, Glass NA, Shah AS, Miller BJ, et al. Acceptance and commitment therapy delivered via a mobile phone messaging robot to decrease postoperative opioid use in patients with orthopedic trauma: randomized controlled trial. J Med Internet Res 2020 Jul 29;22(7):e17750 [FREE Full text] [doi: 10.2196/17750] [Medline: 32723723]
78. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ 2021 Sep 30;374:n2061 [FREE Full text] [doi: 10.1136/bmj.n2061] [Medline: 34593508]
79. Roper C, Grey F, Cadogan E. Co-production: putting principles into practice in mental health contexts. University of Melbourne. 2018. URL: https://healthsciences.unimelb.edu.au/__data/assets/pdf_file/0007/3392215/Coproduction_putting-principles-into-practice.pdf [accessed 2021-03-03]
80. Gamwell KL, Kollin SR, Gibler RC, Bedree H, Bieniak KH, Jagpal A, et al. Systematic evaluation of commercially available pain management apps examining behavior change techniques. Pain 2021 Mar 01;162(3):856-865 [FREE Full text] [doi: 10.1097/j.pain.0000000000002090] [Medline: 33003110]

Abbreviations

mHealth: mobile health
MSS: mean sum score
REDCap: Research Electronic Data Capture

Edited by A Mavragani; submitted 26.06.22; peer-reviewed by F Howard, A Civit, S Bartels; comments to author 15.08.22; revised version received 05.09.22; accepted 06.09.22; published 10.11.22.

Please cite as:
Magee MR, Gholamrezaei A, McNeilage AG, Sim A, Dwyer L, Ferreira ML, Darnall BD, Glare P, Ashton-James CE
A Digital Video and Text Messaging Intervention to Support People With Chronic Pain During Opioid Tapering: Content Development Using Co-design
JMIR Form Res 2022;6(11):e40507
URL: https://formative.jmir.org/2022/11/e40507
doi:10.2196/40507
PMID:36355415

Original Paper

Mental Health Outcomes for Youths With Public Versus Private Health Insurance Attending a Telehealth Intensive Outpatient Program: Quality Improvement Analysis

Kate Gliske[1], PhD; Katie R Berry[1], PhD; Jaime Ballard[2], PhD; Michelle Evans-Chase[3], PhD; Phyllis L Solomon[3], PhD; Caroline Fenkel[1], MSS, LCSW, DSW

[1]Charlie Health, Bozeman, MT, United States

[2]Center For Applied Research and Educational Improvement, University of Minnesota, Saint Paul, MN, United States

[3]School of Social Policy and Practice, University of Pennsylvania, Philadelphia, PA, United States

Corresponding Author:
Kate Gliske, PhD
Charlie Health
233 E Main St.
Suite 401
Bozeman, MT, 59715
United States
Phone: 1 952 334 1411
Email: kate.gliske@charliehealth.com

Abstract

Background: COVID-19 exacerbated a growing mental health crisis among youths and young adults, worsened by a lack of existing in-person options for high-acuity care. The emergence and growth of remote intensive outpatient programs (IOPs) is a solution to overcome geographic limitations to care. However, it remains unclear whether remote IOPs engender equivalent clinical outcomes among youths with public insurance (eg, Medicaid) versus private insurance (eg, commercial) given the disparities found in previous research on place-based treatment in both clinical and engagement outcomes.

Objective: This analysis sought to establish, as part of ongoing quality improvement efforts, whether engagement and clinical outcomes among adolescents and young adults attending remote IOP treatment differed between youths with public and those with private insurance. The identification of disparities by payer type was used to inform programmatic decisions within the remote IOP system for which this quality improvement analysis was conducted.

Methods: Pearson chi-square analyses and independent 2-tailed *t* tests were used to establish that the 2 groups defined by insurance type were equivalent on clinical outcomes (depression, suicidal ideation, and nonsuicidal self-injury [NSSI]) at intake and compare changes in clinical outcomes. McNemar chi-square analyses and repeated-measure 2-tailed *t* tests were used to assess changes in clinical outcomes between intake and discharge in the sample overall. In total, 495 clients who attended the remote IOP for youths and young adults in 14 states participated in ≥7 treatment sessions, and completed intake and discharge surveys between July 2021 and April 2022 were included in the analysis.

Results: Overall, the youths and young adults in the remote IOP attended a median of 91% of their scheduled group sessions (mean 85.9%, SD 16.48%) and reported significantly fewer depressive symptoms at discharge (t_{447}=12.51; P<.001). McNemar chi-square tests of change indicated significant reductions from intake to discharge in suicidal ideation (N=470, χ^2_1=104.4; P<.001), with nearly three-quarters of youths who reported active suicidal ideation at intake (200/468, 42.7%) no longer reporting it at discharge (142/200, 71%), and in NSSI (N=430, χ^2_1=40.7; P<.001), with more than half of youths who reported NSSI at intake (205/428, 47.9%) reporting lower self-harm at discharge (119/205, 58%). No significant differences emerged by insurance type in attendance (median public 89%, median private 92%; P=.10), length of stay (t_{416}=−0.35; P=.73), or reductions in clinical outcomes (depressive symptom severity: t_{444}=−0.87 and P=.38; active suicidal ideation: N=200, χ^2_1=0.6 and P=.49; NSSI frequency: t_{426}=−0.98 and P=.33).

Conclusions: Our findings suggest that youths and young adults who participated in remote IOP had significant reductions in depression, suicidal ideation, and NSSI. Given access to the same remote high-acuity care, youths and young adults on both public and private insurance engaged in programming at comparable rates and achieved similar improvements in clinical outcomes.

(JMIR Form Res 2022;6(11):e41721) doi:10.2196/41721

KEYWORDS

telehealth; telepsychiatry; telemedicine; intensive outpatient; remote outpatient; mental health; quality improvement; routine outcome monitoring; mental health treatment; patient outcome; outpatient program; youth; young adult; depression; suicidal ideation; health outcome; outcome monitoring

Introduction

Background

The US Surgeon General recently declared youth mental health a state of emergency [1]. Nearly half of youths in the United States report feeling persistently sad or hopeless, and 9% report attempting suicide [2]. Left untreated, depression can lead to school dropout, unemployment, substance abuse, violence, and death [3-6]. It is critical for youths with depression to access effective treatments. When depression is severe and suicidality is high, intensive treatment can be life-saving [7]. However, less than half of US counties have facilities with programs for youths with severe mental health needs [8]. Telehealth may help address this critical shortage of services for youths with severe needs as telehealth intensive psychotherapy services can be accessed remotely from any location.

For telehealth to effectively address this gap, it must be able to serve youths of many demographic groups. However, previous research suggests that telehealth in intensive psychiatric services has poorer reach among youths on public insurance than among youths on private insurance [9], and previous research on intensive psychiatric telehealth among adults found poorer outcomes among those who were publicly insured [10]. Most public health insurance for youths is designed for low-income families, “making it imperative to monitor trends in access to services, including appointment attendance, among Medicaid-insured psychiatrically vulnerable youth” [9]. Research on disparities in reach and outcomes by insurance type is critical to ensure that services reach and meet the needs of all youths.

This paper presents the findings from ongoing quality improvement efforts conducted by *Charlie Health*, a national provider of remote intensive outpatient programs (IOPs) for adolescents and young adults whose program staff collect measurement-based clinical outcome data to assess treatment efficacy and meet the quality assurance reporting requirements of payers and providers as part of routine outcome monitoring. Routine outcome monitoring allows for the rapid translation of findings back into the refinement of a clinical care model and can be particularly useful for understanding what is or is not working in a newer treatment modality such as group telehealth [11,12]. The goal of the analysis reported in this paper was to determine (1) whether clinical outcomes improved during treatment and (2) whether programming provided equivalent engagement and clinical outcomes for all clients regardless of payer type. This outcome monitoring is essential to assess whether remote IOPs are engaging and effective for all clients and better understand and inform how the clinical treatment team approaches differences in barriers to care.

Clinical Outcomes

A primary diagnosis for which many of the youths at Charlie Health seek treatment is major depressive disorder and comorbid behavioral health issues associated with it, such as suicidal ideation and self-harm. As such, 3 of the primary clinical outcomes regularly assessed and analyzed by Charlie Health include depression, suicidal ideation, and nonsuicidal self-injury (NSSI).

Depression is one of the most common mental health disorders among youths and young adults, affecting nearly 1 in 5 [13]. It is linked to a host of difficulties across the life span, including unemployment, substance use disorders, and suicidality [3,4,14,15]. Suicide is the second most common cause of death among youths and young adults [6]. Among high school students, 19% have seriously considered suicide, and 16% have made a suicide plan [16]. Suicidal ideation and attempts are also linked to other high-risk behaviors such as substance and tobacco use [17,18]. NSSI commonly co-occurs with depression among youths and is a significant predictor of future suicide attempts [5]. Some methods of NSSI cause significant physical damage and health risks, and NSSI significantly increases the likelihood of hospitalization among youths with depression and suicidality [19].

Intensive Outpatient Services

Overview

Partial hospitalization, day treatment, and IOPs are important service offerings to address rising mental health needs and prevent emergency room visits and referrals to inpatient care for youths [20]. Moreover, these treatments are crucial service options given that emergency rooms are past capacity and there exists a current shortage of residential services for youths [21]. Furthermore, a growing research base on youth IOPs demonstrates significant symptom reduction and functional improvement [22-24] while being a more cost-effective alternative to inpatient or residential treatment [25,26].

Despite growing evidence of the effectiveness of IOPs in mitigating mental health severity, these programs are not accessible to all youths. Nationally, less than half (45%) of youths with a mental health diagnosis received treatment in the previous year [27] and, in a National Survey of Mental Health Treatment Facilities, only 63% of US counties had a facility providing mental health outpatient treatment for youths and less than half of US counties had a facility with programs for youths

with severe mental health needs [8]. Even among young people who find intensive care, long drives and transportation challenges can lead to drop out [28]. Transportation and scheduling barriers are even more common among youths and families with low income [29]. Therefore, telehealth intensive services may address the critical need to tackle barriers such as the lack of local services and transportation challenges.

Providing IOPs via Telehealth

Overview

For intensive outpatient services to meet the growing mental health needs of youths, they must be both *engaging* and *effective*. Services must be *engaging* to youths and young adults for them to remain in treatment long enough to receive an adequate dose. This requires structuring treatment to address pressing barriers such as transportation and scheduling challenges as well as building strong relationships with both clients and their families. Services must then be *effective* in reducing symptoms and improving functioning.

Telehealth IOPs may address transportation and scheduling barriers that could contribute to engagement, and meta-analyses document that telehealth psychiatric care has equivalent outcomes to those of in-person care [30,31]. However, there is a critical need to investigate potential disparities in engagement and effectiveness by demographic factors [32]. Preliminary research has led to a call for greater attention to treatment disparities by insurance type [9,10,32,33]. Such research is critical for informing equitable care. If there are disparities in engagement, it may be that services have inequitable barriers to participation. If there are disparities in outcomes, there may be differences in either the quality of services by insurance type or in the stressors faced by families by insurance type that require more effective tailoring of services.

Telehealth IOPs, Engagement, and Insurance Type

Preliminary research suggests that telehealth attendance among youths with public insurance is lower than among privately insured youths. A study of psychiatric intensive outpatient services that transitioned to telehealth at the start of the COVID-19 pandemic found that adolescents' attendance was higher only for privately insured youths but not for publicly insured youths [9]. Another investigation of claim data among publicly insured children similarly found that those with the lowest income had overall lower rates of telehealth use [34]. It is notable that research on adults shows no difference in attendance and retention for telehealth between publicly and privately insured individuals, although publicly insured adults had lower attendance for in-person services [10]. This suggests that, for children, attendance to telehealth sessions may be influenced by different barriers from those for adults and that children on publicly funded insurance plans may face other obstacles to attend.

Telehealth IOPs, Clinical Outcomes, and Insurance Type

In a recent study of publicly and privately insured adults in remote IOP services, clients with commercial insurance showed significantly greater improvement than those with public insurance despite having comparable treatment engagement [10]. No studies to our knowledge have investigated the differences in outcomes among youths attending telehealth IOP services. However, 3 studies have evaluated differences by insurance type in in-person intensive programs for youths. In each case, youths with private insurance had significantly greater improvements than youths with public insurance. A study of a transdiagnostic child partial hospitalization program found greater reductions in depression and emotional symptoms among children with private insurance relative to those with public insurance [32]. In the only study identified that collected data from multiple reporters, adolescents in an IOP treating self-harm reported no differences by insurance type, but parents of youths with private insurance reported significantly greater behavior improvement than parents of youths with public insurance [28]. A recent study of an IOP for self-harm examined treatment differences by insurance type with the largest sample to date (n=1327 [35]), finding a significantly greater increase in functioning for youths and young adults on private insurance than for those on public insurance. In each of these cases, the authors note that time and the financial commitments of the intensive program may cause greater stressors for youths and families on public insurance than for those on private insurance.

Aims and Clinical Implications

Charlie Health collects clinical outcome data from all clients to track changes in client needs as well as to iteratively inform organization-level processes. The analysis includes comparisons among client subgroups such as those defined by gender, sexual orientation, and—as reported in this paper—insurance type. This approach to quality improvement allows the program to identify the need for differential approaches that could increase engagement and improve patient outcomes. The ultimate goal of all Charlie Health quality improvement efforts is to assess and improve the quality of services for all youths and young adults in the program. Given the existing literature reviewed, the analysis presented in this paper sought to explore (1) whether depression, suicidality, and self-harm improved during treatment and (2) whether exposure to the same program engendered equivalent program engagement and changes in clinical outcomes for all clients regardless of insurance type (public vs private).

Methods

Overview

All youths attending Charlie Health are exposed to the same clinical program regardless of insurance type, thus allowing the program to make direct comparisons of engagement and outcomes across the 2 client groups in response to the same treatment. This paper represents one such analysis conducted to better understand the similarities and differences in outcomes and program engagement between clients with 2 different types of insurance: public and private. Data from self-reported intake and discharge assessments along with administrative program engagement tracking metrics (ie, rate of attendance and length of stay [LOS]) were used to explore differences between the 2 insurance subgroups.

Ethics Approval

This project was reviewed and determined by the University of Pennsylvania Institutional Review Board to qualify as quality improvement, indicating that these activities are not human participant research.

Client and Program Characteristics

Charlie Health clients come from many geographic regions across the United States, including rural, urban, and suburban communities where there is variable access to mental health services. Relatedly, the program accepts both public and private insurance in 10 of the 12 states in which it operates, removing a common access barrier to services for clients of varying socioeconomic backgrounds. Clients generally present with high-acuity primary and co-occurring mental and behavioral health needs. Approximately half of the clients have recent emergency room visits or inpatient stays related to mental health issues, whereas some clients are initiating mental health treatment for the first time and others are stepping up from a lower level of care. The client population comprises a myriad of marginalized cultural and social identities related to gender and sexual orientation.

Data for clients discharged from treatment between July 21, 2021, and April 28, 2022, were considered for inclusion in the analysis. Client cases were included if they met a minimum engagement threshold defined as ≥7 treatment sessions (approximately 21 hours) and 2 weeks in care and completed both intake and discharge surveys. Engagement criteria were based on neurological evidence suggesting that structural changes occur in response to cognitive and behavioral therapies after approximately 18 hours [36-39]. The resulting 495 cases included all client cases who met these criteria regardless of whether the client completed treatment. Reasons for client discharge before completing treatment vary and include discontinuing treatment against clinician advice, disengagement in treatment sessions, transfer to a higher level of care, and insurance denial.

The Charlie Health IOPs are all telehealth, offering a standard 9 hours of group therapy and an additional 1 hour of individual and 1 hour of family therapy per week. Group sessions are offered during the morning, afternoon, and evening hours to meet the demands of variable work, family, and school schedules. Charlie Health also offers additional optional family programs that include parent support groups that provide optional breakout sessions for parents of clients who are members of the lesbian, gay, bisexual, transgender, and queer community. When clients are admitted to the program, they are placed on a group track that is informed by their mental and behavioral health challenges and demographic characteristics. The program is multimodal in its therapeutic approach, offering evidence-based interventions specific to client needs. For instance, clients with a significant history of trauma are placed in trauma-informed cognitive behavioral therapy groups, and clients who are at high risk of suicide are placed in dialectical behavioral therapy groups. In addition to evidence-based interventions, clients attend general therapeutic processes and experiential groups (ie, mindfulness, creative arts, and music therapy).

Data Collection Procedures

Clinical outcome data on depression, self-harm, and suicidal risk and behaviors were collected at intake and discharge. Assessments were performed during the clients' first and last IOP sessions by program staff using a Qualtrics (Qualtrics International Inc) link to the assessment survey. To ensure completeness of the data collected, clients who did not attend their closing IOP group were sent a small financial incentive to complete their discharge survey via an emailed link. All intake and discharge assessment data were downloaded, deidentified, and uploaded to a secure cloud-based folder that was shared with the University of Pennsylvania assessment team, which conducted analyses monthly.

Measures

Client demographic characteristics were collected at treatment intake (age) and discharge (age, gender identity, and sexual orientation). Treatment episode data came from administrative records that track admission and discharge dates, the number of weeks in treatment, the total number of sessions scheduled and attended, and the discharge reason (eg, treatment completion, leaving against clinical advice, insurance denial, or transferring to a higher or lower level of care).

Treatment Engagement

Charlie Health tracks program engagement administratively (vs self-report), monitoring client attendance by session, day, and week as well as overall LOS and type of discharge (ie, treatment completed or discharged for other reasons [insurance denial, referral to a higher level of care, or disengagement from treatment]). The data used to assess treatment engagement differences by insurance type in this analysis included attendance rate measured by the proportion of treatment sessions attended versus those scheduled for the client, LOS measured as the number of weeks between intake and discharge, and type of discharge (treatment complete vs discharged for other reasons).

Patient Health Questionnaire Modified for Adolescents

The Patient Health Questionnaire modified for adolescents (PHQ-A) is a screening tool for depression that is administered before and after treatment. The PHQ-A is a 9-item self-report measure that classifies clients into 5 depression severity categories based on their score from 0 to 27: minimal (0-4), mild (5-9), moderate (10-14), moderately severe (15-19), and severe (20-27) [40]. The instructions ask clients to rate how bothered they have been by symptoms of depression over the past 2 weeks (ie, *Feeling down, depressed, irritable, or hopeless* or having *Thoughts that you would be better off dead, or of hurting yourself in some way*). Responses are rated on a 4-point Likert scale from 0 (not at all) to 3 (nearly every day) wherein the sum scores range from 0 to 27. The PHQ-A has been established as an excellent diagnostic screening tool for depression in adolescent samples [41,42]. Reliability in this sample was excellent at admission and discharge, with a Cronbach α of .91 at both time points.

Suicide Risk

Suicidality was measured by the program using the 5-item Ask Suicide-Screening Questions (ASQ) Toolkit. The ASQ was developed as a suicidal risk screening toolkit for health care providers wherein, if clients responded *yes* to any of items 1 through 4, they would be screened as *positive* for suicide risk and asked an additional question about current suicidal ideation. The ASQ 1 and ASQ 2 ask clients to reflect on passive suicidal ideation (*thinking you would be better off dead* and *felt that you or your family would be better off if you were dead)*, the ASQ 3 asks about active ideation (*In the past week, have you had thoughts of killing yourself*), and the ASQ 4 asks about previous attempts (*Have you ever tried to kill yourself?).* The 4 items were validated in a pediatric sample (ages 10-21 years) wherein the measure correctly identified 96.9% of the sample that screened *positive* for suicide risk [43].

NSSI Measure

NSSI was assessed by the program using Criterion A of the Alexian Brothers Assessment of Self-Injury (ABASI) scale [44]. The ABASI was created to measure clinical severity of NSSI behaviors. The scale is broken up into 4 criteria that mirror the Diagnostic and Statistical Manual of Mental Disorders criteria for NSSI disorders. For this assessment, the 21-item Criterion A subscale was used to assess self-harm frequency and clinical severity. The subscale asks clients to report the number of days they engaged in any one of 21 different types of self-harm (ie, *Cut yourself enough to tear the skin and/or bleed* or *scratched, rubbed, or pinched at your skin to the point of bruising or bleeding*) over the previous 30 days. Washburn et al [44] suggest a cutoff score of 5 total days for any one NSSI type to be considered clinically significant NSSI.

Data Preparation

Before the analysis, 8 new variables were created, equivalence between insurance groups was assessed, and tests of normality of continuous variables were conducted.

Newly Created Variables of Outcome Change and Engagement

Depression

In total, 2 change variables were computed using the continuous PHQ-A scores. The first change variable was created by subtracting the raw PHQ-A intake score from the raw score at discharge, resulting in a continuous change score. To create the second change variable, continuous PHQ-A scores were recoded to classify clients into symptom severity categories based on their scores at intake and discharge, with scores of 0 to 4 indicating 1 (minimal), scores of 5 to 9 indicating 2 (mild), scores of 10 to 14 indicating 3 (moderate), scores of 15 to 19 indicating 4 (moderately severe), and scores of 20 to 27 indicating 5 (severe). Next, a change variable representing whether symptom severity improved between intake and discharge was created by subtracting the discharge severity classification variable from the intake severity classification variable and recoded with 1=*improved* (decreased at least one severity classification), 0=*not improved* (stayed the same or increased in severity classification), and −1=*worsened* (increased 1 severity class). Finally, to assess for changes in depression severity classification, the change variable was recategorized to classify clients as 1=*improved* (dropped 1 severity classification) or 0=*not improved* (stayed the same or worsened in severity class).

Suicide Risk

For this evaluation, 2 change variables were computed to assess outcomes by group. The first was a screening status change variable that categorized clients based on improvement wherein clients were classified as *improved* if their screening status changed from positive to negative across time and classified as *not improved* if their status remained positive or transitioned from negative to positive. First, a composite score was created for intake and discharge ASQ items 1 to 4. Then, 2 screening status variables were created wherein clients were categorized as 1=*positive* if their score was >0 and 0=*negative* if their composite score was equal to 0. Next, a change variable was created by subtracting the discharge screening variable from the intake screening variable. This variable was then recoded to reflect changes in screening status, which classified clients as 1=*improved* if their screening status change variable was equal to 1 and 0=*not improved* if their change score was either 0 (indicating no change) or −1 (indicating a negative prescreen status and positive postscreen status).

To assess for changes in active suicidal ideation, a change variable for the ASQ 3 was computed by subtracting the discharge assessment ASQ 3 score (1=*yes* and 0=*no*) from the intake response score to categorize clients as 1=*improved* (“yes” at intake and “no” at discharge) or 0=*not improved* (“yes” at intake and discharge or “no” at intake and “yes” at discharge).

NSSI Variable

In total, 2 change variables were created to explore differences in change in NSSI. First, scores on each of the 21 types of NSSI were recoded into dichotomous variables where 1=*met the criteria* (score of ≥5 days) and 0=*did not meet the criteria* (score of <5 days). A composite score for the 21 dichotomous items was used to create criteria variables for intake and discharge where 0=*did not meet the criteria* and 1=*met the criteria* (a score of >1 would indicate that the client met the criteria for at least one subtype). To assess change, a difference variable was created by subtracting the discharge criteria variable from the intake criteria variable. Finally, this variable was transformed into a dichotomous variable where 1=*improved* (difference score of 1) or 0=*not improved* (difference score of 0 or −1).

The second change variable created reflects a total change in frequency across subtypes of NSSI. To calculate this variable, 2 sum scores were computed across subtypes of NSSI at intake and discharge. The change variable was computed by subtracting the discharge frequency score from the intake frequency score. The instructions asked clients to reflect on their frequency of self-harm over the past month; therefore, a cap of 30 was used on total scores to improve the interpretability of the total score.

Program Engagement

To explore differences in engagement by group, 2 variables were created that reflected attendance rate and LOS. The attendance rate variable was created by dividing the total number of sessions attended by the client by the total number of sessions

scheduled. The LOS variable was computed by subtracting the clients' discharge date from their admission date and dividing by 7 to calculate the total number of weeks attended by the client.

Missing Data

The electronic survey used to collect client responses does not use a force response mechanism; as such, the sample sizes for each of the resultant analyses may differ slightly. The number of clients who met the eligibility criteria was 495; however, given the authors' decision not to use imputation, the range of sample sizes in the subsequent analyses ranged from 459 to 495 (data coverage of ≥92.7% for each analysis).

Test of Normality

To inform decisions about the use of parametric versus nonparametric tests, the distributions for all continuous variables included in this analysis were evaluated for normality. The PHQ-A, NSSI frequency scores, and LOS were normally distributed; however, both age and attendance rate were positively skewed. Therefore, nonparametric tests were determined to be most appropriate for the analysis of the latter 2 variables.

Equivalence of Insurance Groups at Intake

A series of tests was conducted to assess the equivalence of the 2 insurance groups (public insurance vs private insurance) at intake that could plausibly explain differences at discharge. Independent *t* tests (2-tailed) were conducted using the PHQ-A intake scores and NSSI frequency. Differences at intake in depression symptom severity classification were also tested using the chi-square test of independence.

Data Analysis Strategy

Descriptive statistics were used to describe the sample as a whole by demographic and baseline clinical characteristics. Demographic characteristics included age and self-identified gender and sexual orientation. Clinical characteristics comprised baseline clinical assessment scores on the PHQ-A, ABASI, and ASQ.

Program Engagement

Program engagement was calculated for the sample as a whole. Differences between insurance groups were analyzed using a Mann-Whitney *U* test of median differences for attendance, an independent-sample *t* test (2-tailed) of mean differences for LOS, and chi-square analysis for discharge type.

Program Effectiveness

Before testing group differences in clinical outcomes by insurance type, changes in outcomes between intake and discharge across all clients were tested for each outcome variable. Differences by insurance type were then tested for those outcomes and were found to show significant improvement between time points.

Depression

A repeated-measure 2-tailed *t* test was run on the sample as a whole, and an independent-sample *t* test (2-tailed) was run on the continuous PHQ-A change variable. A McNemar chi-square test was used to assess the number of clients who moved from the moderate to severe depression classification at intake to a lower severity class at discharge. Two chi-square analyses were run on the dichotomous PHQ-A improvement variable: a McNemar chi-square test for the sample as a whole and a Pearson chi-square test to compare the 2 groups.

Suicide Risk and Ideation

McNemar chi-square analyses were run to explore changes in suicide risk and active suicidal ideation across groups in a subsample of youths who screened positive at intake. A Pearson chi-square test of independence was run to compare group differences within the subsample.

NSSI Variable

To test changes in the sample as a whole, a McNemar chi-square test was used between intake and discharge on the screening variable, and a repeated-measure 2-tailed *t* test was run for NSSI frequency. To explore between-group differences, a Pearson chi-square analysis was run on the NSSI improvement variable, and an independent-sample *t* test (2-tailed) was used to test changes in NSSI frequency. The chi-square analyses of between-group differences in NSSI were run on a subsample of clients who met the criteria at intake (205/459, 44.6%), excluding those who did not meet the criteria at intake.

Results

Equivalence of Insurance Groups at Intake

Differences in PHQ-A scores were nonsignificant between the public (mean 13.28, SD 7.58) and private (mean 14.35, SD 7.95) insurance groups ($t_{475}=-1.297$; $P=.20$). Similarly, no significant differences were found in NSSI frequency between the public (mean 12.40, SD 12.40) and private (mean 11.77, SD 11.77) insurance groups ($t_{455}=0.49$; $P=.63$). The results of a chi-square test on significant differences indicated no significant difference in severity classification by insurance type (N=477, $\chi^2_1=6.3$; $P=.18$). The results of a chi-square analysis conducted with suicide risk status also indicated no significant differences between the groups at intake (N=491, $\chi^2_1=0.0$; $P=.98$). Similarly, no significant differences were found on the ASQ 3 (N=490, $\chi^2_1=3.2$; $P=.07$). Finally, differences in NSSI clinical severity status were tested using chi-square analysis. The findings indicated no difference between the groups concerning the proportion of clients who met the criteria versus those who did not (N=459, $\chi^2_1=0.1$; $P=.74$). Given that no significant differences were found at admission on clinical characteristics, the 2 insurance groups were assumed to be equivalent; consequently, these scores were not included as covariates in the main analyses. See Table 1 for additional details.

Table 1. Chi-square clinical differences at intake.

	Insurance group		Chi-square (*df*)	*P* value
	Public	Private		
Depression severity			6.3 (4)[a]	.18
Minimal				
Participants, n (%)	19[b] (15.8)	52[c] (14.6)		
Standardized residual	0.30	−0.20		
Mild				
Participants, n (%)	26[b] (21.7)	55[c] (15.4)		
Standardized residual	1.20	−0.70		
Moderate				
Participants, n (%)	20[b] (16.7)	67[c] (18.8)		
Standardized residual	−0.40	0.20		
Moderately severe				
Participants, n (%)	26[b] (21.7)	61[c] (17.1)		
Standardized residual	0.90	−0.50		
Severe				
Participants, n (%)	29[b] (24.2)	122[c] (34.2)		
Standardized residual	−1.50	0.80		
Suicide risk status			0.001 (1)[d]	.98
Positive				
Participants, n (%)	88[e] (71)	260[f] (70.8)		
Standardized residual	0.00	0.00		
Negative				
Participants, n (%)	36[e] (29)	107[f] (29.2)		
Standardized residual	0.00	0.00		
ASQ 3[g]			3.23 (1)[h]	.07
No				
Participants, n (%)	62[i] (50.4)	219[f] (59.7)		
Standardized residual	−1.00	0.60		
Yes				
Participants, n (%)	61[i] (49.6)	148[f] (40.3)		
Standardized residual	1.20	−0.70		
NSSI[j] criteria status			0.11 (1)[k]	.74
Met criteria				
Participants, n (%)	59[l] (48.4)	157[m] (46.6)		
Standardized residual	0.20	−0.10		
Did not meet criteria				
Participants, n (%)	63[l] (51.6)	180[m] (53.4)		
Standardized residual	−0.20	0.10		

[a]N=477.

[b]N=120.
[c]N=357.
[d]N=491.
[e]N=124.
[f]N=367.
[g]ASQ 3: Ask Suicide-Screening Questions 3.
[h]N=490.
[i]N=123.
[j]NSSI: nonsuicidal self-injury.
[k]N=459.
[l]N=122.
[m]N=337.

Client Demographic Characteristics

There were a total of 1461 clients admitted to the program during the data collection period, of whom 995 (68.1%) met the engagement criteria of attending ≥7 sessions. Of those who did not meet the engagement criteria, 62.9% (293/466) were discharged within the first week of treatment at Charlie Health (≤3 sessions). Of the 995 remaining clients, 500 (50.3%) did not complete a discharge survey, resulting in 495 client cases included in the analysis.

The average age of the clients was 16.7 years, and over three-quarters of the sample were adolescents (385/495, 77.8%). The age range of this sample was 11 to 35 years (wherein clients aged between 11 and 25 years comprised 477/495, 96.4% of the sample; the remaining were between the ages of 26 and 35 years). Over half of the clients identified as women (249/495, 50.3%), almost a quarter (113/495, 22.8%) self-identified as a nonbinary, and 20.1% (99/492) self-identified as transgender. Most clients (321/495, 64.8%) identified as members of the lesbian, gay, bisexual, transgender, and queer community, with 34.9% (173/495) identifying as heterosexual. Approximately half of the clients (248/494, 50.2%) reported an admission to a higher level of care or a visit to an emergency room in the 30 days before IOP admission. The only significant difference in demographic factors between groups was transgender identity such that a significantly greater proportion of clients with public insurance identified as transgender compared with clients with private insurance (183/495, 37% vs 86/495, 17.4%). The ratio of public to private insurance among Charlie Health clients (370/495, 74.7% have private insurance and 125/495, 25.3% have public insurance) mirrors that of the national population aged <65 years (67.7% have private insurance and 20.7% have public insurance) [45].

Program Engagement

Table 2 presents the results of the Mann-Whitney *U* test and independent-sample 2-tailed *t* tests that were run to determine if there were differences in the attendance rate and LOS between insurance types. The median attendance of the sample as a whole was 91% (mean 85.9%, SD 16.48%). The results of the Mann-Whitney *U* test were nonsignificant, indicating no significant difference in attendance rate between clients with public (123/491, 25.1%; median 89%) and private (368/491, 74.9%; median 92%) insurance types (U=24,858.5; z score 1.65; P=.10).

The average LOS for the sample as a whole (N=491) was 10.8 weeks (SD 4.60). The results of the independent-sample *t* test (2-tailed) comparing LOS across insurance groups indicated no statistically significant difference between clients with public (123/491, 25.1%; mean 10.54) and private (368/491, 74.9%; mean 10.88) insurance (t_{489}=−0.72; P=.47).

Finally, significant differences were assessed on discharge reason by insurance type. The results of the chi-square analysis indicated no significant differences between the groups, wherein 67.2% (84/125) of clients with Medicaid insurance completed treatment and 71.2% (259/364) of clients with commercial insurance completed treatment (N=489, χ^2_1=0.7; P=.43).

Table 2. Client length of stay (LOS) and attendance by insurance type.

	Values	Mean difference
LOS , mean (SD)		
Public	10.54 (4.79)	−0.34
Private	10.88 (4.54)	N/A[a]
Attendance, median; mean rank[b]		
Public	89%; 227.90	N/A
Private	86.55%; 252.02	N/A

[a]N/A: not applicable.
[b]P=.10 (2-tailed).

Depression

The repeated-measure 2-tailed *t* test indicated that, across the sample as a whole, clients scored significantly lower on the PHQ-A at discharge compared with at intake (t_{447}=12.51; *P*<.001). The independent-sample *t* test (2-tailed) comparing changes in PHQ-A score by insurance type found no significant difference between the groups (t_{444}=−0.87; *P*=.38), suggesting that clients improved significantly from intake to discharge in depression severity regardless of insurance type. The McNemar chi-square test assessing change in symptom severity among the whole sample between intake and discharge indicated that a significant number (163/192, 84.9%) of those who scored moderate to severe at intake (N=650, χ^2_1=17.6; *P*<.001) scored in a lower category at discharge, 83.2% (99/119) of clients who were classified as "moderately severe" at intake moved down at least one classification at discharge (N=434, χ^2_1=12.3; *P*<.001), and 73.2% (82/112) of clients who were classified as "moderate" at intake dropped at least one classification level by discharge (N=296, χ^2_1=9.7; *P*<.001). Pearson chi-square tests comparing changes in symptom severity between the 2 insurance groups indicated no significant differences (N=446, χ^2_1=0.1; *P*=.71). See Table 3 for more details.

Table 3. Chi-square analysis of difference in depression severity by insurance type (N=446).

	Insurance group		Total	Chi-square (*df*)	*P* value
	Public (n=111)	Private (n=335)			
Improved between intake and discharge				0.1 (1)	.71
Not improved					
Participants, n (%)	47 (42.3)	135 (40.3)	182 (40.8)		
Standardized residual	0.30	−0.10	N/A[a]		
Improved					
Participants, n (%)	64 (57.7)	200 (59.7)	264 (59.2)		
Standardized residual	−0.20	0.10	N/A		

[a]N/A: not applicable.

Suicide Risk

The McNemar chi-square test assessing the number of clients who improved in symptom severity between intake and discharge across the sample as a whole found that, of the clients who screened positive at intake (330/470, 70.2%), significantly fewer (115/330, 34.8%) screened positive at discharge (*P*<.001). The results of the Pearson chi-square analysis indicated no significant differences in the number of clients who improved between intake and discharge by insurance type (N=330, χ^2_1=0.1; *P*=.89). See Table 4 for more details.

The McNemar chi-square test assessing the number of clients who improved in active suicidal ideation symptoms between intake and discharge across the sample as a whole found that, of the clients who screened positive at intake (201/470, 42.8%), significantly fewer (92/201, 45.8%) reported active suicidal ideation at discharge (N=470, χ^2_1=19.2; *P*<.001). The results of the Pearson chi-square analysis indicated no differences in the number of clients who improved between intake and discharge between the 2 insurance types (N=200, χ^2_1=0.6; *P*=.49). In other words, both groups improved similarly on active suicidal ideation from intake to discharge (see Table 4 for additional details).

Table 4. Chi-square analysis of differences in suicide risk by insurance type.

Screening status between intake and discharge	Insurance group		Total	Chi-square (*df*)	*P* value
	Public	Private			
Suicidal ideation				0.06 (1)[a]	.89
Not improved					
Participants, n (%)	28 (33.7)	87 (35.2)	115 (34.8)		
Standardized residual	−0.20	0.10	N/A[b]		
Improved					
Participants, n (%)	55 (66.3)	160 (64.8)	215 (65.2)		
Standardized residual	0.10	−0.10	N/A		
Total, n (%)	83 (100)	247 (100)	330 (100)		
Active suicidal ideation				0.56 (1)[c]	.49
Not improved					
Participants, n (%)	19 (32.8)	39 (27.5)	58 (29)		
Standardized residual	0.50	−0.30	N/A		
Improved					
Participants, n (%)	39 (67.2)	103 (72.5)	142 (71)		
Standardized residual	−0.30	0.20	N/A		
Total, n (%)	58 (100)	142 (100)	200 (100)		

[a]N=330.

[b]N/A: not applicable.

[c]N=200.

NSSI Variable

The repeated-measure 2-tailed *t* test run on all clients (n=430) indicated a significant reduction in NSSI frequency from intake (mean 12.09, SD 12.34) to discharge (mean 6.08, SD 9.38; t_{429}=10.41; *P*<.001) across the 2 groups. The independent-sample *t* test (2-tailed) evaluating differences between insurance groups in changes in frequency (total days) of NSSI was not significant (t_{426}=−0.98; *P*=.33). Thus, clients appear to have improved in NSI frequency across the sample regardless of insurance type.

The McNemar chi-square test comparing changes in symptom severity among those clients who met NSSI criteria at intake (205/428, 47.9%) indicated a significant change such that 58% (119/205) did not meet the criteria at discharge (N=430, χ^2_1=40.7; *P*<.001). The results of the Pearson chi-square analysis indicated no significant difference between the public and private health insurance groups in the changes experienced between intake and discharge in NSSI symptom severity (N=205, χ^2_1=0.9; *P*=.35). See Table 5 for additional details.

Table 5. Chi-square analysis of differences in nonsuicidal self-injury by insurance type (N=205).

	Insurance group		Total	Chi-square (*df*)	*P* value
	Public (n=55)	Private (n=150)			
Improved between intake and discharge				0.9 (1)	.35
Not improved					
Participants, n (%)	26 (47.3)	60 (40)	86 (42)		
Standardized residual	0.60	−0.40	N/A[a]		
Improved					
Participants, n (%)	29 (52.7)	90 (60)	119 (58)		
Standardized residual	−0.50	0.30	N/A		

[a]N/A: not applicable.

Discussion

Principal Findings

The aims of the quality improvement analysis reported in this paper were to (1) assess whether clinical outcomes improved during treatment and (2) assess for differences in engagement or outcomes between adolescents and young adults engaging in remote IOPs on either public or private insurance. The findings of this analysis support the effectiveness of remote IOP treatment in reducing clinical symptoms across all clients, including reduced depression, suicidality, and self-harm. The findings indicate reduced symptoms regardless of insurance type; no differences among youth clients by private or public insurance type emerged across the clinical outcomes tested, evincing similar reductions in depressive symptoms, suicidal ideation, and NSSI. Similarly, youths on public and private insurance were equivalently engaged in treatment, attending treatment for comparable lengths of time and both groups attending nearly all scheduled group sessions (median 91%; mean 85.9%, SD 16.48%) during their treatment stay.

Across the entire sample and regardless of insurance type, youths in remote IOP reported significantly fewer depression symptoms at discharge, with nearly 60% (264/446, 59.2%) of patients evidencing a clinically significant reduction. Of the youths who were actively suicidal at intake, 71% (142/200) no longer reported the same at discharge. Finally, more than half (119/205, 58%) of those who met the criteria for clinical NSSI at intake no longer met the criteria for NSSI at discharge, indicating a significant decline in self-harming behavior. These results provide preliminary support for both the effectiveness of remote IOPs for adolescents and youths with complex mental health needs and the comparable effectiveness of group telehealth among publicly and privately insured patients.

These findings contrast with previous research comparing telehealth IOP outcomes by insurance type that found disparities in outcomes for adults [10] as well as disparities in in-person IOP outcomes for children [32], youths [28], and young adults [35]. In studies addressing youths and young adults, these different findings may be explained by the handling of dropout as these studies appear to have included all clients who initiated treatment [28,35]. It may be that there are higher dropout levels among clients on public insurance and that the resulting effect on outcomes from a smaller dose of treatment is inadequate for symptom improvement. However, a study of a child partial hospitalization program used a nearly identical methodology including only those who stayed for 2 weeks and completed both intake and discharge assessments, finding greater reductions in depression and emotional symptoms for youths on public insurance [32]. As their study focused on children aged 7 to 13 years, it may be that there are differences in treatment outcomes and disparities between children and young adolescents and the adolescents and young adults in our sample. This study also found that families with greater recent stressors had smaller treatment gains even after controlling for insurance type. The authors suggest that families with state-funded insurance may have had "additional stressors that created barriers to treatment use, such as difficulties with transportation and inadequate social supports" [32]. It may be that the telehealth option of these services posed lower stressors on the family, leading to more equitable outcomes.

This quality improvement analysis also found that program attendance among clients who met the inclusion criteria (at least 2 weeks of programming and 7 sessions attended) did not significantly differ in their engagement, both in the total number of weeks and rate of attendance to sessions. This finding also runs counter to recent research that found that, when adolescents were transitioned from in-person to telehealth IOP services during the COVID-19 pandemic, publicly insured and lower-income clients had significantly lower attendance rates [9]. The differences in the findings may be due to the different handling of dropout. In their study, Childs et al [9] measured attendance to all scheduled appointments. It may be that families with public insurance have greater initial barriers and higher early dropout rates. The comparable attendance and outcomes among those who engage suggests that telehealth IOP services can be accessible and effective across insurance types.

Previous studies have identified childcare, transportation, and scheduling challenges as primary barriers to attendance for low-income families [46,47]. Telehealth obviates some of these structural barriers to accessing intensive mental health treatment. Evening and weekend hours can further address barriers to attendance among low-income families [48] and particularly among families with public insurance. Children with public insurance are less likely to have a usual source of care during nighttime or weekend hours and are more likely to experience a delay in receiving care because they could not go when services were open or because of transportation challenges; these differences are significant even after controlling for health, demographic, and socioeconomic differences [49]. Charlie Health has carefully designed its scheduling for accessibility to all clients; services are offered during morning, afternoon, and evening hours to match demanding schedules. The next step in this quality improvement analysis would be to explore the outliers—those adolescents and young adults who had lower than average engagement and clinical improvement—to better understand barriers that might affect this subpopulation of clients who are within the reach of influence of Charlie Health.

Recent systemic changes likely influenced the findings of this analysis in the desired direction. For example, as a result of the pandemic, access to technology and the internet increased in many areas, which, in turn, has facilitated access to telehealth care regardless of income or resources [50]. Ultimately, the analyses conducted in this study suggest that, given access to the same remote high-acuity care, youths and young adults engaged in programming at comparable rates and achieved similar improvements in clinical outcomes regardless of insurance type.

Strengths and Limitations

These findings should be interpreted within the limitations of the available data, the most notable of which is that the inclusion criteria restricted cases to only those of clients who met the engagement threshold. This likely introduced a selection bias wherein exploring differences in early disengagement may have revealed significant differences between insurance groups. In

this project, demographic data were not collected at intake, so it was not possible to assess differences in dropout by demographics. However, the results of this assessment were meant to inform and assess the program from which the client cases came, which is why the survey collection criteria predicated participation in at least 2 weeks of programming to be eligible for discharge surveys. However, future quality improvement efforts should be made to explore potential differences in early engagement that may necessitate program improvement. Future work should also assess a larger range of demographic factors. Notably, rural populations have less access to the internet than urban populations, although the divide is narrowing [51]. Studies are needed to assess the role of geographic location in telehealth IOP engagement.

There are also inherent limitations to solely relying on client self-report data in assessing meaningful clinical change. The data used in the pre-post analyses in this study were collected at 2 distinct points: intake and discharge. Thus, there is no way of knowing if there were specific facilitators or barriers to engagement and treatment effectiveness that emerged throughout the course of treatment. Furthermore, the timing of survey dissemination may have influenced clients' responses in either direction—for instance, it has been noted by clinical staff that clients report a range of intense positive and negative emotions at the time of intake and discharge. Thus, responses about clinical symptoms may be influenced by heightened emotions elicited at these 2 points (minimizing or exaggerating clinical severity). However, for the purposes of this preliminary quality improvement analysis, the available data proved sufficient to explore high-level differences in engagement and outcomes by insurance type. Future analyses would be strengthened by the inclusion of observational data provided by program staff or treatment-involved family members. Additional data points throughout the course of treatment may also provide a more balanced and nuanced narrative of treatment experiences beyond what can be inferred from the pre- and postsurvey data.

A notable strength of the quality improvement analysis conducted in this study was the ability to compare outcomes and engagement between clients using public and private insurance who participated in the same program at the same time. Research has long noted that publicly insured clients have significantly fewer treatment options compared with privately insured clients [52,53]. Thus, comparing outcomes between populations with private and public health insurance precludes the investigation of the moderating influence of variable program quality as publicly insured clients have fewer options and likely fewer quality options. Poorer comparative outcomes that disfavor publicly insured clients may consequently affirm the damaging stereotypes that this population is less capable of clinical improvement [46]. However, these analyses remove the variable influence of program quality and suggest that socioeconomic factors are not deterministic of treatment outcomes, further impressing the importance of providing equitable access to quality treatment regardless of insurance type.

Implications

The finding that youths improved regardless of insurance type has direct implications for practice. Previous research suggests that low socioeconomic status adversely affects the likelihood of youths benefiting from mental health treatment [46]. However, the findings of this assessment demonstrate equitable outcomes among youths of varying socioeconomic status using health insurance as a proxy. A notable strength of the program setting is the variable times offered for groups, which addresses a barrier to services frequently shared by families with lower incomes or on public insurance [48,49]. Furthermore, the provision of services on a web-based platform may remove some of the common barriers to service attendance that caregivers report related to transportation, childcare, and time off work [47]. Mental health intensive outpatient service providers considering expansion to publicly insured clients should consider variable times for service offers and remote alternatives.

Conclusions

Given that these analyses investigated services from a multistate psychiatric care provider for youths and young adults, this assessment is larger in scope than previous investigations of outcomes by insurance type among telehealth intensive psychiatric services. This study contributes to the currently limited evidence base on disparities by insurance type for telehealth intensive psychiatric services. The results suggest that adolescents and young adults on public and private insurance engage in remote IOPs at similar rates, achieving comparable improvements in depressive symptoms, NSI, and suicidal ideation. This suggests that, when given access to the same quality of intensive care in a remote, flexible scheduling format, youths and young adults on either public or private insurance have equal engagement and outcomes. Remote IOPs lead to reduced symptoms for youths with mental health needs across insurance types at a time when such services are needed by millions of adolescents and young adults who do not frequently have access to care because of geographic or financial limitations.

Acknowledgments

This quality improvement initiative was funded by Charlie Health, Inc.

Data Availability

Owing to the quality improvement nature of this study, participants did not agree to their data being shared publicly, so supporting data are not available.

Authors' Contributions

KG and KRB developed the concept of the study and cowrote the first draft. KRB conducted analyses, and MEC verified the findings and made suggestions for alternative statistical tests. JB helped research and write the introduction and literature review. All authors contributed to critical review and editing of the manuscript.

Conflicts of Interest

CF is the founder and chief clinical officer of Charlie Health. KG and KRB are employees of and hold equity in Charlie Health. JB and MEC report consulting fees from Charlie Health. PLS has no disclosures to report.

References

1. U.S. surgeon general issues advisory on youth mental health crisis further exposed by COVID-19 pandemic. U.S. Department of Health and Human Services. 2021 Dec 7. URL: https://www.hhs.gov/about/news/2021/12/07/us-surgeon-general-issues-advisory-on-youth-mental-health-crisis-further-exposed-by-covid-19-pandemic.html [accessed 2022-07-29]
2. Jones SE, Ethier KA, Hertz M, DeGue S, Le VD, Thornton J, et al. Mental health, suicidality, and connectedness among high school students during the COVID-19 pandemic - adolescent behaviors and experiences survey, United States, January-June 2021. MMWR Suppl 2022 Apr 01;71(3):16-21 [FREE Full text] [doi: 10.15585/mmwr.su7103a3] [Medline: 35358165]
3. Clayborne ZM, Varin M, Colman I. Systematic review and meta-analysis: adolescent depression and long-term psychosocial outcomes. J Am Acad Child Adolesc Psychiatry 2019 Jan;58(1):72-79. [doi: 10.1016/j.jaac.2018.07.896] [Medline: 30577941]
4. Groenman AP, Janssen TW, Oosterlaan J. Childhood psychiatric disorders as risk factor for subsequent substance abuse: a meta-analysis. J Am Acad Child Adolesc Psychiatry 2017 Jul;56(7):556-569. [doi: 10.1016/j.jaac.2017.05.004] [Medline: 28647007]
5. Asarnow JR, Porta G, Spirito A, Emslie G, Clarke G, Wagner KD, et al. Suicide attempts and nonsuicidal self-injury in the treatment of resistant depression in adolescents: findings from the TORDIA study. J Am Acad Child Adolesc Psychiatry 2011 Aug;50(8):772-781 [FREE Full text] [doi: 10.1016/j.jaac.2011.04.003] [Medline: 21784297]
6. Facts about Suicide. Centers for Disease Control and Prevention. 2022. URL: https://www.cdc.gov/suicide/facts/index.html [accessed 2022-07-29]
7. McCauley E, Berk MS, Asarnow JR, Adrian M, Cohen J, Korslund K, et al. Efficacy of dialectical behavior therapy for adolescents at high risk for suicide: a randomized clinical trial. JAMA Psychiatry 2018 Aug 01;75(8):777-785 [FREE Full text] [doi: 10.1001/jamapsychiatry.2018.1109] [Medline: 29926087]
8. Cummings JR, Wen H, Druss BG. Improving access to mental health services for youth in the United States. JAMA 2013 Feb 13;309(6):553-554 [FREE Full text] [doi: 10.1001/jama.2013.437] [Medline: 23403677]
9. Childs AW, Bacon SM, Klingensmith K, Li L, Unger A, Wing AM, et al. Showing up is half the battle: the impact of telehealth on psychiatric appointment attendance for hospital-based intensive outpatient services during COVID-19. Telemed J E Health 2021 Aug;27(8):835-842. [doi: 10.1089/tmj.2021.0028] [Medline: 33999738]
10. Gannon JM, Brar JS, Zawacki S, Painter T, O'Toole K, Chengappa KN. From office-based treatment to telehealth: comparing clinical outcomes and patient participation in a psychiatric intensive outpatient program with a large transdiagnostic sample. Telemed J E Health 2022 Aug;28(8):1126-1133. [doi: 10.1089/tmj.2021.0421] [Medline: 34964653]
11. Boswell JF, Kraus DR, Miller SD, Lambert MJ. Implementing routine outcome monitoring in clinical practice: benefits, challenges, and solutions. Psychother Res 2015;25(1):6-19. [doi: 10.1080/10503307.2013.817696] [Medline: 23885809]
12. Lambert MJ, Whipple JL, Kleinstäuber M. Collecting and delivering progress feedback: a meta-analysis of routine outcome monitoring. Psychotherapy (Chic) 2018 Dec;55(4):520-537. [doi: 10.1037/pst0000167] [Medline: 30335463]
13. Major Depression. National Institute of Mental Health. 2022. URL: https://www.nimh.nih.gov/health/statistics/major-depression [accessed 2022-07-29]
14. Esmaeelzadeh S, Moraros J, Thorpe L, Bird Y. Examining the association and directionality between mental health disorders and substance use among adolescents and young adults in the U.S. and Canada-a systematic review and meta-analysis. J Clin Med 2018 Dec 13;7(12):543 [FREE Full text] [doi: 10.3390/jcm7120543] [Medline: 30551577]
15. Maughan B, Collishaw S, Stringaris A. Depression in childhood and adolescence. J Can Acad Child Adolesc Psychiatry 2013 Feb;22(1):35-40 [FREE Full text] [Medline: 23390431]
16. Youth Risk Behavior Survey: Data Summary and Trends Report, 2009-2019. Centers for Disease Control and Prevention. 2019. URL: https://www.cdc.gov/healthyyouth/data/yrbs/pdf/YRBSDataSummaryTrendsReport2019-508.pdf [accessed 2022-07-29]
17. Lee JW, Kim BJ, Lee CS, Cha B, Lee SJ, Lee D, et al. Association between suicide and drinking habits in adolescents. Soa Chongsonyon Chongsin Uihak 2021 Oct 01;32(4):161-169 [FREE Full text] [doi: 10.5765/jkacap.210024] [Medline: 34671189]

18. Sampasa-Kanyinga H, Dupuis LC, Ray R. Prevalence and correlates of suicidal ideation and attempts among children and adolescents. Int J Adolesc Med Health 2017 Apr 01;29(2):pp. 20150053. [doi: 10.1515/ijamh-2015-0053] [Medline: 26556839]
19. Berona J, Horwitz AG, Czyz EK, King CA. Psychopathology profiles of acutely suicidal adolescents: associations with post-discharge suicide attempts and rehospitalization. J Affect Disord 2017 Feb;209:97-104 [FREE Full text] [doi: 10.1016/j.jad.2016.10.036] [Medline: 27894037]
20. Gonzalez K, Patel F, Cutchins LA, Kodish I, Uspal NG. Advocacy to address emergent pediatric mental health care. Clin Pediatr Emerg Med 2020 Jun;21(2):100778. [doi: 10.1016/j.cpem.2020.100778]
21. Ibeziako P, Kaufman K, Scheer KN, Sideridis G. Pediatric mental health presentations and boarding: first year of the COVID-19 pandemic. Hosp Pediatr 2022 Sep 01;12(9):751-760. [doi: 10.1542/hpeds.2022-006555] [Medline: 35578918]
22. Cook MN, Crisostomo PS, Simpson TS, Williams JD, Wamboldt MZ. Effectiveness of an intensive outpatient program for disruptive children: initial findings. Community Ment Health J 2014 Feb;50(2):164-171. [doi: 10.1007/s10597-012-9588-9] [Medline: 23299227]
23. Öst LG, Ollendick TH. Brief, intensive and concentrated cognitive behavioral treatments for anxiety disorders in children: a systematic review and meta-analysis. Behav Res Ther 2017 Oct;97:134-145. [doi: 10.1016/j.brat.2017.07.008] [Medline: 28772195]
24. Sperling J, Boger K, Potter M. The impact of intensive treatment for pediatric anxiety and obsessive-compulsive disorder on daily functioning. Clin Child Psychol Psychiatry 2020 Jan;25(1):133-140. [doi: 10.1177/1359104519871338] [Medline: 31464534]
25. Delaney RK, Sisco-Taylor B, Fagerlin A, Weir P, Ozanne EM. A systematic review of intensive outpatient care programs for high-need, high-cost patients. Transl Behav Med 2020 Oct 12;10(5):1187-1199. [doi: 10.1093/tbm/ibaa017] [Medline: 33044534]
26. Gregory ST, Kay B, Riemann BC, Goodman WK, Storch EA. Cost-effectiveness of treatment alternatives for treatment-refractory pediatric obsessive-compulsive disorder. J Anxiety Disord 2020 Jan;69:102151. [doi: 10.1016/j.janxdis.2019.102151] [Medline: 31864218]
27. Costello EJ, He JP, Sampson NA, Kessler RC, Merikangas KR. Services for adolescents with psychiatric disorders: 12-month data from the National Comorbidity Survey-Adolescent. Psychiatr Serv 2014 Mar 01;65(3):359-366 [FREE Full text] [doi: 10.1176/appi.ps.201100518] [Medline: 24233052]
28. James S, Freeman KR, Mayo D, Riggs ML, Morgan JP, Schaepper MA, et al. Does insurance matter? Implementing dialectical behavior therapy with two groups of youth engaged in deliberate self-harm. Adm Policy Ment Health 2015 Jul;42(4):449-461. [doi: 10.1007/s10488-014-0588-7] [Medline: 25199812]
29. Slaunwhite AK. The role of gender and income in predicting barriers to mental health care in Canada. Community Ment Health J 2015 Jul;51(5):621-627. [doi: 10.1007/s10597-014-9814-8] [Medline: 25563485]
30. Reay RE, Looi JC, Keightley P. Telehealth mental health services during COVID-19: summary of evidence and clinical practice. Australas Psychiatry 2020 Oct;28(5):514-516 [FREE Full text] [doi: 10.1177/1039856220943032] [Medline: 32722963]
31. Venturo-Conerly KE, Fitzpatrick OM, Horn RL, Ugueto AM, Weisz JR. Effectiveness of youth psychotherapy delivered remotely: a meta-analysis. Am Psychol 2022 Jan;77(1):71-84. [doi: 10.1037/amp0000816] [Medline: 34807635]
32. Yaptangco M, Preddy TM, Musella KE, Parade SH, Umaschi S, Walters A. Family stress moderates emotional and behavioral symptoms in a child partial hospital setting. J Child Fam Stud 2019 May 22;28(10):2762-2771. [doi: 10.1007/s10826-019-01456-y]
33. Philippe TJ, Sikder N, Jackson A, Koblanski ME, Liow E, Pilarinos A, et al. Digital health interventions for delivery of mental health care: systematic and comprehensive meta-review. JMIR Ment Health 2022 May 12;9(5):e35159 [FREE Full text] [doi: 10.2196/35159] [Medline: 35551058]
34. Sen BP, Sharma P, Brisendine A, Blackburn J, Morrisey M, Menachemi N, et al. Disparities in telehealth utilization in a population of publicly insured children during the COVID-19 pandemic. Popul Health Manag 2022 Apr;25(2):178-185. [doi: 10.1089/pop.2021.0343] [Medline: 35442789]
35. Slesinger NC, Hayes NA, Washburn JJ. Understanding predictors of change in a day treatment setting for non-suicidal self-injury. Psychol Psychother 2021 Apr;94 Suppl 2:517-535. [doi: 10.1111/papt.12295] [Medline: 32662182]
36. Garrett A, Cohen JA, Zack S, Carrion V, Jo B, Blader J, et al. Longitudinal changes in brain function associated with symptom improvement in youth with PTSD. J Psychiatr Res 2019 Jul;114:161-169 [FREE Full text] [doi: 10.1016/j.jpsychires.2019.04.021] [Medline: 31082658]
37. Nahman-Averbuch H, Schneider VJ, Chamberlin LA, Kroon Van Diest AM, Peugh JL, Lee GR, et al. Alterations in brain function after cognitive behavioral therapy for migraine in children and adolescents. Headache 2020 Jun;60(6):1165-1182. [doi: 10.1111/head.13814] [Medline: 32323877]
38. Roberts H, Jacobs RH, Bessette KL, Crowell SE, Westlund-Schreiner M, Thomas L, et al. Mechanisms of rumination change in adolescent depression (RuMeChange): study protocol for a randomised controlled trial of rumination-focused cognitive behavioural therapy to reduce ruminative habit and risk of depressive relapse in high-ruminating adolescents. BMC Psychiatry 2021 Apr 23;21(1):206 [FREE Full text] [doi: 10.1186/s12888-021-03193-3] [Medline: 33892684]

39. Tang YY, Lu Q, Fan M, Yang Y, Posner MI. Mechanisms of white matter changes induced by meditation. Proc Natl Acad Sci U S A 2012 Jun 26;109(26):10570-10574 [FREE Full text] [doi: 10.1073/pnas.1207817109] [Medline: 22689998]
40. 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] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
41. Gilbody S, Richards D, Brealey S, Hewitt C. Screening for depression in medical settings with the Patient Health Questionnaire (PHQ): a diagnostic meta-analysis. J Gen Intern Med 2007 Nov;22(11):1596-1602 [FREE Full text] [doi: 10.1007/s11606-007-0333-y] [Medline: 17874169]
42. Johnson JG, Harris ES, Spitzer RL, Williams JB. The patient health questionnaire for adolescents: validation of an instrument for the assessment of mental disorders among adolescent primary care patients. J Adolesc Health 2002 Mar;30(3):196-204. [doi: 10.1016/s1054-139x(01)00333-0] [Medline: 11869927]
43. Horowitz LM, Bridge JA, Teach SJ, Ballard E, Klima J, Rosenstein DL, et al. Ask Suicide-Screening Questions (ASQ): a brief instrument for the pediatric emergency department. Arch Pediatr Adolesc Med 2012 Dec;166(12):1170-1176 [FREE Full text] [doi: 10.1001/archpediatrics.2012.1276] [Medline: 23027429]
44. Washburn JJ, Potthoff LM, Juzwin KR, Styer DM. Assessing DSM-5 nonsuicidal self-injury disorder in a clinical sample. Psychol Assess 2015 Mar;27(1):31-41. [doi: 10.1037/pas0000021] [Medline: 25265415]
45. Cha AE, Cohen RA. Demographic variation in health insurance coverage:United States, 2020. Natl Health Stat Report 2022 Feb(169):1-15. [Medline: 35166656]
46. Eruchalu CN, Pichardo MS, Bharadwaj M, Rodriguez CB, Rodriguez JA, Bergmark RW, et al. The expanding digital divide: digital health access inequities during the COVID-19 pandemic in New York city. J Urban Health 2021 Apr;98(2):183-186 [FREE Full text] [doi: 10.1007/s11524-020-00508-9] [Medline: 33471281]
47. Wolf ER, O'Neil J, Pecsok J, Etz RS, Opel DJ, Wasserman R, et al. Caregiver and clinician perspectives on missed well-child visits. Ann Fam Med 2020 Jan;18(1):30-34 [FREE Full text] [doi: 10.1370/afm.2466] [Medline: 31937530]
48. Ofonedu ME, Belcher HM, Budhathoki C, Gross DA. Understanding barriers to initial treatment engagement among underserved families seeking mental health services. J Child Fam Stud 2017 Mar;26(3):863-876 [FREE Full text] [doi: 10.1007/s10826-016-0603-6] [Medline: 28584498]
49. Kenney GM, Coyer C. National Findings on Access to Health Care and Service Use for Children Enrolled in Medicaid or CHIP - MACPAC Contractor Report No 1. Medicaid and CHIP Payment and Access Commission (MACPAC). 2012 Mar. URL: https://www.macpac.gov/wp-content/uploads/2012/03/National-Findings-on-Access-to-Health-Care-and-Service-Use-for-Children-Enrolled-in-Medicaid-or-CHIP.pdf [accessed 2022-09-14]
50. Castro-Ramirez F, Al-Suwaidi M, Garcia P, Rankin O, Ricard JR, Nock MK. Racism and poverty are barriers to the treatment of youth mental health concerns. J Clin Child Adolesc Psychol 2021;50(4):534-546. [doi: 10.1080/15374416.2021.1941058] [Medline: 34339320]
51. Vogels EA. Some digital divides persist between rural, urban and suburban America. Pew Research Center. 2021 Aug 19. URL: https://www.pewresearch.org/fact-tank/2021/08/19/some-digital-divides-persist-between-rural-urban-and-suburban-america/ [accessed 2022-09-14]
52. Holgash K, Heberlein M. Physician acceptance of new Medicaid patients: what matters and what doesn't. Health Affairs Forefront. 2019 Apr 10. URL: https://www.healthaffairs.org/do/10.1377/forefront.20190401.678690/full/ [accessed 2022-07-29]
53. Cama S, Malowney M, Smith AJ, Spottswood M, Cheng E, Ostrowsky L, et al. Availability of outpatient mental health care by pediatricians and child psychiatrists in five U.S. cities. Int J Health Serv 2017 Oct;47(4):621-635. [doi: 10.1177/0020731417707492] [Medline: 28474997]

Abbreviations

ABASI: Alexian Brothers Assessment of Self-Injury
ASQ: Ask Suicide-Screening Questions
IOP: intensive outpatient program
LOS: length of stay
NSSI: nonsuicidal self-injury
PHQ-A: Patient Health Questionnaire modified for adolescents

Edited by A Mavragani; submitted 05.08.22; peer-reviewed by J Blackburn, B Holtz, A AL-Asadi; comments to author 12.09.22; revised version received 21.09.22; accepted 14.10.22; published 10.11.22.

Please cite as:
Gliske K, Berry KR, Ballard J, Evans-Chase M, Solomon PL, Fenkel C
Mental Health Outcomes for Youths With Public Versus Private Health Insurance Attending a Telehealth Intensive Outpatient Program: Quality Improvement Analysis
JMIR Form Res 2022;6(11):e41721
URL: https://formative.jmir.org/2022/11/e41721
doi:10.2196/41721
PMID:36355428

Original Paper

Recruitment and Retention in Remote Research: Learnings From a Large, Decentralized Real-world Study

Sophia Xueying Li[1], MHI; Ramzi Halabi[1], PhD; Rahavi Selvarajan[1], BEng; Molly Woerner[2], BA; Isabell Griffith Fillipo[2], BA; Sreya Banerjee[3], PhD; Brittany Mosser[2], MSW; Felipe Jain[3], MD; Patricia Areán[2], PhD; Abhishek Pratap[1,4,5,6,7], PhD

[1]Krembil Centre for Neuroinformatics, Centre for Addiction and Mental Health, Toronto, ON, Canada

[2]Department of Psychiatry, University of Washington, Seattle, WA, United States

[3]Depression Clinical and Research Program, Department of Psychiatry, Massachusetts General Hospital, Harvard Medical School, Boston, MA, United States

[4]Department of Psychiatry, University of Toronto, Toronto, ON, Canada

[5]Vector Institute for Artificial Intelligence, Toronto, ON, Canada

[6]Kings College London, London, United Kingdom

[7]Department of Biomedical Informatics and Medical Education, University of Washington, Seattle, WA, United States

Corresponding Author:
Abhishek Pratap, PhD
Krembil Centre for Neuroinformatics, Centre for Addiction and Mental Health
250 College Street, 12th floor
Toronto, ON, M5T 1R8
Canada
Phone: 1 416 535 8501
Email: Abhishek.Pratap@camh.ca

Abstract

Background: Smartphones are increasingly used in health research. They provide a continuous connection between participants and researchers to monitor long-term health trajectories of large populations at a fraction of the cost of traditional research studies. However, despite the potential of using smartphones in remote research, there is an urgent need to develop effective strategies to reach, recruit, and retain the target populations in a representative and equitable manner.

Objective: We aimed to investigate the impact of combining different recruitment and incentive distribution approaches used in remote research on cohort characteristics and long-term retention. The real-world factors significantly impacting active and passive data collection were also evaluated.

Methods: We conducted a secondary data analysis of participant recruitment and retention using data from a large remote observation study aimed at understanding real-world factors linked to cold, influenza, and the impact of traumatic brain injury on daily functioning. We conducted recruitment in 2 phases between March 15, 2020, and January 4, 2022. Over 10,000 smartphone owners in the United States were recruited to provide 12 weeks of daily surveys and smartphone-based passive-sensing data. Using multivariate statistics, we investigated the potential impact of different recruitment and incentive distribution approaches on cohort characteristics. Survival analysis was used to assess the effects of sociodemographic characteristics on participant retention across the 2 recruitment phases. Associations between passive data-sharing patterns and demographic characteristics of the cohort were evaluated using logistic regression.

Results: We analyzed over 330,000 days of engagement data collected from 10,000 participants. Our key findings are as follows: first, the overall characteristics of participants recruited using digital advertisements on social media and news media differed significantly from those of participants recruited using crowdsourcing platforms (Prolific and Amazon Mechanical Turk; $P<.001$). Second, participant retention in the study varied significantly across study phases, recruitment sources, and socioeconomic and demographic factors ($P<.001$). Third, notable differences in passive data collection were associated with device type (Android vs iOS) and participants' sociodemographic characteristics. Black or African American participants were significantly less likely to share passive sensor data streams than non-Hispanic White participants (odds ratio 0.44-0.49, 95% CI 0.35-0.61; $P<.001$). Fourth, participants were more likely to adhere to baseline surveys if the surveys were administered immediately after enrollment.

Fifth, technical glitches could significantly impact real-world data collection in remote settings, which can severely impact generation of reliable evidence.

Conclusions: Our findings highlight several factors, such as recruitment platforms, incentive distribution frequency, the timing of baseline surveys, device heterogeneity, and technical glitches in data collection infrastructure, that could impact remote long-term data collection. Combined together, these empirical findings could help inform best practices for monitoring anomalies during real-world data collection and for recruiting and retaining target populations in a representative and equitable manner.

(JMIR Form Res 2022;6(11):e40765) doi:10.2196/40765

KEYWORDS

participant recruitment; participant retention; decentralized studies; active and passive data collection; retention; adherence; compliance; engagement; smartphone; mobile health; mHealth; sensor data; clinical research; data sharing; recruitment; mobile phone

Introduction

Background

Smartphones offer an unprecedented anytime-anywhere medium for researchers to engage with and assess health-related behaviors in large populations in real-world settings [1,2]. As of 2020, the rate of smartphone ownership in the United States has reached over 80% [3]. The large-scale, high-frequency daily use of such devices coupled with increasingly multimodal onboard sensing capabilities offers an effective approach for conducting large-scale health research [4,5]. The adoption of digital health tools to develop and deploy digitally augmented trials has been rising steadily since the first fully remote decentralized trial in 2011 [6-8]. Recent studies have shown the benefits of remote monitoring using smartphones for assessing real-world behavior [9,10], for managing chronic pain [11], cancer care [12], diabetes [13], Parkinson symptom severity [14], and cardiovascular health [15] and for the delivery of remote interventions [16]. The COVID-19 pandemic has further accelerated this growth, enabling over 220 digitally augmented trials in 2021 alone [17,18].

Using smartphones for health research can also help achieve operational efficiency by relying less on traditional research facilities or intermediaries for data collection, which require in-person contact between the study participants and the research team [6,19,20]. Researchers can communicate asynchronously and synchronously with participants and assess their health by actively and passively collecting individualized real-world data [4,21,22]. Active data are defined as data generated through effortful participation (eg, completing a survey). In contrast, passive data are collected without direct input from participants (eg, the number of daily steps estimated through onboard sensors) [23]. Such scalable remote observational models [6,20] could help investigators to understand people's day-to-day experiences of living with a health condition [4] and the relationship between individualized real-world behavior and health outcomes [22].

Challenges in Remote Participant Recruitment and Retention

However, despite the promise of decentralized health research, several challenges related to the representation and inclusiveness of recruitment and the retention of target populations have surfaced [21,24,25], resulting in sparse, unbalanced, and nonrepresentative real-world data collection [21]. Typically, decentralized studies recruit from various web-based sources such as social media (Facebook [26] and Reddit [27]), crowdsourced platforms (Prolific [28]; Amazon Mechanical Turk, MTurk [29]; Centiment [30]; and CloudResearch [31]), and partnerships with patient registries or advocacy groups [32,33]. Although these recruitment channels have shown the potential to reach and recruit large populations remotely [34-36], the long-term and uniform retention of remote participants has been challenging. Recent findings show that retention rates vary from 1% to 50% [24], with monetary incentives being able to significantly improve long-term retention [10]. With large-scale open recruitment approaches, including the use of financial incentives, the risk of enrolling gamers or malicious actors increases [37].

With large studies using multiple web-based sources to reach and recruit participants remotely, there is a need to assess the impact that such strategies have on the characteristics of the enrolled cohorts and their retention in the studies. In addition, further research is needed to understand how variations in study participation incentives (eg, time and frequency of payments) and differences between Android and iOS operating systems [38] affect long-term data collection in decentralized studies.

Objectives

To investigate some of these challenges in collecting health data through smartphones in real-world settings, we examined the recruitment, retention, and passive data-sharing patterns of more than 10,000 participants in a large, decentralized research study. Specifically, we evaluated the following three key questions: (1) Does combining different recruitment and incentive distribution approaches lead to a heterogeneous cohort with varying characteristics? (2) Can the participant retention and uniformity of data collection in remote studies be affected by cohort heterogeneity? (3) What are the factors that can affect passive data collection in real-world settings?

Methods

Ethics Approval

This study was approved and monitored by the Institutional Review Board of the University of Washington (STUDY00004997) and the Department of Defense Human

Research Protection Office; the approval for the study was granted on February 11, 2020.

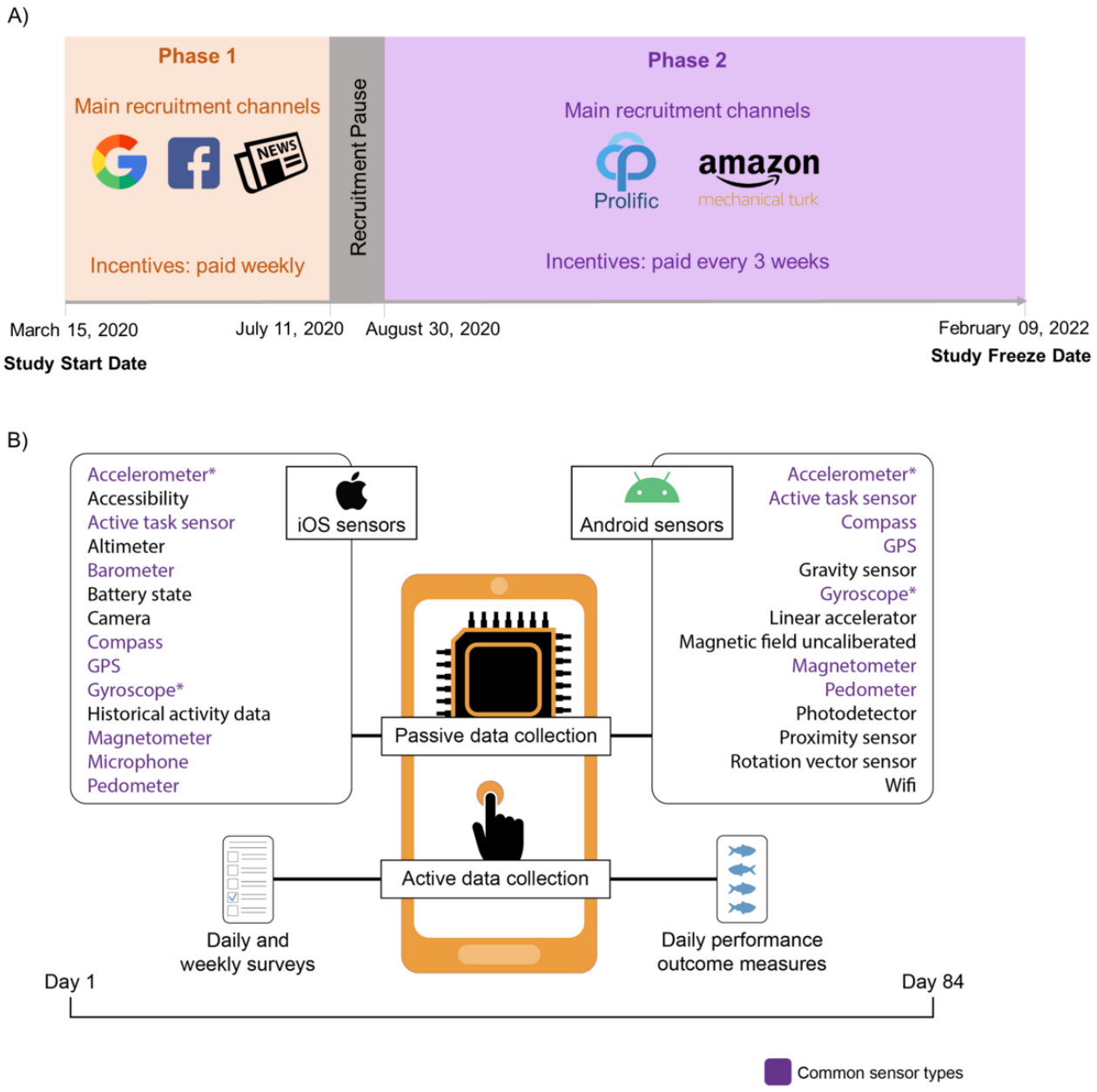
Study Overview

The participants in the Warfighter Analytics Using Smartphones for Health (WASH) study were volunteers who lived in the United States and agreed to engage in a 12-week smartphone-based study. The primary goal of the study was to understand the real-world factors that could help with the early prediction of cold, influenza, and the impact of traumatic brain injury on daily functioning. The eligible participants were individuals aged ≥19 years, English speakers, residents of the United States, owners, and primary users of iPhone or Android smartphones with internet access. The potential participants were required to complete an eligibility screener before consenting, and those who did not meet the inclusion criteria were not permitted to complete subsequent procedures.

Recruitment

Participant recruitment started on March 15, 2020, with rolling enrollment until January 4, 2022. The participants for the study were recruited in 2 phases, using different recruitment and incentive distribution approaches (Figure 1). Participants could receive up to US $90 for completing the baseline survey and 12 weeks of follow-up surveys. The final participation incentive was determined on the basis of the number of complete surveys. Participants were not informed about the financial breakdown during the consent process; however, additional details regarding when they would receive compensation and how much compensation they would receive were provided upon request.

Figure 1. (A) Schematic representation of different study recruitment and participation incentive distribution approaches during phases 1 and 2. Participants recruited during phase 1 were paid weekly (12 times) starting their first day in the study. Anyone who had participated up to October 3, 2020 (who was recruited close to the recruitment pause date) still received weekly payments. Participants recruited during phase 2 were paid every 3 weeks (a total of 4 times) starting their first day in the study. (B) Details of smartphone-based active and passive data collected through the study app during the study observation period (84 days). *Indicates that sharing of accelerometer and gyroscope was made mandatory on August 28, 2020.



Participation Incentives

Phase 1 (March 15, 2020, to July 11, 2020)

Participants were primarily recruited by placing advertisements on social media platforms that directed potential participants to a study recruitment website. Press releases in local news outlets also served as a recruitment source [39,40]. Participants recruited during this phase were paid weekly on the basis of the days a participant completed all daily surveys. The amount received per day increased throughout the 12 weeks (eg, approximately US $4 in weeks 2 to 4, approximately US $6 in weeks 5 to 8, and US $7 in weeks 9 to 11, with more significant payments made for weeks 1 and 12 because of higher incentives rewarding those who completed the baseline survey in week 1 and the exit survey in week 12). However, a significant increase in study enrollment in June 2020, which seemed to be inconsistent with planned recruitment, led the study team to pause enrollment on July 11, 2020. The analysis of participant activity during this period indicated that some malicious actors were engaged in the study. Further details on the assessment of malicious actors can be found in the study by Bracken et al [37].

Phase 2 (August 30, 2020, to the Data Freeze Date, February 9, 2022)

Recruitment resumed on August 30, 2020, after implementing additional strategies to stop fraudulent attempts to join the study,

such as disallowing the autofilling of surveys in the study app, changing recruitment sources, and changing incentive payment frequency [37]. Participants were recruited from 2 web-based recruitment platforms, Prolific and MTurk, from January 4 to December 28, 2021, and from May 15 to December 21, 2021. Prolific is a web-based research platform that includes several safeguards for preserving data quality [41-44]; minimizes gamers or malicious actors; and has been shown to be reliable, efficient, and affordable for remote data collection for behavioral research [45]. Similar to Prolific, MTurk is another web-based crowdsourcing platform regularly used in health research to recruit study participants to complete tasks such as data processing, problem-solving, and surveys [46]. In phase 2, participants were paid every 3 weeks, with the first participation incentive payment taking up to 5 weeks. The change in the payment schedule was implemented for allowing sufficient time to execute procedures intended to identify malicious actors.

Active Surveys

Assessments were divided into 1 longer baseline survey and brief daily assessments. The baseline survey assessing participants' health history, mood, physical activity, and phone use was administered 24 hours after consent was obtained in phase 1 of the study. However, in phase 2, the baseline health survey was administered immediately after consent was obtained. In both phases, the participants were administered the same scheduled health-related surveys twice daily for 12 weeks. The survey asked participants about their mood, physical activity, and phone use.

Sensor-Based Data Collection

Sensor-based data were collected actively and passively from participants through the study app. Participants completed performance outcome measures [47] such as standing and walking tests and sharing voice recordings. The participants were also asked to allow the study app to collect passive data from their smartphones. Passive data included, but were not limited to, device movement and orientation; actual and relative location; the device's status (eg, active use or connected to a data network); and local environmental information such as ambient light, temperature, and humidity. Participants had the option to not share the passive data and remain in the study. However, all participants enrolled in the study on or after August 28, 2020 (before the start of phase 2), were required to allow the study app to passively collect the accelerometer and gyroscope sensor data from their smartphones.

Data Access

Overview

All the data collected from the participants were deidentified. The data collected through the app were encrypted on the phone and stored on secure servers, separate from any identifiable information. Raw data, such as image, proximity, voice, and actual location data of participants, were stored separately from all other sensor data and were not shared with the research team. For this study, data from the enrolled participants between the study launch date (March 15, 2020) and the data freeze date (February 9, 2022) were used for analysis.

Data Cleaning

Before analysis, data from 6788 suspected malicious actors were removed based on the rules for flagging such actors that were defined in the study by Bracken et al [37]. Test data collected before the study launch date on March 15, 2020, were removed. If a survey was submitted more than once, we used the most recent submission to assess the participant's compliance in the study. If participant responses had values outside the expected range of valid values, they were marked as invalid data.

Data Harmonization

To investigate participant retention in the study, we classified the data collected by the study app into two broad categories: (1) survey data, representing any active survey data shared by participants through the study app, and (2) sensor data, representing passive continuous sensor data gathered by the study app without active input from participants as well as active sensor data collected during a performance outcome assessment (eg, walking test data collected from accelerometers and gyroscopes).

Statistical Analysis

Overview

Statistical analyses were performed using data from 10,768 participants after excluding 6788 malicious actors from the data set (6788/17,556, 38.66%). Descriptive analyses of recruitment and cohort characteristics for categorical variables were based on frequencies and percentages. Levels of categorical baseline variables that contained <5% of the cohort were omitted or combined with other levels that contained <5% of the cohort to reduce data sparsity in the analysis. We used median values with the 25th and 75th percentile (IQR) for summarizing continuous variables that were not normally distributed. The differences in cohort characteristics were compared using bivariate analysis methods. The chi-square test was used for testing statistically significant differences between categorical variables; the Fisher exact test was used when table cell counts were <5, and the Mann-Whitney *U* test was used for continuous variables. We used the logistic regression model to assess any statistically significant association between patterns of passive data sharing and participants' sociodemographic characteristics and technical variables. These included race, ethnicity, age, sex, education level, income level, device type, and recruitment phase. Specifically, we compared 3 data-sharing patterns of participants sharing at least 25% (2/8), 50% (4/8), or 75% (6/8) of the 8 common passive data streams between Android and iOS devices. The 95% CIs and *P* values were computed using a Wald *Z* distribution approximation.

We adjusted *P* values by using false discovery rate correction to correct for multiple comparisons across different sensor types. The analyses were conducted using R (version 4.1.1). Statistical significance was assumed when the false discovery rate–corrected *P* value was <.05.

Retention Analysis

To examine overall retention in the study, we used the univariate Kaplan-Meier survival curves [48], which were tested for

statistically significant differences using the nonparametric log-rank test [49]. A participant's last day in the study was determined by the last day of their data sharing. To assess the difference in retention between active and passive data sharing, we also computed study retention for active and passive data streams separately. We used right-censored data for the Kaplan-Meier estimator, given that participants could have continued to use the study app beyond the end of the study period (84 days).

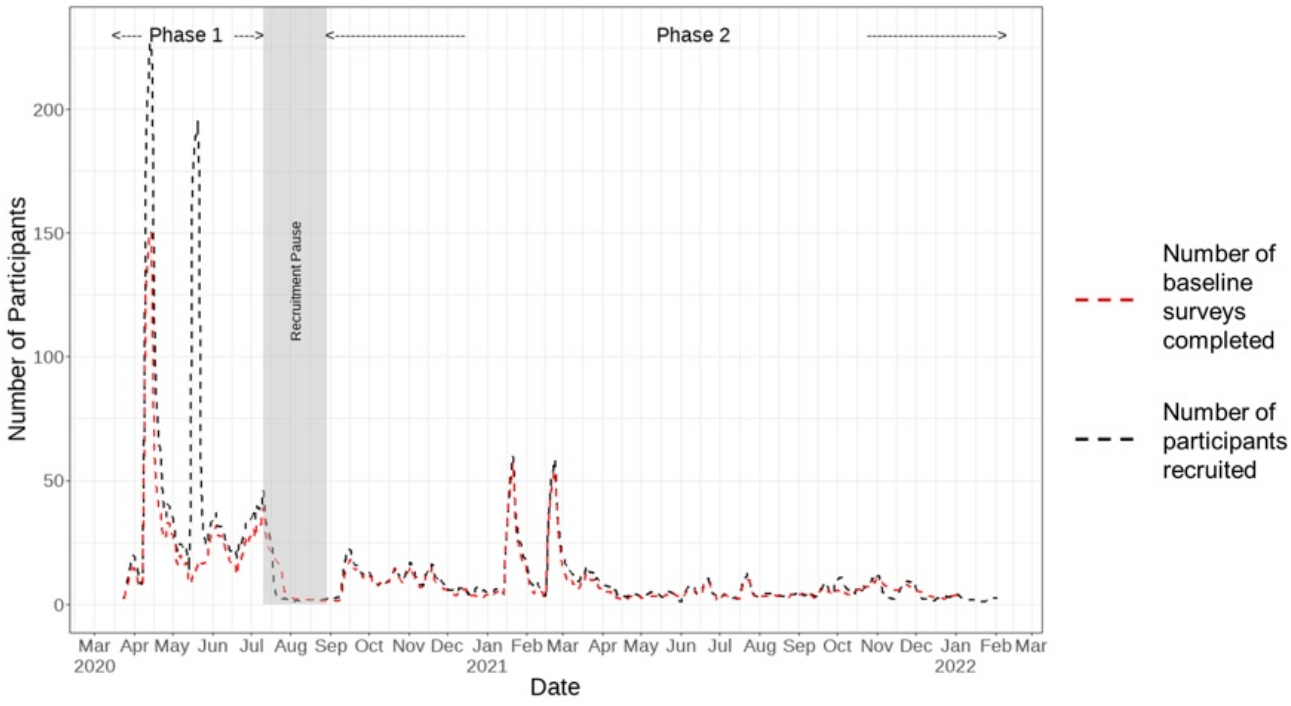
To assess the joint effect of multiple variables of interest, including sociodemographics, on participants' retention in the study, we initially used a multivariate Cox proportional hazards (CoxPH) model [50]. However, one of the key assumptions for CoxPH models (the effect of covariates should not change over time) tested using the Schoenfeld individual test was not met [51]. Multimedia Appendix 1 presents test statistics showing that the CoxPH model assumption is not being met. With the underlying retention data not supporting the CoxPH model assumption, we used a nonparametric log-rank test [52] to assess the statistically significant impact of individual variables on retention within each phase. We cross-compared the median retention for each level of a variable of interest across the 2 study phases.

Results

Recruitment

As of the data freeze date (February 9, 2022), the study recruited 10,768 participants. Most participants (6494/10,768, 60.3%) were recruited during phase 1, and the remaining (4274/10,768, 39.69%) were recruited during phase 2 (see the Methods section). A significant proportion of participants, most notably in phase 1, did not complete the baseline survey (phase 1: 3135/6494, 48.27%, vs phase 2: 918/4274, 21.47%). Figure 2 compares the recruitment rate of the study with the baseline survey submission rate over time. The number of baseline surveys completed generally was in line with the number of participants recruited during the study period. Recruitment peaked in mid-April and mid-May 2020 for phase 1 and in mid-January and early March 2021 for phase 2. However, during phase 1, between May and July 2020, the number of baseline surveys completed was significantly lower than the number of recruited participants, which explains the large proportion of missing baseline data in phase 1. We further assessed the effect of missing baseline surveys on participant engagement in the study (see the Retention Analysis section). Additional statistics on missingness and invalid data entries in the baseline surveys are summarized in Multimedia Appendix 2.

Figure 2. Comparison of the 7-day moving average between the number of participants recruited (black) and the number of baseline surveys completed (red) during the study period. Gray shaded area shows the study recruitment pause phase from July 11, 2020, to August 29, 2020.



Cohort Characteristics

Most of the participants who completed the baseline sociodemographic survey were female (3817/6574, 58.06%). The median age was 30 (IQR 24-40) years, with a larger proportion of participants being aged 19 to 29 years (2949/6267, 47.05%). The non-Hispanic White population was the largest (3938/6677, 58.97%), followed by the Asian (931/6677, 13.94%) and Hispanic or Latino (783/6677, 11.72%) populations. Most participants were iOS users (5883/10,583, 55.58%). Table 1 summarizes the sociodemographic characteristics of the overall cohort.

The population recruited in phase 2 had a higher proportion of younger adults (aged 19 to 29 years; 1685/3194, 52.75%) and a lower proportion of older adults (aged ≥60 years; 94/3194, 2.94%) than that recruited in phase 1 (P<.001; Table 1). A higher proportion of Black or African American participants were recruited in phase 2 (phase 1: 267/3342, 7.98%; phase 2:

456/3339, 13.65%; *P*<.001). Notably, a larger proportion of participants (1942/3308, 58.71%) with lower levels of annual income (≤US $49,999) were recruited in phase 2 than in phase 1 (1062/2483, 42.77%; *P*<.001). The proportion of Android versus iOS users also varied across the recruitment phases. A significantly higher proportion of iOS users (*P*<.001) were recruited in phase 1 (3958/5883, 67.27%) than in phase 2 (1925/5883, 32.72%). Multimedia Appendix 3 further compares the sociodemographic characteristics of Android and iOS users across the 2 recruitment phases.

Table 1. Characteristics of the overall study cohort (N=10,768) along with comparison of participants recruited between phase 1 (n=6494) and phase 2 (n=4274).

	Overall cohort	Participants recruited during phase 1	Participants recruited during phase 2	Test statistics, chi-square (*df*)	*P* value (phase 1 vs phase 2)
Age (years), n (%)	6267 (58.21)	3073 (47.32)	3194 (74.73)	235.29 (4)	<.001
19-29	2949 (47.05)	1264 (41.14)	1685 (52.8)	__[a]	—
30-39	1637 (26.12)	739 (24)	898 (28.11)	—	—
40-49	804 (12.82)	459 (14.9)	345 (10.8)	—	—
50-59	490 (7.81)	318 (10.3)	172 (5.4)	—	—
≥60	387 (6.37)	293 (9.51)	94 (2.96)	—	—
Missing and invalid data[b]	4501	3421	1080	—	—
Sex, n (%)	6574 (61.13)	3304 (50.92)	3270 (76.54)	15.25 (1)	<.001
Female	3817 (58.11)	1997 (60.41)	1820 (55.73)	—	—
Male	2757 (41.9)	1307 (39.64)	1450 (44.37)	—	—
Missing and invalid data[b]	4194	3190	1004	—	—
Race, n (%)	6681 (62.03)	3342 (51.57)	3339 (78.15)	101.02 (4)	<.001
Non-Hispanic White	3938 (58.95)	1953 (58.44)	1985 (59.41)	—	—
Asian	931 (13.93)	487 (14.67)	444 (13.32)	—	—
Hispanic, Latino, or Spanish	783 (11.72)	424 (12.75)	359 (10.81)	—	—
Black or African American	723 (10.82)	267 (8.02)	456 (13.77)	—	—
Other	306 (4.61)	211 (6.32)	95 (2.85)	—	—
Missing and invalid data[b]	4087	3152	935	—	—
Marital status, n (%)	6681 (62.03)	3341 (51.42)	3341 (78.21)	134.02 (3)	<.001
Single	3312 (49.65)	1439 (43.14)	1873 (56.13)	—	—
Married or in a domestic partnership	2821 (42.22)	1549 (46.47)	1273 (38.12)	—	—
Divorced	410 (6.11)	275 (8.28)	135 (4.03)	—	—
Other	138 (2.16)	78 (2.39)	60 (1.81)	—	—
Missing and invalid data[b]	4087	3153	933	—	—
Income level (US $), n (%)	5793 (53.85)	2483 (38.24)	3310 (77.47)	245.48 (4)	<.001
<25,000	1736 (30.05)	599 (24.11)	1137 (34.42)	—	—
25,000 to 49,999	1268 (21.91)	463 (18.64)	805 (24.33)	—	—
50,000 to 74,999	886 (15.37)	349 (14.15)	537 (16.21)	—	—
75,000 to 99,999	710 (12.33)	343 (13.85)	367 (11.14)	—	—
≥100,000	1193 (20.62)	729 (29.41)	464 (14.05)	—	—
Missing and invalid data[b]	4975	4011	964	—	—
Level of education, n (%)	6677 (62.04)	3340 (51.43)	3337 (78.11)	35.34 (2)	<.001
High school or lower	868 (13.09)	448 (13.41)	420 (12.65)	—	—
College	3881 (58.16)	1827 (54.71)	2054 (61.62)	—	—
Graduate school	1928 (28.91)	1065 (31.94)	863 (25.93)	—	—
Missing and invalid data[b]	4091	3154	937	—	—

[a]Not available.

[b]The proportion is based on the number of participants who completed the baseline survey, and missing and invalid data are presented in Multimedia Appendix 2.

Passive Data Sharing

The number of data modalities that were passively collected by the study app varied across the Android (31 data modalities) and iOS (14 data modalities) operating systems. The variation in the number of passive data modalities available across Android and iOS devices is because of the available onboard sensors and data collection restrictions across the two operating systems [38]. Of the 31 Android passive data streams, 18 (58%) were shared by at least 50% of the Android users across the 2 study phases (Table 2). In contrast, 86% (12/14) of the distinct passive data streams were shared by at least 50% of the participants using iOS devices. Multimedia Appendix 4 summarizes data-sharing proportions per sensor stratified across Android and iOS devices. None of the participants with iOS devices shared passive data from the camera or barometer. Similarly, participants with Android devices did not share any data from some passive data streams, including temperature, camera, and humidity (Multimedia Appendix 4). This variation in passive data sharing could also be linked to the heterogeneity and nonavailability of specific sensors in some devices. It is worth noting that phase 2 of the study required participants to share accelerometer and gyroscope data passively. However, a small yet notable proportion of the cohort recruited in phase 2 did not share accelerometer (503/4089, 12.31%) and gyroscope (856/4089, 20.89%) data.

In addition, across the 8 passive data streams that were common between Android and iOS devices, the participants' passive data sharing was linked to sociodemographic characteristics and device type. In total, 3 data-sharing patterns of participants sharing at least 2 (25%), 4 (50%), or 6 (75%) of the total 8 passive data streams were tested. Across all 3 data-sharing patterns, Black or African American participants were found to be statistically significantly less likely to share passive sensor data than non-Hispanic White participants (odds ratio [OR] 0.44-0.49, 95% CI 0.35-0.61; $P<.001$). Furthermore, participants sharing ≥75% (6/8) of the passive data streams were more likely to be iOS device users (OR 1.98, 95% CI 1.71-2.28; $P<.001$) and earning more than US $25,000 per year (OR 1.27-1.55, 95% CI 1.06-1.93; $P<.001$). Multimedia Appendix 5 provides further details on the association between participants' sociodemographic characteristics and passive data sharing.

Table 2. Comparison of the impact of individual sociodemographic variables on the median retention (95% CI) of participants (in days) in the Warfighter Analytics Using Smartphones for Health study across 2 phases.

	Phase 1		Phase 2	
	Retention median (95% CI)	*P* value	Retention median (95% CI)	*P* value
Data streams	37 (37-37)	<.001	51 (49-53)	<.001
Passive	37 (37-37)		44 (43-46)	
Active	36 (36-36)		47 (44-49)	
Baseline data missingness		<.001		<.001
Yes	36 (36-37)		59 (57-62)	
No	37 (37-38)		19 (16-24)	
Age (years)		.01		<.001
19-29	36 (34-37)		59 (55-63)	
30-39	36 (34-37)		59 (53-64)	
40-49	37 (36-38)		60 (50-66)	
50-59	37 (37-38)		69 (61-79)	
≥60	38 (37-38)		83 (78-N/A[a])	
Race or ethnicity		<.001		.50
Asian	39 (38-49)		60 (54-69)	
Black or African American	6 (4-10)		68 (63-72)	
Hispanic, Latino, or Spanish	20 (14-25)		58 (50-65)	
Non-Hispanic White	37 (37-38)		57 (54-61)	
Other	5 (3-17)		55 (35-67)	
Income level (US $)		<.001		.56
<25,000	37 (36-38)		61 (56-65)	
25,000 to 49,999	34 (31-36)		60 (55-66)	
50,000 to 74,999	36 (33-37)		62 (56-68)	
75,000 to 99,999	24 (19-30)		55 (49-64)	
>100,000	24 (19-30)		55 (50-61)	
Level of education		<.001		.001
High school or lower	5 (4-10)		50 (46-56)	
College	38 (37-38)		60 (57-64)	
Graduate school	36 (35-37)		62 (57-67)	
Device type		<.001		<.001
Android	22 (17- 27)		59.5 (56-63)	
iOS	37 (37-37)		49 (46-52)	

[a]N/A: not available.

Participant Retention

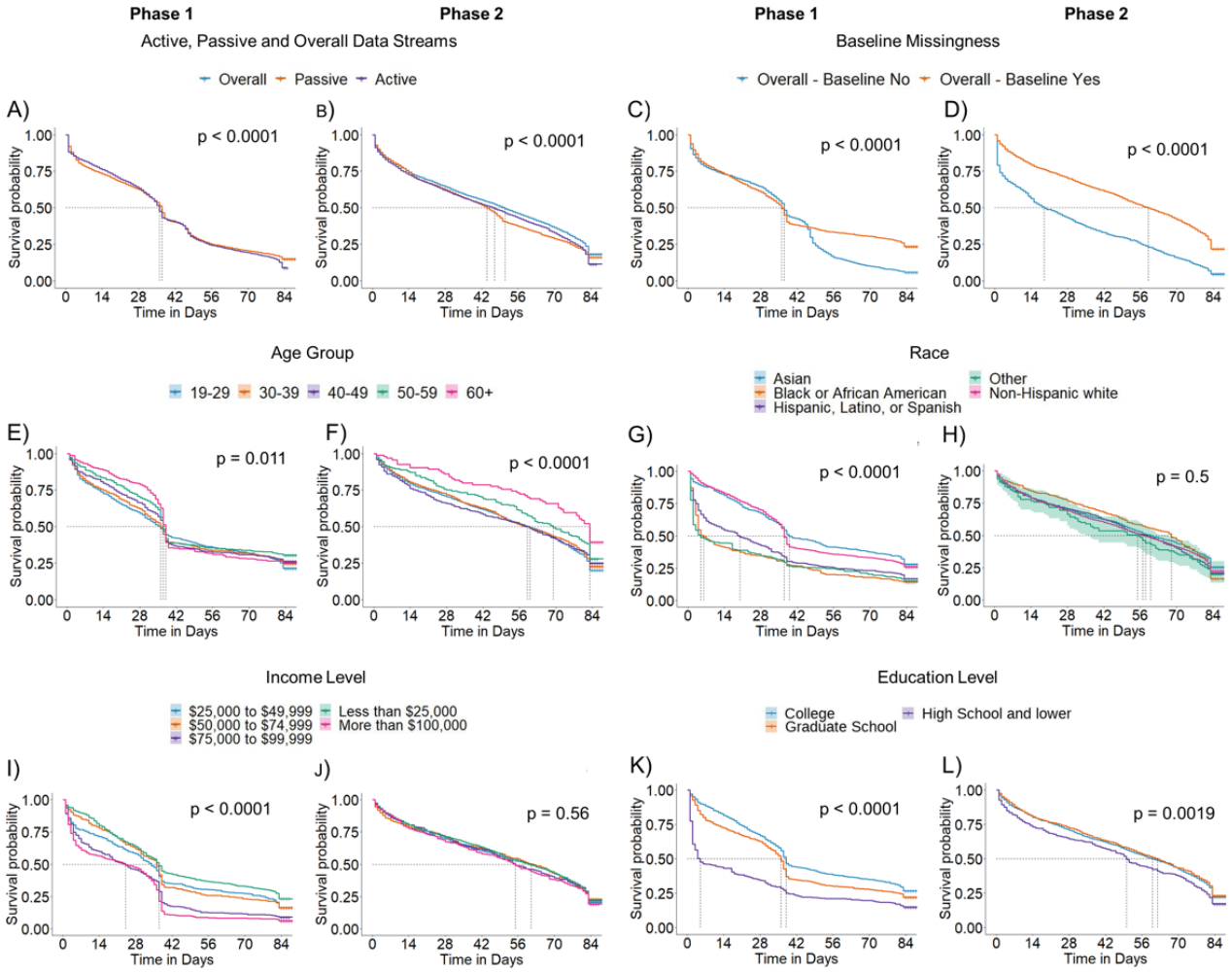
The median retention time of the overall cohort was 38 days, within the 84-day study observation period. No meaningful difference was observed in cohort retention across the active (median 37 days) and passive (median 38 days) data streams (Multimedia Appendix 6). The sensitivity analysis of participant retention also showed no significant difference in median survival across the active and passive data streams (Multimedia Appendix 7). Consequently, all subsequent retention analyses were conducted by combining the active and passive data streams.

Notable differences in retention were observed across the population recruited between phases 1 and 2. Participants recruited in phase 2 had a significantly higher median retention (+14 days) than those recruited in phase 1 (phase 1: median 37 days; phase 2: median 51 days; *P*<.001; Figures 3A and 3B). Older participants (≥60 years), recruited in both phases, remained engaged in the study for the longest duration (phase 1 and phase 2 median retention 38 days and 83 days,

respectively) relative to the younger cohort (Figures 3E and 3F).

It is worth noting that certain characteristics, including socioeconomic factors, distinctly impacted participant retention across the cohorts recruited in phases 1 and 2 (Table 2). Participants who completed the baseline survey administered immediately after enrollment in phase 2 were retained for a significantly longer period (with median values of baseline survey: yes 59 days vs no 19 days in phase 2; Figure 3D). However, the same trend was not observed for participants recruited from social media platforms in phase 1. Similarly, in phase 1, the non-Hispanic White population was retained in the study for a significantly longer time (median 37 days) than the Hispanic or Latino population (median 20 days; Figure 3G). No meaningful differences were observed among non-Hispanic White and Hispanic, Latino, or Spanish populations in phase 2 (Figure 3H). Education level mainly impacted retention in phase 1. Participants reporting high school or lower education levels had the shortest retention (median 5 days) than other participants (median ≥36 days) in phase 1. Such a large difference in retention because of educational level was not seen in the population recruited in phase 2 from crowdsourcing platforms (Figures 3I-3K). Participants' self-reported income was also found to be significantly associated with retention in phase 1 only. Participants with incomes of <US $49,999 were retained longer than participants earning >US $100,000 (phase 1: US $49,999 vs US $100,000 median retention 34 days vs 24 days, respectively; P<.001; Figures 3I and 3J). We also noticed a dramatic difference in median participant retention between Android and iOS users enrolled in phase 1 (iOS 37 days and Android 22 days; P<.001). Table 2 and Multimedia Appendix 8 provide additional results and details on the survival analysis.

Figure 3. Study retention patterns across the 2 recruitment phases using Kaplan-Meier survival curves. (A)-(B) Cohort retention stratified by active (purple), passive (orange), and overall (ie, active or passive; blue) data streams. (C)-(D) Difference in retention based on completion of the baseline survey; cohort retention by (E)-(F) age group, (G)-(H) race or ethnicity, (I)-(J) income level, and (K)-(L) education level. The shaded region shows the 95% CIs based on the survival model fit.



Discussion

Principal Findings

Our results from the analysis of over 330,000 days of engagement data collected from over 10,000 participants in real-world settings showed that combining different recruitment and incentive distribution approaches can yield heterogeneous cohorts. To the best of our knowledge, this is one of the first studies to empirically assess real-world differences in participants' sharing of multimodal passive data collected from iOS and Android devices using a bring your own device (BYOD) approach.

Overall, there were 5 key learnings. First, recruiting participants using different media, for example, digital advertisements on social media and web-based newspapers or crowdsourcing platforms, could result in heterogeneous subcohorts with varying characteristics. Second, participant engagement could vary significantly based on the recruitment source (eg, social media vs crowdsourced platforms) and incentive distribution approaches. Third, passive data collection could be substantially affected by technical variations in Android and iOS devices and the sociodemographic demographics of the cohort. Fourth, there is a greater likelihood of participants completing baseline health surveys if they are administered immediately after consent or enrollment. Fifth, monitoring patterns in real-world data collection at the study level could reveal technical glitches that could help guide contextual data filtering and cohort selection, leading to more reliable evidence generation. We now contextualize our principal findings to inform strategies to recruit, retain, and monitor trends in remote data collection to help collect real-world health data in a representative and equitable manner.

Combining Recruitment Platforms Could Yield Heterogeneous Real-world Cohorts

Notable differences were observed between the demographic and socioeconomic characteristics of participants recruited from web-based advertisements in social media and newspapers (phase 1) versus crowdsourcing platforms (phase 2). This indicates that combining multiple web-based recruitment sources could yield heterogeneous cohorts, resulting in nonuniform data collection. Future remote studies should assess the potential impact of combining the real-world data obtained from participants enrolled through different recruitment media. Furthermore, the web-based advertisement–based open enrollment approach in phase 1 while successfully recruiting a large cohort quickly also resulted in a significant proportion of bad actors joining the study to receive monetary incentives [37]. On the other hand, crowdsourcing platforms (MTurk and Prolific) were slower in recruiting participants, but their retention was notably higher than that of participants recruited using social media advertisements in phase 1. Indeed, as we have noted in an earlier paper, news outlets and social media recruitment are more likely to attract malicious actors [37] and, as we demonstrate here, less-committed research participants. However, despite the benefits of paid crowdsourcing platforms in effectively reaching and recruiting participants, researchers should carefully consider other factors that could influence the findings of a study [53-57] when recruiting participants from such platforms. These include (1) the primary motivation to remain engaged in remote studies, which may be tied to monetary incentives linked to task completion, and (2) the recruited population may not be representative of the general population [58] or of target health conditions. The characteristics of recruited participants may also vary substantially across recruitment platforms. (3) Nonnaivety-recruited people could be routine participants in research, which could impact the assessment of the actual underlying effect and (4) assessment of the fitness for the purpose of crowdsourced workers for a particular task or study [59-61].

Changes in the Incentive Distribution Can Have an Impact on Recruitment and Retention

By increasing the interval at which participants were paid, a significant reduction was observed in the number of malicious actors joining the study in phase 2. Furthermore, keeping the total incentive paid the same, participants who received less compensation weekly (phase 1) remained engaged in the study for a significantly shorter duration than those receiving a larger sum every 3 weeks (phase 2). Although higher retention in phase 2 cannot solely be attributed to a change in incentive distribution (because of a lack of randomization), it is indicative of a potentially interesting behavioral economics model [62] that addresses the perceived burden of participants with episodic but more significant rewards. The value of incentives relative to the study burden also varied by socioeconomic characteristic. In both phases, participants in lower-income groups engaged for longer, likely driven by the incentives, than those in higher-income groups, a finding evident in other studies [63]. Past research has shown that incentives can be an effective way to retain such participants, as small incentives could constitute a way of dealing with monetary barriers [64]. However, researchers should use such incentive-based engagement strategies in a noncoercive manner [65,66] so that potential study participants are not unduly influenced to join and share their data in a research study.

Assessing Patterns in Real-world Data Collection Can Reveal Underlying Technical Issues

The evaluation of day-to-day study-level data revealed several patterns indicating transient technical glitches in data collection that, if unaddressed, could bias downstream evidence generation. First, a significant drop in the relative rate of enrollment for baseline survey completion was observed in phase 1 (Figure 2). This could be indicative of a technical glitch in the data collection system or an attempt by a large number of bad actors to join the study to gain monetary incentives (if applicable). Second, active and passive data collection patterns varied notably across the study recruitment phases. For example, we identified 2 periods during the second phase of the study, when the study app collected no passive data despite the completion of active tasks by participants (Multimedia Appendix 9). This is likely a technical glitch in passive data collection that could severely impact the passive data density for the participants who were active during this period. Understanding the context and period in which the data are missing can guide cohort and data selection for a reliable and unbiased downstream analysis. Third, a small but substantial subset of participants recruited in phase 2 did not share the 2 mandatory passive data streams, accelerometer (503/4089, 12.3%) and gyroscope (856/4089, 20.9%), but continued to remain active in the study (Table 2). Near real-time comparison of data modalities shared by participants can help the study teams triage participants who do not meet the required inclusion criteria per the approved study protocol. Fourth, the retention analysis stratified by technical variables (eg, device type) revealed latent idiosyncratic patterns. We observed a notable trend in retention for the participants recruited in phase 1 (Multimedia Appendix 8; Figure 3G). Participants with iOS devices showed a dramatic drop in retention around day 37 compared with a gradual decline for

those with Android devices. There may be several plausible reasons for this significant yet idiosyncratic retention pattern, seen only in phase 1 of the study. The sociodemographic characteristics varied significantly between the iOS and Android cohorts in phase 1 compared with phase 2 (Multimedia Appendix 3). In addition, there could have been a bug in the iOS app around week 6 (days 35-42) that could have impacted participant experience and data sharing in phase 1. Taken together, these findings show an urgent need to prioritize real-time monitoring of data collection in real-world settings while the study is in progress. This also provides a just-in-time intervention opportunity to understand, document, and fix the root cause, preventing lower-quality data collection.

Passive Data Collection Can Vary Substantially in BYOD Studies

Passive data collected from the participants' own devices showed that the onboard sensors available across Android and iOS devices can vary substantially. Even for the common passive data streams available on both Android and iOS devices, there can be substantial differences in the sharing of multiple passive data streams linked to participants' sociodemographic characteristics and device types. For example, in this study, Black individuals or African Americans were significantly less likely to share multimodal passive sensor data (Multimedia Appendix 5). Researchers should expect a high degree of heterogeneity in passive sensor data streams in large BYOD studies and consider the impact of device heterogeneity on data collection, analysis, and evidence generation [38,67-70].

Impact of Participants' Sociodemographic Characteristics on Retention

Older participants (aged ≥60 years) were retained in the study for the longest duration. This finding is consistent with a previous large cross-study comparison of retention [63]. However, the impact of sociodemographic characteristics on participant retention was considerably different between the cohorts recruited using social media advertisements (phase 1) versus crowdsourcing platforms (phase 2). The relative difference in median retention within individual categories (eg, non-Hispanic White vs Hispanic or Latino) was remarkably higher and aligned with prior research [63] in the phase 1 cohort than the cohort recruited in phase 2 (Table 2). This indicates a significant discrepancy in how sociodemographic characteristics may affect participant retention based on recruitment sources. Our findings offer evidence that the population recruited from web-based crowdsourcing platforms shows more homogeneous engagement in research studies than the general population, a behavior that is likely driven by underlying motivation and monetary incentives.

In addition, the demographic composition of the United States is becoming increasingly multiethnic and pluralistic, and it is projected that there will be no majority racial or ethnic groups by 2060 [71]. The sociodemographic characteristics of the enrolled cohort together with nonuniform participant attrition show that large observational studies may not enroll and collect health outcomes from a diverse and representative population uniformly. Future studies should emphasize enrolling diverse populations, such as an All of Us cohort [72], and retaining a diverse sample throughout the study period to ensure that their learnings apply to diverse populations. In addition, some of the challenges in recruiting a diverse cohort have been identified to be related to participants' perceptions, trust, and willingness to enroll and share their data with researchers, governments, and academic institutions [46,71].

Timing of Administration of Baseline Surveys May Impact Completion Rates

The engagement data showed that the timing of administration of the baseline survey could be linked to survey completion rates. The missingness rates of the baseline survey were notably different between the 2 phases (phase 1: 3135/6494, 48.27%; phase 2: 918/4274, 22.47%; Figures 3C and 3D). This indicates that participants were more likely to complete the baseline assessments if they were administered immediately after consent or enrollment (phase 2). This is likely due to a higher level of engagement when enrolling for the study than at subsequent time points, when attention may be captured by other activities. This finding is aligned with some prior research in which participants were more likely to engage with a mobile health app within 24 hours if prompts were provided when participants are most receptive [73]. Moreover, Bidargaddi et al [73] revealed that the degree of engagement is also influenced by other contexts, such as the time of day and the day of the week. These results could help us understand the importance of time of administering an assessment and its impact on data quality in research studies.

Limitations and Future Directions

The analysis of participant recruitment and retention data from the WASH study should be interpreted within the context of certain limitations. First, large-scale, fully remote data collection started close to the declaration of the COVID-19 pandemic in the United States, which is known to have changed our behavior and interaction with technology and devices [74]. Indeed, Inverso et al [75] showed higher engagement rates during the COVID-19 pandemic because of an increased reliance on technology during the lockdown. The WASH study began recruitment on March 15, 2020, shortly after the World Health Organization declared COVID-19 a pandemic on March 11, 2020. Therefore, we did not have pre-post pandemic data to account for the potential impact of the pandemic on participant engagement with technology and devices. Second, the original purpose of the WASH study was to use the study app to detect cold and influenza symptoms. Thus, participants were not randomized among different recruitment platforms, incentive distribution frequency, and timing of baseline surveys that varied between phases 1 and 2 of the study. Consequently, our findings are not causal or linked to the impact of 1 factor on participant recruitment and retention between phases. For example, this analysis compares the population characteristics of those recruited from web-based crowdsourcing platforms (phase 2) compared with participants enrolling based on social media and local advertisements (open enrollment phase 1) as a whole. We were not able to explore within-phase recruitment differences; that is, between those recruited from social media versus those recruited from local news media advertisements. This is mainly because of the limited information available in the study data,

which does not allow for such differences to be investigated. Further research studies using a randomized design are needed to investigate the impact of individual changes in recruitment and retention strategies and their effectiveness for use in decentralized research. Third, we could not control for the participants' previous experience in crowdsourcing platforms and research tasks, which can be a confounder [41] depending on the nature of the assessment. Future research studies should assess participants' prior participation in similar or other research studies to assess any differential impact on primary outcomes. Fourth, in phase 1, participant recruitment via press releases was centered in the Greater Seattle area, which may not be representative of the population of the United States. In addition, because of the high proportion of missingness in the baseline geolocation data, we could not determine the geospatial representativeness of the cohort. Future studies should prioritize collecting high-level geolocation data, such as the state, city, or zip code, to help assess the geospatial representativeness of the study cohort. Fifth, we could not account for all the underlying within-study differences in the outcomes; for instance, the probable technical glitches concerning the steep drop in participant engagement at the participant level on day 36 in phase 1 and fluctuations in sensor data collection or management in phase 2 (Multimedia Appendix 9). These technical issues could have impacted the participants' willingness to remain engaged and increased the perceived burden of participants who were active in the study at the time of technical glitches. Sixth, despite our filtering out bad actors, some could still have been successfully enrolled by creating multiple accounts or using multiple devices. We suggest that future digital health research studies specifically report and compare the impact of different temporal recruitment and incentive strategies on enrolled cohorts' characteristics and engagement metrics as well as fraudulent enrollments to allow for future replication and the establishment of a set of guidelines for successful methods of participant recruitment and retention.

Acknowledgments

The data from this study were funded by the SC1712603 *Warfighter Analytics Using Smartphones for Health (principal investigator: Bracken)*. This material is based on work supported by the United States Air Force and the Defense Advanced Research Projects Agency under contract FA8750-18-C-0056, titled "Health and Injury Prediction and Prevention Over Complex Reasoning and Analytic Techniques Integrated on a Cellphone App (HIPPOCRATIC App)." The views, opinions, and findings expressed are those of the author and should not be interpreted as representing the official views or policies of the Department of Defense or the United States Government.

The authors thank all the study participants for their time in the study and for sharing their data for research purposes. The authors would also like to acknowledge Calvin Herd for helping download the study engagement data and Aditi Surendra for proofreading the manuscript and helping in the design of the study data collection illustration.

FJ was supported by funding from the National Institute on Aging (K76AG064390) and has received honoraria from the Benson-Henry Institute for Mind Body Medicine at the Massachusetts General Hospital and salary support from the Massachusetts General Hospital Clinical Trials Network Institute, which receives funding from the National Institutes of Health and pharmaceutical companies. SXL, RS, RH, and AP were funded for this work in part by the Centre for Addiction and Mental Health Discovery Fund and the Krembil Foundation, Canada.

Data Availability

The individual-level engagement data used for the study will be available upon reasonable request from the corresponding author. The complete code used for data loading and analysis is available through an open-source GitHub code repository [28].

Conflicts of Interest

At the time of this manuscript's acceptance for publication, Biogen employs AP. However, all analysis, manuscript writing, and initial submission were completed before AP joined Biogen.

Multimedia Appendix 1
Test statistics from the Schoenfeld test used to assess the Cox proportional hazards model assumption.
[PDF File (Adobe PDF File), 40 KB - formative_v6i11e40765_app1.pdf]

Multimedia Appendix 2
Missing data analysis.
[PDF File (Adobe PDF File), 47 KB - formative_v6i11e40765_app2.pdf]

Multimedia Appendix 3
Sociodemographic characteristics of participants in Warfighter Analytics Using Smartphones for Health (WASH) Study stratified by device ownership (iOS and Android).

[PDF File (Adobe PDF File), 54 KB - formative_v6i11e40765_app3.pdf]

Multimedia Appendix 4
The proportion of participants who shared sensor data by device type, Android and iOS, across the 2 study phases. *Sensor types that were mandatory to share. **Permission is required to share the data on Android devices.
[XLSX File (Microsoft Excel File), 22 KB - formative_v6i11e40765_app4.xlsx]

Multimedia Appendix 5
Odds ratios and 95% CIs from a logistic regression model showing the association between participants' sociodemographic characteristics and passive data sharing. The results from 3 separate models comparing passive data-sharing patterns using 2 (25%), 4 (50%), and 6 (75%) of the 8 passive data streams that were common between Android and iOS devices are shown below. *P<.051. **P<.001. ***P<.0001.
[PDF File (Adobe PDF File), 75 KB - formative_v6i11e40765_app5.pdf]

Multimedia Appendix 6
Overall participant retention in the Warfighter Analytics Using Smartphones for Health (WASH) study.
[PDF File (Adobe PDF File), 27 KB - formative_v6i11e40765_app6.pdf]

Multimedia Appendix 7
Sensitivity analysis of participant retention by extending the observation window (84 days) by 2 weeks.
[PDF File (Adobe PDF File), 32 KB - formative_v6i11e40765_app7.pdf]

Multimedia Appendix 8
Additional survival curves.
[PDF File (Adobe PDF File), 484 KB - formative_v6i11e40765_app8.pdf]

Multimedia Appendix 9
Patterns in Study App Data Collection.
[PDF File (Adobe PDF File), 317 KB - formative_v6i11e40765_app9.pdf]

References

1. Cao J, Lim Y, Sengoku S, Guo X, Kodama K. Exploring the shift in international trends in mobile health research from 2000 to 2020: bibliometric analysis. JMIR Mhealth Uhealth 2021 Sep 08;9(9):e31097 [FREE Full text] [doi: 10.2196/31097] [Medline: 34494968]
2. Lunn MR, Lubensky M, Hunt C, Flentje A, Capriotti MR, Sooksaman C, et al. A digital health research platform for community engagement, recruitment, and retention of sexual and gender minority adults in a national longitudinal cohort study--the PRIDE study. J Am Med Inform Assoc 2019 Aug 01;26(8-9):737-748 [FREE Full text] [doi: 10.1093/jamia/ocz082] [Medline: 31162545]
3. Penetration rate of smartphones in selected countries 2021. Statista. 2022. URL: https://www.statista.com/statistics/539395/smartphone-penetration-worldwide-by-country/ [accessed 2022-03-21]
4. De Brouwer W, Patel CJ, Manrai AK, Rodriguez-Chavez IR, Shah NR. Empowering clinical research in a decentralized world. NPJ Digit Med 2021 Jul 01;4(1):102 [FREE Full text] [doi: 10.1038/s41746-021-00473-w] [Medline: 34211085]
5. Badawy R, Hameed F, Bataille L, Little MA, Claes K, Saria S, et al. Metadata concepts for advancing the use of digital health technologies in clinical research. Digit Biomark 2019 Oct 7;3(3):116-132 [FREE Full text] [doi: 10.1159/000502951] [Medline: 32175520]
6. Van Norman GA. Decentralized clinical trials: the future of medical product development?*. JACC Basic Transl Sci 2021 Apr;6(4):384-387 [FREE Full text] [doi: 10.1016/j.jacbts.2021.01.011] [Medline: 33997523]
7. Roehr B. Pfizer launches virtual clinical trial. BMJ 2011 Jun 14;342:d3722. [doi: 10.1136/bmj.d3722] [Medline: 21672995]
8. Pfizer Conducts First "Virtual" Clinical Trial Allowing Patients to Participate Regardless Of Geography. Pfizer. 2011 Jun 7. URL: https://www.pfizer.com/news/press-release/press-release-detail/pfizer_conducts_first_virtual_clinical_trial_allowing_patients_to_participate_regardless_of_geography [accessed 2022-04-19]
9. Pratap A, Atkins DC, Renn BN, Tanana MJ, Mooney SD, Anguera JA, et al. The accuracy of passive phone sensors in predicting daily mood. Depress Anxiety 2019 Jan;36(1):72-81 [FREE Full text] [doi: 10.1002/da.22822] [Medline: 30129691]
10. Nickels S, Edwards MD, Poole SF, Winter D, Gronsbell J, Rozenkrants B, et al. Toward a mobile platform for real-world digital measurement of depression: user-centered design, data quality, and behavioral and clinical modeling. JMIR Ment Health 2021 Aug 10;8(8):e27589 [FREE Full text] [doi: 10.2196/27589] [Medline: 34383685]

11. Lee JL, Cerrada CJ, Vang MK, Scherer K, Tai C, Tran JL, et al. The DiSCover project: protocol and baseline characteristics of a decentralized digital study assessing chronic pain outcomes and behavioral data. medRxiv 2021 Jul 18. [doi: 10.1101/2021.07.14.21260523]
12. Sundquist S, Batist G, Brodeur-Robb K, Dyck K, Eigl BJ, Lee DK, et al. CRAFT-a proposed framework for decentralized clinical trials participation in Canada. Curr Oncol 2021 Sep 30;28(5):3857-3865 [FREE Full text] [doi: 10.3390/curroncol28050329] [Medline: 34677247]
13. Larbi D, Randine P, Årsand E, Antypas K, Bradway M, Gabarron E. Methods and evaluation criteria for apps and digital interventions for diabetes self-management: systematic review. J Med Internet Res 2020 Jul 06;22(7):e18480 [FREE Full text] [doi: 10.2196/18480] [Medline: 32628125]
14. Omberg L, Chaibub Neto E, Perumal TM, Pratap A, Tediarjo A, Adams J, et al. Remote smartphone monitoring of Parkinson's disease and individual response to therapy. Nat Biotechnol 2022 Apr;40(4):480-487. [doi: 10.1038/s41587-021-00974-9] [Medline: 34373643]
15. Mayfield JJ, Chatterjee NA, Noseworthy PA, Poole JE, Ackerman MJ, Stewart J, COVID-19 Early Treatment Team, et al. Implementation of a fully remote randomized clinical trial with cardiac monitoring. Commun Med (Lond) 2021 Dec 20;1:62 [FREE Full text] [doi: 10.1038/s43856-021-00052-w] [Medline: 35604806]
16. Badawy SM, Radovic A. Digital approaches to remote pediatric health care delivery during the COVID-19 pandemic: existing evidence and a call for further research. JMIR Pediatr Parent 2020 Jun 25;3(1):e20049 [FREE Full text] [doi: 10.2196/20049] [Medline: 32540841]
17. Farah J, Vasey J, Kallenbach L, Caplea G. Conducting Research at the Point of Care. Veradigm® Allscripts Healthcare. 2021. URL: https://veradigm.com/img/resource-conducting-research-at-point-of-care-whitepaper.pdf [accessed 2022-03-21]
18. Hillman A, Castañeda R. Who and what is at the crest of the clinical trial decentralisation wave? Clinical Trials Arena. 2022 Feb 3. URL: https://www.clinicaltrialsarena.com/analysis/dct-adoption-tracker-who-and-what-is-at-the-crest-of-the-trial-decentralisation-wave/ [accessed 2022-04-19]
19. Inan OT, Tenaerts P, Prindiville SA, Reynolds HR, Dizon DS, Cooper-Arnold K, et al. Digitizing clinical trials. NPJ Digit Med 2020 Jul 31;3:101 [FREE Full text] [doi: 10.1038/s41746-020-0302-y] [Medline: 32821856]
20. Khozin S, Coravos A. Decentralized trials in the age of real-world evidence and inclusivity in clinical investigations. Clin Pharmacol Ther 2019 Jul;106(1):25-27. [doi: 10.1002/cpt.1441] [Medline: 31013350]
21. Ferrar J, Griffith GJ, Skirrow C, Cashdollar N, Taptiklis N, Dobson J, et al. Developing digital tools for remote clinical research: how to evaluate the validity and practicality of active assessments in field settings. J Med Internet Res 2021 Jun 18;23(6):e26004 [FREE Full text] [doi: 10.2196/26004] [Medline: 34142972]
22. Hilderbrand A, Zangrilli M, Stinson M. Decentralized clinical trial case study: five-stage process for recruiting and completing a site-less clinical study in less time and lower cost than traditional methods. Am J Health Res 2021 Dec;9(6):213-217. [doi: 10.11648/j.ajhr.20210906.11]
23. Torous J, Kiang MV, Lorme J, Onnela JP. New tools for new research in psychiatry: a scalable and customizable platform to empower data driven smartphone research. JMIR Ment Health 2016 May 05;3(2):e16 [FREE Full text] [doi: 10.2196/mental.5165] [Medline: 27150677]
24. Goodday SM, Karlin E, Alfarano A, Brooks A, Chapman C, Desille R, Stress And Recovery Participants, et al. An alternative to the light touch digital health remote study: the stress and recovery in frontline COVID-19 health care workers study. JMIR Form Res 2021 Dec 10;5(12):e32165 [FREE Full text] [doi: 10.2196/32165] [Medline: 34726607]
25. Meyerowitz-Katz G, Ravi S, Arnolda L, Feng X, Maberly G, Astell-Burt T. Rates of attrition and dropout in app-based interventions for chronic disease: systematic review and meta-analysis. J Med Internet Res 2020 Sep 29;22(9):e20283 [FREE Full text] [doi: 10.2196/20283] [Medline: 32990635]
26. Kapp JM, Peters C, Oliver DP. Research recruitment using Facebook advertising: big potential, big challenges. J Cancer Educ 2013 Mar;28(1):134-137. [doi: 10.1007/s13187-012-0443-z] [Medline: 23292877]
27. Shatz I. Fast, free, and targeted: Reddit as a source for recruiting participants online. Soc Sci Comput Rev 2016 May 19;35(4):537-549. [doi: 10.1177/0894439316650163]
28. Prolific. URL: https://www.prolific.co/ [accessed 2022-10-27]
29. Amazon Mechanical Turk. URL: https://www.mturk.com/ [accessed 2022-10-27]
30. Centiment. URL: https://www.centiment.co/ [accessed 2022-10-27]
31. CloudResearch. Prime Research Solutions. URL: https://www.cloudresearch.com/ [accessed 2022-10-27]
32. PatientsLikeMe. URL: https://www.patientslikeme.com/ [accessed 2022-06-22]
33. The Michael J. Fox Foundation for Parkinson's Research. URL: https://www.michaeljfox.org/homepage [accessed 2022-06-22]
34. Myers TL, Augustine EF, Baloga E, Daeschler M, Cannon P, Rowbotham H, 23andMe Research Team, et al. Recruitment for remote decentralized studies in Parkinson's disease. J Parkinsons Dis 2022;12(1):371-380 [FREE Full text] [doi: 10.3233/JPD-212935] [Medline: 34744053]
35. Anguera JA, Jordan JT, Castaneda D, Gazzaley A, Areán PA. Conducting a fully mobile and randomised clinical trial for depression: access, engagement and expense. BMJ Innov 2016 Jan;2(1):14-21 [FREE Full text] [doi: 10.1136/bmjinnov-2015-000098] [Medline: 27019745]

36. Sun S, Folarin AA, Ranjan Y, Rashid Z, Conde P, Stewart C, RADAR-CNS Consortium. Using smartphones and wearable devices to monitor behavioral changes during COVID-19. J Med Internet Res 2020 Sep 25;22(9):e19992 [FREE Full text] [doi: 10.2196/19992] [Medline: 32877352]
37. Bracken BK, Wolcott J, Potoczny-Jones I, Mosser BA, Griffith-Fillipo IR, Arean PA. Detection and remediation of malicious actors for studies involving remote data collection. In: Proceedings of the 15th International Joint Conference on Biomedical Engineering Systems and Technologies. 2022 Presented at: BIOSTEC '22; February 9-11, 2022; Virtual p. 377-383. [doi: 10.5220/0010805500003123]
38. Nishiyama Y, Ferreira D, Sasaki W, Okoshi T, Nakazawa J, Dey AK, et al. Using iOS for inconspicuous data collection: a real-world assessment. In: Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers. 2020 Sep Presented at: UbiComp-ISWC '20; September 12-17, 2020; Virtual p. 261-266. [doi: 10.1145/3410530.3414369]
39. Boyle A. UW Medicine seeks 25,000 volunteers to try outbreak-predicting smartphone app. GeekWire. 2020 Apr 7. URL: https://www.geekwire.com/2020/uw-medicine-seeks-25000-volunteers-try-outbreak-predicting-smartphone-app/ [accessed 2022-06-10]
40. UW Medicine recruiting for app to predict next outbreak. University of Washington Medicine. 2020 Apr 7. URL: https://newsroom.uw.edu/news/uw-medicine-recruiting-app-predict-next-outbreak [accessed 2022-06-10]
41. Palan S, Schitter C. Prolific.ac—a subject pool for online experiments. J Behav Exp Finance 2018 Mar;17:22-27. [doi: 10.1016/j.jbef.2017.12.004]
42. Peer E, Brandimarte L, Samat S, Acquisti A. Beyond the Turk: alternative platforms for crowdsourcing behavioral research. J Exp Soc Psychol 2017 May;70:153-163. [doi: 10.1016/j.jesp.2017.01.006]
43. Why participants get banned. Prolific. URL: https://www.prolific.co/blog/why-participants-get-banned [accessed 2022-06-30]
44. Bradley P. Bots and data quality on crowdsourcing platforms. Prolific. 2018 Aug 10. URL: https://www.prolific.co/blog/bots-and-data-quality-on-crowdsourcing-platforms [accessed 2022-07-01]
45. Lettmann H, Lumsden J. Prolific's participant pool – the present and the future. Prolific. 2018 Sep 18. URL: https://www.prolific.co/blog/prolifics-participant-pool-its-present-and-its-future [accessed 2022-06-10]
46. Pratap A, Allred R, Duffy J, Rivera D, Lee HS, Renn BN, et al. Contemporary views of research participant willingness to participate and share digital data in biomedical research. JAMA Netw Open 2019 Nov 01;2(11):e1915717 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.15717] [Medline: 31747031]
47. Coravos A, Goldsack JC, Karlin DR, Nebeker C, Perakslis E, Zimmerman N, et al. Digital medicine: a primer on measurement. Digit Biomark 2019 May 9;3(2):31-71 [FREE Full text] [doi: 10.1159/000500413] [Medline: 32095767]
48. Rich JT, Neely JG, Paniello RC, Voelker CC, Nussenbaum B, Wang EW. A practical guide to understanding Kaplan-Meier curves. Otolaryngol Head Neck Surg 2010 Sep;143(3):331-336 [FREE Full text] [doi: 10.1016/j.otohns.2010.05.007] [Medline: 20723767]
49. Hazra A, Gogtay N. Biostatistics series module 9: survival analysis. Indian J Dermatol 2017;62(3):251-257 [FREE Full text] [doi: 10.4103/ijd.IJD_201_17] [Medline: 28584366]
50. Kumar D, Klefsjö B. Proportional hazards model: a review. Reliab Eng Syst Saf 1994 Jan;44(2):177-188. [doi: 10.1016/0951-8320(94)90010-8]
51. Hess KR. Graphical methods for assessing violations of the proportional hazards assumption in Cox regression. Stat Med 1995 Aug 15;14(15):1707-1723. [doi: 10.1002/sim.4780141510] [Medline: 7481205]
52. Bland JM, Altman DG. The logrank test. BMJ 2004 May 01;328(7447):1073 [FREE Full text] [doi: 10.1136/bmj.328.7447.1073] [Medline: 15117797]
53. Pennington CR, Jones AJ, Tzavella L, Chambers CD, Button KS. Beyond online participant crowdsourcing: the benefits and opportunities of big team addiction science. Exp Clin Psychopharmacol 2022 Aug;30(4):444-451. [doi: 10.1037/pha0000541] [Medline: 35025584]
54. Khare R, Good BM, Leaman R, Su AI, Lu Z. Crowdsourcing in biomedicine: challenges and opportunities. Brief Bioinform 2016 Jan;17(1):23-32 [FREE Full text] [doi: 10.1093/bib/bbv021] [Medline: 25888696]
55. Gleibs IH. Are all "research fields" equal? Rethinking practice for the use of data from crowdsourcing market places. Behav Res Methods 2017 Aug;49(4):1333-1342 [FREE Full text] [doi: 10.3758/s13428-016-0789-y] [Medline: 27515317]
56. Lovett M, Bajaba S, Lovett M, Simmering MJ. Data quality from crowdsourced surveys: a mixed method inquiry into perceptions of Amazon's Mechanical Turk Masters. Appl Psychol 2018;67(2):339-366. [doi: 10.1111/apps.12124]
57. Berry C, Kees J, Burton S. Drivers of data quality in advertising research: differences across MTurk and professional panel samples. J Advert 2022 Jun 27;51(4):515-529. [doi: 10.1080/00913367.2022.2079026]
58. Walters K, Christakis DA, Wright DR. Are Mechanical Turk worker samples representative of health status and health behaviors in the U.S.? PLoS One 2018 Jun 7;13(6):e0198835 [FREE Full text] [doi: 10.1371/journal.pone.0198835] [Medline: 29879207]
59. Anderson CA, Allen JJ, Plante C, Quigley-McBride A, Lovett A, Rokkum JN. The MTurkification of social and personality psychology. Pers Soc Psychol Bull 2019 Jun;45(6):842-850. [doi: 10.1177/0146167218798821] [Medline: 30317918]
60. Edgar J, Murphy J, Keating M. Comparing traditional and crowdsourcing methods for pretesting survey questions. SAGE Open 2016 Oct 01;6(4):215824401667177 [FREE Full text] [doi: 10.1177/2158244016671770]

61. Tahaei M, Vaniea K. Recruiting participants with programming skills: a comparison of four crowdsourcing platforms and a CS student mailing list. In: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems. 2022 Apr Presented at: CHI '22; April 29-May 05, 2022; New Orleans, LA, USA p. 1-15. [doi: 10.1145/3491102.3501957]
62. Kamenica E. Behavioral economics and psychology of incentives. Annu Rev Econ 2012 Sep 01;4(1):427-452. [doi: 10.1146/annurev-economics-080511-110909]
63. Pratap A, Neto EC, Snyder P, Stepnowsky C, Elhadad N, Grant D, et al. Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. NPJ Digit Med 2020 Feb 17;3:21 [FREE Full text] [doi: 10.1038/s41746-020-0224-8] [Medline: 32128451]
64. Galárraga O, Sosa-Rubí SG. Conditional economic incentives to improve HIV prevention and treatment in low-income and middle-income countries. Lancet HIV 2019 Oct;6(10):e705-e714 [FREE Full text] [doi: 10.1016/S2352-3018(19)30233-4] [Medline: 31578955]
65. Singer E, Bossarte RM. Incentives for survey participation when are they "coercive"? Am J Prev Med 2006 Nov;31(5):411-418. [doi: 10.1016/j.amepre.2006.07.013] [Medline: 17046413]
66. Ambuehl S, Ockenfels A. The ethics of incentivizing the uninformed. A vignette study. SSRN J 2016 Dec 22:1-5. [doi: 10.2139/ssrn.2888554]
67. Nishiyama Y, Ferreira D, Eigen Y, Sasaki W, Okoshi T, Nakazawa J, et al. IOS crowd–sensing won't hurt a bit!: AWARE framework and sustainable study guideline for iOS platform. In: Proceedings of the 8th International Conference on Distributed, Ambient and Pervasive Interactions. 2020 Presented at: HCII '20; July 19–24, 2020; Copenhagen, Denmark p. 223-243. [doi: 10.1007/978-3-030-50344-4_17]
68. Russell C, McCarthy M, Cappelleri JC, Wong S. Choosing a mobile sensor technology for a clinical trial: statistical considerations, developments and learnings. Ther Innov Regul Sci 2021 Jan;55(1):38-47. [doi: 10.1007/s43441-020-00188-2] [Medline: 32557010]
69. Demanuele C, Lokker C, Jhaveri K, Georgiev P, Sezgin E, Geoghegan C, et al. Considerations for conducting bring your own "device" (BYOD) clinical studies. Digit Biomark 2022 Jul 4;6(2):47-60 [FREE Full text] [doi: 10.1159/000525080] [Medline: 35949223]
70. Cho PJ, Yi J, Ho E, Shandhi MM, Dinh Y, Patil A, et al. Demographic imbalances resulting from the bring-your-own-device study design. JMIR Mhealth Uhealth 2022 Apr 08;10(4):e29510 [FREE Full text] [doi: 10.2196/29510] [Medline: 34913871]
71. Yancey AK, Ortega AN, Kumanyika SK. Effective recruitment and retention of minority research participants. Annu Rev Public Health 2006;27:1-28. [doi: 10.1146/annurev.publhealth.27.021405.102113] [Medline: 16533107]
72. All of Us Research Program Investigators, Denny JC, Rutter JL, Goldstein DB, Philippakis A, Smoller JW, et al. The "All of Us" research program. N Engl J Med 2019 Aug 15;381(7):668-676 [FREE Full text] [doi: 10.1056/NEJMsr1809937] [Medline: 31412182]
73. Bidargaddi N, Almirall D, Murphy S, Nahum-Shani I, Kovalcik M, Pituch T, et al. To prompt or not to prompt? A microrandomized trial of time-varying push notifications to increase proximal engagement with a mobile health app. JMIR Mhealth Uhealth 2018 Nov 29;6(11):e10123 [FREE Full text] [doi: 10.2196/10123] [Medline: 30497999]
74. Vargo D, Zhu L, Benwell B, Yan Z. Digital technology use during COVID-19 pandemic: a rapid review. Human Behav and Emerg Tech 2021 Jan;3(1):13-24. [doi: 10.1002/hbe2.242]
75. Inverso H, Abadula F, Morrow T, LeStourgeon L, Parmar A, Streisand R, et al. Pivoting during a pandemic: lessons learned from transitioning a multisite randomized controlled trial to a remote protocol in response to COVID-19. Transl Behav Med 2021 Dec 14;11(12):2187-2193 [FREE Full text] [doi: 10.1093/tbm/ibab103] [Medline: 34427687]

Abbreviations

BYOD: bring your own device
CoxPH: Cox proportional hazards
MTurk: Amazon Mechanical Turk
OR: odds ratio
WASH: Warfighter Analytics Using Smartphones for Health

Edited by T Leung, A Mavragani; submitted 05.07.22; peer-reviewed by K Riekert, M Eder; comments to author 02.08.22; revised version received 02.09.22; accepted 05.10.22; published 14.11.22.

Please cite as:
Li SX, Halabi R, Selvarajan R, Woerner M, Fillipo IG, Banerjee S, Mosser B, Jain F, Areán P, Pratap A
Recruitment and Retention in Remote Research: Learnings From a Large, Decentralized Real-world Study
JMIR Form Res 2022;6(11):e40765
URL: https://formative.jmir.org/2022/11/e40765
doi:10.2196/40765
PMID:36374539

Original Paper

Measurement of Vital Signs Using Lifelight Remote Photoplethysmography: Results of the VISION-D and VISION-V Observational Studies

Emily Heiden[1*], BM, MRCP (Resp); Tom Jones[1*], MRCP; Annika Brogaard Maczka[2], BSc; Melissa Kapoor[2], BSc, DPhil; Milan Chauhan[1], BSc; Laura Wiffen[1], BMBS, MRCP; Helen Barham[3], BSc, PhD; Jeremy Holland[4,5], MA, PhD; Manish Saxena[6], MBBS, MSc; Simon Wegerif[4], BSc; Thomas Brown[1], MBChB, PhD; Mitch Lomax[7], BSc, MSc, PhD; Heather Massey[7], BSc, MSc, PhD; Shahin Rostami[8], BSc, PhD; Laurence Pearce[4], BSc, MBA; Anoop Chauhan[1], MBChB, MBE, PhD

[1]Portsmouth Hospitals University NHS Trust, Portsmouth, United Kingdom

[2]Mind Over Matter Medtech Ltd, London, United Kingdom

[3]The Text Doctor, Wantage, United Kingdom

[4]Xim Ltd, Southampton, United Kingdom

[5]Centre for Business Innovation Limited, Cambridge, United Kingdom

[6]Barts NIHR Biomedical Research Centre, Queen Mary University of London, London, United Kingdom

[7]School of Sport, Health & Exercise Science, University of Portsmouth, Portsmouth, United Kingdom

[8]Data Science Lab, Polyra Limited, Bournemouth, United Kingdom

[*]these authors contributed equally

Corresponding Author:
Melissa Kapoor, BSc, DPhil
Mind Over Matter Medtech Ltd
Kemp House
London, EC1V 2NX
United Kingdom
Phone: 44 7881 927063
Email: melissa@mind-medtech.com

Abstract

Background: The detection of early changes in vital signs (VSs) enables timely intervention; however, the measurement of VSs requires hands-on technical expertise and is often time-consuming. The contactless measurement of VSs is beneficial to prevent infection, such as during the COVID-19 pandemic. Lifelight is a novel software being developed to measure VSs by remote photoplethysmography based on video captures of the face via the integral camera on mobile phones and tablets. We report two early studies in the development of Lifelight.

Objective: The objective of the Vital Sign Comparison Between Lifelight and Standard of Care: Development (VISION-D) study (NCT04763746) was to measure respiratory rate (RR), pulse rate (PR), and blood pressure (BP) simultaneously by using the current standard of care manual methods and the Lifelight software to iteratively refine the software algorithms. The objective of the Vital Sign Comparison Between Lifelight and Standard of Care: Validation (VISION-V) study (NCT03998098) was to validate the use of Lifelight software to accurately measure VSs.

Methods: BP, PR, and RR were measured simultaneously using Lifelight, a sphygmomanometer (BP and PR), and the manual counting of RR. Accuracy performance targets for each VS were defined from a systematic literature review of the performance of state-of-the-art VSs technologies.

Results: The VISION-D data set (17,233 measurements from 8585 participants) met the accuracy targets for RR (mean error 0.3, SD 3.6 vs target mean error 2.3, SD 5.0; n=7462), PR (mean error 0.3, SD 4.0 vs mean error 2.2, SD 9.2; n=10,214), and diastolic BP (mean error −0.4, SD 8.5 vs mean error 5.5, SD 8.9; n=8951); for systolic BP, the mean error target was met but not the SD (mean error 3.5, SD 16.8 vs mean error 6.7, SD 15.3; n=9233). Fitzpatrick skin type did not affect accuracy. The VISION-V data set (679 measurements from 127 participants) met all the standards: mean error −0.1, SD 3.4 for RR; mean error 1.4, SD 3.8 for PR; mean error 2.8, SD 14.5 for systolic BP; and mean error −0.3, SD 7.0 for diastolic BP.

Conclusions: At this early stage in development, Lifelight demonstrates sufficient accuracy in the measurement of VSs to support certification for a Level 1 Conformité Européenne mark. As the use of Lifelight does not require specific training or equipment, the software is potentially useful for the contactless measurement of VSs by nonclinical staff in residential and home care settings. Work is continuing to enhance data collection and processing to achieve the robustness and accuracy required for routine clinical use.

International Registered Report Identifier (IRRID): RR2-10.2196/14326

(JMIR Form Res 2022;6(11):e36340) doi:10.2196/36340

KEYWORDS

general practice; vital signs/methods; vital signs/standards; photoplethysmography; remote photoplethysmography; remote photoplethysmography; Lifelight; contactless; software; algorithm development; algorithm; blood pressure; health monitoring; health technology; remote monitoring

Introduction

The regular measurement of vital signs (VSs) is an integral component of clinical care, as changes in VS often occur a few hours before an adverse event [1], providing an opportunity for intervention. However, the recording of VSs is often inadequate, such that clinical deterioration often goes unnoticed or is not detected in time to treat effectively [2]. In response to this challenge, the National Early Warning Score (NEWS) has been developed as a systematic approach to identify and respond to patients at risk of deterioration in health care settings based on the scoring of respiratory rate (RR), oxygen saturation, temperature, systolic blood pressure (SBP), pulse rate (PR), and level of consciousness [3]. The Recognise Early Soft Signs, Take Observations, Respond, Escalate (RESTORE2) system for use in care homes incorporates the NEWS alongside observations of soft signs to identify potential deterioration in clinical conditions [4]; however, this requires staff to be trained in the measurement of VSs.

VS measurement following discharge, for example, after surgery, is also important to identify deterioration. A European study of 193 readmitted patients identified marked deteriorations in PR (23%) and RR (28%) but only small changes (1%-2%) in blood pressure (BP) and oxygen saturation [5]. However, another study of 725 patients reported that, while 53% followed at least 10 of the recommended steps necessary for accurate BP measurement at home, only 1% followed all 15 recommendations [6]. Thus, home measurement of VSs is important—RR and PR in particular—but requires several pieces of equipment (BP monitor, pulse oximeter) and for patients to be educated in best practices.

Digital health technologies, such as wireless smart patches that measure PR and RR and finger clip BP monitors, have the potential to improve the ease and accuracy of VS measurement (Table 1). Photoplethysmography (PPG; the basis of pulse oximetry) enables the rapid and simultaneous measurement of VSs by detecting changes in the light reflected from the skin surface due to volumetric changes in the blood vessels. PPG has been used to measure PR [7,8], oxygen saturation [9], BP [10,11], and RR [7,12]. The COVID-19 pandemic has increased interest in using remote technology as a way to monitor patients with nonserious symptoms to reduce the burden on health care facilities, making them available for high-risk groups and the seriously affected, and to monitor patients with other medical conditions, thereby avoiding the risk of SARS-CoV-2 infection associated with visits to health care facilities [13]. Contactless technology is also potentially useful in situations where current care cannot be readily used, such as in mental health settings [14].

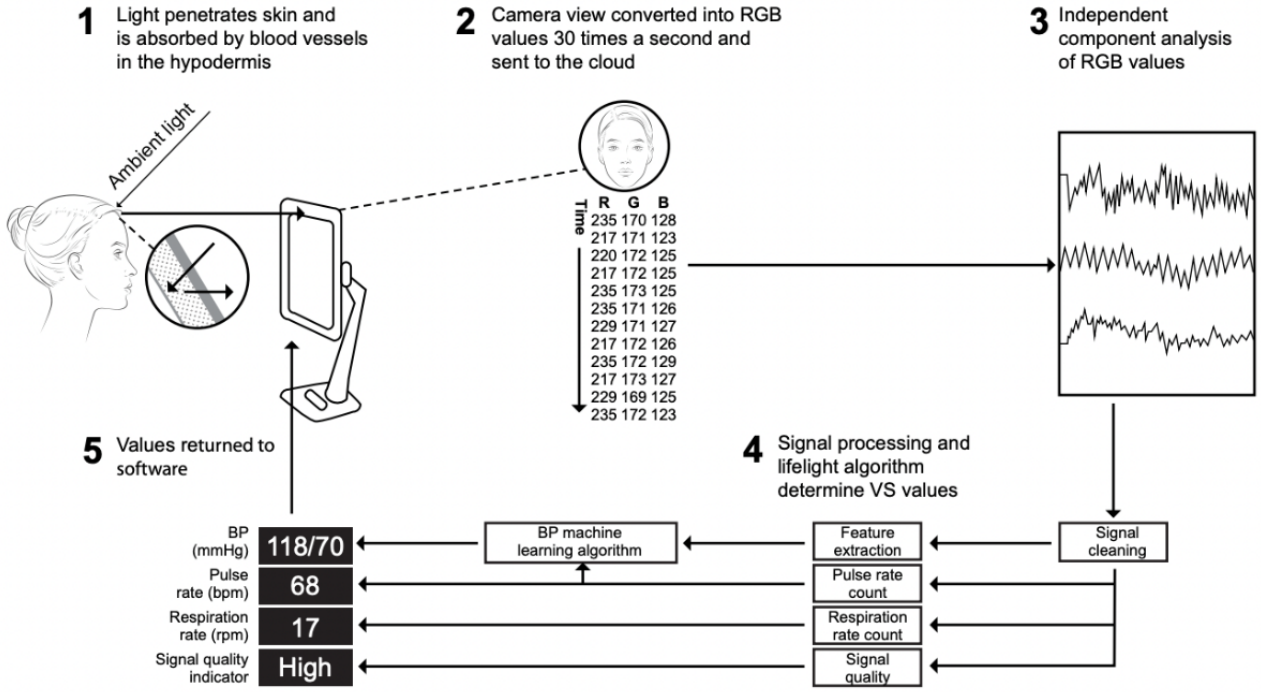
Lifelight (Xim Ltd) is a novel software being developed as a medical device for the measurement of VSs by remote PPG (rPPG), based on live video capture of the face using the integral camera on smart devices (eg, laptops or smartphones). The software captures the average color of multiple regions of interest 30 times every second for 60 seconds; subtle changes in coloration are used to determine VSs (Figure 1).

Table 1. Accuracy performance targets for Lifelight.

Key innovative technology and vital sign	Target accuracy, mean error (SD)	Basis[a]	References
Wireless smart patches			
Pulse rate (beats per minute)	2.2 (9.2)	Weighted average of performance of 3 devices	[15-17]
Respiratory rate (respirations per minute)	2.3 (5.0)	Weighted average of performance of 4 devices	[15,16,18,19]
Finger photoplethysmography monitor			
Systolic blood pressure (mmHg)	6.7 (15.3)	Weighted average of performance of 6 devices	[20-25]
Diastolic blood pressure (mmHg)	5.5 (8.9)	Weighted average of performance of 6 devices	[20-25]

[a]Relevance and quality scores: 3.0-3.16 for pulse rate, 2.83-3.16 for respiratory rate, 2.83-3.5 for systolic/diastolic blood pressure.

Figure 1. Use of remote photoplethysmography in the Lifelight software. BP: blood pressure; bpm: beats per minute; RGB: red, green, blue; rpm: respirations per minute; VS: vital sign.



Here, we report early validation steps of the Lifelight software for the measurement of PR, RR, SBP, and diastolic BP (DBP; in line with the intended purpose of Lifelight) [26]. As Lifelight is a contactless technology, there are no relevant International Organization for Standardization (ISO) standards for validation. The most similar devices are wireless smart patches and finger clip BP monitors. We therefore developed a novel methodology for validation, informed by the standards for pulse oximetry [27] and noninvasive sphygmomanometers [28], and a validation study of the pyroelectric polymer for measuring RR [29]. A rigorous systematic literature review was performed according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines to identify the performance of relevant devices for each VS. Quality and relevance scores were used to weight the findings (average score out of 4 for methodological quality: study design, sample size, method of comparison; scientific validity; and relevance to Lifelight's intended purpose). These performance targets (Table 1) were calculated in preparation for a Conformité Européenne (CE)–marking audit. As per criterion 1 of the standard for noninvasive sphygmomanometers [28], the mean error and SD of the Lifelight measurements are compared with standard of care (SOC) measurements recorded concurrently. The targets in Table 1 have been approved by the UK Health Research Authority (HRA) for the ongoing Vital Sign Comparison Between Lifelight and Standard of Care (VISION) Acute study (NCT04589923).

Here, we report two early studies in the development of Lifelight. The objective of the Vital Sign Comparison Between Lifelight and Standard of Care: Development (VISION-D) study (NCT04763746) was to collect RR, PR, and BP measurements simultaneously by using the current SOC manual methods and the Lifelight software to iteratively refine the software algorithms. The objective of the Vital Sign Comparison Between Lifelight and Standard of Care: Validation (VISION-V) study (NCT03998098) was to validate the use of Lifelight software to accurately measure VSs.

Methods

VISION-D

VISION-D was a prospective observational study conducted over 12 months during 2018 and 2019 [26], involving 8585 inpatients, outpatients, and healthy volunteers aged >3 years. There were no exclusion criteria to ensure local representation in age, sex, health condition, and skin tone and the inclusion of a wide range of VS values within and outside normal healthy ranges. The sample size was expected to exceed 2000 volunteers but was not formally prespecified, as it would depend on the incremental improvement in accuracy of the Lifelight system. The study continued until acceptable accuracy was achieved through machine learning. The sponsor kept the study teams informed on progress.

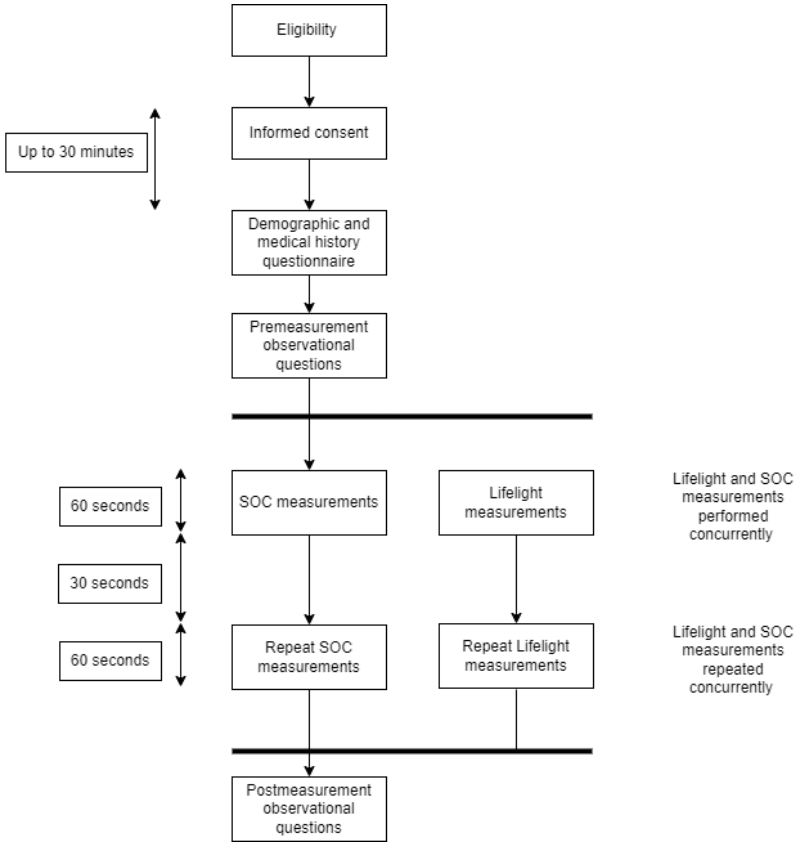
The study was conducted at Queen Alexandra Hospital, Portsmouth Hospitals University National Health Service (NHS) Trust in accordance with Good Clinical Practice and was approved by the HRA (Integrated Research Application System number: 242581). All participants gave written informed consent.

Measurements were taken by trained nursing staff and clinical trial assistants. PR and BP were measured with a standard clinical automatic sphygmomanometer (Welch Allyn Connex Spot Monitor) on one arm, allowing both to be measured simultaneously, rather than also using an electrocardiogram to record PR. RR was determined via the manual counting of observed inspirations over 60 seconds. The Lifelight software was run on a sixth generation Apple iPad, held approximately

1 meter from the participant and angled toward their face. Measurement started and stopped automatically, and the data were sent to a secure database without being displayed (to prevent clinical interpretation or analysis). Two sets of measurements were taken by two staff members during the same 60-second period and then repeated, giving 4 sets in total (Figure 2). Pre- and postmeasurement observations were made of background luminosity, temperature, the use of makeup, and facial features.

Transmitted data were encrypted and stored in a secure database. No identifiable data were stored. Only data for adults (≥18 years) were reported. The data were used to train the software algorithms: the ensemble machine learning algorithm Extra Trees [30] was used for BP, and the filtering of the Fourier-transformed space followed by shape feature counting was used for PR and RR.

Figure 2. Data collection in VISION-D. SOC: standard of care; VISION-D: Vital Sign Comparison Between Lifelight and Standard of Care: Development.

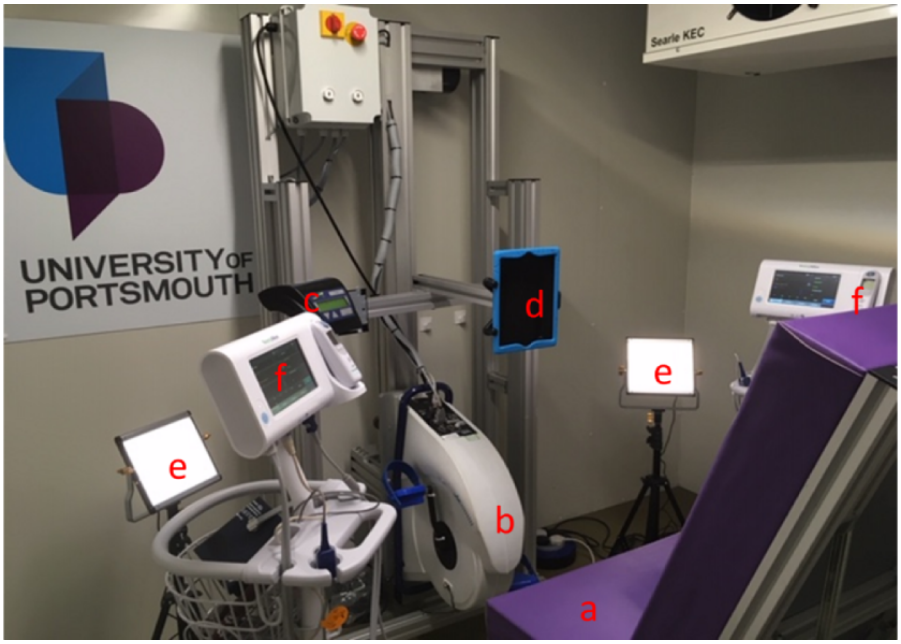


VISION-V

VISION-V (n=127) was conducted at the School of Sport, Health and Exercise Science at the University of Portsmouth, United Kingdom, during 2019. Measurements were performed as in VISION-D but in a normobaric hypoxic chamber (Figure 3). VSs were measured 3 times in each participant by two observers who were blinded to their device readings and to each other's readings. Data collection was overseen by an independent supervisor. The study was conducted in accordance with Good Clinical Practice and approved by the HRA (Integrated Research Application System number: 258187). All participants gave written informed consent.

In addition to standard VSs measurement, healthy participants aged 18-39 years exercised on a recumbent cycle ergometer (maximum intensity 200 W) to generate a wide range of PR and RR values, per the laboratory's standard operating procedure and under the advice of the independent medical officer. The exercise intensity and hypoxic environment were individually titrated to induce ≥80% oxygen desaturation. VSs were measured immediately after each exercise bout.

Figure 3. The environmental chamber used for hypoxic exercise testing and normoxic blood pressure (BP) evaluation: (a) recumbent chair; (b) ergometer; (c) ergometer control panel used for cycle exercise; (d) iPad running Lifelight software; (e) photographic lights to supplement chamber lighting; (f) Welch Allyn Connex Spot Monitors used for measuring BP, pulse rate, and oxygen saturation.



Ethics Approval

Ethical approval for VISION-V was granted by the London-Dulwich Research Ethics Committee (reference 19/LO/0427). The Medicines and Healthcare products Regulatory Agency issued a notice of no objection for the medical device to be used in VISION-V (reference CI/2018/0078). VISION-D was approved by the HRA and Health and Care Research Wales (reference 18-NS-0047). All participants provided written informed consent.

Statistical Analysis

In VISION-D, the enrolled set comprised all recruited participants; the full analysis set (FAS) comprised those for whom VS measurements are included. Reasons for exclusion were an age of <18 years, incorrect or incomplete data entry, physiologically implausible data (determined by the clinical investigator), and low signal quality (pulse signal quality indicator <0.85; eg, because of excessive movement or insufficient light).

To ensure the accuracy of sphygmomanometers over a clinically useful range, the ISO standard for SBP requires that ≥5% of measurements are ≤100 mmHg, ≥5% are ≥160 mmHg, and ≥20% are ≥140 mmHg [28]. For DBP, ≥5% of measurements should each be ≤60 mmHg and ≥100 mmHg, and ≥20% should be ≥85 mmHg [28]. We therefore analyzed similar BP subgroups constructed using data randomly selected from the full data set; the distribution was calculated by up-weighting all SBP/DBP bands not meeting the minimum percentages to become ≥5% or ≥25% of the subgroup as appropriate and down-weighting bands exceeding the minimum percentages.

A subgroup was also created using the Fitzpatrick Skin Type Scale [31], comprising ≥5% each in groups 1 and 4-6 and ≥20% each for groups 2 and 3, with up- and down-weighting as described for the BP subgroup.

The primary analysis in both studies assessed the performance of Lifelight against the SOC measurements; an accuracy target was deemed to be met if mean error and SD for Lifelight measurements at least equaled the target (Table 1). Heat maps were generated for the VISION-D data, as the large amount of data rendered a scatter plot unclear. Scatter plots were developed for the smaller VISION-V data set (which was insufficient for a heat map).

Linear regression was used to assess the impact of skin tone on the accuracy of Lifelight for measuring each VS, using the Fitzpatrick skin tones as the exploratory variable.

Results

VISION-D

The enrolled set comprised 8585 participants; 60%-67% were included in individual VS analyses, and 17,233 measurements were collected, of which 43%-59% were included in the individual VS analyses (FAS). Demographic details are provided in Table 2. There were no protocol deviations or adverse events.

The performance targets were met for all measurements except SBP in the FAS and the BP subgroup (Table 3). Heat maps of the reference method (SOC manual measurement) versus the test measurement are shown in Figure 4. Values for RR fell within a narrow range, distorting the appearance of the heat map. Amplifying the proportion of DBP data at extreme values slightly reduced the accuracy whereas amplifying the proportion of SBP data at extreme values had little effect (analysis not shown).

Table 2. Characteristics of the final analysis set in VISION-D[a].

Characteristic	Participants[b]	Value, range	Value, mean (SD)
Sex		N/A[c]	N/A
Female, n (%)	5649 (65.8)	N/A[c]	N/A
Male, n (%)	2936 (34.2)	N/A	N/A
Age (years), n	8585	4-96	49.7 (17.1)
Fitzpatrick skin tone, n (%)[d]		N/A	N/A
1			
RR[e]	197 (3.44)		
PR[f]	170 (3.30)		
SBP[g]	158 (2.98)		
DBP[h]	149 (2.89)		
2			
RR	2967 (51.81)		
PR	2646 (51.41)		
SBP	2733 (51.54)		
DBP	2663 (51.69)		
3			
RR	2292 (40.02)		
PR	2068 (40.18)		
SBP	2162 (40.77)		
DBP	2101 (40.78)		
4			
RR	189 (3.30)		
PR	185 (3.59)		
SBP	175 (3.30)		
DBP	169 (3.28)		
5			
RR	28 (0.49)		
PR	29 (0.56)		
SBP	23 (0.43)		
DBP	22 (0.43)		
6			
RR	5 (0.09)		
PR	2 (0.04)		
SBP	5 (0.09)		
DBP	5 (0.10)		
Unassigned			
RR	49 (0.86)		
PR	47 (0.91)		
SBP	50 (0.94)		
DBP	43 (0.83)		

Characteristic	Participants[b]	Value, range	Value, mean (SD)
PR (beats per minute)	5727	32-183	73.6 (12.2)
RR (respirations per minute)	5147	8-23	16.1 (2.8)
SBP (mmHg)	5303	71-223	130.7 (19.4)
DBP (mmHg)	5152	46-136	79.9 (9.2)

[a]VISION-D: Vital Sign Comparison Between Lifelight and Standard of Care: Development.

[b]Values for sex and age data are for the enrolled set; data for vital signs are for the full analysis set; only data from adults (>18 years) were analyzed.

[c]N/A: not applicable.

[d]Not all measurements contributed to each vital sign (eg, not everyone with a Fitzpatrick skin tone of 1 had respiratory rate, pulse rate, systolic blood pressure, and diastolic blood pressure analyzed because of exclusions, such as poor signal quality).

[e]RR: respiratory rate.

[f]PR: pulse rate.

[g]SBP: systolic blood pressure.

[h]DBP: diastolic blood pressure.

Table 3. Performance of Lifelight in VISION-D[a].

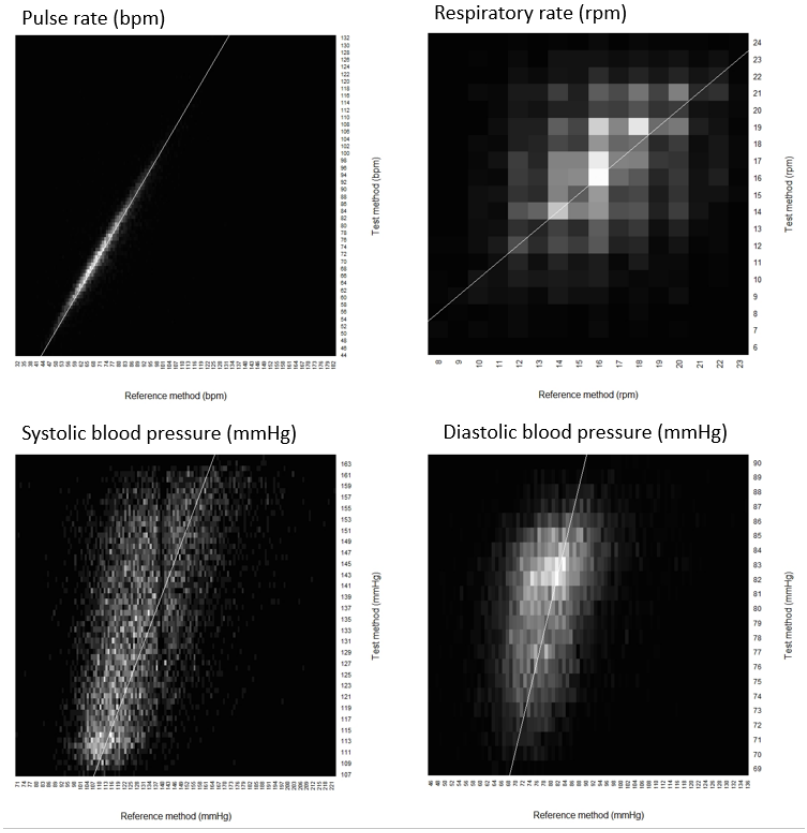
	Measurements (n=17,233), n (%)	Accuracy, mean error (SD)		
		Target	Full analysis set	BP[b] subgroup
Pulse rate (beats per minute)	10,214 (59)	2.2 (9.2)	0.3 (4.0)	N/A[c]
Respiratory rate (respirations per minute)	7462 (43)	2.3 (5.0)	0.3 (3.6)	N/A
Systolic BP (mmHg)	9233 (54)	6.7 (15.3)	3.5 (16.8)	4.1 (17.0)
Diastolic BP (mmHg)	8951 (52)	5.5 (8.9)	−0.4 (8.5)	−1.0 (10.0)

[a]VISION-D: Vital Sign Comparison Between Lifelight and Standard of Care: Development.

[b]BP: blood pressure.

[c]N/A: not applicable.

Figure 4. Heat maps for correlation between Lifelight (test methods) and standard of care (reference) measurements in VISION-D. The density of white points illustrates the extent of overlap. Correlation coefficients (R^2) were 0.89 for pulse rate, 0.11 for respiratory rate, 0.30 for systolic blood pressure, and 0.15 for diastolic blood pressure. The line of identity (y=x) illustrates when the Lifelight (test) measurement provided was the same as the reference. bpm: beats per minute; VISION-D: Vital Sign Comparison Between Lifelight and Standard of Care: Development.



For the Fitzpatrick subgroup, the performance targets were met for PR, RR, and DBP; for SBP, the standard was met for mean error but not SD (Table 4). The regression analysis for skin tone showed only small changes in error between one Fitzpatrick group and the next, with similar changes in error for the FAS and Fitzpatrick subgroup (Table 5).

As data accumulated, signal processing was used to improve the accuracy of PR and RR measurement and machine learning for BP. The SBP SD decreased from 22 to 14 mmHg over the 12-month study. The proportion of measurements ≤5 mmHg of the reference doubled from 15% to 30%, and the percentage of measurements ≤10 mmHg increased from 30% to 50%.

Table 4. Performance of Lifelight in VISION-D[a] in the Fitzpatrick subgroup.[b]

	Eligible measurements, n	Target, mean error (SD)	Accuracy, mean error (SD)
Pulse rate (beats per minute)	6700	2.2 (9.2)	0.3 (4.0)
Respiratory rate (respirations per minute)	4520	2.3 (5.0)	0.4 (3.7)
Systolic blood pressure (mmHg)	5152	6.7 (15.3)	3.6 (16.6)
Diastolic blood pressure (mmHg)	4960	5.5 (8.9)	−0.4 (8.5)

[a]VISION-D: Vital Sign Comparison Between Lifelight and Standard of Care: Development.

[b]The subgroup comprised ≥5% each in groups 1 and 4-6 and ≥20% each for groups 2 and 3.

Table 5. Regression analysis for skin tone.

	Full analysis set		Fitzpatrick subgroup	
	Measurements, n	Change in error[a]	Measurements, n	Change in error
Pulse rate (beats per minute)	10,131	0.45	6700	0.46
Respiratory rate (respirations per minute)	7391	0.11	4520	0.13
Systolic blood pressure (mmHg)	9148	0.5	5152	1.1
Diastolic blood pressure (mmHg)	8870	−0.8	4960	−0.8

[a]Values are the change in error from one Fitzpatrick group to the next in the full analysis set and the Fitpatrick subgroup, comprising ≥5% each in groups 1 and 4-6 and ≥20% each for groups 2 and 3.

VISION-V

Characteristics of the FAS (n=125) are presented in Table 6. There were no protocol deviations or adverse events. For the different VSs, 61%-83% of measurements were eligible for the performance analysis (Table 7). The scatter plots showed good correlations between Lifelight and SOC measurements of VSs (Figure 5). The performance targets for the FAS were met for all VSs (Table 8).

Table 6. Reference data for the full analysis set in VISION-V[a] (n=125 participants).

Characteristic	Value	Value, mean (SD)
Sex, n (%)		N/A[b]
Male	55 (44)	
Female	70 (56)	
Age (years), range	18-66	30.2 (11.6)
Pulse rate (beats per minute), range	47-127	79.4 (15.2)
Respiratory rate (respirations per minute), range	6-27	14.8 (3.9)
Systolic blood pressure (mmHg), range	96-176.5	122.9 (15.9)
Diastolic blood pressure (mmHg), range	62-109.5	77.8 (7.8)

[a]VISION-V: Vital Sign Comparison Between Lifelight and Standard of Care: Validation.

[b]N/A: not applicable.

Table 7. Eligible measurements in VISION-V[a].

	Participants included in analyses, n	Eligible measurements[b], n (%)
Pulse rate	72	308 (83)
Respiratory rate	67	143 (61)
Systolic blood pressure	115	115 (69)
Diastolic blood pressure	113	113 (64)

[a]VISION-V: Vital Sign Comparison Between Lifelight and Standard of Care: Validation.

[b]Complete measurement sets where the photoplethysmography signal quality was adequate to measure vital signs using Lifelight. Lifelight measurements were eligible for the analysis if the photoplethysmography signal quality was ≥0.85 and the measurement set was complete (one set each for pulse rate, systolic blood pressure, and diastolic blood pressure.

Figure 5. Scatter plots with correlation lines for the vital signs measured with Lifelight (test) versus standard of care (reference) method. Correlation coefficients (R2) were 0.94 for pulse rate (PR), 0.30 for respiratory rate (RR), 0.21 for systolic blood pressure (SBP), and 0.17 for diastolic blood pressure (DBP); 95% Bland-Altman limits of agreement were −5.9 and 8.8 bpm for PR, −6.9 and 6.6 for RR, −25.6 and 31.2 for SBP, and −14.1 and 13.4 for DBP. bpm: beats per minute; rpm: respirations per minute.

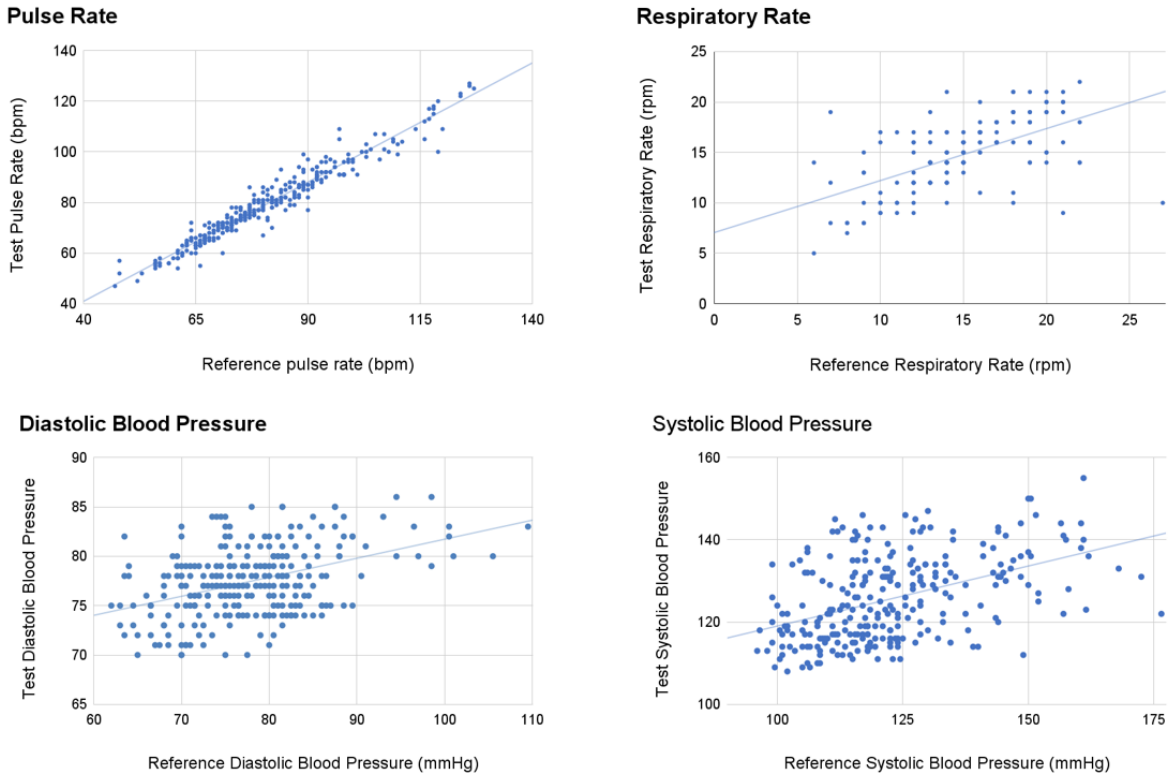


Table 8. Accuracy of Lifelight in VISION-V[a].

	Target accuracy, mean error (SD)	Accuracy in VISION-V, mean error (SD)
Pulse rate (beats per minute)	2.2 (9.2)	1.4 (3.8)
Respiratory rate (respirations per minute)	2.3 (5.0)	−0.1 (3.4)
Systolic blood pressure (mmHg)	6.7 (15.3)	2.8 (14.5)
Diastolic blood pressure (mmHg)	5.5 (8.9)	–0.3 (7.0)

[a]VISION-V: Vital Sign Comparison Between Lifelight and Standard of Care: Validation.

Discussion

Principal Results

VISION-D and VISION-V demonstrate the accuracy of the Lifelight software in the simultaneous contactless measurement of VSs, based on more than 17,000 measurements. The predefined performance targets were met for PR, RR, and DBP in VISION-D; for SBP, the mean error was met but not SD (Table 3). All targets were met in VISION-V (Table 8). On the basis of these data, Lifelight achieved Level 1 CE mark certification as a medical device [14].

The use of mobile devices for measurement of VSs presents some challenges compared with controlled laboratory scenarios, for example [32]. To mitigate some of these challenges, we have compared VSs that were measured simultaneously by using SOC methods and Lifelight. We believe VISION-D to be the largest study to date to measure VSs using rPPG. As there are currently no standards for contactless measurement of VSs, we developed performance targets in discussion with the HRA and a CE-marking auditor (Table 1). The HRA has accepted these targets for the ongoing VISION-Acute study (NCT04589923), and the targets can therefore be considered applicable to the VISION-V and VISION-D studies. The software algorithms were refined continuously by using data collected during VISION-D, and the final algorithms were used in VISION-V. The accuracy targets (set before data analysis) were met in VISION-V. Although participants in VISION-V had a wide range of VS values, the ISO distribution criteria for BP were not met, likely because the participants were from a healthy population (few had hypotension or hypertension). However, amplifying the proportion of participants with high/low BP in VISION-D did not affect the accuracy of SBP measurement, but the SD for the DBP performance target was no longer met.

Although the accuracy targets were met for RR, values recorded by Lifelight were 10 to 20 respirations per minute (rpm), whereas the reference values were 5 to 22 rpm, indicating that

there may be some loss of accuracy at the slower rates. An RR above 22 rpm is clinically important but was not captured in the VISION-V and VISION-D studies, likely because the participants were mostly healthy. This is being addressed in both the VISION-Acute and the VISION: Multisite Development (VISION-MD) studies (NCT04589923 and NCT04763746, respectively), which are enrolling a broader range of patients with VS values outside the normal range, including some who are critically ill, to improve the accuracy of Lifelight for clinical use.

Our substantial database from VISION-D includes medical history, temperature, light (lux meter), Fitzpatrick skin tone, facial tattoos, birthmarks, facial hair, etc, which can be used to explore potential interference factors (in contrast to the Medical Information Mart for Intensive Care PPG database of patients who are critically ill [33]).

Limitations and Future Work

Studies in 2018 and 2019 demonstrated the potential of PPG to detect changes in cardiovascular activity and the measurement of BP [34-36]; a recent study claims to meet the ISO standards for BP measurement (ISO 81060-2), based on 225 measurements in 85 volunteers [37]. The accuracy in VISION-V was also within the ISO 81060-2 standard (5 ± 8 mmHg) for DBP, and it was within the mean error but not within the SD for SBP, although this ISO relates to the cuff-based measurement of BP. In addition, based on the mean error in VISION-D, the performance of Lifelight was comparable to that reported in the literature for most of the devices on which the standards were based. BP is inherently more complex to measure than PR and RR, in terms of the data form and machine learning and because reference measurements are less accurate.

As with any recording device, signal quality may be compromised if the participant moves excessively or light levels are insufficient. The proportion of eligible measurements ranged from 61% for RR to 83% for PR. Ineligible measurements were largely due to the inadequate quality (blurring) of the video recordings. Higher-resolution video recording is being used in the current VISION studies (described in more detail below), which is expected to provide a cleaner and more robust signal. However, Lifelight is easy to use, and measurements can be repeated within a minute in the event of a poor signal.

Skin type is a potential source of error with PPG devices, as melanin absorbs green light, potentially increasing errors in measurements in dark-skinned individuals compared with light-skinned individuals [38]. However, skin type does not affect the accuracy of Lifelight: the performance targets were met for RR, PR, and DBP in the Fitzpatrick subgroup, and mean error was met for SBP but not SD. Moreover, amplifying the proportion of participants with light and dark tones did not affect accuracy. Bent and colleagues [38] also reported that Fitzpatrick skin type had no significant effect on the accuracy of PR measurements by wearable optical heart rate sensors; however, this was a small study (n=53). Although the Fitzpatrick Skin Type Scale is the current gold standard [31], its use has been criticized because of racial bias, weak correlation with skin color, and broad within-group variations in skin tone. Spectrocolorimetry, which uses multiple variables to categorize skin tone objectively, has been proposed as an alternative [31], which may be incorporated into later studies to confirm our findings.

The accuracy of the Lifelight algorithms will further improve with continuing data collection. The ongoing VISION-MD study is collecting data from a wide range of participants, including patients who are critically ill, which will be used for algorithm development and then testing. Higher-resolution video data are being collected in this study, and the algorithms are focusing on smaller but higher-quality regions of interest.

Comparison With Prior Work

The use of rPPG offers several advantages in addition to the rapid and contactless measurement of RR, PR, and BP simultaneously. There is no need for calibration, servicing, cleaning several pieces of equipment, or specialist training. Such advantages are particularly useful in residential care. Indeed, Lifelight has been piloted with the Hampshire Hospitals NHS Foundation Trust as part of a telemedicine service during the COVID-19 pandemic [39]. Care teams found the software easy to use and care was improved, as residents did not need to travel and VSs could be recorded easily by a known carer; clinicians' travelling time was also reduced. Another study of remote VS monitoring in residential care reported that 87% of emergency department attendances were avoided [40]. Less tangible but valuable benefits include reduced anxiety among staff and residents, particularly the fear of hospitalization [40].

Notably, RR is often missed from VS monitoring or measured inaccurately [41], but changes in RR can be a harbinger of physiological conditions such as hypoxia, hypercapnia, and acidosis [42]. During the COVID-19 pandemic, patients at risk in England were provided with pulse oximeters, with instructions to seek medical help if oxygen saturation fell below 92% [43]. However, changes to RR indicate increased ventilation and precede reductions in oxygen saturation [41], thus giving an earlier indication of clinical deterioration. A PPG device to record RR would therefore be invaluable in this situation. The COVID-19 pandemic also highlighted the importance of contactless VS measurement [13].

Conclusion

This preliminary evaluation of Lifelight demonstrates sufficient accuracy in the measurement of VSs to support Level 1 CE mark certification, with further work ongoing to develop Lifelight into a robust method for measurement of VSs in daily clinical use. As the use of Lifelight does not require specific training or equipment, the software is potentially useful for the contactless measurement of VSs by nonclinical staff in residential and home care settings.

Acknowledgments

Xim Ltd won a Small Business Research Initiative (SBRI) Phase II contract to conduct Vital Sign Comparison Between Lifelight and Standard of Care: Development (VISION-D) and an Innovate UK SMART grant to fund Vital Sign Comparison Between Lifelight and Standard of Care: Validation. VISION-D was supported by SBRI Healthcare, a National Health Service England and National Health Service Improvement initiative supported by the Academic Health Science Networks.

The authors would like to thank all the volunteers who took part in the two studies. The Lifelight Trials Group includes Sharon Allard, Dr Mark Lyons-Amos, Bethany Armstead, Rosalynn Austin, Rebecca Baker, Dr Michelle Baker Moffat, Armida Balawon, Debbi Barnes, Sonia Baryschpolec, Sean Beech, Selina Begum, Lauren Bell, Helen Blackman, Marie Broadway, Kate Burrows, Philippa Copnall, Zoe Daly, Joanne Dash, Mini David, Teresa Day, Jacqueline Denham, Rodrigo Dias, Alison Dimmer, Gemma Dixon, Tracey Dobson, Catherine Edwards, Carole Fogg, Dr Jim Forrer, Francis Galera, Zoe Garner, Andrew Gribbin, Elizabeth Hawes, Serena Howe, Karen Hudson, Amanda Hungate, Victoria Hunter, Jo Kerr, Adam Kiddle, Arjun Kumar, Shanqin Liu, Beverley Longhurst, Sharon McCready, Shoid Miah, Maria Moon, Kirsty Parker, Gina Pelletier, Connie Petronzio, David Petronzio, Michelle Pople, Benildo Jr Quiros, Deidre Rodgers, Dr Mike Sadler, Kerrie Scott, Josh Sephton, Samantha Smith, Bruce Stanley, Nina Szarazova, Nick Thorne, Monika Thwaites, Sarah Tronk, Catherine Tuffrey, Marcus Tuke, Charlotte Turner, James Turner, Lewis Valaitis, Dr Lieke van Putten, Lyn Vinall, Marie White, Melanie Willcox, Jonathon Winter, Carole Wragg, and Kim Wren.

Authors' Contributions

The Vital Sign Comparison Between Lifelight and Standard of Care: Validation (VISION-V) protocol was designed by The Clinical Trials Company. The Vital Sign Comparison Between Lifelight and Standard of Care: Development (VISION-D) protocol was designed by MK, TJ, EH, and AC. Both protocols were approved by LP. Data collection was led by the University of Portsmouth for VISION-V (ML and HM) and by Portsmouth Hospitals National Health Service Trust for VISION-D (TJ, EH, MC, LW, TB, and AC). Data analysis and interpretation for both studies were performed by Mind Over Matter Medtech Ltd (ABM and MK), with support from Xim Ltd (JH, SW, and LP), and with support from MS and SR for VISION-D. The paper was developed and revised by HB, and it has been extensively reviewed and approved by all authors.

Conflicts of Interest

LP is a shareholder of XIM Ltd.

References

1. Buist MD, Jarmolowski E, Burton PR, Bernard SA, Waxman BP, Anderson J. Recognising clinical instability in hospital patients before cardiac arrest or unplanned admission to intensive care. A pilot study in a tertiary-care hospital. Med J Aust 1999 Jul 05;171(1):22-25. [doi: 10.5694/j.1326-5377.1999.tb123492.x] [Medline: 10451667]
2. Brekke IJ, Puntervoll LH, Pedersen PB, Kellett J, Brabrand M. The value of vital sign trends in predicting and monitoring clinical deterioration: a systematic review. PLoS One 2019;14(1):e0210875 [FREE Full text] [doi: 10.1371/journal.pone.0210875] [Medline: 30645637]
3. NEWS frequently asked questions. NHS England. 2021. URL: https://www.england.nhs.uk/ourwork/clinical-policy/sepsis/nationalearlywarningscore/news-frequently-asked-questions [accessed 2022-09-04]
4. West Hampshire CCG and Wessex Patient Safety Collaborative. RESTORE2. Yorkshire and Humber AHSN. 2021. URL: https://www.yhahsn.org.uk/wp-content/uploads/2020/05/RESTORE2-Training-Slides-4_3-WPSC-200401.03-1586963980-1.pdf [accessed 2022-09-04]
5. Panday RSN, Subbe CP, van Galen LS, Kellett J, Brabrand M, Nickel CH, et al. Changes in vital signs post discharge as a potential target for intervention to avoid readmission. Acute Med 2018;17(2):77-82. [Medline: 29882557]
6. Flacco ME, Manzoli L, Bucci M, Capasso L, Comparcini D, Simonetti V, et al. Uneven accuracy of home blood pressure measurement: a multicentric survey. J Clin Hypertens (Greenwich) 2015 Aug;17(8):638-643. [doi: 10.1111/jch.12552] [Medline: 25880129]
7. Johansson A, Oberg PA, Sedin G. Monitoring of heart and respiratory rates in newborn infants using a new photoplethysmographic technique. J Clin Monit Comput 1999 Dec;15(7-8):461-467. [doi: 10.1023/a:1009912831366] [Medline: 12578044]
8. Poh M, Poh YC. Validation of a standalone smartphone application for measuring heart rate using imaging photoplethysmography. Telemed J E Health 2017 Aug;23(8):678-683. [doi: 10.1089/tmj.2016.0230] [Medline: 28140834]
9. Aoyagi T, Miyasaka K. Pulse oximetry: its invention, contribution to medicine, and future tasks. Anesth Analg 2002 Jan;94(1 Suppl):S1-S3. [Medline: 11900029]
10. Elgendi M, Fletcher R, Liang Y, Howard N, Lovell NH, Abbott D, et al. The use of photoplethysmography for assessing hypertension. NPJ Digit Med 2019;2:60. [doi: 10.1038/s41746-019-0136-7] [Medline: 31388564]

11. Radha M, de Groot K, Rajani N, Wong CCP, Kobold N, Vos V, et al. Estimating blood pressure trends and the nocturnal dip from photoplethysmography. Physiol Meas 2019 Feb 26;40(2):025006. [doi: 10.1088/1361-6579/ab030e] [Medline: 30699397]
12. Nilsson L, Johansson A, Kalman S. Monitoring of respiratory rate in postoperative care using a new photoplethysmographic technique. J Clin Monit Comput 2000;16(4):309-315. [doi: 10.1023/a:1011424732717] [Medline: 12578078]
13. Kumar SS, Dashtipour K, Abbasi QH, Imran MA, Ahmad W. A review on wearable and contactless sensing for COVID-19 with policy challenges. Front Comm Net 2021 May 19;2:636293. [doi: 10.3389/frcmn.2021.636293]
14. Lifelight First for monitoring vital signs. National Institute for Health and Care Excellence. 2020 Apr 08. URL: https://www.nice.org.uk/advice/mib213/chapter/The-technology [accessed 2022-09-14]
15. Breteler MJM, KleinJan EJ, Dohmen DAJ, Leenen LPH, van Hillegersberg R, Ruurda JP, et al. Vital signs monitoring with wearable sensors in high-risk surgical patients: a clinical validation study. Anesthesiology 2020 Mar;132(3):424-439 [FREE Full text] [doi: 10.1097/ALN.0000000000003029] [Medline: 31743149]
16. Breteler MJM, Huizinga E, van Loon K, Leenen LPH, Dohmen DAJ, Kalkman CJ, et al. Reliability of wireless monitoring using a wearable patch sensor in high-risk surgical patients at a step-down unit in the Netherlands: a clinical validation study. BMJ Open 2018 Feb 27;8(2):e020162 [FREE Full text] [doi: 10.1136/bmjopen-2017-020162] [Medline: 29487076]
17. Murali S, Rincon F, Cassina T, Cook S, Goy JJ. Heart rate and oxygen saturation monitoring with a new wearable wireless device in the intensive care unit: pilot comparison trial. JMIR Biomed Eng 2020 Apr 22;5(1):e18158. [doi: 10.2196/18158]
18. Lee PJ. Clinical evaluation of a novel respiratory rate monitor. J Clin Monit Comput 2016 Apr;30(2):175-183. [doi: 10.1007/s10877-015-9697-4] [Medline: 25900144]
19. Li T, Divatia S, McKittrick J, Moss J, Hijnen NM, Becker LB. A pilot study of respiratory rate derived from a wearable biosensor compared with capnography in emergency department patients. Open Access Emerg Med 2019;11:103-108. [doi: 10.2147/OAEM.S198842] [Medline: 31123424]
20. Alfano G, Fontana F, Cappelli G. Noninvasive blood pressure measurement in maintenance hemodialysis patients: comparison of agreement between oscillometric and finger-cuff methods. Nephron 2017;136(4):309-317. [doi: 10.1159/000457812] [Medline: 28468005]
21. Allan PD, O'Donnell T, Tzeng Y. Agreement between finger plethysmography- and brachial oscillometry-derived blood pressure measurements. Clin Physiol Funct Imaging 2018 May;38(3):439-446. [doi: 10.1111/cpf.12435] [Medline: 28444896]
22. Becker S, Spiesshoefer J, Brix T, Tuleta I, Mohr M, Emdin M, et al. Validity of transit time-based blood pressure measurements in patients with and without heart failure or pulmonary arterial hypertension across different breathing maneuvers. Sleep Breath 2020 Mar;24(1):221-230. [doi: 10.1007/s11325-019-01848-w] [Medline: 31049851]
23. Gratz I, Deal E, Spitz F, Baruch M, Allen IE, Seaman JE, et al. Continuous non-invasive finger cuff CareTaker® comparable to invasive intra-arterial pressure in patients undergoing major intra-abdominal surgery. BMC Anesthesiol 2017 Mar 21;17(1):48 [FREE Full text] [doi: 10.1186/s12871-017-0337-z] [Medline: 28327093]
24. Tanioku T, Yoshida A, Aratani Y, Fujii K, Kawamata T. Validation of noninvasive continuous arterial pressure measurement by ClearSight System™ during induction of anesthesia for cardiovascular surgery. BMC Anesthesiol 2020 Jul 20;20(1):176 [FREE Full text] [doi: 10.1186/s12871-020-01091-x] [Medline: 32690040]
25. Wang Z, Chen G, Lu K, Zhu Y, Chen Y. Investigation of the accuracy of a noninvasive continuous blood pressure device in different age groups and its ability in detecting hypertension and hypotension: an observational study. BMC Anesthesiol 2019 Dec 05;19(1):223 [FREE Full text] [doi: 10.1186/s12871-019-0899-z] [Medline: 31805877]
26. Jones TL, Heiden E, Mitchell F, Fogg C, McCready S, Pearce L, et al. Developing the accuracy of vital sign measurements using the Lifelight software application in comparison to standard of care methods: observational study protocol. JMIR Res Protoc 2021 Jan 28;10(1):e14326 [FREE Full text] [doi: 10.2196/14326] [Medline: 33507157]
27. BS EN ISO 80601-2-61 - Medical electrical equipment. Particular requirements for basic safety and essential performance of pulse oximeter equipment. The British Standards Institution. 2019. URL: https://landingpage.bsigroup.com/LandingPage/Undated?UPI=000000000030316719 [accessed 2022-09-15]
28. ISO 81060-2:2018 - Non-invasive sphygmomanometers — Part 2: Clinical investigation of intermittent automated measurement type. International Organization for Standardization. 2019. URL: https://www.iso.org/standard/73339.html [accessed 2022-09-09]
29. Brookes CN, Whittaker JD, Moulton C, Dodds D. The PEP respiratory monitor: a validation study. Emerg Med J 2003 Jul;20(4):326-328 [FREE Full text] [doi: 10.1136/emj.20.4.326] [Medline: 12835340]
30. Geurts P, Ernst D, Wehenkel L. Extremely randomized trees. Mach Learn 2006 Mar 2;63(1):3-42. [doi: 10.1007/s10994-006-6226-1]
31. Colvonen PJ. Response to: investigating sources of inaccuracy in wearable optical heart rate sensors. NPJ Digit Med 2021 Feb 26;4(1):38. [doi: 10.1038/s41746-021-00408-5] [Medline: 33637822]
32. Casalino G, Castellano G, Zaza G. Evaluating the robustness of a contact-less mHealth solution for personal and remote monitoring of blood oxygen saturation. J Ambient Intell Humaniz Comput 2022 Jan 14:1-10 [FREE Full text] [doi: 10.1007/s12652-021-03635-6] [Medline: 35043065]
33. Johnson AEW, Pollard TJ, Shen L, Lehman LH, Feng M, Ghassemi M, et al. MIMIC-III, a freely accessible critical care database. Sci Data 2016 May 24;3:160035. [doi: 10.1038/sdata.2016.35] [Medline: 27219127]

34. Barszczyk A, Lee K. Measuring blood pressure: from cuff to smartphone. Curr Hypertens Rep 2019 Oct 10;21(11):84. [doi: 10.1007/s11906-019-0990-3] [Medline: 31599362]
35. Liu J, Luo H, Zheng PP, Wu SJ, Lee K. Transdermal optical imaging revealed different spatiotemporal patterns of facial cardiovascular activities. Sci Rep 2018 Jul 12;8(1):10588. [doi: 10.1038/s41598-018-28804-0] [Medline: 30002447]
36. Wei J, Luo H, Wu SJ, Zheng PP, Fu G, Lee K. Transdermal optical imaging reveal basal stress via heart rate variability analysis: a novel methodology comparable to electrocardiography. Front Psychol 2018;9:98. [doi: 10.3389/fpsyg.2018.00098] [Medline: 29472879]
37. Yang D, Xiao G, Wei J, Luo H. Preliminary assessment of video-based blood pressure measurement according to ANSI/AAMI/ISO81060-2: 2013 guideline accuracy criteria: Anura smartphone app with transdermal optimal imaging technology. Blood Press Monit 2020 Oct;25(5):295-298. [doi: 10.1097/MBP.0000000000000467] [Medline: 32842022]
38. Bent B, Goldstein BA, Kibbe WA, Dunn JP. Investigating sources of inaccuracy in wearable optical heart rate sensors. NPJ Digit Med 2020;3:18. [doi: 10.1038/s41746-020-0226-6] [Medline: 32047863]
39. Remote monitoring of vital signs. NHS Transformation Directorate. 2020 Oct 08. URL: https://www.nhsx.nhs.uk/ai-lab/explore-all-resources/understand-ai/remote-monitoring-vital-signs/ [accessed 2022-09-04]
40. Implementing remote monitoring in care homes: 3 lessons learned. Current Health. 2021. URL: https://www.currenthealth.com/resources/blog/implementing-remote-monitoring-in-care-homes/ [accessed 2022-09-04]
41. Kelly C. Respiratory rate 1: why measurement and recording are crucial. Nurs Times 2018;114:23-24 [FREE Full text]
42. Rolfe S. The importance of respiratory rate monitoring. Br J Nurs 2019 Apr 25;28(8):504-508. [doi: 10.12968/bjon.2019.28.8.504] [Medline: 31002547]
43. COVID-19 guidance note: COVID oximetry @home. NHS England. 2020. URL: https://www.england.nhs.uk/coronavirus/publication/novel-coronavirus-covid-19-standard-operating-procedure-covid-oximetry-home/ [accessed 2022-09-04]

Abbreviations

BP: blood pressure
CE: Conformité Européenne
DBP: diastolic blood pressure
FAS: full analysis set
HRA: Health Research Authority
ISO: International Organization for Standardization
NEWS: National Early Warning Score
NHS: National Health Service
PPG: photoplethysmography
PR: pulse rate
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RESTORE2: Recognise Early Soft Signs, Take Observations, Respond, Escalate
rpm: respirations per minute
rPPG: remote photoplethysmography
RR: respiratory rate
SBP: systolic blood pressure
SOC: standard of care
VISION: Vital Sign Comparison Between Lifelight and Standard of Care
VISION-D: Vital Sign Comparison Between Lifelight and Standard of Care: Development
VISION-MD: Vital Sign Comparison Between Lifelight and Standard of Care: Multisite Development
VISION-V: Vital Sign Comparison Between Lifelight and Standard of Care: Validation
VS: vital sign

Edited by A Mavragani; submitted 24.01.22; peer-reviewed by J Kellett, N Georgi; comments to author 11.05.22; revised version received 29.06.22; accepted 29.06.22; published 14.11.22.

Please cite as:
Heiden E, Jones T, Brogaard Maczka A, Kapoor M, Chauhan M, Wiffen L, Barham H, Holland J, Saxena M, Wegerif S, Brown T, Lomax M, Massey H, Rostami S, Pearce L, Chauhan A
Measurement of Vital Signs Using Lifelight Remote Photoplethysmography: Results of the VISION-D and VISION-V Observational Studies
JMIR Form Res 2022;6(11):e36340
URL: https://formative.jmir.org/2022/11/e36340
doi:10.2196/36340
PMID:36374541

Original Paper

Acceptability to and Engagement With a Virtual Sickle Cell Trait Education Program (SCTaware): Single-Center Prospective Study

Chase M Beeman[1], BS; Mary Ann Abrams[1,2], MPH, MD; Kristin N Zajo[3], MS; Joseph Stanek[3,4], MS; Sarah H O'Brien[1,3], MD; Peter Chan[5], PhD; Yvette Shen[5], MFA; Ben McCorkle[6], PhD; Latrice Johnson[1], BA; Deena Chisolm[1], PhD; Toyetta Barnard-Kirk[2], MSW; John D Mahan[2], MD; Alexandra Martinez-Mendez[1]; Whitney L Phillips[1], BA; Susan E Creary[1,3], MSc, MD

[1]Center for Child Health Equity and Outcomes Research, Nationwide Children's Hospital, The Ohio State University, Columbus, OH, United States

[2]Primary Care Pediatrics, Nationwide Children's Hospital, Columbus, OH, United States

[3]Division of Hematology, Oncology & Bone and Marrow Transplant, Nationwide Children's Hospital, Columbus, OH, United States

[4]Biostatistics Resource, Nationwide Children's Hospital, Columbus, OH, United States

[5]Department of Design, The Ohio State University, Columbus, OH, United States

[6]Department of English, The Ohio State University, Columbus, OH, United States

Corresponding Author:
Susan E Creary, MSc, MD
Center for Child Health Equity and Outcomes Research
Nationwide Children's Hospital
The Ohio State University
700 Children’s Dr
Columbus, OH, 43205
United States
Phone: 1 614 722 3569
Fax: 1 614 722 3544
Email: susan.creary@nationwidechildrens.org

Abstract

Background: Public health programs are tasked with educating the community on health topics, but it is unclear whether these programs are acceptable to learners. Currently, these programs are delivered via a variety of platforms including in-person, virtually, and over the telephone. Sickle cell trait (SCT) education for parents of children with this trait is one of many education programs provided by the Ohio Department of Health. The novel SCTaware videoconference education program was developed by a research team after central Ohio’s standard program transitioned from in-person to telephone-only education during the COVID-19 pandemic.

Objective: Our objectives were to investigate the acceptability of the format and engagement with the SCTaware education and assess parental worry about having a child with SCT before and after receiving SCTaware.

Methods: This was a single-center, prospective study of English-speaking parents of children <3 years of age identified to have hemoglobin S trait by newborn screening. Parents who *previously* received SCT education by telephone, were able to be contacted, and had access to an electronic device capable of videoconferencing were eligible to complete surveys *after receiving the virtual SCTaware education program*. The SCTaware educator also completed a survey to assess participant engagement. Data were summarized descriptively and a McNemar test was used to compare parental worry before and after receiving SCTaware.

Results: In total, 55 participants completed follow-up surveys after receiving standard SCT telephone education and then completing SCTaware. Most (n=51) participants reported that the SCTaware content and visuals were very easy to understand (n=47) and facilitated conversation with the educator (n=42). All of them said the visuals were respectful and trustworthy, helped them understand content better, and that their questions were addressed. Nearly two-thirds (62%, n=34) reported that the pictures appeared very personal and applied to them. The educator noted most participants (n=45) were engaged and asked questions despite having to manage distractions during their education sessions. Many participants (n=33) reported some level of worry following telephone-only education; this was significantly reduced after receiving SCTaware (P<.001).

Conclusions: Our results suggest that SCTaware is acceptable and engaging to parents. While telephone education may make SCT education more accessible, these findings suggest that many parents experience significant worry about their child with SCT after these sessions. A study to evaluate SCTaware's effectiveness at closing parents' SCT knowledge gaps is ongoing.

(JMIR Form Res 2022;6(11):e38780) doi:10.2196/38780

KEYWORDS

virtual education; remote education; internet-based; health education; hematology; patient education; sickle cell; genetic; child; parenting; sickle cell trait; public health education; acceptability; Hemoglobin S-trait; screening; newborn; eHealth; digital health; telemedicine; telehealth

Introduction

Public health programs are tasked with educating communities on health topics such as diabetes, asthma, and heart disease. The Ohio Department of Health (ODH) oversees many such education programs, the format of which can vary from in-person to telephone to virtual. One example is the sickle cell trait (SCT) education program for parents of newborns identified as having hemoglobin S trait through newborn screening (NBS) [1].

Nearly 3 million people in the United States have SCT, and approximately 2000 infants are born annually with sickle cell disease (SCD) [2], a chronic blood disorder that can lead to pain, stroke, and early mortality [3]. Individuals with SCT are typically asymptomatic, but to make informed reproductive decisions, they must be knowledgeable about their SCT status, SCD, and if their reproductive partner also has SCT. This is pertinent since two parents with SCT have a 25% chance of having a child with SCD and a 50% chance of having a child with SCT. Despite universal NBS that reliably identifies infants with SCT, >80% of individuals of childbearing age with SCT do not know their status [4,5]. This suggests that public health programs that notify and educate parents of infants with SCT are not as well received, accessible, or effective as needed in practice. There have been studies to suggest that using videoconferencing to deliver genetic counseling information to other populations is a method that can increase the level of satisfaction and accessibility to counseling [6], but this format has not been studied in SCT.

The ODH supported one-on-one in-person SCT education of parents of children with SCT by a trained educator prior to 2020, which increased many parents' knowledge and was well received, but approximately one-third of parents who were eligible to receive this education in central Ohio did not attend these sessions, likely due to difficulties in transportation and access. With the COVID-19 pandemic, ODH shifted to a telephone-only program. While changing to this format has potential benefits of reducing transportation barriers and increasing parental access to SCT education, it also adds challenges, including difficulty building rapport between the educator and parent and the inability to use supporting visual materials to explain SCT. Its effectiveness, acceptability, and impact on parents' level of worry about having a child with SCT and/or SCD have not been studied. Given work that suggests parents with a trait for a genetic disease report that effective education about their trait decreases their anxiety, increases their preparedness, and increases their sense of control about the potential of having a child with a genetic disease [7], studies are needed to evaluate the effectiveness of novel SCT education programs.

According to Centers for Disease Control and Prevention guidance on providing health education curricula, effective education not only presents clear and understandable information but is also engaging and acceptable to learners in that it addresses their values, attitudes, and beliefs [8]. To overcome some limitations of both in-person and telephone-only education, we developed SCTaware, a health literacy (HL)–informed, videoconferencing-delivered SCT education program. SCTaware allows for at-home access to SCT education with visual materials and the opportunity to build rapport between parent and educator to facilitate engagement, question-asking, and understanding, and to reduce parental worry. Therefore, the aim of this study was to assess the acceptability of the content and format of this program and parental engagement with SCTaware. We also aimed to assess perceptions of worry among parents before and after receiving SCTaware.

Methods

SCTaware

SCTaware was developed by a multidisciplinary research team after evaluation of the existing in-person SCT program in central Ohio [9]. SCTaware is a one-on-one videoconferencing education that is delivered by a trained educator to parents of children with SCT identified by NBS. This program focuses on the reproductive implications of having SCT to encourage parents who do not know their SCT status to get tested. It includes having a trained educator provide the content, core knowledge objectives, a plain language–talking guide to support the educator, HL-based communication strategies to increase learner engagement and foster participation (eg, teach-back), and HL-informed and culturally sensitive visuals to support the verbal content [10].

Study Design

This was a single-center, institutional review board–approved, prospective study of parents of children identified to have hemoglobin S trait by NBS. English-speaking adult biological parents of children <3 years of age with hemoglobin S trait who received telephone education were identified from the telephone educator's schedule within the electronic mearents who were able to be contacted via telephone by the virtual educator and member of the study team were eligible to participate if they did not have SCD or a child with SCD, had not received SCT education in central Ohio for another child, were not (or partner

was not) currently pregnant, and had access to an electronic device capable of videoconferencing

Study Procedures

Parents who consented to participate were asked to complete surveys after receiving the SCTaware content. In addition to completing a demographics survey and reporting the type of device they used to receive SCTaware, acceptability and level of worry were assessed using the Education Effectiveness Survey (EES). This survey includes an 18-item Likert scale of multiple-choice items and one open-ended item. The EES was developed using a modified version of the education satisfaction survey that was used in a prior SCT study [9]. It was used to evaluate parents' satisfaction with the virtual educator and visual materials, to assess if SCTaware addressed parents' learning barriers, to gauge parents' level of worry (eg, not worried, a little worried, very worried) about having a baby with SCT before and after receiving SCTaware, and to obtain parent input on the best methods to provide additional SCT education after SCTaware. It also allowed participants to provide open-ended comments about their experience.

After each session, the virtual educator also completed a survey to assess parent engagement and distractions, to quantify the number of questions parents asked, to assess parents' ability to teach-back key content, and to record the time it took to complete the session. The 5-item survey included 4 multiple-choice responses and 1 open-ended question. This survey was pilot-tested with the SCTaware development team. To assess reliability over time, 6 education sessions were also randomly observed by an additional SCTaware development team member who also completed the survey.

Statistical Analysis

Data were summarized descriptively. Frequency and percentage for qualitative variables and mean or median and IQR were calculated for quantitative variables. A McNemar test was used to compare parental worry before and after receiving SCTaware. A P value <.05 was considered statistically significant. Analyses were completed using the base R statistical package (R Foundation for Statistical Computing).

Ethics Approval

Documentation of verbal informed consent was required prior to study participation (Clinical Trials number: NCT03984500). This study was approved by the Nationwide Children's Hospital Institutional Review Board (STUDY00000122).

Results

Participants

Of the 391 parents of children with SCT who received telephone education between October 2020 and October 2022, 154 (39%) were able to be contacted, and 86 (56%) consented to participate. Of these, 60 (70%) completed the SCTaware education session (Table 1), and 55 (64%) completed the EES.

Table 1. Characteristics of participants completing SCTaware education.

Participant characteristic	Participants (N=60), n (%)
Female sex	59 (98)
Age (years)	
18-24	7 (12)
25-39	51 (85)
40-64	2 (3)
Race	
Black	46 (77)
White	8 (13)
Multiracial	5 (8)
Other	1 (2)
Ethnicity	
Not Hispanic or Latino	57 (95)
Hispanic or Latino	3 (5)
Language spoken at home	
English	50 (83)
Other	10 (17)
Type of electronic device used for education session	
Smartphone (Android or iPhone)	37 (62)
Computer	14 (23)
Tablet or iPad	4 (7)
Missing	5 (8)
Type of internet connection used for education session	
Wi-Fi	47 (78)
Data plan	7 (12)
Hotspot	1 (2)
Missing	5 (8)

SCTaware Education

The mean length of the SCTaware education session was 34.4 (median 30, IQR 10) minutes. A total of 14 parents reported having learning barriers (problems hearing, seeing, reading, understanding English, or other) that make it hard for them to learn new things.

Understandability of Content

Nearly all participants (n=54, 98%) reported the words used by the educator were somewhat to very easy to understand. All participants agreed that they had a chance to ask questions, that these questions were answered, and that they were given the opportunity to teach-back key concepts to ensure that they understood.

All participants agreed or strongly agreed that the SCTaware visuals helped them understand SCT. Most participants (n=47, 85%) reported that the visuals used were very easy to understand, appeared professional (eg, resectful, trustworthy; n=55, 100%), told helpful information (n=42, 76%), and facilitated conversation with the educator (n=42, 76%). Nearly two-thirds (n=34, 62%) reported that the pictures appeared very personal and applied to them, and 27% (n=15) felt that the visuals were actionable, in that they told the participant what to do (eg, directing a participant how to get tested for SCT).

Education Format Preferences

Most participants either preferred the virtual session (n=29, 53%) or having the combination of the telephone education and then the virtual session (n=21, 38%). A few (n=5, 9%) reported they preferred the telephone-only education to the virtual session. Those who preferred the virtual or the combination of virtual and telephone-only education reported that the virtual format “made it like we (Educator and Parent) were face to face,” and that “it allowed me (Parent) to feel present in the conversation and learning.” They also reported that it allowed the educator “to show them things,” and that “I (Parent) was able to learn at my own pace.” Participants reported their preferences about potential methods to receive additional SCT information after the SCTaware session (Table 2).

Table 2. Participant preferences for how to receive additional information on sickle cell trait.

Participant preferences	Participants (n=55), n (%)
A sickle cell trait website on the internet	26 (46)
Another education session by smartphone, computer, tablet, or iPad	25 (45)
Mailed brochure or booklet	17 (30)
Another telephone call	13 (23)
A sickle cell trait mobile app	9 (16)
Video/DVD	3 (5)
Group session	2 (4)
Do not want any more information	5 (9)

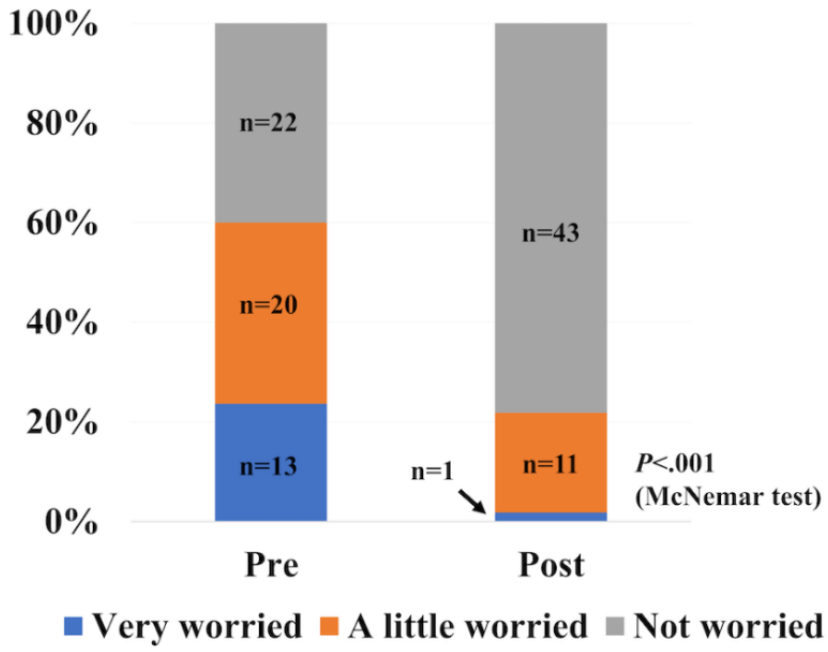
Educator Perceptions on Engagement

The educator reported that most (n=45) of the participants who completed SCTaware were highly engaged in the education; however, 10 did not use their camera on their electronic device, making engagement assessment among these participants challenging. The educator noted that most (82%, n=49) asked at least one question and 40% (n=24) asked more than three questions. The educator also noted that distractions, frequently a young child crying, were common among parents (65%, n=39), but that many distracted participants were still "fairly engaged despite being distracted and made an effort to indicate listening." The observers' and educator's assessments of perceived engagement aligned for all 6 randomly observed sessions. Perceptions of level of distraction were aligned for 3 sessions and were closely aligned for the other 3 sessions.

Parent Perceptions of Worry

Of the 55 participants who completed the EES, 60% (n=33) reported some level of worry about having a child with SCT after receiving telephone-only education and before receiving SCTaware, 24% (n=13) of those being very worried. After receiving SCTaware, 55% (n=30) of parents had less worry; only 5% (n=3) had increased worry. Only 2% (n=1) remained very worried after receiving SCTaware (Figure 1).

Figure 1. Participants' self-reported worry about having a child with SCT (Sickle cell trait) before and after SCTaware.



Participants' Open-ended Feedback on SCTaware

A few parents reported that the virtual educator "explained things very well" and "got me to think about getting blood work done to see if I'm a trait carrier," that the education was "very enlightening," and that "the educator was clear and explained everything in a systematic way that made it easy to understand." One participant also reported that the teach-back method was "cool...but at times it made me feel like a kid being told to listen up." However, this same individual also felt that teach-back "helped things to stick a little better [so] it's a thin line between good and unsure."

Discussion

Principal Findings

With an increase in reliance on digital technology use, especially to provide education [11], virtual education has emerged as a method with the potential to close knowledge gaps on important public health topics. Similar studies on genetic counseling provided through videoconferencing services have found that this method of delivery increases accessibility by reducing travel burdens and increasing convenience [6]; however, acceptability of and engagement with these programs among learners in public health education settings has not been extensively studied. This study demonstrated that the SCTaware virtual education program is highly acceptable to parents of young children with SCT. Most virtually educated parents reported that the visuals facilitated understanding of the content and conversation with the educator. This study also suggests that many parents have some level of worry about having a child with SCT after receiving education about it by telephone. This may indicate that telephone-only education is not adequate to fully address parental concerns or may leave parents feeling unable to ask additional questions that may arise after their telephone education session. This may also be because while telephone education is convenient, it may not support learning goals, may not facilitate adequate psychosocial support between educator and learner since educators are not face-to-face with parents, and may inadequately support learning since visual materials cannot be shared. Fortunately, most participants reported a reduction of worry after receiving SCTaware. Additionally, our educator reported high engagement with participants who completed SCTaware, with 82% (n=49) asking at least 1 question, and 40% (n=24) asking more than 3. These findings and parents' open-ended feedback may reflect the effectiveness of our intentional use of HL strategies, such as putting responsibility on the educator to promote comfort with asking questions and to ensure parent understanding through teach-back. The parents' positive reports of teach-back underscore its value to confirm understanding. The comment regarding both positive and unsure aspects of teach-back reflects the importance of using it correctly, that is, the educator taking responsibility for being clear [12].

This study found that 62% (n=34) of participants reported the pictures appeared very personal and applied to them. This was lower than expected when compared to the high rating of understandability of visuals by participants. This may reflect efforts to address cultural sensitivity and diversity such that the figures depicted in the visuals had a generic appearance. Additional research is needed to better identify optimal depiction of human figures in SCTaware and similar materials to promote both individual connection as well as inclusiveness. Nevertheless, participants reported high satisfaction, affirming the visuals were very easy to understand, professional, helpful, and facilitated conversation with the educator.

Notably, we observed that participants often faced distractions and interruptions, most of which were out of the parents' control (eg, a crying child) during their virtual education session. These interruptions did not necessarily impact parental engagement as assessed through the educator survey. However, divided attention while learning has been shown to impair long-term memory retention and decrease one's ability to apply learned knowledge to new context [13]. This may impact long-term knowledge gain about SCT. A potential strategy to reduce the impact of distractions on knowledge gain could be to ask parents to find a quiet and interruption-free space prior to the education. However, we recognize that this may not always be possible, especially since parents of children with SCT are educated shortly after their child's birth. Alternatively, encouraging parents to attend to distractions and then re-engaging them may shorten the distraction and allow them to remain attentive to the material when it is being presented.

Comparison With Prior Work

We observed many parents had high levels of worry after receiving telephone-only SCT education, but this level of concern lessened after SCTaware. This is consistent with prior literature that suggests that effective genetic counseling for abnormal hemoglobin traits can lower anxiety among family members [14]. Future studies that test the effectiveness of virtual education at increasing SCT knowledge and reducing worry, potentially with a follow-up telephone call, are warranted, considering that many parents reported preference for a combination of virtual and telephone education, and nearly a quarter preferred to receive additional SCT information via telephone.

In contrast to our prior study of in-person SCT education [9], where parents rarely asked questions, we found that most parents who received SCTaware asked 1 or more questions. This is notable because question-asking can be used to confirm that learners are engaged with material and is an opportunity for the educator to correct misunderstandings when these may not have been otherwise identified. Furthermore, since parental information-seeking behavior, including collaborative question-asking between parents and genetic counselors, is associated with enhanced knowledge among parents of children with cystic fibrosis [15], this increase in question-asking may ultimately positively impact parents' SCT knowledge.

Limitations

This study has a few limitations. First, since only English-speaking parents were recruited via telephone, we were unable to assess SCTaware's acceptability among those who could not be contacted, were not proficient in English, who did not have the ability to access the virtual education, or who declined participation. It is possible that parents who participated were more motivated, worried, and engaged in learning about SCT and were, therefore, more likely to report high acceptability of SCTaware. While access to the technology required to complete SCTaware could limit the applicability of our program, we suspect this will be less of a limitation with time, since digital technology is becoming more ubiquitous [16]. Development and assessment of acceptable materials to serve parents who do not speak English as a primary language is also an important next step to increasing accessibility to SCTaware. Since genetic counseling sessions that utilize an interpreter have been shown to reduce the number of questions asked by patients and the overall levels of interaction between patients and providers [17],

training of future educators and interpreters to ensure thoroughness of education and acceptability by non-English speaking parents is vital, especially since most parents who have children with SCT worldwide may not be proficient in English.

Second, while mothers and fathers were eligible, mothers were primarily recruited. This is likely because mothers were listed as the primary contact for telephone SCT education in the electronic medical record. This finding is consistent with the literature. Studies have found that fewer than a third of fathers attended pretesting cancer genetic counseling appointments, yet findings suggest that attendance of both parents in a genetic counseling session may result in parents feeling more informed [18]. Future research is needed to identify how to reduce this disparity and to investigate if and how mothers communicate their SCT knowledge to their child's father.

Third, SCTaware was provided by a single educator who was aware that some sessions were going to be randomly observed by members of the study team and evaluated by participants. This may have impacted how the educator provided the education, and future studies need to consider if ongoing evaluation is needed to ensure that the SCTaware program is consistently delivered. Additionally, the acceptability of the virtual program may be related to this educator's ability to connect with parents and may not be generalizable. Careful selection criteria for future educators and standardized training will be needed to effectively disseminate SCTaware to a larger audience.

Lastly, survey responses may have been biased. For example, participants reported both their pre- and postlevel of worry about having a child with SCT after receiving SCTaware. Postsession impressions may have affected their rating of presession worry. Additionally, participants had an ongoing relationship with the educator as they completed the study activities, which may have influenced their responses. The eucator and participants were also aware when an additional team member observed the education session, which may have impacted the parent's engagement. Finally, it is necessary to determine if SCTaware closes knowledge gaps and if the program results in increased parental testing.

Conclusions

Virtual HL-informed SCT education was found to be highly acceptable as a method for distributing information about a common and important public health topic. This suggests that virtual education in the public health setting may be a promising intervention to increase accessibility to other public health information, especially among populations who may have challenges attending in-person education sessions. Future research is needed to determine if this format increases accessibility, whether it closes knowledge gaps and leads to actionable behavior, and if it is applicable to other health topics.

Conflicts of Interest

None declared.

References

1. Newborn Screening Ohio Department of Health. Ohio Department of Health. URL: https://odh.ohio.gov/about-us [accessed 2022-09-27]
2. Centers for Disease Control and Prevention. 2019. URL: https://www.cdc.gov/ncbddd/sicklecell/data.html [accessed 2022-09-27]
3. Powars D, Chan L, Hiti A, Ramicone E, Johnson C. Outcome of sickle cell anemia: a 4-decade observational study of 1056 patients. Medicine (Baltimore) 2005 Nov;84(6):363-376. [doi: 10.1097/01.md.0000189089.45003.52] [Medline: 16267411]
4. Treadwell MJ, McClough L, Vichinsky E. Using qualitative and quantitative strategies to evaluate knowledge and perceptions about sickle cell disease and sickle cell trait. J Natl Med Assoc 2006 May;98(5):704-710. [Medline: 16749645]
5. Gustafson SL, Gettig EA, Watt-Morse M, Krishnamurti L. Health beliefs among African American women regarding genetic testing and counseling for sickle cell disease. Genet Med 2007 May;9(5):303-310 [FREE Full text] [doi: 10.1097/gim.0b013e3180534282] [Medline: 17505208]
6. Bradbury A, Patrick-Miller L, Harris D, Stevens E, Egleston B, Smith K, et al. Utilizing remote real-time videoconferencing to expand access to cancer genetic services in community practices: A multicenter feasibility study. J Med Internet Res 2016 Feb 01;18(2):e23 [FREE Full text] [doi: 10.2196/jmir.4564] [Medline: 26831751]
7. Schneider JL, Goddard KAB, Davis J, Wilfond B, Kauffman TL, Reiss JA, et al. "Is It Worth Knowing?" Focus group participants' perceived utility of genomic preconception carrier screening. J Genet Couns 2016 Feb;25(1):135-145 [FREE Full text] [doi: 10.1007/s10897-015-9851-7] [Medline: 26093606]
8. Characteristics of an effective health education curriculum 2013. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/healthyschools/sher/characteristics/index.htm [accessed 2022-09-27]
9. Creary S, Adan I, Stanek J, O'Brien SH, Chisolm DJ, Jeffries T, et al. Sickle cell trait knowledge and health literacy in caregivers who receive in-person sickle cell trait education. Mol Genet Genomic Med 2017 Nov;5(6):692-699. [doi: 10.1002/mgg3.327] [Medline: 29178654]
10. Abrams M, Zajo K, Beeman C, O'Brien SH, Chan P, Shen Y, et al. A health literate approach to address health disparities: a virtual program for parents of children with sickle cell trait. J Commun Healthc 2022;15(2):112-120. [doi: 10.1080/17538068.2022.2026056] [Medline: 36275941]

11. Selwyn N, Facer K. The sociology of education and digital technology: past, present and future. Oxford Rev Education 2014 Jul 31;40(4):482-496. [doi: 10.1080/03054985.2014.933005]
12. Abrams M, Rita S, Kurtz-Rossi S, Nielsen G. Always Use Teach-back! toolkit.: Institute for Healthcare Advancement; 2012. URL: http://www.teachbacktraining.org/home [accessed 2022-09-27]
13. Schmidt S. Distracted learning: big problem and golden opportunity. J Food Sci Education 2020;19(4):278-291.
14. Kladny B, Williams A, Gupta A, Gettig EA, Krishnamurti L. Genetic counseling following the detection of hemoglobinopathy trait on the newborn screen is well received, improves knowledge, and relieves anxiety. Genet Med 2011 Jul;13(7):658-661 [FREE Full text] [doi: 10.1097/GIM.0b013e31821435f7] [Medline: 21546841]
15. Dillard JP, Shen L, Robinson JD, Farrell PM. Parental information seeking following a positive newborn screening for cystic fibrosis. J Health Commun 2010 Dec;15(8):880-894 [FREE Full text] [doi: 10.1080/10810730.2010.522226] [Medline: 21170789]
16. Smith A. U.S. smartphone use in 2015. Pew Research Center. 2015. URL: https://www.pewresearch.org/internet/2015/04/01/us-smartphone-use-in-2015/ [accessed 2022-09-27]
17. Ault R, Morales A, Ault R, Spitale A, Martinez GA. Communication pitfalls in interpreted genetic counseling sessions. J Genet Couns 2019 Aug;28(4):897-907. [doi: 10.1002/jgc4.1132] [Medline: 31112366]
18. Demarco TA, Nusbaum RH, Peshkin BN, Patenaude AF, Schneider KA, Garber JE, et al. Prevalence and correlates of mothers and fathers attending pretest cancer genetic counseling together. Patient Educ Couns 2010 Jan;78(1):29-33 [FREE Full text] [doi: 10.1016/j.pec.2009.05.013] [Medline: 19545972]

Abbreviations

EES: education effectiveness survey
HL: health literacy
ODH: Ohio Department of Health
NBS: newborn screening
SCD: sickle cell disease
SCT: sickle cell trait

Edited by A Mavragani; submitted 15.04.22; peer-reviewed by T Mayo-Gamble, A Johnson; comments to author 28.05.22; revised version received 01.07.22; accepted 04.07.22; published 17.11.22.

Please cite as:
Beeman CM, Abrams MA, Zajo KN, Stanek J, O'Brien SH, Chan P, Shen Y, McCorkle B, Johnson L, Chisolm D, Barnard-Kirk T, Mahan JD, Martinez-Mendez A, Phillips WL, Creary SE
Acceptability to and Engagement With a Virtual Sickle Cell Trait Education Program (SCTaware): Single-Center Prospective Study
JMIR Form Res 2022;6(11):e38780
URL: https://formative.jmir.org/2022/11/e38780
doi:10.2196/38780
PMID:36394943

Original Paper

Postpartum Migraine Headache Coding in Electronic Health Records of a Large Integrated Health Care System: Validation Study

Jiaxiao Shi[1*], PhD; Michael J Fassett[2,3], MD; Vicki Y Chiu[1], MS; Chantal C Avila[1], MA; Nehaa Khadka[1], MPH; Brittany Brown[1], MPH; Pooja Patel[1], MPH; Nana Mensah[1], MPH, PhD; Fagen Xie[1], PhD; Morgan R Peltier[4], PhD; Darios Getahun[1,5*], MPH, MD, PhD

[1]Department of Research and Evaluation, Kaiser Permanente Southern California, Pasadena, CA, United States

[2]Department of Obstetrics and Gynecology, Kaiser Permanente West Los Angeles Medical Center, Los Angeles, CA, United States

[3]Department of Clinical Sciences, Kaiser Permanente Bernard J Tyson School of Medicine, Pasadena, CA, United States

[4]Department of Psychiatry, Jersey Shore University Medical Center, Neptune, NJ, United States

[5]Department of Health Systems Science, Kaiser Permanente Bernard J Tyson School of Medicine, Pasadena, CA, United States

[*]these authors contributed equally

Corresponding Author:
Darios Getahun, MPH, MD, PhD
Department of Research and Evaluation
Kaiser Permanente Southern California
100 S Los Robles Avenue
Pasadena, CA, 91101
United States
Phone: 1 626 564 5658
Email: Darios.T.Getahun@kp.org

Abstract

Background: Migraine is a common neurological disorder characterized by repeated headaches of varying intensity. The prevalence and severity of migraine headaches disproportionally affects women, particularly during the postpartum period. Moreover, migraines during pregnancy have been associated with adverse maternal outcomes, including preeclampsia and postpartum stroke. However, due to the lack of a validated instrument for uniform case ascertainment on postpartum migraine headache, there is uncertainty in the reported prevalence in the literature.

Objective: The aim of this study was to evaluate the completeness and accuracy of reporting postpartum migraine headache coding in a large integrated health care system's electronic health records (EHRs) and to compare the coding quality before and after the implementation of the International Classification of Diseases, 10th revision, Clinical Modification (ICD-10-CM) codes and pharmacy records in EHRs.

Methods: Medical records of 200 deliveries in all 15 Kaiser Permanente Southern California hospitals during 2 time periods, that is, January 1, 2012 through December 31, 2014 (International Classification of Diseases, 9th revision, Clinical Modification [ICD-9-CM] coding period) and January 1, 2017 through December 31, 2019 (ICD-10-CM coding period), were randomly selected from EHRs for chart review. Two trained research associates reviewed the EHRs for all 200 women for postpartum migraine headache cases documented within 1 year after delivery. Women were considered to have postpartum migraine headache if either a mention of migraine headache (yes for diagnosis) or a prescription for treatment of migraine headache (yes for pharmacy records) was noted in the electronic chart. Results from the chart abstraction served as the gold standard and were compared with corresponding diagnosis and pharmacy prescription utilization records for both ICD-9-CM and ICD-10-CM coding periods through comparisons of sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), as well as the summary statistics of *F*-score and Youden *J* statistic (*J*). The kappa statistic (κ) for interrater reliability was calculated.

Results: The overall agreement between the identification of migraine headache using diagnosis codes and pharmacy records compared to the medical record review was strong. Diagnosis coding (*F*-score=87.8%; *J*=82.5%) did better than pharmacy records (*F*-score=72.7%; *J*=57.5%) when identifying cases, but combining both of these sources of data produced much greater accuracy in the identification of postpartum migraine cases (*F*-score=96.9%; *J*=99.7%) with sensitivity, specificity, PPV, and NPV of

100%, 99.7%, 93.9%, and 100%, respectively. Results were similar across the ICD-9-CM (*F*-score=98.7%, *J*=99.9%) and ICD-10-CM coding periods (*F*-score=94.9%; *J*=99.6%). The interrater reliability between the 2 research associates for postpartum migraine headache was 100%.

Conclusions: Neither diagnostic codes nor pharmacy records alone are sufficient for identifying postpartum migraine cases reliably, but when used together, they are quite reliable. The completeness of the data remained similar after the implementation of the ICD-10-CM coding in the EHR system.

(JMIR Form Res 2022;6(11):e42955) doi:10.2196/42955

KEYWORDS

migraine headache; validation; diagnosis; pharmacy; postpartum; medical record; health plan; electronic health record; coding; pharmacy record; diagnostic code; EHR system

Introduction

Migraine is a common neurological disorder characterized by repeated headaches of varying intensity [1]. Often presenting in one side of the head, these headaches are typically accompanied by nausea, vomiting, and extreme sensitivity or intolerance to light and sound that lasts for 4 to 72 hours [1,2]. Globally, migraines were ranked as the sixth most prevalent cause of diseases [3] and the third cause of disabilities [4]. The prevalence of migraine increases initially with age, peaks at about 30-39 years, and then gradually declines over time [3]. Migraine headaches are more prevalent among women of reproductive ages (15-49 years), with 25% of women affected worldwide [3]. Women also experience more severe symptoms with a longer duration of attack, recurrent headaches, and higher migraine-related disability than men [5]. Moreover, migraines during pregnancy have been linked with preeclampsia or postpartum stroke [6,7]. Migraines during the postpartum period are particularly common, with reported prevalence rates ranging from 5% to 55% [8]. However, due to the lack of a validated instrument for uniform case ascertainment, there is uncertainty in the reported prevalence in the literature.

In 2009, the Health Information Technology for Economic and Clinical Health (HITECH) Act was enacted to promote the adoption of electronic health record (EHR) systems among health care providers [9]. Consequently, many health care institutions implemented EHR systems in their facilities [10]. The Kaiser Permanente Southern California (KPSC) integrates EHRs with comprehensive inpatient and outpatient clinical records, and prescription medication history can be used for pharmacoepidemiologic research, including examining potential risk factors and triggers for postpartum migraine and its effects on the lives of the affected individuals.

Currently, many health care systems use the official International Classification of Diseases-Clinical Modification (ICD-CM) coding systems to classify diagnoses and procedures in the EHR [11]. In KPSC settings, ICD-9-CM coding was transitioned to the ICD-10-CM coding system after October 1, 2015, which provides increased specificity and details for many health conditions [11]. However, the accuracy of postpartum migraine headache data in the EHRs, including the impact of the ICD-9-CM/ICD-10-CM transition, has not been fully examined, leaving uncertainty around the validity of case ascertainment methods. Furthermore, whether such accuracy improves when postpartum migraine headache diagnosis (Dx) codes are used in conjunction with automated pharmacy records (Rx) has not been elucidated. Therefore, we evaluated the reliability and accuracy of postpartum migraine headache Dx codes with and without the supplemental use of Rx and whether the ICD-10-CM coding system has improved or decreased the accuracy of postpartum migraine headache case ascertainment.

Methods

Study Setting

This study was conducted using EHR data from KPSC. The KPSC health care system comprises over 4.8 million members, 15 hospitals, and 236 medical offices throughout southern California. Prenatal and postnatal care to the member patients is provided as outpatient care at KPSC. Although most members receive their care at KPSC hospitals and <10% utilize contracting hospitals, all diagnostic, procedural, and Rx data are captured and maintained by the KPSC EHR since its full implementation in 2008. Furthermore, the characteristics of KPSC members closely reflect the Californian population [12].

Ethics Approval

This study was approved by the KPSC institutional review board (approval 13114), and informed consent was waived, as the study was low risk and strictly involved the use of internal EHR data, that is, access permission was given to authorized personnel only when needed.

Cohort and Sample Selection

Data were obtained retrospectively from women who delivered live infants at the KPSC health system during 2 distinct time periods: (1) January 1, 2012 to December 31, 2014 (ICD-9-CM period) and (2) January 1, 2017 to December 31, 2019 (ICD-10-CM period). We carefully selected the 2 time periods to investigate the medical coding accuracy before and after the implementation of ICD-10-CM in the KPSC system. For each time period, we randomly sampled 25 cases in each of the following 4 strata based on EHRs: (1) those without any Dx codes or Rx for migraine headaches (neither Dx nor Rx), (2) those with only Dx codes for migraine headaches (Dx only), (3) those with only Rx selected a priori as a treatment for migraine headaches (Rx only), and (4) those with both Dx and Rx (Dx+Rx). Thus, a total of 200 individual deliveries were selected for this validation study via the stratified sampling scheme mentioned above. The accuracy for each of the 4 strata, either case ascertainment (Dx only, Rx only, or Dx+Rx) or

noncase ascertainment (neither Dx nor Rx), was expected to be around 85%. A sample size of 25 would provide less than 15% one-sided margin for a 90% CI of the accuracy for each stratum.

Defining Postpartum Migraine Headache

In this study, postpartum headache was defined based on documented clinical records within the first 12 months after delivery. In particular, primary clinical Dx codes, coded by medical coders from the clinical data management team (See Multimedia Appendix 1 for ICD-9-CM/ICD-10-CM codes) and pharmacy dispense records from all inpatient and outpatient services during the postpartum periods of the index pregnancy were used to ascertain its diagnosis (see Multimedia Appendix 1 for the medication list).

Chart Abstraction Process

Two trained research associates reviewed EHRs for documentation of a diagnosis or medication for migraine headache during the postpartum period. Women were considered to have postpartum migraine headache if either a mention of migraine headache or a prescription for the treatment of migraine headache was noted in the chart during the first year of the postpartum period. Research associates confirmed any medical diagnosis (yes/no) and prescription for treatment (yes/no) of migraine headaches for all 200 women. To ensure data quality and consistency of chart reviews between the 2 abstractors, interrater reliability assessment was performed on randomly selected charts stratified by the 4 strata (DX only, Rx only, Dx+Rx, and neither Dx nor Rx). Discrepant cases in clinical utilizations were adjudicated by the study investigators with clinical expertise (MJF and DG). The postpartum migraine headache cases abstracted through this process served as the gold standard.

Maternal Characteristics

Maternal characteristics for KPSC births included maternal age (<20, 20-29, 30-34, and ≥35 years), race/ethnicity (categorized as non-Hispanic White [hereafter referred to as White], non-Hispanic Black [hereafter referred to as Black], Hispanic, Asian/Pacific Islander, and others/unknown), educational attainment (less than high school, high school graduate, some college, bachelor's/associate degree, and master's degree or higher), household income in USD (<$30,000, $30,000-$49,999, $50,000-$69,999, $70,000-$89,999, and ≥$90,000), timing of prenatal care (early or first trimester and none or late initiation), and self-reported smoking (yes/no). Gestational age at birth, reported in completed weeks, was derived from clinical estimates.

We obtained the characteristics of all births of State of California residents during the same time periods by using publicly available data that has been posted on the Centers for Disease Control and Prevention Wonder website [13]. Both the KPSC EHR and the Centers for Disease Control and Prevention Wonder website provided information on maternal characteristics, including maternal age, race/ethnicity, educational attainment, timing of prenatal care, smoking during pregnancy, and gestational age at delivery (in completed weeks of gestation). Data on median household income was estimated based on census tracts for KPSC deliveries.

Statistical Analysis

We described the characteristics of our study population, all women who delivered in KPSC hospitals, and the State of California residents with live births during 2012-2014 and 2017-2019 by using frequency distributions. We also calculated the kappa statistic (κ), which estimates agreement between the 2 abstractors. As mentioned above, abstracted chart reviews for migraine headache cases were set as the gold standard. We compared findings from the manual chart review with corresponding diagnosis and prescription medication utilization records for ICD-9-CM and ICD-10-CM coding periods through sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). These performance measurements were reported as weighted percentages with corresponding 95% CIs using normalized sampling weights (W_i, i=1,2,3,4), derived by dividing the number of deliveries in each stratum by the multiplication of the total study population and the number of samples from the corresponding stratum. To evaluate the overall performance, we also reported the summary statistics of F-score and Youden J statistic (J), which are composite measurements of sensitivity and PPV (F-score) or sensitivity and specificity (J). All analyses were conducted using SAS statistical software version 9.4 (SAS Institute Inc).

Results

An overview of the patient characteristics for our sample, study population, and California state birth population is shown in Table 1. Overall, 157,501 deliveries from all KPSC hospitals were obtained during the 2 periods, with 72,471 for 2012-2014 and 85,030 for 2017-2019. Compared with California State birth data, KPSC deliveries had slightly higher percentages for maternal age over 30 years (1,465,998/2,874,396, 51% vs 96,157/157,501, 61.05%, respectively), education with college and above (1,605,882/2,874,396, 55.87% vs 114,297/157,501, 72.57%, respectively), and early prenatal care initiation (first trimester; 2,386,232/2,874,396, 83.02% vs 147,017/157,501, 93.34%, respectively). There were some discrepancies for race/ethnicity, likely due to the unknown information for 30.07% (864,311/2,847,396) of the state population versus only 3.67% (5784/157,501) for the KPSC population. Despite the stratified sampling, our sample of 200 deliveries was broadly representative of the KPSC study population with respect to maternal age, early prenatal care initiation, prenatal smoking status, and gestational age. However, Hispanics and lower household income populations (<US $69,999) were slightly oversampled.

Table 1. Characteristics of the women who delivered in all Kaiser Permanente Southern California hospitals and in the State of California (2012-2014 and 2017-2019).

Characteristics	Charts reviewed (n=200)[a], n (%)	Study population (N=157,501), n (%)	California state (N=2,874,396)[b], n (%)
Maternal age (years)			
<20	6 (3)	4665 (2.96)	1,44,945 (5.04)
20-29	76 (38)	56,679 (35.99)	1,263,453 (43.96)
30-34	71 (35.50)	54,810 (34.80)	843,010 (29.33)
≥35	47 (23.50)	41,347 (26.25)	622,988 (21.67)
Race/ethnicity			
White	39 (19.50)	39,219 (24.90)	372,037 (12.94)
Black	15 (7.50)	10,862 (6.90)	68,195 (2.37)
Hispanic	118 (59)	78,853 (50.07)	1,356,354 (47.19)
Asian/Pacific Islander	25 (12.50)	22,783 (14.47)	213,499 (7.43)
Others/unknown	3 (1.50)	5784 (3.67)	864,311 (30.07)
Educational attainment			
Less than high school	5 (2.50)	4355 (2.77)	435,360 (15.15)
High school graduate	49 (24.50)	35,411 (22.48)	694,118 (24.15)
Some college	54 (27)	32,616 (20.71)	558,288 (19.42)
Bachelor's/associate degree	61 (30.50)	54,293 (34.47)	729,896 (25.39)
Master's degree/above	27 (13.50)	27,388 (17.39)	317,698(11.05)
Missing	4 (2)	3438 (2.18)	139,036 (4.84)
Household income (USD)[c]			
<$30,000	3 (1.50)	5194 (3.30)	N/A[d]
$30,000-$49,999	50 (25)	39,969 (25.38)	N/A
$50,000-$69,999	73 (36.50)	47,864 (30.39)	N/A
$70,000-$89,999	45 (22.50)	32,486 (20.63)	N/A
≥$90,000	29 (14.50)	31,925 (20.27)	N/A
Missing	0 (0)	63 (0.04)	N/A
Timing of prenatal care			
First trimester	186 (93)	147,017 (93.34)	2,386,232 (83.02)
No or late care	14 (7)	9860 (6.26)	442,493 (15.39)
Missing	0 (0)	624 (0.40)	45,671 (1.59)
Smoking during pregnancy	8 (4)	6420 (4.08)	46,977 (1.63)
Gestational age (weeks)			
<34	10 (5)	3412 (2.17)	66,099 (2.30)
34-36	18 (9)	9865 (6.26)	180,352 (6.27)
37+	171 (85.50)	144,192 (91.55)	2,624,620 (91.31)
Missing	1 (0.50)	32 (0.02)	3325 (0.12)

[a]Sample is based on data from Kaiser Permanente Southern California electronic health records of 2012-2014 and 2017-2019.

[b]Data from the natality information of Centers for Disease Control and Prevention webpage [13] (accessed on January 9, 2022).

[c]Household income is not available for the California state data.

[d]N/A: not applicable.

The overall κ between our abstractors was 100%. Table 2 gives the distribution and their sample sizes for the 4 strata (Dx+Rx, Dx only, Rx only, and neither Dx nor Rx) among our overall study population and by the 2 time periods. The majority (150,801/157,501, 95.75%) of the women did not have any record of diagnosis or pharmacy usage, while only 1131 (0.72%) women had Rx indicating obtaining prescription medication but without a diagnosis of postpartum migraine headache. Regardless of the uneven numbers among the 4 strata, we sampled 50 cases from each stratum. The corresponding chart review results and the normalized weights are provided in Table 2. For those with both Dx codes and Rx, our chart abstractors confirmed all 50 had a postpartum migraine headache, and those with neither were all confirmed to be noncases. The sample with postpartum migraine headache Dx codes only had 88% (44/50) of true positives, while having prescription Rx only resulted in 96% (48/50) of true positives.

Based on these comparisons against the chart review results, the performance measurements of our EHR are shown in Table 3. The overall weighted sensitivity, specificity, PPV, and NPV for postpartum migraine headaches were 100%, 99.7%, 93.9%, and 100%, respectively, if the postpartum migraine cases were determined by having either Dx or Rx. The corresponding *F*-score and Youden *J* statistic were 96.9% and 99.7%, respectively. Compared with using Dx codes alone (either Dx+Rx or Dx only) for postpartum migraine case identification, using Rx alone (either Dx+Rx or Rx only) had poorer performance, especially on the sensitivity (82.7% for Dx codes alone vs 57.5% for Rx alone). Results were similar across the ICD-9-CM and ICD-10-CM coding time periods.

Table 2. Frequencies of the study population and chart review results by diagnosis codes and pharmacy records.

Postpartum migraine headache case ascertainment method	Study population			Sample		
	Overall (N=157,501), n (%)	2012-2014 (n=72,471), n (%)	2017-2019 (n=85,030), n (%)	Charts (n=200)	True cases by chart review (n=142), n (%)	Normalized sampling weight
Diagnosis codes + pharmacy records	2531 (1.61)	1261 (1.74)	1270 (1.49)	50	50 (100)	0.000321
Diagnosis codes only	3038 (1.93)	1488 (2.05)	1550 (1.82)	50	44 (88)	0.000386
Pharmacy records only	1131 (0.72)	393 (0.54)	738 (0.87)	50	48 (96)	0.000144
Neither diagnosis codes nor pharmacy records	150,801 (95.75)	69,329 (95.66)	81,472 (95.82)	50	0 (0)	0.019149

Table 3. Weighted performance measurements for postpartum migraine headache based on data sources before and after implementation of the International Classification of Diseases, 10th revision, Clinical Modification codes in the Kaiser Permanente Southern California system in 2015 (n=200).

Weighted performance measurements (%)	Overall			2012-2014			2017-2019		
	ICD-CM-9/10[a] Dx[b]	Rx[c]	ICD-CM-9/10 Dx or Rx	ICD-CM-9/10 Dx	Rx	ICD-CM-9/10 Dx or Rx	ICD-CM-9/10 Dx	Rx	ICD-CM-9/10 Dx or Rx
Sensitivity (95% CI)	82.7 (77.7-87.8)	57.5 (48.5-66.5)	100 (100-100)	83.4 (76.5-90.3)	55.4 (42.7-68)	100 (100-100)	82 (74.6-89.5)	59.8 (46.7-72.9)	100 (100-100)
Specificity (95% CI)	99.8 (99.6-100)	100 (99.9-100)	99.7 (99.5-99.9)	99.9 (99.8-100)	100 (99.9-100)	99.9 (99.7-100)	99.6 (99.2-100)	100 (99.9-100)	99.6 (99.2-100)
PPV[d] (95% CI)	93.5 (88.3-98.6)	98.8 (97-100)	93.9 (89.5-98.2)	97.8 (93.5-100)	98.8 (96.3-100)	97.5 (93.7-100)	89.1 (79.8-98.3)	98.8 (96.3-100)	90.3 (82.5-98.1)
NPV[e] (95% CI)	99.3 (99-99.6)	98.3 (97.6,99)	100 (100-100)	99.3 (98.9-99.7)	98.1 (97-99.2)	100 (100-100)	99.3 (98.9-99.7)	98.4 (97.5-99.4)	100 (100-100)
F-score[f]	87.8	72.7	96.9	90	71	98.7	85.4	74.5	94.9
Youden *J* statistic[f]	82.5	57.5	99.7	83.3	55.4	99.9	81.6	59.8	99.6

[a]ICD-CM-9/10: International Classification of Diseases, Clinical Modification, 9th/10th revision.

[b]Dx: diagnosis codes.

[c]Rx: pharmacy records.

[d]PPV: positive predictive value.

[e]NPV: negative predictive value.

[f]*F*-score and Youden *J* statistic are statistics that capture the overall performance of a dichotomous diagnostic test.

Discussion

This validation study was performed to determine the accuracy of identifying postpartum migraine headache cases by using data abstracted from the EHR of a large health care system with a sociodemographically diverse patient population. To our knowledge, the accuracy of the data on postpartum migraine headache using both Dx and Rx has not been validated in EHR data or the extent that the transition of ICD-9-CM to the ICD-10-CM coding system has impacted postpartum migraine headache case ascertainment. Our study showed that having either Dx or Rx had high sensitivity (100%) and specificity (99.7%) for case ascertainment. The overall performance measured by the *F*-score and Youden *J* statistic were 96.9% and 99.7%, respectively. Such findings were similar across the ICD-9-CM and ICD-10-CM coding time periods.

Over the past few decades, EHRs have become important data sources for pharmacoepidemiologic studies and have become standard among health care providers as part of the American Recovery and Reinvestment Act of 2009 (specifically, the HITECH Act) [9]. Although the EHR is a highly sophisticated information management and care delivery system that ensures quality care by providing access to comprehensive patient information and the latest best practice research, its completeness and reliability for pharmacoepidemiologic studies have been questioned due to various reasons, including clinical knowledge, attention to details, communication between providers and coders, coding procedures, and others [14-16].

At KPSC, the process of coding and coding rules of the medical diagnosis recorded in patients' EHRs is carried out by highly skilled medical coders from KPSC's clinical data management team. Furthermore, the individual medical coder's accuracy has been evaluated critically for consistency. Nevertheless, we may still need to develop reliable disease-specific algorithms by using diagnosis coding in combination with other EHRs, including but not limited to pharmacy and utilization records [17]. Therefore, we performed this validation study to evaluate (1) the accuracy of postpartum migraine headache case identification in EHRs and (2) the impact of implementing the ICD-10-CM coding system in extracting data on patients with postpartum migraine headache in the health system. The findings of this study suggested that the validity of EHR data for the identification of postpartum migraine headache cases differed depending on the EHR data sources (Dx and Rx). The accuracy of case ascertainment based on Dx was higher than that of case ascertainment based on Rx in this study. We speculate that the low performance of pharmacy data on case ascertainment may be driven by lactating patients declining migraine medication to avoid infant exposure through breastmilk. Others may have opted to use over-the-counter pain medications rather than prescription medications. It is also possible that a given prescription may be an appropriate drug for conditions other than migraine. However, postpartum migraine headache cases can be identified with high accuracy if Dx codes and prescription medications for the conditions are used together. Furthermore, the transition from the ICD-9-CM coding system to the

ICD-10-CM coding system had minimal impact on the overall accuracy of ascertaining postpartum migraine headache cases.

The main strength of this study is the development of a valid and reliable method of postpartum migraine headache case ascertainment by using EHR data extracted from a large insured and socioeconomically diverse southern California population, which is likely generalizable to other health care settings with similar EHR database systems. In addition, in this study, the review process includes the entire medical health record, not limited to clinical Dx codes and prescription Rx, for determining the accuracy of postpartum migraine headache. A potential limitation of this study was the use of medical record abstractors who were not blind to the source of the data. Although it is possible that this could have biased the study in unforeseen ways, a previous study that evaluated the agreement between masked and unmasked medical records abstractors reported no impact of this [18].

In conclusion, our findings suggest that postpartum migraine headache is not reliably coded in the EHRs. The use of Rx along with clinical Dx codes improves the identification of true postpartum migraine cases more than the use of clinical Dx codes alone, and the transition to the ICD-10-CM diagnosis coding system had no impact.

Acknowledgments

This research was funded by the Garfield Memorial Fund (project RNG211070) for DG. The opinions expressed are solely the authors' responsibility and do not necessarily reflect the official views of the funding agency. The Garfield Memorial Fund study team would like to thank Kaiser Permanente members who contributed the electronic health information to this study. The authors thank Ms Sole Cardoso for her technical support.

Conflicts of Interest

The Garfield Memorial Fund study team led the design of the study and interpretation of the results. The authors have no competing interests. DG receives research support from National Institute of Health (NIH), National Institute of Child Health and Human Development (NICHD), National Institute of Environmental Health Sciences, Department of Health and Human Services, Patient Centered Outcomes Research Institute, Bayer AG, and Hologic, Inc. MJF receives research support from NIH, NICHD, Bayer AG, and Hologic, Inc. JS conducted the analyses, which were reviewed by the study team members.

Multimedia Appendix 1
Diagnosis codes and medication list for identifying postpartum migraine headache.
[DOCX File , 14 KB - formative_v6i11e42955_app1.docx]

References

1. No authors listed. Headache Classification Committee of the International Headache Society (IHS) The International Classification of Headache Disorders, 3rd edition. Cephalalgia 2018 Jan;38(1):1-211. [doi: 10.1177/0333102417738202] [Medline: 29368949]
2. Lucas S. Migraine and Other Headache Disorders: ACOG Clinical Updates In Women's Health Care Primary and Preventive Care Review Summary Volume XVIII, Number 4. Obstet Gynecol 2019 Jul;134(1):211. [doi: 10.1097/AOG.0000000000003322] [Medline: 31241597]
3. GBD 2016 DiseaseInjury IncidencePrevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet 2017 Sep 16;390(10100):1211-1259 [FREE Full text] [doi: 10.1016/S0140-6736(17)32154-2] [Medline: 28919117]
4. Steiner TJ, Stovner LJ, Vos T. GBD 2015: migraine is the third cause of disability in under 50s. J Headache Pain 2016 Dec;17(1):104 [FREE Full text] [doi: 10.1186/s10194-016-0699-5] [Medline: 27844455]
5. Vetvik KG, MacGregor EA. Sex differences in the epidemiology, clinical features, and pathophysiology of migraine. Lancet Neurol 2017 Jan;16(1):76-87. [doi: 10.1016/S1474-4422(16)30293-9] [Medline: 27836433]
6. Bushnell CD, Jamison M, James AH. Migraines during pregnancy linked to stroke and vascular diseases: US population based case-control study. BMJ 2009 Mar 10;338:b664 [FREE Full text] [doi: 10.1136/bmj.b664] [Medline: 19278973]
7. Facchinetti F, Allais G, Nappi RE, D'Amico R, Marozio L, Bertozzi L, et al. Migraine is a risk factor for hypertensive disorders in pregnancy: a prospective cohort study. Cephalalgia 2009 Mar;29(3):286-292. [doi: 10.1111/j.1468-2982.2008.01704.x] [Medline: 19220309]
8. Klein AM, Loder E. Postpartum headache. Int J Obstet Anesth 2010 Oct;19(4):422-430. [doi: 10.1016/j.ijoa.2010.07.009] [Medline: 20833030]
9. The Health Information Technology for Economic and Clinical Health Act (HITECH) Act ARRA Components. URL: https://www.hipaasurvivalguide.com/hitech-act-text.php [accessed 2022-09-01]
10. Adler-Milstein J, Jha AK. HITECH Act Drove Large Gains In Hospital Electronic Health Record Adoption. Health Aff (Millwood) 2017 Aug 01;36(8):1416-1422. [doi: 10.1377/hlthaff.2016.1651] [Medline: 28784734]

11. ICD-10: international statistical classification of diseases and related health problems: tenth revision. World Health Organization. URL: https://tinyurl.com/58etdccn [accessed 2022-01-04]
12. Koebnick C, Langer-Gould AM, Gould MK, Chao CR, Iyer RL, Smith N, et al. Sociodemographic characteristics of members of a large, integrated health care system: comparison with US Census Bureau data. Perm J 2012;16(3):37-41 [FREE Full text] [doi: 10.7812/TPP/12-031] [Medline: 23012597]
13. Natality information - live births. CDC WONDER. URL: https://wonder.cdc.gov/natality.html [accessed 2022-06-01]
14. Weiskopf NG, Weng C. Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. J Am Med Inform Assoc 2013 Jan 01;20(1):144-151 [FREE Full text] [doi: 10.1136/amiajnl-2011-000681] [Medline: 22733976]
15. Bello NA, Zhou H, Cheetham TC, Miller E, Getahun DT, Fassett MJ, et al. Prevalence of Hypertension Among Pregnant Women When Using the 2017 American College of Cardiology/American Heart Association Blood Pressure Guidelines and Association With Maternal and Fetal Outcomes. JAMA Netw Open 2021 Mar 01;4(3):e213808 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.3808] [Medline: 33787907]
16. O'Malley KJ, Cook KF, Price MD, Wildes KR, Hurdle JF, Ashton CM. Measuring diagnoses: ICD code accuracy. Health Serv Res 2005 Oct;40(5 Pt 2):1620-1639 [FREE Full text] [doi: 10.1111/j.1475-6773.2005.00444.x] [Medline: 16178999]
17. Getahun D, Shi JM, Chandra M, Fassett MJ, Alexeeff S, Im TM, et al. Identifying Ectopic Pregnancy in a Large Integrated Health Care Delivery System: Algorithm Validation. JMIR Med Inform 2020 Nov 30;8(11):e18559 [FREE Full text] [doi: 10.2196/18559] [Medline: 33141678]
18. Reisch LM, Fosse JS, Beverly K, Yu O, Barlow WE, Harris EL, et al. Training, quality assurance, and assessment of medical record abstraction in a multisite study. Am J Epidemiol 2003 Mar 15;157(6):546-551. [doi: 10.1093/aje/kwg016] [Medline: 12631545]

Abbreviations

Dx: diagnosis
EHR: electronic health record
HITECH: Health Information Technology for Economic and Clinical Health
ICD-9-CM: International Classification of Diseases, 9th revision, Clinical Modification
ICD-10-CM: International Classification of Diseases, 10th revision, Clinical Modification
KPSC: Kaiser Permanente Southern California
NPV: negative predictive value
PPV: positive predictive value
Rx: pharmacy records

Edited by A Mavragani; submitted 25.09.22; peer-reviewed by KM Kuo, T Karen; comments to author 19.10.22; revised version received 31.10.22; accepted 01.11.22; published 17.11.22.

Please cite as:
Shi J, Fassett MJ, Chiu VY, Avila CC, Khadka N, Brown B, Patel P, Mensah N, Xie F, Peltier MR, Getahun D
Postpartum Migraine Headache Coding in Electronic Health Records of a Large Integrated Health Care System: Validation Study
JMIR Form Res 2022;6(11):e42955
URL: https://formative.jmir.org/2022/11/e42955
doi:10.2196/42955
PMID:36394937

Original Paper

Extraction and Quantification of Words Representing Degrees of Diseases: Combining the Fuzzy C-Means Method and Gaussian Membership

Feng Han[1], MSc; ZiHeng Zhang[1], MSc; Hongjian Zhang[2], MSc; Jun Nakaya[3], MD; Kohsuke Kudo[4], MD; Katsuhiko Ogasawara[2], MBA, PhD

[1]Graduate School of Medicine, Hokkaido University, Sapporo, Japan

[2]Graduate School of Health Sciences, Medical Management and Informatics, Hokkaido University, Sapporo, Japan

[3]Division of Advanced Diagnostic Imaging Development, Graduate School of Medicine, Hokkaido University, Sapporo, Japan

[4]Department of Diagnostic Imaging, Faculty of Medicine, Hokkaido University, Sapporo, Japan

Corresponding Author:
Katsuhiko Ogasawara, MBA, PhD
Graduate School of Health Sciences
Medical Management and Informatics
Hokkaido University
N12 W5
Sapporo, 060-0812
Japan
Phone: 81 011 706 3409
Email: oga@hs.hokudai.ac.jp

Abstract

Background: Due to the development of medical data, a large amount of clinical data has been generated. These unstructured data contain substantial information. Extracting useful knowledge from this data and making scientific decisions for diagnosing and treating diseases have become increasingly necessary. Unstructured data, such as in the Marketplace for Medical Information in Intensive Care III (MIMIC-III) data set, contain several ambiguous words that demonstrate the subjectivity of doctors, such as descriptions of patient symptoms. These data could be used to further improve the accuracy of medical diagnostic system assessments. To the best of our knowledge, there is currently no method for extracting subjective words that express the extent of these symptoms (hereinafter, “degree words”).

Objective: Therefore, we propose using the fuzzy c-means (FCM) method and Gaussian membership to quantify the degree words in the clinical medical data set MIMIC-III.

Methods: First, we preprocessed the 381,091 radiology reports collected in MIMIC-III, and then we used the FCM method to extract degree words from unstructured text. Thereafter, we used the Gaussian membership method to quantify the extracted degree words, which transform the fuzzy words extracted from the medical text into computer-recognizable numbers.

Results: The results showed that the digitization of ambiguous words in medical texts is feasible. The words representing each degree of each disease had a range of corresponding values. Examples of membership medians were 2.971 (atelectasis), 3.121 (pneumonia), 2.899 (pneumothorax), 3.051 (pulmonary edema), and 2.435 (pulmonary embolus). Additionally, all extracted words contained the same subjective words (low, high, etc), which allows for an objective evaluation method. Furthermore, we will verify the specific impact of the quantification results of ambiguous words such as symptom words and degree words on the use of medical texts in subsequent studies. These same ambiguous words may be used as a new set of feature values to represent the disorders.

Conclusions: This study proposes an innovative method for handling subjective words. We used the FCM method to extract the subjective degree words in the English-interpreted report of the MIMIC-III and then used the Gaussian functions to quantify the subjective degree words. In this method, words containing subjectivity in unstructured texts can be automatically processed and transformed into numerical ranges by digital processing. It was concluded that the digitization of ambiguous words in medical texts is feasible.

(JMIR Form Res 2022;6(11):e38677) doi:10.2196/38677

KEYWORDS

medical text; fuzzy c-means; cluster; algorithm; machine learning; word quantification; fuzzification; Gauss; radiology; medical report; documentation; text mining; data mining; extraction; unstructured; free text; quantification; fuzzy; diagnosis; diagnostic; EHR; support system

Introduction

Owing to the development of medical data, several electronic medical reports such as clinical records have been created, which provide a large amount of clinical data to the medical professional [1]. They have been shown to have capabilities such as contributing to health care knowledge discovery processes, for example, disease phenotyping and diagnosis [2,3], identification of new associations [4], development of disease surveillance systems [5], and health care monitoring systems [6]. Electronic medical reports are categorized into structured and free-text formats [7]. Where unstructured clinical notes contain rich subjective information [8-10]. A radiology report records a patient's condition created by a health care professional, such as a doctor, and contains medical evaluation information [11]. Although, with the advent of natural language processing (a branch of artificial intelligence applicable to unstructured textual data), clinical text mining is increasingly used in various health domains, there is still relatively little text-based machine learning research available compared to the more common structured numerical data-based analysis [12].

The traditional method involves manual extraction of symptoms. Hyun et al [13] extracted symptoms by manually defining a dictionary of symptoms. Jagannatha and Yu [14,15] and Fodeh et al [16] manually labeled and extracted symptoms. Matheny et al [17] set rules to extract symptoms. Chen and Sarkar [7] and Tamang et al [18] used a combination of several of these methods to extract symptoms. The disadvantage of these methods is that they require considerable time and labor costs, and are error-prone [11].

Reátegui and Ratté [11] and Wu et al [19] extracted symptoms using the entity naming and extraction method. It improves the disadvantages of the manual extraction method by using a computer to automatically extract named entities from a large amount of text to extract symptoms, but unstructured data needs to be associated with entities to extract symptoms. Therefore, there is a risk of errors in modeling and naming entity relationships, which can affect the results.

De Silva et al [20] used the machine learning least absolute shrinkage and selection operator (LASSO) or ridge regularization method to extract symptoms from unstructured text (Marketplace for Medical Information in Intensive Care III [MIMIC-III]) and to predict mortality in patients with diabetes. Since LASSO uses a machine learning approach, it does not require association with entities and saves manual time and labor cost.

Notably, these methods were limited to the extraction of symptoms [11,14-17,21,22] and did not extract words that express the degrees of these symptoms (hereinafter, "degree words"). These degree words also contain information about doctors and patients if extracted from unstructured data that can provide a basis for judgment in medical diagnostic systems, and therefore, it is necessary to process these data to further improve the judgment accuracy of the medical diagnostic system [22]. However, unlike structured data, these degree terms include the subjectivity of doctors. They do not have a unified standard and cannot be processed directly by using a computer [21].

In 1965, Zadeh [23] of the University of California, Berkeley published the first paper on fuzzy theory, submitting for the first time the concept of fuzzy sets from the perspective of set theory to describe things with fuzziness and to describe things in everyday life by fuzzy logical reasoning, similar to the human thinking patterns and the probability theory proposed by Zadeh [24], describing the difference between randomness and probability, which is considered as the second milestone in the development of fuzzy mathematics. The emergence of fuzzy theory has provided a solid theoretical foundation and effective tools for the wide application of fuzzy mathematics in pattern recognition and other fields.

The concept of fuzzy theory emphasizes the use of fuzzy logic to describe things in real life and to make up for the shortcomings of classical logic (binary logic), which cannot describe things with unclear boundaries. Human natural language is vague in its presentation, and it is difficult to fully describe real-world problems using the dichotomy of "right or wrong" and "good or bad." Therefore, fuzzy theory uses the definition of a fuzzy set to define and quantify the membership function (membership rank) of the degree to which an event belongs to this set, and solves different problems by quantifying the membership rank (membership value) [25].

Fuzzy c-means (FCM) is an unsupervised soft computing technology developed by Dunn [26] in 1973 and improved by Bezdek et al [27] in 1981. Unlike the hard clustering method, the soft clustering method uses fuzzy sets [23], which can solve the problem of text ambiguity better. In fuzzy sets, membership indicates the matching degree between the element and the set, with membership values ranging from 0 to 1. Furthermore, the concept of membership is extended in the FCM method, wherein the membership matrix represents the membership values of the elements in multiple clusters. FCM is one of the most commonly used methods to solve fuzzy problems. Compared with other clustering methods, it is more flexible and can represent the degree of data affiliation more appropriately [28]. The proposed method has two main advantages. First, it is unsupervised; therefore, labeled data is not required during training. Second, it can resolve subjective word ambiguity in fuzzy sets and medical texts such as a description of symptoms.

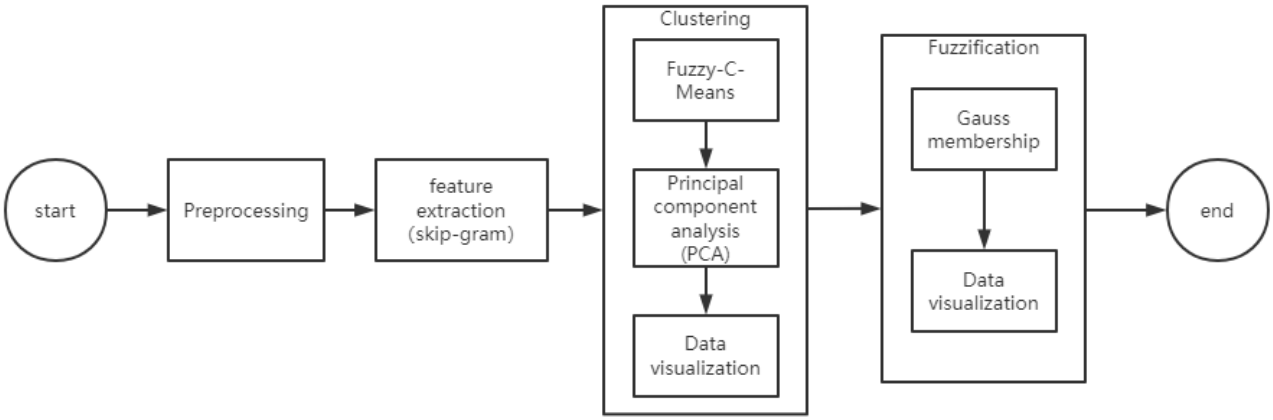
Therefore, we propose using the FCM and Gauss membership methods to quantify the subjective degree words in the English-interpreted report of the MIMIC-III data set.

Methods

Overview

The overall structure of the proposed method is shown in Figure 1. It contains five parts: preprocessing, feature extraction, clustering, digitization, and visualization. All the calculation methods used in this experiment were implemented in Python. First, the raw data are preprocessed. Next, a skip-gram is used to extract word features and convert them into computer-processable data. Thereafter, the processed data are clustered. Subsequently, clustered data are digitized. Finally, the results of clustering and digitization are visualized to make the data structure easier to understand.

Figure 1. Overall structure of the proposed method.



Data Set

The data set used in this study was an interpretation report of 522,279 physicians from the MIMIC-III data set [29]. Five lung diseases (atelectasis, pneumonia, pneumothorax, pulmonary edema, and pulmonary embolus) were selected from Normalized Clinical Knowledge (NCK) [30] as keywords for the searched MIMIC-III data set. The search results are presented in Table 1. The number of texts used in the experiment was 381,091.

Table 1. Search results for 5 diseases.

Keywords	Texts (N=381,091), n
Atelectasis	119,765
Pneumonia	72,553
Pneumothorax	138,785
Pulmonary edema	47,364
Pulmonary embolus	2624

Preprocessing

The input data set was preprocessed using normalization and Python code using the following tasks:

- Text classification: Filter text according to database keywords from NCK and classifications of various diseases
- Merge similar text: Merge all data for the same illness into one file to create a corpus of illnesses
- Remove special characters: All special characters (punctuation marks, question marks, exclamation marks, etc) are removed from the text and replaced with spaces
- Case exchange: Change all letters to lowercase to reduce vocabulary repetition
- Delete stop words: Delete stop words with no special meaning (eg, am, name)
- Tag words: To extract words, we used the Natural Language Toolkit to tag words.
- Word extraction: Degree words are primarily adjectives and adverbs, and illnesses are primarily nouns; therefore, only adjectives, adverbs, and nouns should be extracted.

Feature Extraction

Feature extraction must be used to convert natural language into computer-processable numerical data. We used word2vec skip-gram [31] to extract features from the collected data. For the word vector to reflect the contextual relationships, the data used in the word vector training were not the extracted words but the text with the stop words removed. The extracted words were used for cluster analysis. Furthermore, to exclude irrelevant words, we removed words with fewer than 10 occurrences. The dimension of the word vector was set to 100 [32]. This implies that each word was represented as a 1 × 100 vector. Text consisting of N words was represented as a matrix size of N × 100. These word vectors can contain the positional relationship and structural information of each word in the text.

Fuzzy C-Means

We used the FCM method to cluster the features. This method allows symptoms of the same disease in different texts to be grouped into the same category. Each element has a membership value for each category. The degree of membership depends on the distance from the element to the center of the cluster, and they are inversely proportional [33]. In this study, the number of clusters for each disease was set to 10. The fuzzy index m setting was set to 2 [34]. The loop was set to 50 times.

Fuzzification

To import the trained model as a basis for evaluation into a medical diagnostic system, digitization must be used to convert the FCM results into a numerical range. In this experiment, the Euclidean distance from each word to its center was used as a reference to digitize the word. The formula for calculating the distance is shown in equation 1.

The more popular membership functions are the triangular membership function and the Gaussian membership function, and since the center of the triangular affiliation function is too steep, this experiment uses the Gaussian membership function for the numerical range conversion. This membership function is given by equation 2. The center of each Gaussian membership function can be determined based on these distances, where x_{1k} is the kth coordinate value of the center and x_{2k} is the kth coordinate value of the element.

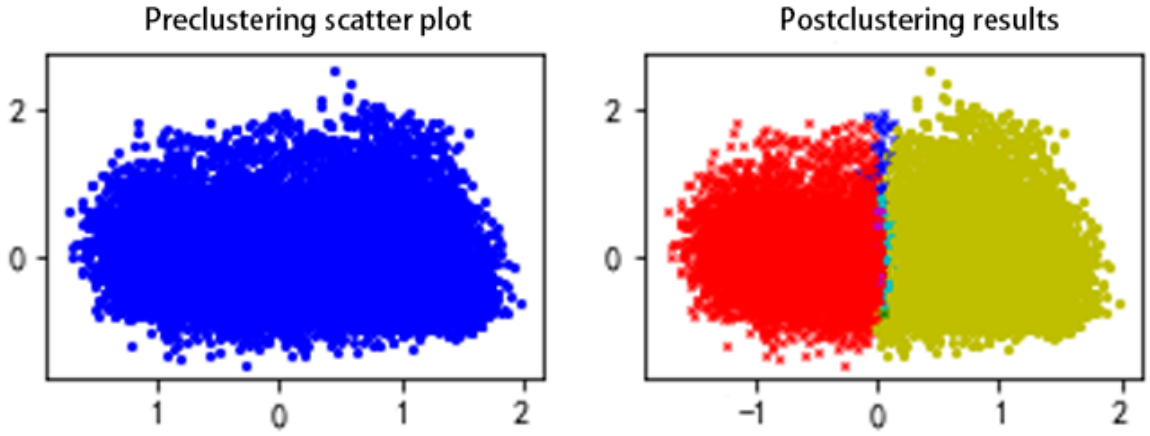
Where σ is the width parameter to control radial range of function. For clustering problems, if the clustering structure in the feature space of sample points is compact, the smaller σ can ensure the effect of the sample points clustering. If the clustering structure is dispersed, the larger σ can help to obtain the explicit membership function distribution. We set the σ value to 0.4.

It is necessary to ensure that the membership function takes values in the range of 0-1 in the fuzzy set [35]. Therefore, we set the boundary c to 0.1.

Principal Component Analysis and Data Visualization

In addition, it is necessary to visualize the results of the two parts (ie, clustering and digitization). To visualize the results after clustering, principal component analysis (PCA) must be used to reduce the dimensions of the data. To make the data easier to understand, we used PCA [36] to project high-dimensional data into a low-dimensional space. PCA can reduce the dimensions of data while retaining key information and increasing readability. The numerical results will be transformed into easy-to-understand 2D plots. Figure 2 is the visualization result of the disease atelectasis.

Figure 2. Visualization result of the disease atelectasis.



Results

This section introduces the results of this study. First, the results of the clustering process, data dimension reduction, and visualization are described. Thereafter, the results of the proposed fuzzy method are presented.

Fuzzy C-Means

Textbox 1 lists the results of clustering different diseases. The [Atelectasis] cluster contained degree words such as “quickly” and “with small,” and symptom words such as “obstruction.” The [Pneumonia] cluster contained degree words such as “small” and “mid,” site words such as “pulmonary,” and symptom words such as “effusions.” The [Pneumothorax] cluster contained symptom words such as “pain” and “fever,” and degree words such as “expansion.” The [Pulmonary edema] cluster contained symptom words such as “nausea” and “neutropenic,” and degree words such as “copious.” The [Pulmonary embolus] cluster contained degree words such as “expansion.” Different diseases will produce different degree words and symptoms. These degree words and symptoms were used to judge diseases. To facilitate understanding of the variability of degree words between diseases, we have chosen the same words for explanation below.

Textbox 1. Clustering results. The italics indicate extracted degree words that are processed through fuzzification afterward.

Atelectasis

"quickly," "over penetrated," "stenosis," "*small*," "qualified," "midthigh," "unlike," "hepatic," "wet massive," "hypodensities," "effusion with," "*high*," "obstruction," "ultra-fast," "tine"...

Pneumonia

"pleural," "lung," "*high*," "effusion," "pulmonary," "tube," "lobe," "lower," "contrast," "atelectasis," "unchanged," "*small*," "upper," "bilateral," "normal," "edema," "effusions," "opacity," "within," "consolidation," "stable," "mild," "increased," "consistent," "unremarkable," "clear," "moderate," "enlarged," "large," "low," "opacification," "well," "mid," "slightly," "improved," "significant," "improvement," "severe," "minimal," "slight," "decreased," "less"...

Pneumothorax

"pain," "fever," "worsening," "cough," "increasing," "expansion," "exacerbation," "vomiting," "*small*," "sputum," "febrile," "*high*," "syncope," "decreasing," "overdose"...

Pulmonary edema

"swelling," "*high*," "nausea," "diarrhea," "rising," "neutropenic," "prematurely," "copious," "malaise," "instability," "small"...

Pulmonary embolus

"thin," "reduced," "*small*," "inflated," "expansion," "*high*," "anomaly," "hyperostosis," "overinflation," "heights"...

Fuzzification

Figure 3 shows a membership diagram of the degree word "low." The x-axis represents the distance from the center. The y-axis represents the membership value. The median Gaussian membership is the distance from this element to the cluster's center. The greater the distance from the center point, the lower the degree of membership. The maximum value at the center point is 1.

Table 2 shows the central values for membership of the same word for the five illnesses. The differences between conditions can be determined by comparing the same vocabulary for different conditions. By comparing the words of each degree, we can identify the presenting characteristics of each symptom and create a table of the possible symptoms of the patient by comparing them to the patient's presenting characteristics.

Figure 3. Membership chart of the degree word low.

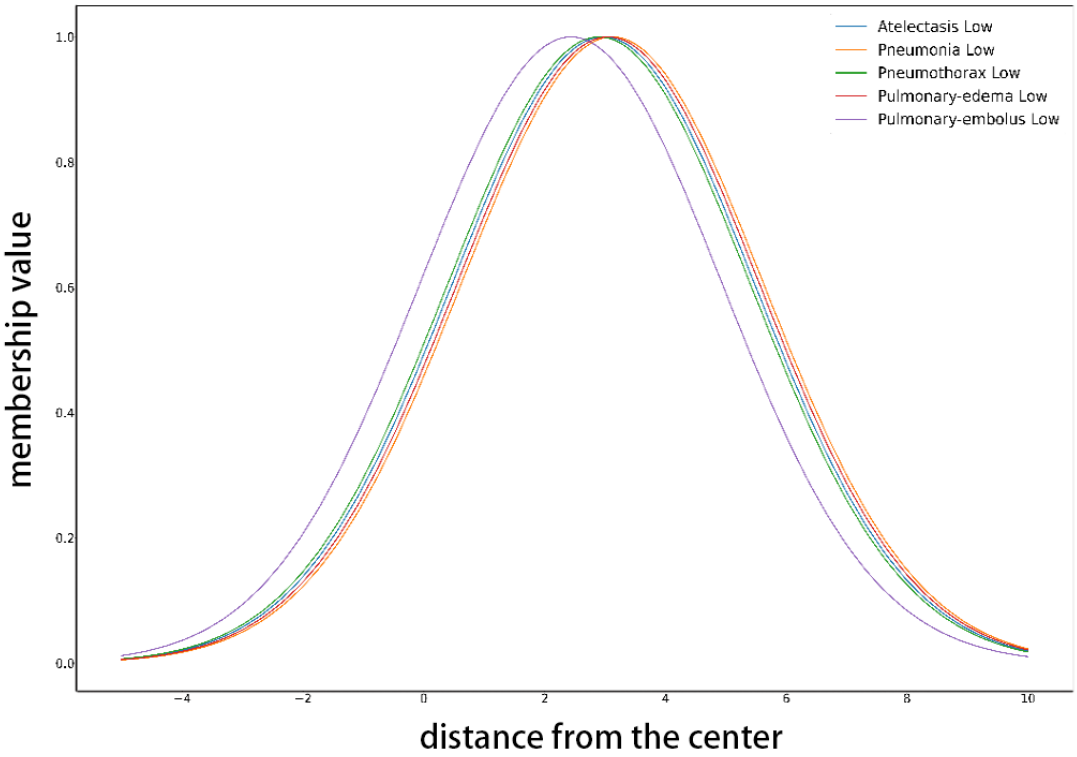


Table 2. Medians of the same word for 5 illnesses.

	Low, median	High, median	Small, median	Large, median
Atelectasis	2.9713	2.7161	2.4547	2.4047
Pneumonia	3.1217	2.9383	2.4354	2.4190
Pneumothorax	2.8998	2.8240	2.3956	2.3285
Pulmonary edema	3.0518	2.9880	2.4539	2.4003
Pulmonary embolus	2.4357	1.7908	2.6419	2.5570

Discussion

Principal Findings

Our proposed method is capable of identifying and extracting degree words from unstructured text (Textbox 1) and transform them into computer-recognizable membership values (Table 2). The results in Textbox 1 show that all extracted words contain the same subjective words (low, high, etc), which allows for an objective evaluation method. These same ambiguous words can be used as a new set of feature values to represent the disorders. That is, the same subjective words can be used to determine a disease. Thereafter, we quantified the extracted subjective words using the Gaussian membership method to transform the features contained in the ambiguous words into numerical features.

For the fuzzification section, we used the Gaussian fuzzy function to transform the fuzzy words extracted from the medical text into numbers. The results in Figure 3 and Table 2 show that it is feasible to transform the features contained in the subjective words into digital features. The digitization of ambiguous words (symptomatology and degree words) can improve the use of medical texts, which is crucial for improving the accuracy and interpretability of medical diagnostic systems. Furthermore, we will verify the specific impact of the quantification results for ambiguous words such as symptom words and degree words on the use of medical texts in subsequent studies to improve the accuracy and interpretability of medical diagnosis systems.

Comparison With Other Methods

Fodeh et al [16] used a random forest classifier to identify clinical notes with pain assessment information. They extracted subjective words describing pain intensity such as mild, moderate, and severe from the EHR and used them to help determine whether the text described pain. The results showed higher accuracy for the method using subjective words. The conclusion that subjective words can be used to further improve the medical diagnostic system is the same as in this study. Unlike this study, they only classified the text by subjective words and did not further use subjective words as a basis for a new condition determination. This study translates the subjective words into features of the disorder by fuzzification.

Matheny et al [17] pointed out the importance of symptoms and proposed a rule-based approach to extract them. This method is compared to this study, which requires using detailed rules, the creation of which is limited to the medical personnel. Manually writing rules requires considerable time and may not be exhaustive. The unsupervised method in this study requires no manual processing and no human involvement, reducing time consumption and extraction omissions. Additionally, while they note the importance of symptoms for the disorder, they do not extract and analyze the subjective words.

The method that uses a dictionary to extract symptoms also requires considerable time to construct and is not general [37]. For different tasks, rules need to be adapted. Additionally, the rule formulation process involves deleting and adding rules and discussing the rule coverage. Therefore, substantial time cost is required. Whereas, the FCM method is not limited to a specific scope and has excellent generality. The rules are summarized and extracted from the text, so they are more broadly applicable.

De Silva et al [20] used the LASSO method to predict mortality for patients with diabetes in MIMIC-III. The prediction results suggest that the clinical text provides a resource for future optimization and personalization of diabetes care. Additionally, the LASSO method has strong generalizability. However, a large amount of data is required to obtain more accurate results, but currently there are far less data with fine-grained labels than unlabeled data in the medical field. This method is equally effective for unlabeled data since it is an unsupervised method, but the impact on the accuracy of the results needs further discussion.

Reátegui and Ratté [11] used a named entity extraction system to extract symptoms. It extracts symptoms by automatically extracting named entities from large amounts of text using a computer, but it can create errors in modeling and named entity relationships that can affect the results. As the methods in this study do not need to be bound to named entities, these errors do not occur.

Wu et al [19] developed a deep neural network (DNN) to generate word vectors from a large, unlabeled, Chinese corpus by unsupervised learning. The results show that the DNN-based approach can capture grammatical features by word vectors and achieves higher performance compared with the traditional conditional random fields approach. The time spent will be substantially reduced by unsupervised learning. It is noted that further performance improvement will be achieved if high-dimensional discrete features are used. However, it is difficult to use DNNs as a basis for diagnostic systems because of their black box effect. However, this study is able to give the basis (affiliation) of the results while outputting the results through the affiliation function. Therefore, it can be used as a reference for medical practitioners.

Limitations and Further Study

This study has some limitations. The first is data. There were abbreviations and misspellings in the data, which were not extracted. The study has confirmed the presence of a large number of abbreviations in unstructured texts that contain equally large amounts of information [17]. We will try to extract and analyze the presence of abbreviated words in the data in a subsequent study. Additionally, the word2vec method we used has the limitation that the vector representation does not change when the words are used in different contexts. Due to the different clinical contexts of diseases, a comparative study using context-sensitive embedding algorithms is also needed. Finally, the results for each word calculated in this experiment need to be verified by the medical diagnostic system, and the method we propose will be verified in future studies.

This study is mainly about the training of the model and discusses whether it is possible to quantify the subjective word and use it as an indicator to determine the disease. In this study, we used data from different conditions to quantify the subjective words contained in them separately. Concerning future research, we envision using the quantified data to process hospital-specific data to give accuracy rates. For clinical application, we will perform disease prediction based on the subjective words in the data to get a list of possible diseases of the patient and give the correlation rate based on the fuzzy graph (Figure 3) generated in this study. Thus, it will assist physicians in diagnosis. Additionally, this method can solve the black box problem.

Conclusion

This study proposes an innovative method for handling subjective words. We used the FCM method to extract the subjective degree words in the English-interpreted report of MIMIC-III and then used the Gaussian functions to quantify the subjective degree words. In this method, words containing subjectivity in unstructured text can be automatically processed and transformed into numerical ranges by digital processing. The results show that membership medians of “low” were 2.971 (atelectasis), 3.121 (pneumonia), 2.899 (pneumothorax), 3.051 (pulmonary edema), and 2.435 (pulmonary embolus). This proves that the digitization of ambiguous words in medical texts is feasible. Our results proved the feasibility of using subjective words in medical texts. This makes it possible for medical researchers to further improve the use of electronic medical texts, thus affecting the accuracy and interpretability of the medical diagnostic system. In future research, we will consider how to extract abbreviations from abbreviation dictionaries and texts, as well as further validate the impact on the accuracy and interpretability of the medical diagnostic system.

Conflicts of Interest

None declared.

References

1. Shickel B, Tighe PJ, Bihorac A, Rashidi P. Deep EHR: a survey of recent advances in deep learning techniques for electronic health record (EHR) analysis. IEEE J Biomed Health Inform 2018 Sep;22(5):1589-1604 [FREE Full text] [doi: 10.1109/JBHI.2017.2767063] [Medline: 29989977]
2. Wu C, Kuo C, Su C, Wang S, Dai H. Using text mining to extract depressive symptoms and to validate the diagnosis of major depressive disorder from electronic health records. J Affect Disord 2020 Jan 01;260:617-623. [doi: 10.1016/j.jad.2019.09.044] [Medline: 31541973]
3. Sung S, Lin C, Hu Y. EMR-based phenotyping of ischemic stroke using supervised machine learning and text mining techniques. IEEE J Biomed Health Inform 2020 Oct;24(10):2922-2931. [doi: 10.1109/JBHI.2020.2976931] [Medline: 32142458]
4. Sijbers A, Schoemaker R, Nauta A, Alkema W. Revealing new leads for the impact of galacto-oligosaccharides on gut commensals and gut health benefits through text mining. Benef Microbes 2020 May 11;11(3):283-302. [doi: 10.3920/BM2019.0105] [Medline: 32299221]
5. Ciofi Degli Atti ML, Pecoraro F, Piga S, Luzi D, Raponi M. Developing a surgical site infection surveillance system based on hospital unstructured clinical notes and text mining. Surg Infect (Larchmt) 2020 Oct;21(8):716-721. [doi: 10.1089/sur.2019.238] [Medline: 32105569]
6. Ali F, El-Sappagh S, Islam SR, Kwak D, Ali A, Imran M, et al. A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion. Inf Fusion 2020 Nov;63:208-222. [doi: 10.1016/j.inffus.2020.06.008]
7. Chen E, Sarkar IN. Mining the electronic health record for disease knowledge. Methods Mol Biol 2014;1159:269-286. [doi: 10.1007/978-1-4939-0709-0_15] [Medline: 24788272]
8. Hashir M, Sawhney R. Towards unstructured mortality prediction with free-text clinical notes. J Biomed Inform 2020 Aug;108:103489 [FREE Full text] [doi: 10.1016/j.jbi.2020.103489] [Medline: 32592755]
9. Banerjee I, Madhavan S, Goldman RE, Rubin DL. Intelligent word embeddings of free-text radiology reports. AMIA Annu Symp Proc 2017;2017:411-420 [FREE Full text] [Medline: 29854105]
10. Korach Z, Yang J, Rossetti SC, Cato KD, Kang MJ, Knaplund C, et al. Mining clinical phrases from nursing notes to discover risk factors of patient deterioration. Int J Med Inform 2020 Mar;135:104053 [FREE Full text] [doi: 10.1016/j.ijmedinf.2019.104053] [Medline: 31884312]

11. Reátegui R, Ratté S. Comparison of MetaMap and cTAKES for entity extraction in clinical notes. BMC Med Inform Decis Mak 2018 Sep 14;18(Suppl 3):74 [FREE Full text] [doi: 10.1186/s12911-018-0654-2] [Medline: 30255810]
12. Demoulin NTM, Coussement K. Acceptance of text-mining systems: the signaling role of information quality. Inf Manage 2020 Jan;57(1):103120. [doi: 10.1016/j.im.2018.10.006]
13. Hyun S, Johnson SB, Bakken S. Exploring the ability of natural language processing to extract data from nursing narratives. Comput Inform Nurs 2009;27(4):215-23; quiz 224 [FREE Full text] [doi: 10.1097/NCN.0b013e3181a91b58] [Medline: 19574746]
14. Jagannatha A, Yu H. Structured prediction models for RNN based sequence labeling in clinical text. Proc Conf Empir Methods Nat Lang Process 2016 Nov;2016:856-865 [FREE Full text] [doi: 10.18653/v1/d16-1082] [Medline: 28004040]
15. Jagannatha A, Yu H. Bidirectional RNN for medical event detection in electronic health records. Proc Conf 2016 Jun;2016:473-482 [FREE Full text] [doi: 10.18653/v1/n16-1056] [Medline: 27885364]
16. Fodeh S, Finch D, Bouayad L, Luther SL, Ling H, Kerns RD, et al. Classifying clinical notes with pain assessment using machine learning. Med Biol Eng Comput 2018 Jul;56(7):1285-1292 [FREE Full text] [doi: 10.1007/s11517-017-1772-1] [Medline: 29280092]
17. Matheny ME, Fitzhenry F, Speroff T, Green JK, Griffith ML, Vasilevskis EE, et al. Detection of infectious symptoms from VA emergency department and primary care clinical documentation. Int J Med Inform 2012 Mar;81(3):143-156. [doi: 10.1016/j.ijmedinf.2011.11.005] [Medline: 22244191]
18. Tamang S, Patel MI, Blayney DW, Kuznetsov J, Finlayson SG, Vetteth Y, et al. Detecting unplanned care from clinician notes in electronic health records. J Oncol Pract 2015 May;11(3):e313-e319 [FREE Full text] [doi: 10.1200/JOP.2014.002741] [Medline: 25980019]
19. Wu Y, Jiang M, Lei J, Xu H. Named entity recognition in Chinese clinical text using deep neural network. Stud Health Technol Inform 2015;216:624-628 [FREE Full text] [Medline: 26262126]
20. De Silva K, Mathews N, Teede H, Forbes A, Jönsson D, Demmer RT, et al. Clinical notes as prognostic markers of mortality associated with diabetes mellitus following critical care: a retrospective cohort analysis using machine learning and unstructured big data. Comput Biol Med 2021 May;132:104305. [doi: 10.1016/j.compbiomed.2021.104305] [Medline: 33705995]
21. Chen L, Song L, Shao Y, Li D, Ding K. Using natural language processing to extract clinically useful information from Chinese electronic medical records. Int J Med Inform 2019 Apr;124:6-12. [doi: 10.1016/j.ijmedinf.2019.01.004] [Medline: 30784428]
22. Raghavan P, Chen JL, Fosler-Lussier E, Lai AM. How essential are unstructured clinical narratives and information fusion to clinical trial recruitment? AMIA Jt Summits Transl Sci Proc 2014;2014:218-223 [FREE Full text] [Medline: 25717416]
23. Zadeh L. Fuzzy sets. Inf Control 1965 Jun;8(3):338-353. [doi: 10.1016/s0019-9958(65)90241-x]
24. Zadeh L. Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets Syst 1978 Jan;1(1):3-28. [doi: 10.1016/0165-0114(78)90029-5]
25. Zimmermann H. Fuzzy set theory. WIREs Computational Statistics 2010 Apr 16;2(3):317-332. [doi: 10.1002/wics.82]
26. Dunn JC. A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. J Cybern 1973 Jan;3(3):32-57. [doi: 10.1080/01969727308546046]
27. Bezdek JC, Ehrlich R, Full W. FCM: The fuzzy c-means clustering algorithm. Comput Geosciences 1984 Jan;10(2-3):191-203. [doi: 10.1016/0098-3004(84)90020-7]
28. Principles of pattern recognition I (introduction and uses). NPTEL. 2014. URL: https://nptel.ac.in/courses/106106046 [accessed 2022-11-10]
29. Johnson AEW, Pollard TJ, Shen L, Lehman LH, Feng M, Ghassemi M, et al. MIMIC-III, a freely accessible critical care database. Sci Data 2016 May 24;3:160035. [doi: 10.1038/sdata.2016.35] [Medline: 27219127]
30. Nakaya J, Koji S, Hiroshi T. Condensed cross-clinical knowledge. Int J Computer Sci Network Secur 2006;6(7A):6-11.
31. Mikolov T, Chen K, Corrado G, Dean J. Efficient estimation of word representations in vector space. arXiv Preprint posted online on September 7, 2013. [doi: 10.48550/arXiv.1301.3781]
32. Wang Y, Liu S, Afzal N, Rastegar-Mojarad M, Wang L, Shen F, et al. A comparison of word embeddings for the biomedical natural language processing. J Biomed Inform 2018 Nov;87:12-20 [FREE Full text] [doi: 10.1016/j.jbi.2018.09.008] [Medline: 30217670]
33. Mantilla L, Yari Y. FCM algorithm: analysis of the membership function influence and its consequences for fuzzy clustering. In: Orjuela-Cañón AD, Lopez J, Arias-Londoño JD, Figueroa-García JC, editors. Applications of Computational Intelligence Third IEEE Colombian Conference, ColCACI 2020, Cali, Colombia, August 7-8, 2020, Revised Selected Papers. Cham: Springer; 2021:119-132.
34. Anam SA, Rahman AMM, Saleheen NN, Arif H. Automatic text summarization using fuzzy C-means clustering. 2018 Presented at: 2018 Joint 7th International Conference on Informatics, Electronics & Vision and 2018 2nd International Conference on Imaging, Vision & Pattern Recognition; June 25-29, 2018; Kitakyushu, Japan. [doi: 10.1109/iciev.2018.8641055]
35. Sadollah A. Introductory chapter: which membership function is appropriate in fuzzy system? In: Sadollah A, editor. Fuzzy Logic Based in Optimization Methods and Control Systems and Its Applications. London: IntechOpen Limited; 2018:1.

36. Hess AS, Hess JR. Principal component analysis. Transfusion 2018 Jul;58(7):1580-1582. [doi: 10.1111/trf.14639] [Medline: 29732564]
37. Iqbal E, Mallah R, Rhodes D, Wu H, Romero A, Chang N, et al. ADEPt, a semantically-enriched pipeline for extracting adverse drug events from free-text electronic health records. PLoS One 2017;12(11):e0187121 [FREE Full text] [doi: 10.1371/journal.pone.0187121] [Medline: 29121053]

Abbreviations

DNN: deep neural network
FCM: fuzzy c-means
LASSO: least absolute shrinkage and selection operator
MIMIC-III: Marketplace for Medical Information in Intensive Care III
NCK: Normalized Clinical Knowledge
PCA: principal component analysis

Edited by A Mavragani; submitted 13.04.22; peer-reviewed by K Ding, B Ru; comments to author 15.09.22; revised version received 29.09.22; accepted 24.10.22; published 18.11.22.

Please cite as:
Han F, Zhang Z, Zhang H, Nakaya J, Kudo K, Ogasawara K
Extraction and Quantification of Words Representing Degrees of Diseases: Combining the Fuzzy C-Means Method and Gaussian Membership
JMIR Form Res 2022;6(11):e38677
URL: https://formative.jmir.org/2022/11/e38677
doi:10.2196/38677
PMID:36399376

Original Paper

Regular Testing of HIV and Sexually Transmitted Infections With Self-Collected Samples From Multiple Anatomic Sites to Monitor Sexual Health in Men Who Have Sex With Men: Longitudinal Study

Ngai Sze Wong[1,2], PhD; Tsz Ho Kwan[2], PhD; Denise P C Chan[2], PhD; Grace C Y Lui[3], MBChB; Shui Shan Lee[2], MD

[1]Jockey Club School of Public Health and Primary Care, The Chinese University of Hong Kong, Hong Kong, China

[2]Stanley Ho Centre for Emerging Infectious Diseases, The Chinese University of Hong Kong, Hong Kong, China

[3]Department of Medicine & Therapeutics, The Chinese University of Hong Kong, Hong Kong, China

Corresponding Author:
Shui Shan Lee, MD
Stanley Ho Centre for Emerging Infectious Diseases
The Chinese University of Hong Kong
2/F, Postgraduate Education Centre
Prince of Wales Hospital, Shatin
Hong Kong
China
Phone: 86 22528812
Fax: 86 263549777
Email: sslee@cuhk.edu.hk

Abstract

Background: Regular HIV and sexually transmitted infection (STI) testing for men who have sex with men (MSM) is an important means of infection prevention, the adoption of which remains suboptimal in the community.

Objective: On the hypothesis that engagement plays an important role in sexual health monitoring, this study aimed to pilot-test internet-based HIV and STI testing with self-sampling to enhance engagement of MSM with regular testing.

Methods: This 1-year cohort study was conducted on HIV-negative MSM aged 18 years or older. A designated website was set up to enable participants to make appointments for baseline and follow-up visits at 3-monthly intervals. On-site blood sampling was performed for HIV and syphilis tests, along with self-collection of pharyngeal swabs, rectal swabs, and urine samples for *Chlamydia trachomatis* (CT) and *Neisseria gonorrhoeae* (NG) testing. Full engagement, as defined by having made at least 3 visits over a 6-12 months' follow-up period, was compared with partial engagement in the bivariable logistic regression model.

Results: Between August 2019 and October 2020, 204 MSM were recruited, after the exclusion of 2 baseline HIV-positive MSM. The majority (189/204, 92.7%) were Chinese, the median age was 31 (IQR 26-39) years, and 58.0% (116/200) had experience with pre-exposure prophylaxis (PrEP) at baseline. Full engagement (146/204, 71.6%) was associated with incident STI during the follow-ups (odds ratio [OR] 4.23, 95% CI 1.63-10.94), seeking a medical referral after STI detection (OR 10.25, 95% CI 3.25-29.79), and a synchronized schedule of HIV and STI testing with PrEP visits (OR 51.85, 95% CI 19.30-139.34). No incident HIV was detected in the follow-up period. At baseline, the overall STI (CT, NG, or syphilis) prevalence was 30%, with CT at 18%, NG at 13%, and syphilis at 5%. During follow-up, the incidences were 59.08/100 person-years (py) for any STI, 33.05/100 py for CT, 29.86/100 py for NG, and 10.4/100 py for syphilis. The detection rates of CT and NG in urine samples were lower than with pharyngeal swabs and rectal swabs. The scores for convenience, confidence of correct sampling, and accuracy of self-sampling were high (7 to 8 out of 10).

Conclusions: Both baseline prevalence and incidence of STI were high among MSM engaged in regular testing. A high degree of engagement in regular STI and HIV testing was positively associated with incident STI, history of health-seeking behaviors, and perceived convenience of self-sampling. Self-sampling could be introduced as a means of enhancing engagement in regular HIV and STI testing.

(JMIR Form Res 2022;6(11):e40996) doi:10.2196/40996

KEYWORDS

HIV testing; STI testing; self-sampling; internet-based testing; men who have sex with men; HIV; monitoring; sex; infection; prevention; community; engagement; cohort study; testing

Introduction

Worldwide, the prevalence of sexually transmitted infection (STI) is high in men who have sex with men (MSM), with a pooled syphilis prevalence of 7.5% (95% CI 7%-8%) [1], rectal gonorrhea prevalence of 6.1%, and rectal chlamydia prevalence of 9% [2], as summarized in systematic reviews. Although syndromic management of STI is advocated to facilitate care for infected persons, especially in limited-resource settings [3], the approach has a shortcoming of ignoring asymptomatic infections, which could be a driver of ongoing transmission [4]. In response to the high STI prevalence and high proportion of asymptomatic infections in MSM, HIV and STI testing at 3 to 6-monthly intervals has been recommended in guidelines from the US Centers for Disease Control and Prevention [5], professional bodies in other countries including Australia [6], and the Scientific Committee on AIDS and STI in Hong Kong [7]. However, adherence of MSM to regular testing, which involves undergoing HIV and STI tests at regular intervals, may not be high. In Australia, adherence of MSM to regular testing was poor, with an HIV and STI retesting rate of less than 40% at 1 year and below 20% within 6 months [8]. Similarly, the proportion of MSM undergoing regular STI testing at 6 months and 1 year was low in Europe [9-11].

Barriers such as cost, inconvenience, previous negative testing results, unawareness, confidentiality concerns, and lack of time might hinder uptake of HIV and STI tests, especially with regards to regular testing [12]. Apart from cost, the required sampling method for testing is an important factor affecting engagement in retesting [12]. Unlike urine, which is easily self-sampled, collection of pharyngeal and rectal swabs is commonly performed by health care workers in clinical settings. Self-sampled testing for HIV, *Chlamydia trachomatis* (CT), and *Neisseria gonorrhoeae* (NG) are preferred by some MSM as it overcomes the barriers of privacy concerns, time, and geographic limitations of services provided [12]. The cost of self-sampled testing for MSM could be lower by eliminating clinic attendance if the delivery cost is nominal. Combining an online service with self-sampling and self-testing may further reduce the burden of an on-site service, reserving the capacity for complex clinical management. The transition from on-site to online services for self-sampling and self-testing of STI has been shown to be effective in previous studies in the United Kingdom [13,14]. Also, a previous study of human papillomavirus (tested via oral fluid, penis, rectum specimens) in China and a study of CT and NG (tested via extragenital samples) in the United Kingdom showed comparable accuracy and consistency between self-sampled and clinician-sampled specimens in MSM [15,16].

In Hong Kong, MSM accounted for less than 20% of attendees at Social Hygiene Clinics, which are the government's STI service. The prevalence of asymptomatic CT and NG infections in MSM attendees in 2014-2015 was 19.6% [17]. The proportion of MSM with newly diagnosed CT and NG infection has increased over time [7]. We hypothesized that engagement plays an important role in the participation of MSM in regular HIV and STI testing. In this study, we aimed to pilot-test internet-based self-sampled HIV and STI testing from multiple anatomic sites to enhance engagement of MSM in regular sexual health monitoring.

Methods

Participants and Study Design

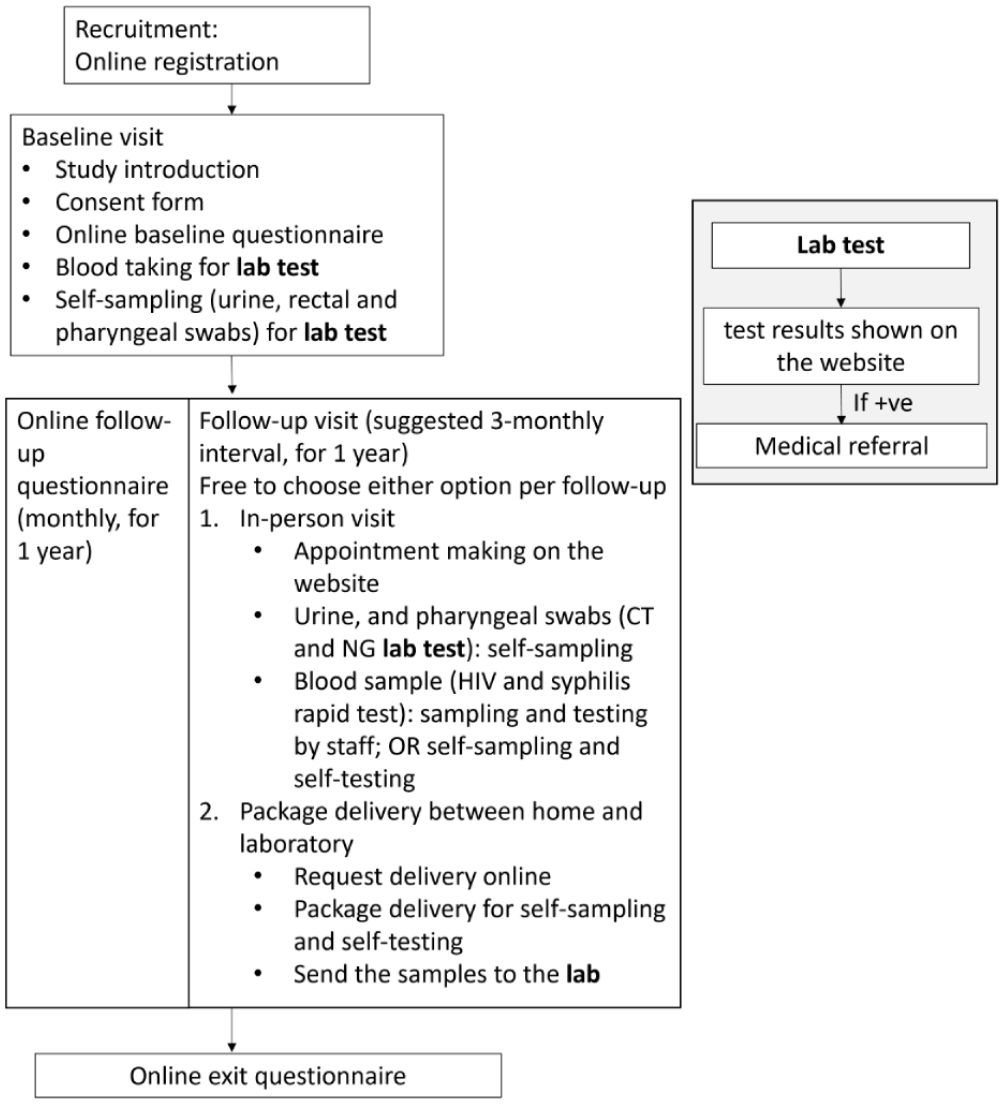
In this 1-year cohort study, MSM were recruited to participate in regular testing of HIV and STIs at 3-monthly intervals. Men aged 18 years or older who had had sex with men in the past 12 months, normally lived in Hong Kong, had a negative HIV test result, and could communicate in written and spoken English or Chinese were eligible to join the study. Recruitment was performed via referral from community-based organizations, online advertisements, peers, and an HIV pre-exposure prophylaxis (PrEP) pilot study conducted by the research team. Written informed consent was obtained from each participant.

In this study, a designated website was created to describe the study procedures and provide materials about HIV and STI prevention. After registration, the same website allowed participants to complete self-administered questionnaires, make appointments for HIV and STI testing visits, and check their HIV and STI test results (Figure 1). Monthly reminders to complete questionnaires and appointment confirmation emails were sent automatically. Help desk service supported by the research team was available via WhatsApp, email, and phone calls.

With written consent at the baseline visit, participants self-collected pharyngeal swabs, rectal swabs, and urine samples for CT and NG testing by following the procedures depicted on an instruction sheet (manufacturer's materials). Blood samples for HIV and syphilis point-of-care (POC) tests were collected by venesection. At follow-up visits, participants followed the same procedures as for the baseline visit, except that finger-prick blood samples were either self-sampled by the participants with or without assistance from research staff or venesection by a doctor. Dried blood spots (DBS) were collected using a Whatman card for HIV and syphilis laboratory testing and archiving. As an alternative to attending follow-up visits, courier delivery of kits for self-sampling for CT and NG testing and HIV and syphilis self-testing was offered. Participants were asked to upload photos of self-testing results to the system or through WhatsApp and return the self-sampled specimens for free to the laboratory through the designated courier service. Test results were inputted to the system when available so the participants could access them. Medical referral was made upon request.

Participants were asked to complete a baseline questionnaire (sociodemographic information, sexual behavioral history, PrEP use, sex networking, HIV and STI testing history, and the latest use of sexual health services), monthly follow-up questionnaires (sexual behavioral history, sex networking, HIV and STI testing history and results, scoring of the self-sampling process if participated in follow-up visits, sexual health service utilization, and PrEP use in the past 1 month), and an exit questionnaire (future plans for HIV and STI testing, preferred channels of testing, history of checking HIV and STI testing results in the system, scoring of the self-sampling process) at the end of the study. Scoring on the self-sampling process ranged between 1 (strongly disagree) and 10 (strongly agree). Chemsex engagement was defined as the use of recreational drugs before or during sex, excluding single use of either poppers or erectile dysfunction agents. Incentives via a HK $25 (US $3.20) voucher were offered to participants who had attended 3 or more visits over a 1-year period.

Figure 1. Study flowchart.



Point-of-Care Testing and Laboratory Testing

The fourth generation Alere HIV Combo test was used for POC HIV testing. Participants never diagnosed with syphilis were tested with the SD BIOLINE Syphilis 3.0 (treponemal test). For participants who previously tested positive or were ever diagnosed with syphilis, their samples were further tested using the CHEMBIO DPP Syphilis Screen & Confirm (treponemal and nontreponemal test). To confirm the detection of HIV and syphilis infection, tests were performed on the collected DBS at the research laboratory. Briefly, a spot was punched from the DBS and suspended in 200 µl of phosphate-buffered saline in the well of a flat-bottomed uncoated microtiter plate. The plate was incubated at 4° C for 2 hours, and 150 µl of the eluate was subsequently used as the specimen for testing. Eluted DBS was separately tested for HIV and syphilis antibody using a commercial ELISA kit (Murex HIV AG/Ab Combo and Murex ICE-Syphilis, DiaSorin, Saluggia VC, Italy, respectively) following the manufacturers' instructions. CT and NG nucleic acid amplification tests were performed on 3 types of specimens, namely, urine, rectal swabs, and pharyngeal swabs. CT and NG were detected using the Aptima Combo 2 Assay (Hologic, Marlborough, MA).

Ethical Approval

Ethical approval from the Joint Chinese University of Hong Kong-New Territories East Cluster Clinical Research Ethics Committee was obtained (approval number: CREC2019.167).

Data Analysis

MSM participants who had made at least 3 visits spanning over at least 180 days were regarded as fully engaged, while the rest were classified as partially engaged. Baseline characteristics and participation experiences during the follow-up period were compared between the 2 engagement groups in bivariable logistic regression models. Scores for the self-sampling processes were summarized, and their changes (dichotomized: 8-10 points for convenience, confidence, and accuracy coded as 1; 3-10 points for discomfort coded as 1) by visit were examined in a binary logistic generalized estimating equation (GEE), following unstructured working correlation matrices. Sensitivity analyses were performed with different cut-off values for the dependent variables in the GEE models.

We estimated the baseline STI prevalence by dividing the number of MSM with a positive STI result (including any sites for CT, any sites for NG, or reactive syphilis or HIV antibody) by the total number of MSM with baseline STI test results. The 95% CI was estimated using a binomial exact test. Incident STI referred to the first positive STI result after a negative result. Time to event was the interval (in months) from the baseline visit to either the date of an incident STI or the last visit, whichever was earlier. Incidence and 95% CI were calculated based on the Poisson distribution assumption. All statistical analyses were performed in SPSS 25 (IBM Corp, Armonk, NY), and complete case analyses were performed.

Results

Characteristics of Participants at Baseline

Between August 10, 2019, and October 9, 2020, a total of 242 participants were registered in the study; 206 MSM had provided consent and attended the baseline visit. Per the eligibility criteria, the 2 MSM who tested HIV-positive at baseline (2/206, 1%; 95% CI 0%-2%) were excluded from the study. Among the 204 successfully recruited MSM, the median age was 31 (IQR 26-39) years, 92.7% (189/204) were local Chinese, 92.5% (185/200) had attained a secondary educational level or higher, and 80.6% (162/201) were employed (Table 1). Most participants (158/204, 77.5%) had a synchronized schedule for HIV and STI testing with the PrEP study visit (ie, co-enrollment in the PrEP study). At baseline, 38.0% (76/200) reported a history of chemsex engagement, 87.0% (174/200) reported a history of group sex, 36.4% (68/187) reported a history of an STI diagnosis, 58.0% (116/200) had experience with PrEP, and 19.6% (39/199) had previously taken post-exposure prophylaxis (PEP) against HIV.

At baseline, the overall STI prevalence was 30% (95% CI 24%-36%), with CT at 18% (95% CI 12%-23%), NG at 13% (95% CI 8%-17%), and syphilis at 5% (95% CI 2%-8%). By anatomic site, the highest proportion of samples with a positive result was rectal swabs, at 12.9% (26/201) for CT, and pharyngeal swabs, at 9.0% (18/199) for NG. The lowest proportion of samples testing positive was in urine samples, at only 3.9% (8/202) for CT and 1.5% (3/201) for NG.

Table 1. Characteristics of recruited men who have sex with men (n=204).

Variables	Results
Sociodemographic variables	
Age (years), median (IQR) (n=203)	31 (26-39)
Recruitment source (n=200), n (%)	
NGOs[a]	69 (34.5)
Other research projects	109 (54.5)
Online	10 (5.0)
Peers	12 (6.0)
Local Chinese (n=204), n (%)	
No	15 (7.4)
Yes	189 (92.6)
Education level (n=200), n (%)	
Secondary or below	15 (7.5)
Postsecondary	185 (92.5)
Employed (n=200), n (%)	
No	38 (19.0)
Yes	162 (81.0)
Monthly income level (HK $[b]; n=200), n (%)	
0	17 (8.5)
1-15,000	40 (20.0)
15,001-30,000	72 (36.0)
30,001-50,000	42 (21.0)
>50,000	29 (14.5)
Channels ever used for seeking sex partners (n=200), n (%)	
Local physical venue	142 (71.0)
Overseas physical venue	79 (39.5)
Virtual venue	186 (93.0)
History of sexual behavior (n=200)	
Group sex, n (%)	
Never	26 (13.0)
Ever	174 (87.0)
Chemsex engagement, n (%)	
Never	124 (62.0)
Ever	76 (38.0)
Sex with a woman, n (%)	
Never	170 (85.0)
Ever	30 (15.0)
STI[c] diagnosis (n=187), n (%)	
Never	119 (63.6)
Ever	68 (36.4)
HIV prevention measures	
Knowledge and usage of PrEP[d] (n=200), n (%)	

Variables	Results
Never heard of	2 (1.0)
Have heard of but never taken	82 (41.0)
Have taken	116 (58.0)
Knowledge and usage of PEP[e] (n=198), n (%)	
Never heard of	3 (1.5)
No	156 (78.8)
Yes	39 (19.7)

[a]NGOs: nongovernmental organizations.

[b]US $1=HK $7.8.

[c]STI: sexually transmitted infection.

[d]PrEP: pre-exposure prophylaxis.

[e]PEP: post exposure prophylaxis.

Self-Sampling Preference for Regular STI Testing at Baseline

From the baseline questionnaire (n=200), 65.5% (131/200) of MSM had heard of self-sampling for STI or HIV testing, and 82.5% (165/200) indicated a willingness to try (Table 2). Most MSM preferred POC testing with samples collected by a trained staff member (165/200, 82.5%), followed by self-testing (103/200, 51.5%) and sampling by health care workers with consequent laboratory testing (64/200, 32.0%). The lowest proportion preferred self-sampling with consequent laboratory testing (42/200, 21.0%). Most MSM supported regular HIV testing (192/194, 99.0%) and STI testing (187/194, 96.4%) at least once every 2 years, with more than one-half expecting to undergo testing once every 3 months (149/194, 76.8% for HIV testing; 129/194, 66.5% for STI testing). Among all sampling methods, finger prick, saliva, and urine were well-accepted (>70% MSM replied each was totally acceptable) by the participants, while urethral sampling was the least acceptable method (58/189, 30.7% of MSM replied it was totally acceptable).

Table 2. Preference of HIV and sexually transmitted infection (STI) testing at baseline (n=200).

Preferences	Results, n (%)
Testing frequency (n=194)	
Appropriate HIV testing frequency	
Test whenever necessary	2 (1.0)
Once every 2 years	1 (0.5)
Once per year	3 (1.5)
Once every 6 months	24 (12.4)
Once every 3 months	149 (76.8)
Once per month	15 (7.7)
Appropriate STI testing frequency	
Test whenever necessary	7 (3.6)
Once every 2 years	2 (1.0)
Once per year	10 (5.2)
Once every 6 months	34 (17.5)
Once every 3 months	129 (66.5)
Once per month	12 (61.9)
Heard of self-sampled HIV and STI test (n=200)	
No	69 (34.5)
Yes	131 (65.5)
Willing to try self-sampling for HIV and STI testing (n=200)	
Yes, even if paid	28 (14.0)
Yes, if free	137 (68.5)
Not decided yet	21 (10.5)
No	14 (7.0)
Type of sampling and testing preferred (n=200)	
Self-test	103 (51.5)
Point-of-care test with samples collected by a trained staff member	165 (82.5)
Self-sampling and tested by a lab	42 (21.0)
Sampled by an HCW[a] and tested by a lab	64 (32.0)
Level of acceptability of the different sampling methods	
Giving blood (n=198)	
Totally acceptable	116 (58.6)
Acceptable	73 (36.9)
Unacceptable	4 (2.0)
Totally unacceptable	5 (2.5)
Finger prick (n=197)	
Totally acceptable	147 (74.6)
Acceptable	44 (22.3)
Unacceptable	1 (0.5)
Totally unacceptable	5 (2.5)
Urine sample (n=199)	
Totally acceptable	153 (76.9)

Preferences	Results, n (%)
Acceptable	42 (21.1)
Unacceptable	0 (0)
Totally unacceptable	4 (2.0)
Saliva sample (n=199)	
Totally acceptable	155 (77.9)
Acceptable	40 (20.1)
Unacceptable	0 (0)
Totally unacceptable	4 (2.0)
Rectal swab (n=193)	
Totally acceptable	94 (48.7)
Acceptable	87 (45.1)
Unacceptable	8 (4.2)
Totally unacceptable	4 (2.1)
Pharyngeal swab (n=190)	
Totally acceptable	99 (52.1)
Acceptable	76 (40.0)
Unacceptable	11 (5.8)
Totally unacceptable	4 (2.1)
Urethral swab (n=189)	
Totally acceptable	58 (30.7)
Acceptable	88 (46.6)
Unacceptable	32 (16.9)
Totally unacceptable	11 (5.8)

[a]HCW: health care worker.

Comparison of Participants by Study Engagement Level

Regarding study visits, among 204 participants, 28 (13.7%) attended the baseline visit only, while 29 (14.2%) visited twice, and 147 (72.1%) attended 3 or more visits including 1 visit with a follow-up period shorter than 180 days (Multimedia Appendix 1). Baseline characteristics were not significantly different between MSM with full (at least 3 visits) and partial (less than 3 visits) engagement (Table 3). MSM with an incident STI in the follow-up period (odds ratio [OR] 4.23, 95% CI 1.63-10.94), who had ever sought a medical referral following STI detection during the study (OR 10.25, 95% CI 3.25-29.79), and who had a synchronized schedule of HIV and STI testing with the PrEP visit (OR 51.85, 95% CI 19.30-139.34) were more likely to be fully engaged than partially engaged in the study.

Table 3. Comparison of participants' baseline characteristics by study engagement level.

Characteristics	Partial engagement (<3 visits; n=58), n (%)	Full engagement (≥3 visits; n=146), n (%)	Bivariable logistic regression analysis	
			OR[a] (95% CI)	*P* value
Sociodemographic variables				.76
Age (years)[b,c]	30 (25-37)	31 (27-39)	1.01 (0.97-1.04)	
Local Chinese[c]				.87
No	4 (6.9)	11 (7.5)	Reference	
Yes	54 (93.1)	135 (92.5)	0.91 (0.28-2.98)	
Education level[d]				.98
Secondary or below	4 (7.4)	11 (7.5)	Reference	
Postsecondary	50 (92.6)	135 (92.5)	0.98 (0.30-3.23)	
Employed[d]				.27
No	13 (24.1)	25 (17.1)	Reference	
Yes	41 (75.9)	121 (82.9)	1.53 (0.72-3.27)	
Monthly income level (HKD)[d,e]				
0	5 (9.3)	12 (8.2)	Reference	—[f]
1-15,000	16 (29.6)	24 (16.4)	0.63 (0.18-2.12)	.45
15,001-30,000	14 (25.9)	58 (39.7)	1.73 (0.52-5.70)	.37
30,001-50,000	10 (18.5)	32 (21.9)	1.33 (0.38-4.71)	.66
>50,000	9 (16.7)	20 (13.7)	0.93 (0.25-3.42)	.91
Channels ever used for seeking sex partners[d]				
Local physical venue	35 (64.8)	107 (73.3)	1.49 (0.76-2.91)	.24
Overseas physical venue	19 (35.2)	60 (41.1)	1.29 (0.67-2.46)	.45
Virtual venue	51 (94.4)	135 (92.5)	0.72 (0.19-2.69)	.63
History of sexual behavior				
History of group sex[d]				.06
Never	11 (20.4)	15 (10.3)	Reference	
Ever	43 (79.6)	131 (89.7)	2.23 (0.95-5.23)	
History of chemsex engagement[d]				.42
Never	31 (57.4)	93 (63.7)	Reference	
Ever	23 (42.6)	53 (36.3)	0.77 (0.41-1.45)	
History of STI[g] diagnosis[h]				.45
Never	34 (68.0)	85 (62.0)	Reference	
Ever	16 (32.0)	52 (38.0)	1.30 (0.65-2.58)	
HIV and STI testing				
Acceptance of regular HIV testing[i]				—
No	2 (3.8)	0 (0)	—	
Yes	51 (96.2)	141 (100)	—	
Acceptance of regular STI testing[i]				.36
No	3 (5.7)	4 (2.8)	Reference	
Yes	50 (94.3)	137 (97.2)	2.06 (0.44-9.51)	

Characteristics	Partial engagement (<3 visits; n=58), n (%)	Full engagement (≥3 visits; n=146), n (%)	Bivariable logistic regression analysis	
			OR[a] (95% CI)	*P* value
Types of sampling and testing preferred[d]				
Self-test	28 (51.9)	75 (51.4)	0.98 (0.53-1.83)	.95
Point-of-care test with samples collected by a trained staff member	49 (90.7)	116 (79.5)	0.39 (0.14-1.08)	.07
Self-sampling and testing by a lab	10 (18.5)	32 (21.9)	1.24 (0.56-2.72)	.60
Sample taken by an HCW[j] and tested by a lab	11 (20.4)	53 (36.3)	2.23 (1.06-4.68)	.03
Research outcomes and participation				
Baseline positive STI result[c]				.37
No	38 (65.5)	105 (71.9)	Reference	
Yes	20 (34.5)	41 (28.1)	0.74 (0.39-1.42)	
Incident STI during the follow-up[k]				.003
No	24 (80.0)	71 (48.6)	Reference	
Yes	6 (20.0)	75 (51.4)	4.23 (1.63-10.94)	
Sought medical referral during the study[c]				<.001
No	54 (93.1)	83 (56.8)	Reference	
Yes	4 (6.9)	63 (43.2)	10.25 (3.52-29.79)	
Synchronized schedule with PrEP[l] study visit[c]				<.001
No	40 (69.0)	6 (4.1)	Reference	
Yes	18 (31.0)	140 (95.9)	51.85 (19.30-139.34)	

[a]OR: odds ratio.
[b]Median (IQR).
[c]n=204.
[d]n=200.
[e]US $1=HK $7.8.
[f]Not applicable.
[g]STI: sexually transmitted infection.
[h]n=187.
[i]n=194.
[j]HCW: health care workers.
[k]n=176.
[l]PrEP: pre-exposure prophylaxis.

Sampling and Testing During Follow-up

During the follow-up periods, of 762 sets of samples collected, 638 sets (83.7%) were blood sampled by staff for HIV and syphilis rapid tests with self-sampling of swabs and urine, 44 sets (5.8%) were self-sampled (DBS, swabs, urine samples) and self-tested for HIV and syphilis at the site, and 80 sets (10.5%) were self-sampled (DBS, swabs, urine samples) and self-tested for syphilis at home with delivery to the laboratory. From visits 2 to 5, there was a decreasing proportion of participants who selected the home delivery option and the self-sampling with self-testing at the site option (Table 4).

Regarding the scoring after self-sampling, the median score was 8 (IQR 5-9) for convenience, 7 (IQR 4-9) for confidence of correct self-sampling, 7 (IQR 5-9) for accurate reflection of his HIV or STI status from self-sampled specimens, and 3 (IQR 2-5) for discomfort during self-sampling at the baseline visit (Table 5). The proportions of high scores (8-10 points in Models A, B, and C; 3-10 points in Model D) were similar between baseline and follow-up visits 2 to 5, without significant changes identified in the GEE models. In the sensitivity analyses, cut-off values of 7-10 and 9-10 for Models A, B, and C and values of 2-10 and 4-10 for Model D showed the same conclusion. In reference to the median scores for visits 1 to 5, 55.3% (83/150) of MSM scored 8 to 10 for convenience, 45.1% (69/153) scored 8 to 10 for confidence in correct sampling, 53.2% (82/154)

scored 8 to 10 for accuracy, and 39.3% (59/150) scored 1 to 2 for discomfort. A high median score for convenience was positively associated with full study engagement (OR 2.68, 95% CI 1.14-6.30) in the logistic regression model, but no other markers had significant associations.

There was no incident HIV infection in the follow-up period. Among MSM with at least 2 rounds of testing performed during a median follow-up period of 12.53 (IQR 9.62-14.71) months, the incidence of any STI was 59.1 per 100 person-years (py; 81/137.11 years, 95% CI 47.2-73.1 per 100 py), with the highest incidence for CT (51/154.29 years; 33.1, 95% CI 24.9-43.1 per 100 py), followed by NG (48/160.77 years; 29.9, 95% CI 22.3-39.3 per 100 py) and syphilis (18/173.08 years; 10.4, 95% CI 6.4-16.1 per 100 py). By anatomic site, the highest proportion of samples testing CT-positive was in rectal swabs (46/539, 8.5%). The highest proportion of samples testing NG-positive was in pharyngeal swabs (37/546, 6.8%). Among urine samples, only 2.5% (14/554) were CT-positive, and 1.6% (9/555) were NG-positive. Comparing the positivity of one or more STI conditions at each or all anatomic sites during the follow-up in bivariable logistic regression analysis, no difference in STI positivity was identified between sampling by staff or participants nor between in-person visits or home delivery. However, no samples tested positive for syphilis in the home delivery group and self-sampling at site group, with the same syphilis rapid test kit (treponemal and nontreponemal test) provided to participants with or without a past history of syphilis.

A total of 113 participants completed the exit questionnaire. They were more likely to be fully engaged (OR 11.50, 95% CI 5.33-24.80) and be co-enrolled in the PrEP study (OR 5.02, 95% CI 2.41-10.46), compared with those who did not complete the exit questionnaire. Their sociodemographic characteristics (locality, employment status, and educational level) were however not significantly different. In the exit questionnaire data from the 113 participants, 103 (91.2%) had checked the testing results from the designated website, and 100 (99%) replied having a plan for regular STI testing (at least once a year) after the study.

Table 4. Distribution of sampling and testing options by participants and staff over visits (n=762).

Visit number	Number of participants who completed the visit	In-person visit, n (%)		Home delivery, n (%)
		Blood sampled by staff, rapid test for HIV and syphilis by staff (n=638)	Blood self-sampled and self-tested for HIV and syphilis at the site (n=44)	Blood self-sampled and self-tested for syphilis at home (n=80)
1	204	204 (100)	0 (0)	0 (0)
2	176	114 (64.8)	29 (16.5)	33 (18.8)
3	147	111 (75.5)	9 (6.1)	27 (18.4)
4	124	107 (86.3)	3 (2.4)	14 (11.3)
5	79	76 (96.2)	1 (1.3)	2 (2.5)
6	20	15 (75.0)	2 (10.0)	3 (15.0)
7	8	7 (87.5)	0 (0)	1 (12.5)
8	3	3 (100)	0 (0)	0 (0)
9	1	1 (100)	0 (0)	0 (0)

Table 5. Post-self-sampling scores (ranging between 1 [strongly disagree] and 10 [strongly agree]) and their changes across visits in binary logistic generalized estimating equation (GEE) models.

Model and visit number	Participants, n	Median (IQR)	High score[a], n (%)	GEE model results	
				Exp(B)	*P* value
Model A: It's convenient					
1	60	8 (5-9)	31 (52)	1	—[b]
2	105	8 (4-10)	57 (54)	1.03	.91
3	97	8 (5-9)	52 (54)	0.93	.81
4	75	8 (6-10)	45 (60)	1.11	.75
5	51	8 (5-9)	31 (61)	1.45	.28
Intercept	—	—	—	1.19	.48
Model B: I am confident to self-sample correctly					
1	62	7 (4-9)	23 (37)	1	—
2	106	8 (5-9)	55 (52)	1.33	.34
3	97	7 (5-9)	46 (47)	1.05	.87
4	76	8 (5-10)	42 (55)	1.46	.24
5	50	8 (5-9)	26 (52)	1.71	.11
Intercept	—	—	—	0.77	.30
Model C: I am confident that self-sampled specimens can accurately reflect my HIV or STI status					
1	61	7 (5-9)	30 (49)	1	—
2	108	8 (5-9)	61 (58)	0.81	.61
3	96	7 (5-9)	47 (49)	0.67	.29
4	76	8 (5-10)	42 (55)	0.78	.51
5	49	8 (6-9)	29 (58)	0.70	.47
Intercept	—	—	—	1.52	.22
Model D: I feel discomfort during self-sampling					
1	57	3 (2-5)	33 (58)	1	—
2	101	3 (2-6)	69 (68)	1.95	.10
3	93	3 (2-6)	57 (61)	1.42	.38
4	75	3 (2-5)	39 (52)	1.11	.81
5	49	4 (2-7)	32 (65)	0.93	.82
Intercept	—	—	—	1.08	.82

[a]The high score was coded 1 as a binary dependent variable in the GEE models: scores ≥8 for models A, B, and C and ≥3 for model D.

[b]Not applicable.

Discussion

Principal Findings

This is a longitudinal study involving a designated website and self-sampling method to enhance regular HIV and STI testing in MSM. Despite the low proportion (21.0%) of participants indicating a preference for self-sampled testing at baseline, the overall completion rate of regular self-sampled testing was high, at 72.1%. Scoring of the self-sampling process by participating MSM was high for its convenience, confidence of correct sampling, and perceived accurate reflection of STI status, and only very mild discomfort was reported.

Comparison With Prior Work

There was a positive association between seeking a medical referral subsequent to the detection of an incident STI and full engagement in regular HIV and STI testing. The detection of new STI may have alerted MSM about their health risk regardless of experiencing symptoms. With treatment as a follow-up action, STI knowledge may be enhanced in this group. The higher STI knowledge may have promoted engagement in frequent STI testing [18]. On the other hand, the compliance rate for MSM with a synchronized schedule of HIV and STI testing with a PrEP visit was higher than regular HIV and STI testing as an isolated component. Among MSM on PrEP for HIV prevention, a recent HIV test result is often a precondition

for a PrEP refill [19]. Bundling STI testing with HIV testing for a PrEP refill, as has been packaged in many PrEP implementation studies [20-22], could be a useful strategy for individuals at higher risk of HIV and STI. Self-sampled testing may further increase uptake by providing the flexibility to undergo testing.

MSM who scored the convenience of self-sampling higher in the first 5 visits were more likely to be fully engaged with regular testing during participation in the study. This is consistent with the higher acceptability of internet-based testing with self-sampling, an approach that saved travel and clinic waiting times, as concluded in a systematic review [23]. Our participants' preferences for self-sampling and self-testing at baseline were not significantly different by socioeconomic factors including educational level, Chinese locality, and employment. Potential barriers for self-sampling include a lack of confidence in correct self-sampling, inaccurate results in the self-sampled samples, and discomfort [23]. Although these factors were not significantly associated with full engagement, more than one-half of MSM scored confidence and accuracy high, and around 40% provided low scores for discomfort.

In this MSM cohort, the prevalence of overall STI was high at baseline, at 31%, which was higher than the baseline STI prevalence in PrEP studies in Amsterdam (17%) [20] and Australia (13.7%) [24], and the baseline CT and NG prevalences in this study (26%) were also higher than in a previous local study of CT and NG with MSM attending an STI clinic (19.6%) [17]. The incidence of any STI in the follow-up period of this study was however lower than that in previous PrEP studies [21,22]. The proportion of positive test results in extragenital site samples was higher than that in urine samples. The rate of detection of CT and NG from multisite samples highlights the insufficiency of collecting urine samples alone for CT and NG testing in MSM. This observation was consistent with previous studies [21,22] and supported by guidelines for CT and NG screening at extragenital sites in MSM [5-7] and results from a cost-effectiveness analysis [25].

Limitations

There were a few limitations in this study. First, a good proportion (77%) of participants had concurrently joined the PrEP study being conducted by the research team, which allowed synchronized scheduling of appointments for the 2 studies. The completion rate of the regular testing study was likely to be affected partly by the PrEP study. Comparison of PrEP-naïve and experienced participants was not performed because of the unbalanced distribution. Second, CT and NG test results from self-collected samples may be affected by the sampling technique of the participants. However, with high positivity rates and high proportion of detectable CT or NG in the test kits, incorrect sampling was unlikely to be a major concern. Studies have shown comparable accuracy and consistency results between self-sampled and clinician-sampled specimens [15,16]. Third, we did not use a validated tool to score self-sampling. As convenience was identified as a significant factor associated with the degree of engagement, further development of an evaluation tool for self-sampling may be useful. Fourth, self-selection bias might have arisen, with MSM at higher risk of HIV or STI being attracted to participate in the study, and recall bias might exist in the questionnaire survey even though it was completed on a monthly basis. Fifth, there was no control group who had samples taken by health care workers nor a non-web-based booking system, so we were unable to directly assess the outcome of the program.

Conclusion

The STI burden for MSM in the community in Hong Kong was high. MSM with an incident STI, seeking medical referrals, and who scored the convenience of self-sampling high were more likely to become fully engaged in a regular testing program. Self-sampling and the functioning of a website could enhance regular HIV and STI testing, supplementing conventional methods for implementing testing.

Acknowledgments

We thank Mr CY Poon and Ms KY So for their assistance in the laboratory work as well as Mr CQ Ye and Ms Mandy Li for project coordination and providing guidance to participants for sample collection. This work was supported by the Council for the AIDS Trust Fund (MSS307R, MSS325R).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Trends of retention in the study.
[PDF File (Adobe PDF File), 12 KB - formative_v6i11e40996_app1.pdf]

References

1. Tsuboi M, Evans J, Davies EP, Rowley J, Korenromp EL, Clayton T, et al. Prevalence of syphilis among men who have sex with men: a global systematic review and meta-analysis from 2000-20. Lancet Glob Health 2021 Aug;9(8):e1110-e1118 [FREE Full text] [doi: 10.1016/S2214-109X(21)00221-7] [Medline: 34246332]

2. Dewart CM, Bernstein KT, DeGroote NP, Romaguera R, Turner AN. Prevalence of rectal chlamydial and gonococcal infections: a systematic review. Sex Transm Dis 2018 May;45(5):287-293 [FREE Full text] [doi: 10.1097/OLQ.0000000000000754] [Medline: 29465688]
3. Guidelines for the management of symptomatic sexually transmitted infections. World Health Organization. 2021 Jul 15. URL: https://www.who.int/publications/i/item/9789240024168 [accessed 2022-05-17]
4. Garrett NJ, McGrath N, Mindel A. Advancing STI care in low/middle-income countries: has STI syndromic management reached its use-by date? Sex Transm Infect 2017 Feb;93(1):4-5 [FREE Full text] [doi: 10.1136/sextrans-2016-052581] [Medline: 27084840]
5. Workowski KA, Bachmann LH, Chan PA, Johnston CM, Muzny CA, Park I, et al. Sexually transmitted infections treatment guidelines, 2021. MMWR Recomm Rep 2021 Jul 23;70(4):1-187 [FREE Full text] [doi: 10.15585/mmwr.rr7004a1] [Medline: 34292926]
6. Australasian Sexual and Reproductive Health Alliance, Australasian Society for HIV, Viral Hepatitis and Sexual Health Medicine. Men who have sex with men. Australian STI Management Guidelines for Use in Primary Care. URL: https://sti.guidelines.org.au/populations-and-situations/men-who-have-sex-with-men/ [accessed 2022-05-12]
7. Scientific Committee on AIDS and STI. Recommendations on STI testing for MSM in Hong Kong. Centre for Health Protection, Department of Health, Hong Kong. 2020 Nov. URL: https://www.chp.gov.hk/files/pdf/sti_msm_rec_nov20.pdf [accessed 2022-05-12]
8. Guy R, Goller JL, Spelman T, El-Hayek C, Gold J, Lim M, et al. Does the frequency of HIV and STI testing among men who have sex with men in primary care adhere with Australian guidelines? Sex Transm Infect 2010 Oct;86(5):371-376. [doi: 10.1136/sti.2009.040972] [Medline: 20460263]
9. Frankis J, Goodall L, Clutterbuck D, Abubakari A, Flowers P. Regular STI testing amongst men who have sex with men and use social media is suboptimal - a cross-sectional study. Int J STD AIDS 2017 May;28(6):573-583. [doi: 10.1177/0956462416636780] [Medline: 26945592]
10. Visser M, Heijne JCM, Hogewoning AA, van Aar F. Frequency and determinants of consistent STI/HIV testing among men who have sex with men testing at STI outpatient clinics in the Netherlands: a longitudinal study. Sex Transm Infect 2017 Sep;93(6):396-403 [FREE Full text] [doi: 10.1136/sextrans-2016-052918] [Medline: 28159917]
11. Vriend HJ, Stolte IG, Heijne JCM, Heijman T, De Vries HJC, Geskus RB, et al. Repeated STI and HIV testing among HIV-negative men who have sex with men attending a large STI clinic in Amsterdam: a longitudinal study. Sex Transm Infect 2015 Jun;91(4):294-299. [doi: 10.1136/sextrans-2014-051786] [Medline: 25504922]
12. Kularadhan V, Gan J, Chow EPF, Fairley CK, Ong JJ. HIV and STI testing preferences for men who have sex with men in high-income countries: a scoping review. Int J Environ Res Public Health 2022 Mar 04;19(5):1 [FREE Full text] [doi: 10.3390/ijerph19053002] [Medline: 35270694]
13. Turner KM, Zienkiewicz AK, Syred J, Looker KJ, de Sa J, Brady M, et al. Web-based activity within a sexual health economy: observational study. J Med Internet Res 2018 Mar 07;20(3):e74 [FREE Full text] [doi: 10.2196/jmir.8101] [Medline: 29514776]
14. Wilson E, Free C, Morris TP, Syred J, Ahamed I, Menon-Johansson AS, et al. Internet-accessed sexually transmitted infection (e-STI) testing and results service: A randomised, single-blind, controlled trial. PLoS Med 2017 Dec;14(12):e1002479 [FREE Full text] [doi: 10.1371/journal.pmed.1002479] [Medline: 29281628]
15. Ni Y, Lu Y, He X, Li Y, Li Y, Guo S, et al. Self-sampled specimens demonstrate comparable accuracy and consistency to clinician-sampled specimens for HPV detection among men who have sex with men in China. Sex Transm Infect 2022 May 09:1. [doi: 10.1136/sextrans-2022-055408] [Medline: 35534231]
16. Wilson JD, Wallace HE, Loftus-Keeling M, Ward H, Davies B, Vargas-Palacios A, et al. Swab-yourself trial with economic monitoring and testing for infections collectively (SYSTEMATIC): Part 1. A diagnostic accuracy and cost-effectiveness study comparing clinician-taken vs self-taken rectal and pharyngeal samples for the diagnosis of gonorrhea and chlamydia. Clin Infect Dis 2021 Nov 02;73(9):e3172-e3180. [doi: 10.1093/cid/ciaa1266] [Medline: 32877521]
17. Chau CT, Ho KM, Ho CK. Screening of asymptomatic Chlamydia trachomatis and Neisseria gonorrhoeae infections among men who have sex with men in Hong Kong. Hong Kong J. Dermatol. Venereol 2016;24:113-119 [FREE Full text] [doi: 10.26226/morressier.56d6be77d462b80296c97913]
18. Wayal S, Reid D, Weatherburn P, Blomquist P, Fabiane S, Hughes G, et al. Association between knowledge, risk behaviours, and testing for sexually transmitted infections among men who have sex with men: findings from a large online survey in the United Kingdom. HIV Med 2019 Sep;20(8):523-533 [FREE Full text] [doi: 10.1111/hiv.12753] [Medline: 31124278]
19. Preexposure prophylaxis for the prevention of HIV infection in the United States - 2021 update: clinical practice guideline. Centers for Disease Control and Prevention. 2021. URL: https://www.cdc.gov/hiv/pdf/risk/prep/cdc-hiv-prep-guidelines-2021.pdf [accessed 2022-05-12]
20. Hoornenborg E, Coyer L, van Laarhoven A, Achterbergh R, de Vries H, Prins M, Amsterdam PrEP Project team in the HIV Transmission Elimination Amsterdam Initiative. Change in sexual risk behaviour after 6 months of pre-exposure prophylaxis use: results from the Amsterdam pre-exposure prophylaxis demonstration project. AIDS 2018 Jul 17;32(11):1527-1532. [doi: 10.1097/QAD.0000000000001874] [Medline: 29762169]

21. Beymer MR, DeVost MA, Weiss RE, Dierst-Davies R, Shover CL, Landovitz RJ, et al. Does HIV pre-exposure prophylaxis use lead to a higher incidence of sexually transmitted infections? A case-crossover study of men who have sex with men in Los Angeles, California. Sex Transm Infect 2018 Sep;94(6):457-462 [FREE Full text] [doi: 10.1136/sextrans-2017-053377] [Medline: 29487172]
22. Kwan TH, Lui GCY, Lam TTN, Lee KCK, Wong NS, Chan DPC, et al. Comparison between daily and on-demand PrEP (pre-exposure prophylaxis) regimen in covering condomless anal intercourse for men who have sex with men in Hong Kong: A randomized, controlled, open-label, crossover trial. J Int AIDS Soc 2021 Sep;24(9):e25795 [FREE Full text] [doi: 10.1002/jia2.25795] [Medline: 34473402]
23. Spence T, Kander I, Walsh J, Griffiths F, Ross J. Perceptions and experiences of internet-based testing for sexually transmitted infections: systematic review and synthesis of qualitative research. J Med Internet Res 2020 Aug 26;22(8):e17667 [FREE Full text] [doi: 10.2196/17667] [Medline: 32663151]
24. Traeger MW, Cornelisse VJ, Asselin J, Price B, Roth NJ, Willcox J, PrEPX Study Team. Association of HIV preexposure prophylaxis with incidence of sexually transmitted infections among individuals at high risk of HIV infection. JAMA 2019 Apr 09;321(14):1380-1390 [FREE Full text] [doi: 10.1001/jama.2019.2947] [Medline: 30964528]
25. Chesson HW, Bernstein KT, Gift TL, Marcus JL, Pipkin S, Kent CK. The cost-effectiveness of screening men who have sex with men for rectal chlamydial and gonococcal infection to prevent HIV Infection. Sex Transm Dis 2013 May;40(5):366-371 [FREE Full text] [doi: 10.1097/OLQ.0b013e318284e544] [Medline: 23588125]

Abbreviations

CT: Chlamydia trachomatis
DBS: dried blood spots
GEE: generalized estimating equation
MSM: men who have sex with men
NG: Neisseria gonorrhoeae
OR: odds ratio
PEP: post-exposure prophylaxis
POC: point-of-care
PrEP: pre-exposure prophylaxis
py: person-years
STI: sexually transmitted infection

Edited by T Leung; submitted 12.07.22; peer-reviewed by K Fergus, CC Hung; comments to author 21.09.22; revised version received 19.10.22; accepted 25.10.22; published 18.11.22.

Please cite as:
Wong NS, Kwan TH, Chan DPC, Lui GCY, Lee SS
Regular Testing of HIV and Sexually Transmitted Infections With Self-Collected Samples From Multiple Anatomic Sites to Monitor Sexual Health in Men Who Have Sex With Men: Longitudinal Study
JMIR Form Res 2022;6(11):e40996
URL: https://formative.jmir.org/2022/11/e40996
doi:10.2196/40996
PMID:36399372

Original Paper

A Digital Platform to Support HIV Case Management for Youth and Young Adults: Mixed Methods Feasibility Study

Connie Fee[1*], MPH; Julia Fuller[2*], BA; Carly E Guss[3,4*], MPH, MD; Elizabeth R Woods[3,4], MPH, MD; Ellen R Cooper[5,6], MD; Urmi Bhaumik[3], MBBS, MS, ScD; Dionne Graham[4,7], PhD; Sandra K Burchett[4,8], MSc, MD; Olivia Dumont[3], BA; Emily B Martey[3], BA; Maria Narvaez[6], BA; Jessica E Haberer[4,9], MS, MD; Dallas Swendeman[10], MPH, PhD; Shelagh A Mulvaney[11], PhD; Vikram S Kumar[2], MD; Jonathan L Jackson[2], MEng; Y Xian Ho[2*], PhD

[1]University of San Francisco, San Francisco, CA, United States
[2]Dimagi, Inc, Cambridge, MA, United States
[3]Division of Adolescent/Young Adult Medicine, Boston Children's Hospital, Boston, MA, United States
[4]Harvard Medical School, Boston, MA, United States
[5]Boston University School of Medicine, Boston, MA, United States
[6]Division of Pediatric Infectious Diseases, Boston Medical Center, Boston, MA, United States
[7]Program for Patient Safety & Quality, Boston Children's Hospital, Boston, MA, United States
[8]Division of Infectious Diseases, Boston Children's Hospital, Boston, MA, United States
[9]Massachusetts General Hospital, Boston, MA, United States
[10]Department of Psychiatry and Biobehavioral Sciences, David Geffen School of Medicine, University of California, Los Angeles, CA, United States
[11]School of Nursing, Vanderbilt University, Nashville, TN, United States
[*]these authors contributed equally

Corresponding Author:
Y Xian Ho, PhD
Dimagi, Inc
585 Massachusetts Ave., Suite 3
Cambridge, MA, 02139-4075
United States
Phone: 1 617 649 2214 ext 63
Email: yho@dimagi.com

Abstract

Background: Advances in medical treatments in recent years have contributed to an overall decline in HIV-related opportunistic infections and deaths in youth; however, mortality and morbidity rates in perinatally and nonperinatally infected adolescents and young adults (AYA) living with HIV remain relatively high today.

Objective: The goal of this project was to assess the use, utility, and cost-effectiveness of *PlusCare*, a digital app for HIV case management in AYA living with HIV. The app supports routine case management tasks, such as scheduling follow-up visits, sharing documents for review and signature, laboratory test results, and between-visit communications (eg, encouraging messages).

Methods: We conducted a single-group mixed methods pre-post study with HIV case management programs in 2 large urban hospitals in the Boston metro area. Case management staff (case managers [CMs], N=20) and AYA living with HIV participants (N=45) took part in the study with access to PlusCare for up to 15 and 12 months, respectively.

Results: The CMs and AYA living with HIV reported mean System Usability Scale scores of 51 (SD 7.9) and 63 (SD 10.6), respectively. Although marginally significant, total charges billed at 1 of the 2 sites compared with the 12 months before app use (including emergency, inpatient, and outpatient charges) decreased by 41% (P=.046). We also observed slight increases in AYA living with HIV self-reported self-efficacy in chronic disease management and quality of life (Health-Related Quality of Life-4) from baseline to the 12-month follow-up (P=.02 and P=.03, respectively) and increased self-efficacy from the 6- to 12-month follow-up (P=.02). There was no significant change in HIV viral suppression, appointment adherence, or medication adherence in this small-sample pilot study.

Conclusions: Although perceived usability was low, qualitative feedback from CMs and use patterns suggested that direct messaging and timely, remote, and secure sharing of laboratory results and documents (including electronic signatures) between

CMs and AYA living with HIV can be particularly useful and have potential value in supporting care coordination and promoting patient self-efficacy and quality of life.

Trial Registration: ClinicalTrials.gov NCT03758066; https://clinicaltrials.gov/ct2/show/NCT03758066

(JMIR Form Res 2022;6(11):e39357) doi:10.2196/39357

KEYWORDS

HIV; case management; youth; young adult; mobile health; mHealth; digital health; mobile phone

Introduction

Background

Adolescents and young adults (AYA) are estimated to account for more than 1 in 5 new HIV diagnoses in the United States [1]. Advances in medical treatments and combination antiretroviral therapy have contributed to an overall decline in HIV-related opportunistic infections and deaths among youth in the United States and other high-income countries [2-5]. However, rates of mortality and morbidity in perinatally and behaviorally infected AYA living with HIV remain high in the United States, partly owing to poor HIV control [6]. By the end of 2019, it was estimated that only 35% of persons aged 13 to 24 years with HIV were virally suppressed and only 33% were retained in care, the lowest percentages for any age group [7]. AYA living with HIV require developmentally appropriate care and case management to achieve retention in care and viral suppression, as well as other developmental and health outcomes, as they become more responsible for their own care and adherence at the same time that they enter the turbulence of adolescent and young adult transitions [8,9].

Case management programs can help AYA living with HIV work toward achieving primary HIV outcomes, such as viral suppression, by providing support for fundamental HIV care needs, in addition to non–HIV-related health and developmental outcomes that may co-occur with or undermine HIV outcomes [10]. The use of case management has been associated with decreased unmet needs for supportive care of AYA living with HIV, such as health insurance, emotional counseling, and higher medication adherence in patients receiving HIV treatment [11]. Case managers (CMs) and other clinical care staff who serve as part of the case management team, such as physicians, nurse practitioners, nurses, and social workers (hereafter collectively referred to as CMs), play a vital role in the front line of care throughout the HIV care continuum, often starting with reporting a positive test result to a patient and then connecting youth to resources that can assist in fulfilling basic needs including insurance, medical care, housing, and transportation, as well as mental health and general emotional support [12-14]. Engagement and retention in care become critical as AYA living with HIV age and transition into adult care, and studies suggest that case management can contribute to the retention of AYA living with HIV in primary care, particularly in high-risk populations [15-19], by providing linkage to care most needed by hard-to-reach populations [14,20-25]. Logistical barriers, such as arrangements for transportation to the hospital or clinic, can prevent AYA living with HIV from receiving the care that they need. A case management team can support the fundamental aspects of AYA living with HIV care but is limited by constraints such as infrequent clinic visits and unreliable means of communication with patients.

Mobile health (mHealth) technology provides an opportunity to overcome these limitations and facilitates case management for AYA living with HIV. Pew research reported an estimated 95% of adolescents in the United States in 2018 owning or having access to a smartphone, and numbers are likely to be higher today [26]. The pervasiveness of mobile technology in the youth today, combined with its unique capabilities to securely handle data, presents a distinctive opportunity to support AYA living with HIV [27]. Although there have been many advances in mHealth to support persons living with HIV/AIDS, few tools have been developed to directly support case management through a shared platform between the clinical care team and AYA living with HIV, a capability that is desired and has the potential to streamline care [28]. Unlike adults living with HIV, AYA living with HIV likely require a more complex, individualized, youth-friendly approach provided by case management programs [29]. We designed and previously found high perceived usability of the prototype of a digital app system, *PlusCare*, to support HIV case management for AYA living with HIV, demonstrating its acceptability and feasibility [30]. Using a user-centered design approach, we developed a mobile and web app that can directly connect AYA living with HIV and members of their case management team by streamlining communication and facilitating comprehensive support to help overcome logistical barriers to care [31].

Objectives

The objective of this study was to assess the user experience of the PlusCare system for AYA living with HIV case management, cost-effectiveness, and potential effects on physical and psychosocial health.

Methods

A mixed methods, nonrandomized, single-group intervention study design (ClinicalTrials.gov NCT03758066) was used.

Study Settings

The study was conducted with 3 HIV/AIDS programs: 2 based at Boston Children's Hospital (BCH), and 1 at Boston Medical Center (BMC). BCH's programs are housed in both adolescent and young adult medicine and infectious diseases and BMC's program belongs to the infectious diseases division. All 3 programs shared a similar focus on HIV counseling, prevention, screening, and linkage to care, retention, and adherence. They used a model of multidisciplinary team-managed care in which each AYA living with HIV is assigned to a designated team consisting of, at minimum, a CM, a medical provider, and a

nurse. Program staff at both sites communicated similarly with patients, relying primarily on phone calls and in-person visits to connect with patients; for example, to share and discuss laboratory results and collect signatures on documents. At BCH, CMs also communicated with patients via the patient portal, and at BMC, SMS text messaging was occasionally used; both modalities were used for less confidential communication. Paper binders and electronic health record systems were used at both sites to document interactions. BCH and BMC are private, not-for-profit academic hospitals. BMC is the largest safety-net hospital in the region; that is, it serves a racially and ethnically diverse patient population with a large proportion of patients either uncovered or covered by Medicaid in commitment to its mission statement to provide health care to all patients, regardless of their ability to pay.

Recruitment and Eligibility

CMs at each site actively involved in the care of AYA living with HIV were identified and referred to the study site research assistant (RA) to confirm eligibility, consent, and enrollment in the study. Patients were eligible to participate if they were aged 13 to 25 years, living with HIV, had reliable access to a smartphone for 12 months, and were at the study site for at least 12 months before enrollment. A waiver of parental consent for participants aged 13 to 17 years was approved by the institutional review board (IRB) to protect participant confidentiality. All participants were recruited and enrolled between January 2019 and June 2019.

Baseline and App Training

CMs from each of the 3 programs participated in a baseline session that took place in January 2019. These sessions were led by study site RAs at each site who had in-depth knowledge and experience with the app. Written informed consent was obtained from each CM at the beginning of each session. A baseline survey was conducted to collect basic demographic data and information on professional roles and experience (eg, job title, years of CM experience, and number of AYA living with HIV managed). In this session, CMs received training on how to use the PlusCare mobile app and register AYA living with HIV participants.

RAs consented to the participation of eligible AYA living with HIV and then completed a baseline survey which included items to collect demographic information and assess technology use, familiarity, and health outcome measures. CMs were responsible for registering consented and enrolled patients in the PlusCare system and setting up the patient's care team contacts. Only the web app was tested with the patients at enrollment to ensure equivalent access agnostic to operating system, although PlusCare could theoretically be accessed as a mobile app via Android devices. A shortcut to the web app was created on the participants' devices for easy access. Baseline training was considered complete after the PlusCare shortcut was installed and the participant successfully demonstrated that they could (1) access the app from their smartphone and (2) receive incoming SMS text messages from the CM.

Intervention

Study tablets with PlusCare preinstalled as a native app were provided to the case management programs at each site for use by CMs. CMs could access PlusCare via the app on the study tablets or the web app from any computer for up to 15 months from the start date of the study intervention to account for rolling patient enrollment. Patients had access to the PlusCare web app for 12 months after enrollment in the study and continued to receive usual care from their respective case management teams during that time, including communication with patients about upcoming visits, discussion about laboratory results, updating insurance information, renewing medications, and assistance with social determinants of health needs, such as food, housing, and transportation.

To assess the naturalistic use of the app, user-suggested modifications during the intervention period were restricted to those determined to be barriers to use. As such, only one modification was implemented and introduced in early August 2019, after all participants were enrolled, to allow CMs and patients to send asynchronous messages to one another within the app. Before this modification, only CMs could send messages to patients. However, patients were unable to send messages in response to CMs.

Outcomes and Measures

In this study, we evaluated (1) use, usability, and acceptability and (2) health and clinical outcomes (CD4 counts, viral load, clinic visits, medication adherence, self-efficacy in chronic disease management, quality of life, and cost savings).

Use, Usability, and Acceptability

The 10-item Likert-scale positive System Usability Scale (SUS) [32] was administered via a survey in the PlusCare app to assess CMs' and patients' perceived usability of and user satisfaction with the app at the end of the study. The SUS has been validated for use with mobile and web apps and scores range from 0 to 100, with higher scores indicating higher perceived usability [33]. A score >68 is considered above the industry average [34].

Data on which features of the PlusCare app were accessed (ie, forms that users submitted) were also collected passively and deidentified weekly use logs were analyzed separately for CMs and patients to explore use trends over the 15- and 12- month period, respectively.

RAs also conducted brief, structured *check-ins* following an interview guide with CMs every 6 weeks via phone or in person to address and log any issues with their own user experience and their patients. At the end of the study period, the RAs at each site also conducted brief structured one-on-one interviews with CMs at their respective sites to provide qualitative feedback on their experiences using the system with their patients. Qualitative notes were collected for analyses (see Multimedia Appendices 1 and 2 for sample guides used for check-ins and interviews).

Health and Clinical Outcomes

Viral load, CD4 counts, and scheduled clinic visits were compared between 1 year before enrollment (pre-enrollment) and 1 year during intervention (postenrollment). Values for

patient laboratory results and medical visit frequency were obtained from electronic health records.

Once a month, patients completed a medication adherence assessment, where they rated their adherence to their prescribed medication for the past month on a 6-point Likert scale ranging from 1 (“very poor”) to 6 (“excellent”) [35,36]. Patients received monthly SMS text messages alerting them to report adherence via survey forms created in the PlusCare app.

Responses to self-efficacy and health-related quality of life assessments, the self-efficacy for managing chronic diseases 6-item scale [37] and modified Centers for Disease Control and Prevention Health-Related Quality of Life-4 (HRQOL-4) healthy days core measures [38], respectively, were similarly collected via survey forms in the PlusCare app at baseline, 6 months, and 12 months. Self-efficacy scores were calculated by averaging patients’ self-reported confidence in keeping various symptoms from interfering with their everyday life and their confidence in engaging in activities that reduced the impact of their illness and the need to visit a clinician.

Outpatient, emergency, and inpatient per-patient charges were extracted from hospital billing data as deidentified data sets for analyses.

Data Analysis

Overview

Our analytical sample included 20 CMs and 45 patients enrolled in the study. The sample size of CMs was a convenience sample determined based on feasibility and the pool of potentially eligible case management staff at the 2 sites. The effect size was calculated using Cohen d=0.35 and the ability to detect a 38% increase in viral load, independent of sample size [39]. All statistical tests were based on 2-tailed alternatives at a significance level of .05 and analyses were performed using open-source R statistical software (R Foundation for Statistical Computing) [40] and SAS software (version 9.4; SAS Inc) [41].

To assess any differences between sites, baseline demographic characteristics for CMs and patients were compared between sites using 2-sample, 2-tailed t tests for averages (ie, age and savviness with technology) and chi-square tests or Fisher exact tests for categorical variables (ie, ethnicity and race, sex, education, smartphone ownership, hours on phone, and health and HIV apps). Missing responses were excluded from the significance tests.

Use, Usability, and Acceptability

Descriptive statistics (mean, median, SD, and frequency) were used to summarize SUS scores collected at the end of the study and use patterns in data collected passively over the intervention period. Qualitative notes from check-ins and end-of-study interviews with CMs were coded and analyzed by 2 coauthors (CF and YXH) using directed content analysis with predefined categories of (1) facilitators, (2) barriers, and (3) recommendations for system implementation. Recurring themes in the responses were iteratively coded, reviewed, and refined [42].

Health and Clinical Outcomes

We used simple linear regression to test whether self-reported medication adherence improved over the course of the intervention. A 2-tailed nonparametric 1-sample paired sign test was used to determine whether to reject or fail to reject the null hypothesis that the mean difference of self-efficacy for managing chronic diseases 6-item scale and Centers for Disease Control and Prevention HRQOL-4 scores between any of the 2 of 3 time points was 0 (baseline vs 6-month, 6-month vs 12-month, and baseline vs 12-month).

Generalized linear mixed modeling was used to determine changes in HIV viral load and CD4 levels in the preintervention and postintervention periods. Owing to the skewed distribution of the outcome, we log transformed the values before modeling. We included random intercepts for each participant to control for the correlation between repeated measurements within the patient.

Cost-effectiveness analyses were performed following an approach used in other community programs [43-46]. We used data on HIV medical visit frequency, emergency department visits, and hospitalizations for the analyses. We based our analyses on potential cost savings resulting from a reduction in hospital visit frequency between the preintervention period (12 months before baseline) and the intervention period from baseline to the 12-month follow-up. We computed preliminary cost-effectiveness of the PlusCare intervention by dividing the cost savings attained by the program cost of the case management program using the net present value (*net present value = present value of cost savings − present value of program costs*) and return on investment (*return on investment = present value of cost savings / present value of program costs*). The cost savings estimate was adjusted for variations in age, sex, and race and ethnicity using a multivariate regression model to reduce treatment costs.

Ethics Approval

IRB approval for all study activities was obtained from and overseen by the Western Institutional Review Board (WCG IRB 120190387) and the IRB at each study site (IRB-P00029517 and H-38254).

Results

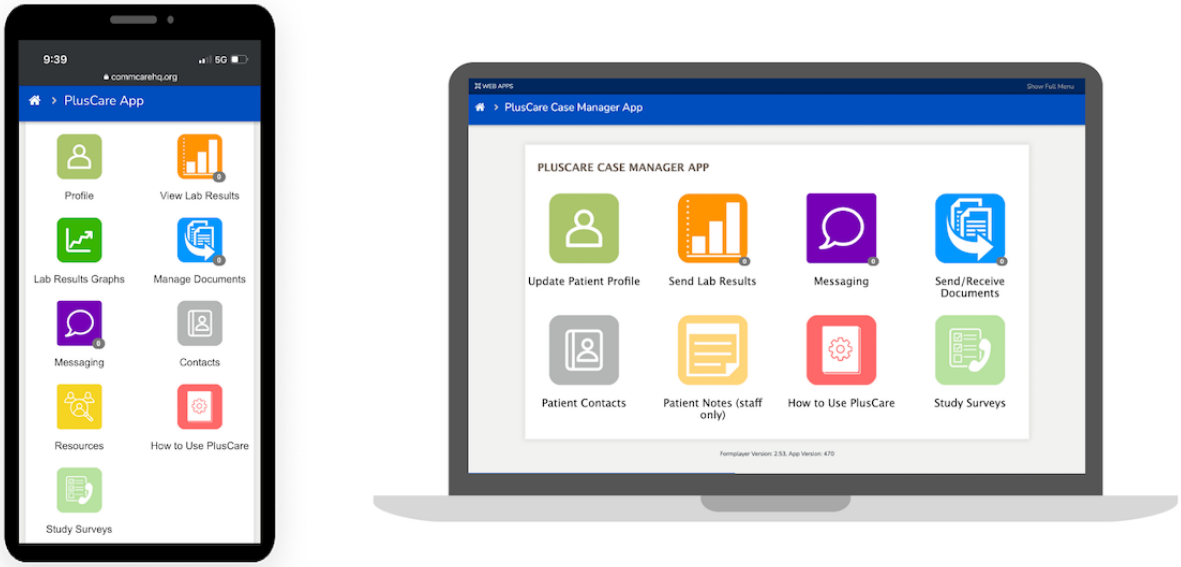
mHealth Study Tool

Overview

PlusCare was designed with 2 separate interfaces connected on a shared platform—one for CMs and one for the AYA living with HIV. The former can access the app via a native app installed on a study tablet or a web app through a computer browser. Patients could access the web app through a shortcut link on their smartphones. Figure 1 shows screenshots of the menu page for the 2 different users: CMs and patients. The app can support critical aspects of AYA living with HIV case management, and the following functionalities were validated through usability testing (corresponding menu items on the patient and CM interfaces shown in Figure 1 are indicated in parentheses):(1) enter, share, and view laboratory results (View

Lab Results, Lab Results Graphs, and Send Lab Results); (2) send, receive, and sign documents via an external document sharing app, Docusign [31] (Manage Documents and Send and Receive Documents); (3) create and update contacts (Contacts and Patient Contacts); (4) schedule, send, and receive SMS text message notifications of app activity, for example, laboratory results, appointment reminders, and encouraging messages (Messaging); and (5) view and go to links to community resources (Resources). In addition, CMs could update patient information registered in the system (Update Patient Profile) and create and share notes between staff (Patient Notes). Both users could access user guides in the system at any time (How to Use PlusCare).

Figure 1. PlusCare dashboard for patient (left) as viewed on a smartphone and case manager as viewed in a web application on a laptop computer (right).



Participant Demographics

A total of 20 CMs were enrolled and trained to use PlusCare at our 2 participating study sites: 11 CMs from BCH and 9 CMs from BMC (Table 1). CM demographics are presented in Table 1. Reported professional roles of individuals performing CM duties included CMs (n=2), medical CMs (n=3), nurse practitioners (n=4), clinical social workers (n=3), physicians (n=2), nurses (n=2), pharmacists (n=1), and health services advocate (n=1). CMs owned either an iPhone (17/20, 85%) or Android (3/20, 15%) smartphone, and 15% (3/20) reported spending >7 hours on their mobile phones per day. On a scale from 1 to 10 for tech savviness, CMs reported an average score of 7.1 (SD 1.2) for their own tech savviness compared with 6.4 (SD 1.9) for a rating of perceived tech savviness in their workplace. An average score of 6.7 (SD 2) was reported for workplace effectiveness using information technology.

A total of 45 AYA living with HIV participants were enrolled across the 2 study sites. Patient demographics at the 2 sites are shown in Table 2. Of the 45 patients, only 8 (18%) were Hispanic and Latino; 34 (76%) were non-Hispanic, and roughly half (23/45, 51%) identified as female. The average age of the AYA living with HIV participant was 20 (SD 2.7) years with education level ranging from 4% (2/20) of individuals with no schooling completed to 9% (4/20) of individuals holding a bachelor's degree. AYA living with HIV owned an iPhone (40/45, 89%) or Android smartphone (5/45, 11%), with 40% (18/45) of participants reporting >7 hours of phone use per day. Of the 45 AYA living with HIV participants, only 55% (11/20) reported using health/HIV apps. An average score of 8 (SD 1.6) was reported for self-rated tech savviness. Of the baseline demographics, only education was found to be statistically different between patients at the 2 sites, with a higher proportion of patients with a high school education or less at BMC (68% vs 35%; P=.04).

Table 1. Case manager (N=20) demographics by site—Boston Medical Center (BMC) and Boston Children's Hospital (BCH).

Demographic	BCH (n=11)	BMC (n=9)	Total	*P* value
Age (years), n (%)				.39
18-24	3 (27)	1 (11)	4 (20)	
25-34	4 (36)	1 (11)	5 (25)	
35-44	0 (0)	2 (22)	2 (10)	
45-54	1 (9)	2 (22)	3 (15)	
≥55	3 (27)	3 (33)	6 (30)	
Ethnicity and race, n (%)				.82
Hispanic and Latino	1 (9)	0 (0)	1 (5)	
Non-Hispanic White	6 (55)	6 (67)	12 (60)	
Non-Hispanic Black	3 (27)	1 (11)	4 (20)	
Non-Hispanic other (including American Indian or Alaska Native, Asian and Pacific Islander, or more than one)	0 (0)	1 (11)	1 (5)	
Unknown	1 (9)	1 (11)	2 (10)	
Sex, n (%)				.19
Female	11 (100)	7 (78)	18 (90)	
Male	0 (0)	2 (22)	2 (10)	
Education—highest level, n (%)				.39
Bachelor's degree	5 (45)	2 (22)	7 (35)	
Master's degree	2 (18)	5 (56)	7 (35)	
Doctoral degree	3 (27)	1 (11)	4 (20)	
Professional degree	1 (9)	1 (11)	2 (10)	
Smartphone ownership, n (%)				>.99
iPhone	9 (82)	8 (89)	17 (85)	
Android	2 (18)	1 (11)	3 (15)	
Hours spent on phone per day, n (%)				.11
0-3	4 (36)	5 (56)	9 (45)	
4-6	2 (18)	6 (67)	8 (40)	
>7	3 (27)	0 (0)	3 (15)	
Tech savvy, mean (SD)				.88
Personal score	7.09 (1.45)	7 (1)	7.05 (1.23)	
Work score	6 (2.05)	6.89 (1.54)	6.4 (1.85)	
Workplace IT use, mean (SD)	6.36 (2.11)	7.11 (1.9)	6.7 (2)	.30
Number of patients				.42
Mean (SD)	37.44 (17.7)	43.7 (24.9)	38.95 (22.4)	
Median (range)	49 (10-54)	43.5 (2-80)	43.5 (2-80)	

Table 2. Patient demographics by site—Boston Medical Center (BMC) and Boston Children's Hospital (BCH; N=45).

Patients	BCH (n=20)	BMC (n=25)	Total	*P* value
Age (years)				.84
Mean (SD)	20.21 (2.5)	20.04 (2.9)	20.11 (2.7)	
Median (range)	20 (16-26)	20 (14-25)	20 (14-26)	
Ethnicity and race, n (%)				.32
Hispanic and Latino	2 (10)	6 (24)	8 (18)	
Non-Hispanic White	3 (15)	1 (4)	4 (9)	
Non-Hispanic Black	9 (45)	15 (60)	24 (53)	
Non-Hispanic other (including American Indian or Alaska Native, Asian and Pacific Islander, or more than one)	4 (20)	2 (8)	6 (13)	
Unknown	2 (10)	1 (4)	3 (7)	
Sex, n (%)				.64
Female	11 (55)	12 (48)	23 (51)	
Male	9 (45)	13 (52)	22 (49)	
Education—highest level, n (%)				.04
High school graduate or less	7 (35)	17 (68)	24 (53)	
Some college credit or more	12 (60)	8 (32)	20 (44)	
N/A[a] (missing)	1 (5)	0 (0)	1 (2)	
Income (US $), n (%)				.13
<20,000	5 (25)	13 (52)	18 (40)	
20,000-49,999	6 (30)	7 (28)	13 (29)	
≥50,000	5 (25)	2 (8)	7 (16)	
N/A (missing)	4 (20)	3 (12)	7 (16)	
Time in years since positive HIV diagnosis, n (%)				.91
1-3	4 (20)	6 (24)	10 (22)	
3-5	1 (5)	5 (20)	6 (13)	
>5	15 (75)	14 (56)	29 (64)	
Smartphone ownership, n (%)				.06
iPhone	20 (100)	20 (80)	40 (89)	
Android	0 (0)	5 (20)	5 (11)	
Hours spent on phone per day, n (%)				.46
0-3	4 (20)	4 (16)	8 (18)	
4-6	10 (50)	9 (36)	19 (42)	
>7	6 (30)	12 (48)	18 (40)	
Health or HIV apps, n (%)				.73
Yes	4 (20)	7 (28)	11 (24)	
No	15 (75)	17 (68)	32 (71)	
N/A (missing)	1 (5)	1 (4)	2 (4)	
Tech savvy				.15
Personal score, mean (SD)	8.3 (1.63)	7.6 (1.50)	7.9 (1.58)	
N/A (missing), n (%)	0 (0)	1 (4)	1 (2)	

[a]N/A: not applicable.

Use, Usability, and Acceptability

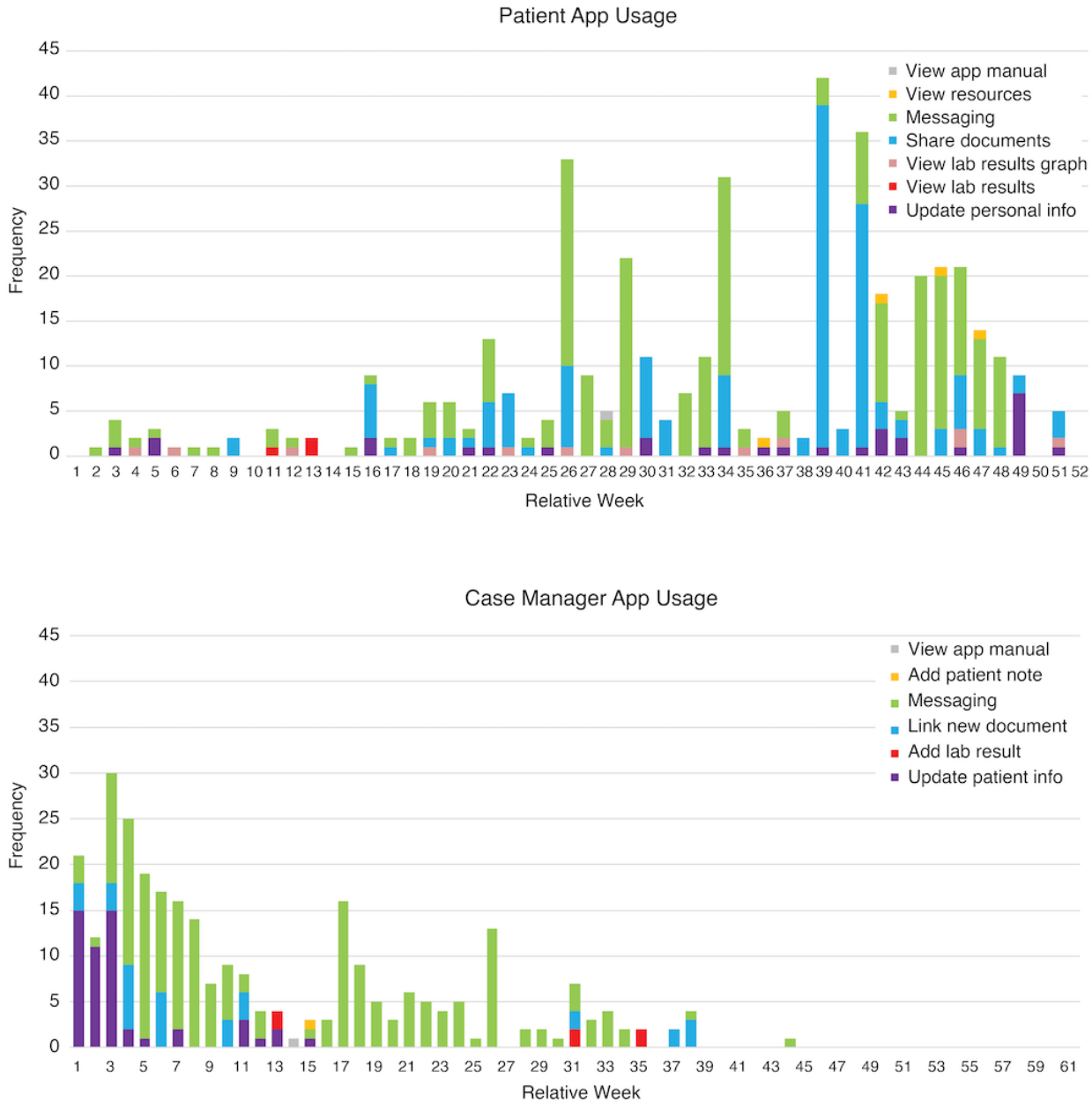
Overview

Average SUS scores for CMs and patients were 51 (SD 7.9; n=13) and 63 (SD 10.6; n=38), respectively. We found intersite differences in overall use as measured by the number of forms submitted through the app; of the 290 forms submitted by CMs and 427 forms submitted by patients at both sites, the majority were submitted by BMC participants (624/717, 87% and 667/717, 93%, respectively). CMs used the asynchronous messaging feature with patients most frequently (197/290, 67.9%), followed by updating patient information (53/290, 18.3%), and uploading documents (32/290, 11%). CMs also used the app to upload laboratory results (6/290, 2.1%), add patient notes (1/290, 0.3%), and access the how-to guide (1/290, 0.3%). Among the 427 forms submitted, patients most commonly sent messages to CMs (228/427, 53.4%) and viewed and shared documents (149/427, 34.9%). Patients also updated their personal information (30/427, 7%) and viewed their laboratory results or laboratory results graphs (15/427, 3.5%); they rarely accessed resources (4/427, 0.9%) and the how-to guide (1/427, 0.2%).

Use data revealed that CM use was highest at the beginning of the intervention period, primarily owing to the enrollment of patients and increased use of messaging starting in week 3; however, use dropped to near 0 approximately after 9 of the 15 months. Patient use data revealed a different pattern—a slow increase in use early in the trial period with spikes of activity in the middle to later weeks, with messaging and document management being the most frequent tasks (Figure 2).

Content analyses of the CM interviews at the end of the study revealed potential facilitators, barriers, and recommendations for improving the usability of the app.

Figure 2. Frequency of patient and case manager app use by feature over week from participant enrollment date (top and bottom, respectively).



Facilitators, Barriers, and Recommendations for Use

Overall, CMs found the app helpful in communicating with patients. Several CMs perceived PlusCare as an alternative way to reach patients and send appointment reminders. At one site, CMs used PlusCare to send positive affirmations to patients via SMS text messages sent through the system. CMs reported that patients seemed to appreciate these messages and remarked that support for bidirectional messaging was particularly helpful for them, consistent with the increase in patients using the messaging feature observed in the latter half of the study after the feature was implemented. Many CMs also reported having the ability to share documents and laboratory results as useful features of PlusCare, although use data did not show frequent sharing of results.

A frequently cited barrier to using PlusCare was the inability of patients to respond or initiate messages. Although 2-way, asynchronous messaging was implemented partway through the study in August in response to the need for this capability, use data showed no remarkable increase in the CM use of the system after implementation. In addition, 22% (4/18) of CMs reported that without a read receipt, it was frustrating to know if a patient received a message. A couple described the experience as sending a message out to a "black hole" or into "oblivion." Consequently, some CMs switched to using alternative methods of contacting patients, such as calling them.

A total of 39% (7/18) of CMs reported that PlusCare was not easy to use or generated extra work. For example, a few CMs found that using the app through the web link (including logging in and uploading documents) was cumbersome and having to manually synchronize the app (a feature inherent in the CommCare platform upon which it is built to enable offline data collection) before each use was a burden. Some CMs also disliked the way notification badges were implemented and indicated that it would have been preferable for the badge activity to reflect the individual's own activity, as opposed to the activity of the CMs as a group.

Other reasons for low to no CM app use included that the app was not relevant to the tasks they performed with the patients and that their patients did not use the app. Some CMs did not have easy access to a computer or tablet with the app installed or linked to and would have preferred a true mobile phone app connection rather than a web link used on the study tablet. Another reason was the preference for other methods of communication (ie, in person or over the phone) because these methods are perceived as more personal and responses are immediate.

Despite these barriers, many CMs remarked that PlusCare has potential and made recommendations for improvement. To address some of the aforementioned barriers, several CMs suggested adding read receipts to messages, updating the messaging feature to facilitate "real-time" or synchronous conversations, and adding the option to send messages to an individual CM or a designated group of CMs. Other recommendations included support for telehealth visits, more support for platform-agnostic features, and interoperability with other systems, such as electronic medical records. A CM felt that the app was a "great idea," but its implementation could be further improved by considering additional patient and provider inputs.

Health and Clinical Outcomes

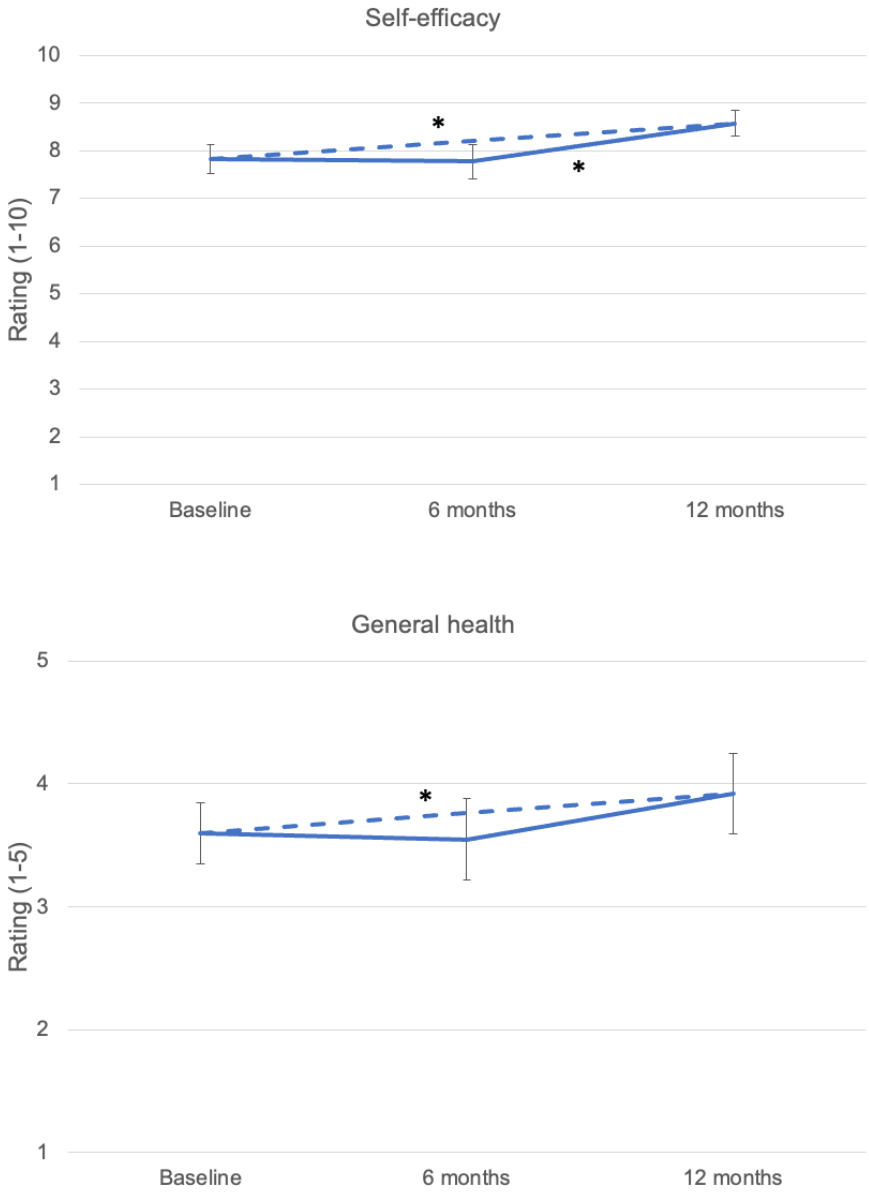
We assessed the changes in viral load, CD4 count, and clinic visits between the preintervention and postintervention periods. There was no significant change in viral load (N=45; mean 32, 95% CI 17.9-60.4 vs mean 24.7, 95% CI 12.1-47.9; P=.30) or CD4 count (N=45; mean 996, 95% CI 857-1157 vs mean 1049, 95% CI 898-1224; P=.18) over time. Although there was an 8% absolute increase in the median appointment adherence rate, defined as actual completed appointments of the total number of expected appointments (completed appointments, no shows, and cancelations) in the postintervention period compared with the preintervention period, this change was not significant (P=.13).

Mean self-reported monthly medication adherence ratings across 12 months were 4.7 (SD 0.23) out of 6. The fitted regression model was medication adherence rating of 4.8+0.01x, and the overall regression was not found to be statistically significant (R^2=0.1; $F_{1,11}$=1.24; P=.29). Although in general, good adherence was reported, qualitative feedback from CMs nonetheless suggested that medication adherence remains one of the biggest problems with the AYA living with HIV they worked with and about one-third of AYA living with HIV at one site reported "medication" or "reminder" when sharing potential outcomes or improvements that could be achieved with better use of technology for HIV management, suggesting there remains potential value for mHealth systems to support this aspect of care.

Patient self-efficacy in chronic disease management and quality of life were assessed at 3 time points over the duration of the study trial (number of patient responses are indicated in parentheses): baseline (n=45), 6 months (n=40), and 12 months (n=38). We found a significant increase in self-efficacy from baseline to 12 months (mean 0.61, SD 0.28; P=.02) and from 6 to 12 months (mean 0.65, SD 0.20; P=.02) but not from baseline to 6 months (Figure 3). Qualitative feedback from several CMs suggested that PlusCare can help AYA living with HIV take more ownership over their health. CMs at BMC noted that sending affirming messages to patients they felt struggled with medication adherence and self-esteem seemed to have a positive effect on patients' perceived self-worth and self-esteem.

Self-rated general health at 12 months was significantly higher than that at baseline (mean 0.32, SD 0.14; P=.03). However, no significant differences were found from baseline to 6 months and from 6 to 12 months (Figure 3).

Figure 3. Average (95% CIs) ratings of self-efficacy in chronic disease management (1=not at all confident to 10=totally confident) and health-related quality of life (1=poor to 5=excellent) among youth living with HIV (top and bottom, respectively). Dashed line represents change between baseline and 12-month time points, solid lines represent comparisons between baseline and 6-month and 6- and 12-month time points. Significant increases in self-efficacy were found from baseline to 12 months (*P*=.02) and from 6 to 12 months (*P*=.02). Self-rated general health at 12 months was significantly higher than that at baseline (*P*=.03). Significant increases in ratings between time points are indicated (*).



We evaluated cost savings by comparing the total per-patient clinic charges during the 12-month pre-enrollment and postenrollment period at each of the 2 sites for emergency department, outpatient, and inpatient visits. Among these charges, the emergency department charges at both sites did not show any decrease between the preintervention and postintervention periods. However, results for inpatient and outpatient charges were mixed. A significant decrease of 65% in inpatient charges was found at BCH (*P*=.03) and a slightly significant decrease of 41% was found in total charges (including emergency, inpatient, and outpatient charges; *P*=.046); however, no decrease was found in outpatient charges alone. No decrease in inpatient charges and small, but not significant (*P*=.24), decrease of 18% in outpatient and 9.5% in total charges (*P*=.37) were found at BMC.

Discussion

Principal Findings

In summary, we developed a digital app, PlusCare, to support HIV case management for AYA living with HIV and tested its use with CMs and AYA living with HIV patients at 2 private, not-for-profit academic hospitals. The app was designed to facilitate sharing and viewing of laboratory results, sending and signing documents, updating contacts, scheduling and sending messages, and viewing community resources. We found that both CMs and AYA living with HIV used the messaging feature most frequently and qualitative feedback from CMs suggested that direct messaging between CMs and AYA living with HIV can be particularly useful. There were no significant changes in HIV viral suppression (ie, viral load and CD4 count), appointment, or medication adherence over the course of the

study. Slight increases in AYA living with HIV self-reported self-efficacy in chronic disease management and quality of life were observed. Results of the cost-saving analyses were mixed and varied by site.

In part owing to the ubiquity of smartphones among adolescents in the United States, mHealth holds great potential in helping AYA living with HIV overcome barriers to receiving care. mHealth interventions for adults have been studied and shown to have a positive effect on retention and clinical outcomes [47]. However, few mHealth tools have been developed to directly serve AYA living with HIV who face even greater challenges while transitioning to adult care [8,9]. Less than 15% of systematic reviews have focused on mHealth tools, specifically on support for AYA living with HIV [47-49]. Today's community-based case management approach has become increasingly interdisciplinary, whereby a team of health professionals may play a role in the collaborative process of working with a given client throughout the HIV care continuum [12,50-52]. Recognizing the need for patient and care team coordination, we developed, tested, and deployed a shared platform for clinical care team members to perform case management duties and the AYA living with HIV they serve.

In this study, higher self-efficacy in chronic disease management and self-reported quality of health were found during the intervention period compared with 1 year prior in a within-participant comparison; however, we found no significant improvements in primary clinical outcomes of interest, that is, viral load or CD4 counts. This may be because our study included mostly healthy patients who were virally suppressed, which limited our ability to detect changes in primary clinical outcomes. No changes in medication adherence were observed, as self-rated medication adherence was moderately high throughout the intervention period and likely affected by other structural barriers not addressed by the app. Trends in the data suggested that PlusCare may be associated with an increase in clinic visit adherence and a decrease in per-patient charges, but neither was found to be significant.

Perhaps most surprisingly, we found lower average perceived usability ratings of the app—particularly for the CM interface—than the industry-average SUS score. There are several possible explanations for these findings. Although the SUS score distribution has been used for benchmarking a vast array of new technologies and more recently shown to be suitable for benchmarking mHealth or digital health apps (excluding apps just for self-monitoring physical activity), it is possible that the benchmark score could be subject to reporting bias in the literature, that is, the file drawer problem, where researchers may be more likely to report only higher usability [34,53]. Low mHealth usability has been found in previous studies in vulnerable, underserved populations based on other usability scales and metrics [54,55]; however, to the best of our knowledge, this is the first time SUS scores have been reported for a digital health app with 2 separate interfaces for 2 types of end users, CMs and AYA living with HIV participants. The usability ratings collected here were also found to be lower than the ratings collected in previous controlled, moderated usability testing conducted on an earlier prototype [30]. Qualitative feedback suggested that the features validated in usability testing may have been subject to limitations realized only when tested in naturalistic, unmoderated settings. As mentioned, a limitation in the handling of messaging between CMs and their patients was addressed earlier in the intervention period, which may have influenced the inverse use patterns, in which CMs used the app more during the early weeks of deployment and patients used the app in the latter half of their participation. We observed a steep decline in CM app use around 3 months after initial use, which is in line with data showing >60% decrease in the number of users 3 months after downloading an app [56]. By contrast, patients were less engaged with the app when they were first enrolled in the study, presumably owing to the inability to send or respond to messages, but they continued to use the app, namely accessing documents and messages, for several months toward the middle and end of their study participation, even when CMs were no longer using the app. This suggested that patients may have found continued utility in the app, but qualitative feedback from CMs suggests that they may have preferred to use other means to communicate with and respond to patients, such as phone calls. CMs reported a strong desire for synchronous 2-way messaging (vs asynchronous) or at least a way for them to be notified if a patient received a message. Improving the utility for CMs could increase the potential for sustained use of PlusCare. There remains need for features that can effectively streamline and simplify CM and AYA living with HIV tasks and better support timely communication when needed.

Another example of enhancing provider-patient communication is the sharing of laboratory results, which even when passively shared, may be best conveyed with an opportunity for immediate follow-up and discussion between provider and patient. CMs suggested that the system was particularly useful in communicating with, including sending positive affirmations to, patients and supporting remote document sharing and sign-off, thereby reducing the need for in-person visits, which has been particularly challenging for patients with longer commutes. This is supported by use data, revealing that messaging and managing documents were the most used features. CM feedback also suggested that the system may be particularly helpful for engaging less-adherent individuals who can benefit from more frequent, directed communication and check-ins supported by a shared platform where multiple CMs can follow-up with a given patient. CMs at one site (BMC) reported that they use a tag team approach where a different CM is scheduled for different time windows to be the point person to follow-up with a nonadherent patient.

The discrepancy in perceived usability and actual use highlights the need to continually assess the acceptability of mHealth tools in real-world settings and to use an iterative design process that incorporates and focuses on system improvements based on direct feedback from users to the extent possible. Opportunities remain for the development of technologies that are responsive to the needs of AYA living with HIV, both in the United States and globally [57-59].

Limitations

This study was designed to explore the naturalistic use of this system in various settings. As such, participants were

encouraged to use the system as needed; however, there were no study requirements to ensure baseline use. Generally, low but variable levels of PlusCare app use were observed at both sites, making it difficult to attribute changes in health and clinical outcomes directly to the use of PlusCare.

Another limitation was the limited qualitative feedback and incomplete survey data from patients compared between the time points. The response rates to 6- and 12-month survey assessments were 89% and 84%, respectively. Qualitative check-ins and end-of-study interviews were conducted only with CMs to probe their experience using the system with their patients, while separate usability feedback was collected from patients via survey. It is possible that direct interviews with patients may have provided additional richer insights into, for example, the observation of inverse trends in use between patients and CMs.

Another limitation was related to the sample of participants in the study. Patients had to have been receiving care for at least a year before enrollment in the study. As such, our study population may have been more committed to managing their HIV care than other AYA living with HIV and thus were generally virally suppressed. Future studies could examine mHealth use across the HIV care continuum by including those who are newly enrolled, as well as those already enrolled in care. In addition, only patients with consistent access to the internet and mobile devices were included, excluding those who could potentially derive the greatest benefit from access to PlusCare.

External and temporal factors, such as changes in case management policies and the increasing age of the patient cohort, may have introduced bias to our results, given the quasi-experimental pre-post study design; thus, we cannot attribute our results solely to the intervention, and it is possible that these other factors may have contributed to any differences reported in the pre-post comparisons. Although it is possible that continuous access to PlusCare over the 12-month period could account for the overall high adherence rates and improvements in general health outcomes, it is difficult to attribute this result to PlusCare without a true comparison to usual care.

Conclusions

Following user-centered design principles, we designed and tested an HIV case management digital app with CMs and AYA living with HIV and found that access to this system correlated with increased self-efficacy in chronic disease management and health-related quality of life. PlusCare was designed to provide remote support for various day-to-day case management functions and decrease the need for in-person visits. The results suggested trends of decreased patient costs; however, there was no significant decrease in clinic visits or outpatient costs. Although it is difficult to attribute our results directly to the use of the PlusCare app given its low use and the lack of a control group, results from this study suggest that digital health tools supporting case management between the CM and AYA living with HIV hold promise in improving patient quality of life and self-efficacy in chronic disease management—a critical goal of care for CMs working with individuals who will one day need to take full ownership of their own health needs. As health care facilities continue to seek ways to reduce the paperwork burden, a shared digital platform, such as the one designed and studied here for AYA living with HIV and the frontline health workers they work directly with, has the potential to streamline and improve AYA living with HIV care across the United States and globally, in areas with increasing connectivity and access to mobile technology among youth and young adults.

Acknowledgments

The authors would like to thank all the case managers and patients who participated in the study and made it possible. The authors also thank Dr Cathryn Samples, Meryn Robinson, and Honora Einhorn for their contributions to the critical early stages of this project. This project was funded by the National Institutes of Health, National Institute of Mental Health, grant R44 MH117956. Drs Woods and Guss were also funded in part by the Health Resources and Services Administration of the US Department of Health and Human Services as part of an MCHB T71MC00009 LEAH training grant. The contents are those of the authors and do not necessarily represent the official views of, nor an endorsement, by the Health Resources and Services Administration, Health and Human Services, or the US government.

Conflicts of Interest

JF, YXH, JLJ, and VSK are employed by the organization (Dimagi, Inc) whose revenue is derived from the open-source platform on which the digital app described here was built and studied. JEH has served as a paid consultant for Merck.

Multimedia Appendix 1
Case manager check-in form.
[PDF File (Adobe PDF File), 93 KB - formative_v6i11e39357_app1.pdf]

Multimedia Appendix 2
Case manager end-of-study interview probes.
[PDF File (Adobe PDF File), 98 KB - formative_v6i11e39357_app2.pdf]

References

1. HIV by Group - HIV in the United States by Age. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/hiv/group/age/youth/index.html [accessed 2017-07-19]
2. Harrison KM, Song R, Zhang X. Life expectancy after HIV diagnosis based on national HIV surveillance data from 25 states, United States. J Acquir Immune Defic Syndr 2010 Jan;53(1):124-130. [doi: 10.1097/QAI.0b013e3181b563e7] [Medline: 19730109]
3. Gortmaker SL, Hughes M, Cervia J, Brady M, Johnson GM, Seage 3rd GR, Pediatric AIDS Clinical Trials Group Protocol 219 Team. Effect of combination therapy including protease inhibitors on mortality among children and adolescents infected with HIV-1. N Engl J Med 2001 Nov 22;345(21):1522-1528. [doi: 10.1056/NEJMoa011157] [Medline: 11794218]
4. Antiretroviral Therapy Cohort Collaboration. Life expectancy of individuals on combination antiretroviral therapy in high-income countries: a collaborative analysis of 14 cohort studies. Lancet 2008 Jul 26;372(9635):293-299 [FREE Full text] [doi: 10.1016/S0140-6736(08)61113-7] [Medline: 18657708]
5. Samji H, Cescon A, Hogg RS, Modur SP, Althoff KN, Buchacz K, North American AIDS Cohort Collaboration on Research and Design (NA-ACCORD) of IeDEA. Closing the gap: increases in life expectancy among treated HIV-positive individuals in the United States and Canada. PLoS One 2013 Dec 18;8(12):e81355 [FREE Full text] [doi: 10.1371/journal.pone.0081355] [Medline: 24367482]
6. Mirani G, Williams PL, Chernoff M, Abzug MJ, Levin MJ, Seage 3rd GR, IMPAACT P1074 Study Team. Changing trends in complications and mortality rates among US youth and young adults with HIV infection in the era of combination antiretroviral therapy. Clin Infect Dis 2015 Dec 15;61(12):1850-1861 [FREE Full text] [doi: 10.1093/cid/civ687] [Medline: 26270680]
7. Monitoring Selected National HIV Prevention and Care Objectives By Using HIV Surveillance Data United States and 6 Dependent Areas, 2019. HIV Surveillance Supplemental Report. Centers for Disease Control and Prevention. 2021 May. URL: https://www.cdc.gov/hiv/library/reports/hiv-surveillance/vol-26-no-2/index.html [accessed 2022-10-30]
8. Tanner AE, Philbin MM, Ma A, Chambers BD, Nichols S, Lee S, Adolescent Trials Network for HIV/AIDS Interventions. Adolescent to adult HIV health care transition from the perspective of adult providers in the United States. J Adolesc Health 2017 Oct;61(4):434-439 [FREE Full text] [doi: 10.1016/j.jadohealth.2017.05.011] [Medline: 28754584]
9. Yehia BR, Stephens-Shields AJ, Fleishman JA, Berry SA, Agwu AL, Metlay JP, et al. The HIV care continuum: changes over time in retention in care and viral suppression. PLoS One 2015 Jun 18;10(6):e0129376 [FREE Full text] [doi: 10.1371/journal.pone.0129376] [Medline: 26086089]
10. Cahill SR, Mayer KH, Boswell SL. The Ryan White HIV/AIDS program in the age of health care reform. Am J Public Health 2015 Jun;105(6):1078-1085. [doi: 10.2105/AJPH.2014.302442] [Medline: 25880940]
11. Katz MH, Cunningham WE, Fleishman JA, Andersen RM, Kellogg T, Bozzette SA, et al. Effect of case management on unmet needs and utilization of medical care and medications among HIV-infected persons. Ann Intern Med 2001 Oct 16;135(8 Pt 1):557-565. [doi: 10.7326/0003-4819-135-8_part_1-200110160-00006] [Medline: 11601927]
12. Johnson RL, Botwinick G, Sell RL, Martinez J, Siciliano C, Friedman LB, et al. The utilization of treatment and case management services by HIV-infected youth. J Adolesc Health 2003 Aug;33(2 Suppl):31-38. [doi: 10.1016/s1054-139x(03)00158-7] [Medline: 12888285]
13. Tenner AD, Trevithick LA, Wagner V, Burch R. Seattle YouthCare's prevention, intervention, and education program: a model of care for HIV-positive, homeless, and at-risk youth. J Adolesc Health 1998 Aug;23(2 Suppl):96-106. [doi: 10.1016/s1054-139x(98)00057-3] [Medline: 9712257]
14. Woods ER, Samples CL, Melchiono MW, Keenan PM, Fox DJ, Harris SK, Boston HAPPENS Program Collaborators. Initiation of services in the Boston HAPPENS Program: human immunodeficiency virus-positive, homeless, and at-risk youth can access services. AIDS Patient Care STDS 2002 Oct;16(10):497-510. [doi: 10.1089/10872910260351276] [Medline: 12442735]
15. Davila JA, Miertschin N, Sansgiry S, Schwarzwald H, Henley C, Giordano TP. Centralization of HIV services in HIV-positive African-American and Hispanic youth improves retention in care. AIDS Care 2013;25(2):202-206. [doi: 10.1080/09540121.2012.689811] [Medline: 22708510]
16. Wohl AR, Garland WH, Wu J, Au CW, Boger A, Dierst-Davies R, et al. A youth-focused case management intervention to engage and retain young gay men of color in HIV care. AIDS Care 2011 Aug;23(8):988-997. [doi: 10.1080/09540121.2010.542125] [Medline: 21390879]
17. Conviser R, Pounds MB. Background for the studies on ancillary services and primary care use. AIDS Care 2002 Aug;14 Suppl 1:S7-14. [doi: 10.1080/09540120220149993] [Medline: 12204138]
18. Tanner AE, Philbin MM, Duval A, Ellen J, Kapogiannis B, Fortenberry JD, Adolescent Trials Network for HIV/AIDS Interventions. "Youth friendly" clinics: considerations for linking and engaging HIV-infected adolescents into care. AIDS Care 2014 Feb;26(2):199-205 [FREE Full text] [doi: 10.1080/09540121.2013.808800] [Medline: 23782040]
19. Philbin MM, Tanner AE, Duval A, Ellen J, Kapogiannis B, Fortenberry JD. Linking HIV-positive adolescents to care in 15 different clinics across the United States: creating solutions to address structural barriers for linkage to care. AIDS Care 2014 Jan;26(1):12-19 [FREE Full text] [doi: 10.1080/09540121.2013.808730] [Medline: 23777542]

20. Cunningham WE, Wong M, Hays RD. Case management and health-related quality of life outcomes in a national sample of persons with HIV/AIDS. J Natl Med Assoc 2008 Jul;100(7):840-847 [FREE Full text] [doi: 10.1016/s0027-9684(15)31379-1] [Medline: 18672562]
21. Blair JM, McNaghten AD, Frazier EL, Skarbinski J, Huang P, Heffelfinger JD. Clinical and behavioral characteristics of adults receiving medical care for HIV infection --- Medical Monitoring Project, United States, 2007. MMWR Surveill Summ 2011 Sep 02;60(11):1-20 [FREE Full text] [Medline: 21881551]
22. Woods ER, Samples CL, Melchiono MW, Harris SK. Boston HAPPENS Program: HIV-positive, homeless, and at-risk youth can access care through youth-oriented HIV services. Semin Pediatr Infect Dis 2003 Jan;14(1):43-53. [doi: 10.1053/spid.2003.127217] [Medline: 12748922]
23. Woods ER, Samples CL, Melchiono MW, Keenan PM, Fox DJ, Chase LH, et al. The Boston HAPPENS program: needs and use of services by HIV-positive compared to at-risk youth, including gender differences. Eval Program Plann 2000 May;23(2):187-198. [doi: 10.1016/S0149-7189(00)00007-0]
24. Katz MH, Cunningham WE, Mor V, Andersen RM, Kellogg T, Zierler S, et al. Prevalence and predictors of unmet need for supportive services among HIV-infected persons: impact of case management. Med Care 2000 Jan;38(1):58-69. [doi: 10.1097/00005650-200001000-00007] [Medline: 10630720]
25. Harris SK, Samples CL, Keenan PM, Fox DJ, Melchiono MW, Woods ER, Boston HAPPENS Program. Outreach, mental health, and case management services: can they help to retain HIV-positive and at-risk youth and young adults in care? Matern Child Health J 2003 Dec;7(4):205-218. [doi: 10.1023/a:1027386800567] [Medline: 14682498]
26. Anderson M, Jiang J. Teens, social media and technology 2018. Pew Research Center. 2018 May 31. URL: https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/ [accessed 2022-10-30]
27. Swendeman D, Rotheram-Borus MJ. Innovation in sexually transmitted disease and HIV prevention: internet and mobile phone delivery vehicles for global diffusion. Curr Opin Psychiatry 2010 Mar;23(2):139-144 [FREE Full text] [doi: 10.1097/YCO.0b013e328336656a] [Medline: 20087189]
28. 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]
29. 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]
30. Guss CE, Woods ER, Cooper ER, Burchett S, Fuller J, Dumont O, et al. 252. Pluscare: a mobile platform designed to increase linkage to care for youth living with HIV/AIDS. J Adolesc Health 2020 Feb 1;66(2):S127-S128. [doi: 10.1016/j.jadohealth.2019.11.255]
31. DocuSign | #1 in Electronic Signature and Agreement Cloud. URL: https://www.docusign.com/ [accessed 2022-05-08]
32. Sauro J, Lewis JR. When designing usability questionnaires, does it hurt to be positive? In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2011 May Presented at: CHI '11; May 7-12, 2011; Vancouver, Canada p. 2215-2224. [doi: 10.1145/1978942.1979266]
33. Sauro J, Lewis JR. Quantifying the User Experience: Practical Statistics for User Research. 2nd edition. Burlington, MA, USA: Morgan Kaufmann; 2016.
34. Hyzy M, Bond R, Mulvenna M, Bai L, Dix A, Leigh S, et al. System usability scale benchmarking for digital health apps: meta-analysis. JMIR Mhealth Uhealth 2022 Aug 18;10(8):e37290 [FREE Full text] [doi: 10.2196/37290] [Medline: 35980732]
35. Lu M, Safren SA, Skolnik PR, Rogers WH, Coady W, Hardy H, et al. Optimal recall period and response task for self-reported HIV medication adherence. AIDS Behav 2008 Jan;12(1):86-94. [doi: 10.1007/s10461-007-9261-4] [Medline: 17577653]
36. Feldman BJ, Fredericksen RJ, Crane PK, Safren SA, Mugavero MJ, Willig JH, et al. Evaluation of the single-item self-rating adherence scale for use in routine clinical care of people living with HIV. AIDS Behav 2013 Jan;17(1):307-318 [FREE Full text] [doi: 10.1007/s10461-012-0326-7] [Medline: 23108721]
37. Lorig KR, Sobel DS, Ritter PL, Laurent D, Hobbs M. Effect of a self-management program on patients with chronic disease. Eff Clin Pract 2001;4(6):256-262. [Medline: 11769298]
38. Measuring Healthy Days: Population Assessment of Health-Related Quality of Life. Centers for Disease Control and Prevention. 2000 Nov. URL: https://www.cdc.gov/hrqol/pdfs/mhd.pdf [accessed 2022-10-30]
39. Belzer ME, Naar-King S, Olson J, Sarr M, Thornton S, Kahana SY, Adolescent Medicine Trials Network for HIV/AIDS Interventions. The use of cell phone support for non-adherent HIV-infected youth and young adults: an initial randomized and controlled intervention trial. AIDS Behav 2014 Apr;18(4):686-696 [FREE Full text] [doi: 10.1007/s10461-013-0661-3] [Medline: 24271347]
40. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria: The R Foundation; 2020. URL: https://www.R-project.org/ [accessed 2022-10-30]
41. SAS/ACCESS Software. SAS Institute. Cary, NC, USA: SAS Institute Inc URL: https://www.sas.com/en_in/software/access.html [accessed 2022-10-30]

42. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
43. Bhaumik U, Norris K, Charron G, Walker SP, Sommer SJ, Chan E, et al. A cost analysis for a community-based case management intervention program for pediatric asthma. J Asthma 2013 Apr;50(3):310-317. [doi: 10.3109/02770903.2013.765447] [Medline: 23311526]
44. Bhaumik U, Sommer SJ, Giller-Leinwohl J, Norris K, Tsopelas L, Nethersole S, et al. Boston children's hospital community asthma initiative: five-year cost analyses of a home visiting program. J Asthma 2017 Mar;54(2):134-142. [doi: 10.1080/02770903.2016.1201837] [Medline: 27624870]
45. Woods ER, Bhaumik U, Sommer SJ, Ziniel SI, Kessler AJ, Chan E, et al. Community asthma initiative: evaluation of a quality improvement program for comprehensive asthma care. Pediatrics 2012 Mar;129(3):465-472. [doi: 10.1542/peds.2010-3472] [Medline: 22351890]
46. Woods ER, Bhaumik U, Sommer SJ, Chan E, Tsopelas L, Fleegler EW, et al. Community asthma initiative to improve health outcomes and reduce disparities among children with asthma. MMWR Suppl 2016 Feb 12;65(1):11-20. [doi: 10.15585/mmwr.su6501a4] [Medline: 26916259]
47. Daher J, Vijh R, Linthwaite B, Dave S, Kim J, Dheda K, et al. Do digital innovations for HIV and sexually transmitted infections work? Results from a systematic review (1996-2017). BMJ Open 2017 Nov 03;7(11):e017604 [FREE Full text] [doi: 10.1136/bmjopen-2017-017604] [Medline: 29101138]
48. Cooper V, Clatworthy J, Whetham J, EmERGE Consortium. mHealth interventions to support self-management in HIV: a systematic review. Open AIDS J 2017 Nov 21;11:119-132 [FREE Full text] [doi: 10.2174/1874613601711010119] [Medline: 29290888]
49. Henny KD, Wilkes AL, McDonald CM, Denson DJ, Neumann MS. A rapid review of eHealth interventions addressing the continuum of HIV care (2007-2017). AIDS Behav 2018 Jan;22(1):43-63 [FREE Full text] [doi: 10.1007/s10461-017-1923-2] [Medline: 28983684]
50. Recommendations for case management collaboration and coordination in federally funded HIV/AIDS programs : a coordinated and collaborative case management environment for the client. Centers for Disease Control and Prevention. 2008 Nov 14. URL: https://stacks.cdc.gov/view/cdc/45790 [accessed 2017-08-31]
51. Sowell RL, Grier J. Integrated case management: the AID Atlanta Model. J Case Manag 1995;4(1):15-21. [Medline: 7795537]
52. McDonald KM, Sundaram V, Bravata D, Lewis R, Lin N, Kraft SA, et al. Closing the Quality Gap: A Critical Analysis of Quality Improvement Strategies (Vol. 7: Care Coordination). Rockville, MD, USA: Agency for Healthcare Research and Quality (US); 2007.
53. Rosenthal R. The file drawer problem and tolerance for null results. Psychol Bull 1979 May;86(3):638-641. [doi: 10.1037/0033-2909.86.3.638]
54. Czaja SJ, Zarcadoolas C, Vaughon WL, Lee CC, Rockoff ML, Levy J. The usability of electronic personal health record systems for an underserved adult population. Hum Factors 2015 May;57(3):491-506 [FREE Full text] [doi: 10.1177/0018720814549238] [Medline: 25875437]
55. 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 [FREE Full text] [doi: 10.1007/s11606-016-3771-6] [Medline: 27418347]
56. Average three month user retention and churn rate of mobile apps worldwide as of 2nd half 2018. Statista Research Department. 2021 Jul 7. URL: https://www.statista.com/statistics/384224/monthly-app-launches-churn/ [accessed 2022-10-30]
57. Mulawa MI, LeGrand S, Hightow-Weidman LB. eHealth to enhance treatment adherence among youth living with HIV. Curr HIV/AIDS Rep 2018 Aug;15(4):336-349 [FREE Full text] [doi: 10.1007/s11904-018-0407-y] [Medline: 29959649]
58. 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]
59. Kharono B, Kaggiah A, Mugo C, Seeh D, Guthrie BL, Moreno M, et al. Mobile technology access and use among youth in Nairobi, Kenya: implications for mobile health intervention design. Mhealth 2022 Jan 20;8:7 [FREE Full text] [doi: 10.21037/mhealth-21-23] [Medline: 35178438]

Abbreviations

AYA: adolescents and young adults
BCH: Boston Children’s Hospital
BMC: Boston Medical Center
CM: case manager
HRQOL-4: Health-Related Quality of Life-4
IRB: institutional review board

mHealth: mobile health
RA: research assistant
SUS: System Usability Scale

Edited by G Eysenbach; submitted 08.05.22; peer-reviewed by 翔 闫, G Sabben; comments to author 02.07.22; revised version received 24.08.22; accepted 08.10.22; published 21.11.22.

Please cite as:
Fee C, Fuller J, Guss CE, Woods ER, Cooper ER, Bhaumik U, Graham D, Burchett SK, Dumont O, Martey EB, Narvaez M, Haberer JE, Swendeman D, Mulvaney SA, Kumar VS, Jackson JL, Ho YX
A Digital Platform to Support HIV Case Management for Youth and Young Adults: Mixed Methods Feasibility Study
JMIR Form Res 2022;6(11):e39357
URL: https://formative.jmir.org/2022/11/e39357
doi:10.2196/39357
PMID:36409541

Original Paper

Patients' and Health Care Providers' Perceptions on mHealth Use After High-Altitude Climate Therapy for Severe Asthma: Mixed Methods Study

Rishi Khusial[1], BSc; Sophia van Koppen[1], MSc, MD; Persijn Honkoop[1], MSc, MD, PhD; Lucia Rijssenbeek-Nouwens[2], MD, PhD; Karin Berthine Fieten[2], PhD; Sascha Keij[1], MSc; Marieke Drijver-Messelink[2], MSc, MD; Jacob Sont[1], PhD

[1]Department of Biomedical Data Sciences, Section Medical Decision Making, Leiden University Medical Center, Leiden, Netherlands
[2]Dutch Asthma Centre Davos, Davos, Switzerland

Corresponding Author:
Rishi Khusial, BSc
Department of Biomedical Data Sciences, Section Medical Decision Making
Leiden University Medical Center
Albinusdreef 2
Leiden, 2300 RC
Netherlands
Phone: 31 7152 64574
Email: r.j.khusial@lumc.nl

Abstract

Background: Asthma is a common chronic disease with various clinical presentations. Although most patients are able to reach good asthma control, some patients are not able to reach sufficient asthma control following the regular treatment guidelines and could be referred to high-altitude climate therapy (HACT). HACT includes environmental trigger avoidance in the alpine climate with multidisciplinary clinical treatment. Patients with severe and difficult-to-control asthma, who are unable to reach asthma control at sea level, can follow a 12-week lung rehabilitation program at 1600 m above sea level. Mobile health (mHealth) tools can be used to enhance self-management in these patients when they return home. For an mHealth system to be effective, it must meet the expectations of the end users.

Objective: In this Davos@home study, we explored the attitudes toward mHealth aimed at supporting the self-management of patients with severe, difficult-to-control asthma who underwent HACT and asthma health care providers.

Methods: In the first stage, interviews with referrers to HACT and focus groups with patients with asthma who participated in or completed HACT were conducted. The data were then analyzed thematically. On the basis of these results, a questionnaire was developed. In the second stage of the study, this questionnaire, combined with the Asthma Control Questionnaire and the Individual Innovativeness Questionnaire, was provided to patients who completed HACT.

Results: In total, 11 interviews and 3 focus groups (n=18, age 47.6, SD 12.1 years, Asthma Control Questionnaire score 2.6, SD 1.0) were conducted. A total of 3 themes were identified: potential goals, useful measurements, and perceived barriers and facilitators. The questionnaire developed in stage 2 included items based on these results. The most agreed-upon goal among the 52 patients who completed the questionnaire was to increase their asthma control (45/52, 86% of the patients).

Conclusions: Different patients reported that they would benefit the most from different functionalities. Therefore, it is important to tailor functionalities to individual (treatment) goals. When developing an mHealth intervention, it is important to allow personalization to avoid overwhelming the users.

(JMIR Form Res 2022;6(11):e26925) doi:10.2196/26925

KEYWORDS

eHealth; mobile health; mHealth; asthma; self-management; home monitoring; mobile phone

Introduction

Background

Asthma is one of the most common chronic diseases worldwide, currently affecting approximately 300 million individuals [1]. Although a majority of the patients can be treated in primary care, some patients have persistent uncontrolled asthma and are referred to specialized asthma care [2]. Within the small group of patients with the most severe, difficult-to-control asthma, standard therapeutic treatments are unable to reach sufficient levels of asthma control [3]. Nonpharmacological interventions, including high-altitude climate therapy (HACT), are proposed for this group of patients [4].

The Davos Dutch Asthma Centre offers patients with severe, difficult-to-control asthma a 12-week high-altitude revalidation program in Switzerland, located 1600 m above sea level [5]. Owing to environmental trigger avoidance in the alpine climate, high-altitude areas may be helpful in to improve asthma control in some patients [6]. A recent study showed improvements in the quality of life and clinical outcomes in patients with severe asthma who completed high-altitude rehabilitation [7]. Following the Dutch guidelines [8], patients are only eligible for HACT when they have uncontrolled asthma (Asthma Control Questionnaire [ACQ] score ≥1.5 [9] and ≥2 exacerbations per year or the use of systemic steroids for ≥6 months per year) despite using high-dose inhalation corticosteroids and long-acting beta-2 agonists [8]. The inhalation technique and medication adherence need to be optimized, and patients are required to have stopped smoking for at least 6 months before referral to HACT is possible. Every year, approximately 80 patients visit the Davos Dutch Asthma Centre.

During their stay in Davos, patients work on different aspects of their rehabilitation, including physiotherapy, asthma education sessions, and self-management skills [10]. On discharge from Davos, the patients are referred back to their pulmonologists in the Netherlands. After discharge, the patients are transferred from 24/7 care to a normal home situation. Between visits to their pulmonologists, patients must manage their asthma themselves. With the help of their health care providers (HCPs), patients engage in self-management [6]. Patients with self-management support after HACT have a lower decline in their asthma control and quality of life compared with patients without self-management support [11]. It is also known that self-management programs have a beneficial effect on asthma control, as reported in a Cochrane review [12]. To improve long-term self-management at home, mobile health (mHealth) systems could be used as self-management support. Using mHealth, the transition to self-management after the patients return from Davos could be softened.

mHealth encompasses the use of mobile devices (such as smartphones, tablets, or wearable devices) to support health care delivery and self-management [13]. Although there are a lot of asthma apps available, not many are clinically validated. During the developmental stage of an mHealth system, it is important to consider the attitudes and expectations of the end users to ascertain the actual use of the system [14]. A user-centered design may allow for the technology to meet the end users' expectations and, therefore, improve asthma-related outcomes [15]. The effective use of an mHealth tool requires a good fit between the system and end users [16].

Objective

The first step in designing a user-centered mHealth self-management support system is to determine the attitudes of patients with asthma and their HCPs toward the functionalities the mHealth system should include [17]. This manuscript describes the first part of the Davos@home study, and it is the first step in designing an mHealth system to assist patients with severe asthma with self-management after completing HACT. The aim of this study was to gain insights into the opinions of patients and HCPs on the role of mHealth after HACT to develop an mHealth system that would match their needs and expectations.

Methods

Overview

This study consisted of 2 stages. In stage 1, interviews with HCPs (pulmonologists and specialized asthma nurse practitioners) and focus groups with patients with asthma were conducted. Stage 2 involved the development and completion of a questionnaire for patients with asthma who completed HACT in Davos to quantify the opinions generated during the focus groups in a larger population.

Stage 1

HCPs from hospitals with frequent referrals to HACT in Davos were contacted by the research team for 30-minute interviews in their hospitals. In total, 10 different hospitals with specialized asthma care in the Netherlands were contacted. From every hospital, an interview with 1 pulmonologist and 1 specialized asthma nurse was planned. The interviews were semistructured (Textbox 1 provides a short interview format and the supplement presents the full interview plan). The goal of the interviews was to gain an insight into the attitudes of HCPs toward mHealth-supported self-management. The interviews were conducted between April and June 2019.

Next, a total of 3 focus groups were scheduled. For each focus group, 4 to 8 patients were invited. Patients were eligible if they were aged ≥18 years at the time of the focus group. The inability to understand or speak Dutch was an exclusion criterion. Focus groups were conducted in June 2019.

Every patient who was available in Davos during the visit of the research team was contacted by the medical staff of the revalidation center to participate in the focus groups and was given an information pamphlet. This resulted in 2 focus groups of patients with asthma who were undergoing HACT in Davos and 1 focus group of patients who had completed the HACT program and already returned to the Netherlands.

Textbox 1. Interview and focus group format.

Interview with health care provider

- Introduction
- Questions relating to
  - Health care providers' current knowledge and use of mobile health apps for asthma care
  - What measurements would be useful for self-management:
    - Environmental factors
    - Physiological factors
    - Behavioral factors
    - Psychological factors
  - Use of alerts and reminders
  - System-based feedback and recommendations
  - Acceptable time investment
  - Privacy
  - Product design
  - Additional functions
  - Barriers preventing use
- Conclusion and summary

Focus group with patients

- Introduction
- Questions relating to
  - Participants' current knowledge and use of mobile health apps for asthma management
  - What measurements would be useful for self-management:
    - Environmental factors
    - Physiological factors
    - Behavioral factors
    - Psychological factors
  - Use of alerts and reminders
  - System-based feedback and recommendations
  - Acceptable time investment
  - Privacy
  - Product design
  - Additional functions
  - Barriers preventing use
- Participants' debrief

Upon arrival, the format of the focus group was explained, informed consent forms were signed, and general participant demographics were collected using a short questionnaire. Only the researchers and patients were present during the focus groups. Each focus group discussion lasted approximately 2 hours with a short break in the middle.

The interviews and focus group discussions used a funnel-type method of questioning (from general to specific questions). Before the interviews and focus groups, the researchers constructed an interview plan with questions on different subjects. Prompts were used to probe further details if necessary. The interviews were audio recorded, and the focus groups were video recorded to assist with transcription.

Interviews and focus groups were conducted by RK or SK. RK (BSc in medicine and BSc in political science) is a male medical student and PhD candidate. SK (at the time BSc in medicine) is a female medical student and research intern. The researchers introduced themselves before the interviews and focus groups.

The focus groups and interviews were transcribed verbatim and underwent thematic analysis [18]. RK and SK coded each transcript independently. After transcription, discrepancies between codes or how parts of the transcripts were coded were discussed between RK and SK until consensus was reached [18]. Next, the codes were grouped into different applicable themes. Data management was supported by the ATLAS.ti qualitative analysis software (version 7).

The thematic analyses of the focus groups and interviews resulted in the construction of a closed-ended questionnaire aimed at quantifying the opinions generated during stage 1.

Stage 2

The web-based questionnaire consisted of 4 parts: baseline patient characteristics, 34 questions based on stage 1 of the study, the ACQ [9], and the Individual Innovativeness Questionnaire (19). The questionnaire generated during stage 1 consisted of questions that participants could rate on a 5-point Likert scale (strongly disagree, disagree, unsure, agree, and strongly agree). Favorable outcomes (agree and strongly agree) were considered as being in agreement with the statement. The Individual Innovativeness Questionnaire is a tool designed to measure individuals' attitudes toward change, and participants can be fitted into 1 of 5 categories (innovators, early adopters, early majority, late majority, and laggards or traditionalists) [19].

The patient federation *Vereniging Nederland Davos* was actively involved in participant recruitment. They made the complete questionnaire available to patients who they knew had completed HACT in Davos in the past. Patients were also recruited through their social media handles, newsletters, and websites. The questionnaire was administered on the web for 2 months. No formal sample size calculations were performed because statistical analyses were not planned.

Ethics Approval

This study was conducted in accordance with the principles of the Declaration of Helsinki (2013). The Medical Ethical Committee of the Leiden University Medical Center offered exemption for the ethics approval of the study protocol (P19.039), as it was not required under Dutch law.

Results

Participants

In total, 11 HCPs from 7 different hospitals were interviewed, and no replies were received from 3 hospitals. From every hospital, 1 pulmonologist was interviewed, and from 4 of these hospitals, 1 nurse practitioner was also interviewed. In addition, 20 patients with asthma were contacted to participate in the focus groups. Of them, 1 patient declined and 1 patient did not show up at the focus group because of (not study related) sickness. In total, 18 patients with asthma participated in 3 separate focus group sessions (n=4, n=6, and n=8 per focus group).

A total of 52 patients completed the questionnaire from stage 2. Table 1 provides a combined overview of the baseline characteristics of the focus group participants and those who completed the questionnaires.

In addition to the described baseline characteristics, we also queried HCPs and focus group patients on the current use of mHealth and eHealth solutions. Overall, 1 HCP used mHealth in relation to immunotherapy, 1 HCP advised a smoking cessation app, 1 HCP used an allergy program (study phase), and 1 HCP had frequent video consultations. Among the focus group patients, 1 patient previously used the Fitbit app, 1 patient used an app named PatientCoach [11] (in the study setting), and 1 patient reported using an asthma diary app.

Table 1. Baseline characteristics of the focus group patients and questionnaire respondents.

	Values
Stage 1 (n=18)	
Sex, n (%)	
Female	9 (50)
Male	9 (50)
Age (years), mean (SD)	49.2 (12.1)
Current smoking status, n (%)	
Yes	0 (0)
No	13 (72)
Previously	5 (28)
ACQ[a] score	
Values, mean (SD)	2.1 (1.0)
Controlled (<0.75), n (%)	1 (6)
Partly controlled (0.75-1.5), n (%)	4 (22)
Uncontrolled (>1.5), n (%)	11 (61)
Stage 2 (n=52)	
Sex, n (%)	
Female	37 (71)
Male	15 (29)
Age (years), mean (SD)	47.2 (14.5)
Current smoking status, n (%)	
Yes	0 (0)
No	44 (85)
Previously	8 (15)
Own a smartphone, n (%)	51 (98)
Visits to high-altitude climate therapy (times), n (%)	
Once	24 (46)
Twice	15 (29)
≥3 times	13 (25)
ACQ score	
Values, mean (SD)	2.6 (1.1)
Controlled (<0.75), n (%)	4 (8)
Partly controlled (0.75-1.5), n (%)	2 (4)
Uncontrolled (>1.5), n (%)	46 (88)
Innovator status, n (%)	
Innovators	8 (15)
Early adopters	18 (35)
Early majority	18 (35)
Late majority	7 (13)
Laggards, traditionalists	1 (2)

[a]ACQ: Asthma Control Questionnaire.

Themes Created in Stage 1

Three distinct themes were identified in stage 1: potential goals, useful measurements, and perceived barriers and facilitators.

Potential Goals

In total, 8 main categories of potential goals were formulated based on the data (Textbox 2).

Patients and HCPs reported that mHealth-supported self-management could improve or maintain disease-specific outcomes, including asthma control, quality of life, exacerbation rate, medication adherence, and lifestyle. The HCPs also mentioned that mHealth could reduce the number of in-person consultations by potentially replacing routine meetings with video calls or chats.

Other potential goals of an mHealth system named in the interviews and focus groups were specifically related to HACT. Patients wanted to maintain the (self-management) skills and knowledge they learned during their revalidation and would appreciate to have the education modules from the revalidation available to them electronically so that they could further train themselves.

Furthermore, patients reported that mHealth could improve their disease awareness with the help of monitoring. Being confronted by ACQ data helps the patients understand the severity of their symptoms and, therefore, could help patients respond to their symptoms sooner. Patients feel that they are so used to being sick that they claim to underestimate the worsening of symptoms, and they believe that monitoring could assist their HCPs in better understanding their disease progression. It is important to manage expectations of the end users. Although HCPs reported that they see mHealth primarily as a self-management support tool for patients, the patients would like to see it aid the HCPs in making treatment decisions.

Textbox 2. Potential goals with example quotes.

Improving asthma control

- "That is in fact the most important thing, so the patient should obtain a better quality of life or have a better asthma control." [HCP3]

Improving the quality of life

- "It (the goal) is very simple, the patient should benefit from it [...] having a better quality of life" [HCP3]

Reducing exacerbation rate

- "If the app and the patient together have prevented an exacerbation, well yes that would be of course very nice." [HCP7]

Improving medication adherence

- "Because that is it eventually, if you talk about what might be the biggest issue with asthma, it is noncompliance with therapy." [HCP2]

Sustaining or inducing lifestyle changes

- "You know what might be a more important goal? eHealth/mHealth could induce lifestyle changes in a patient." [HCP1]

Limiting consultations

- "And perhaps that it is also more convenient for us [HCPs], and for the patient, that it may save outpatient visits." [HCP1]

Retaining high-altitude climate therapy–learned skills and education

- "What is also important [sic], is that you keep remembering what the advises were from Davos that you received when you were send back home." [Patient 112]

Increasing disease awareness

- "If you are confronted with the data and think, we are not doing so well, that you can intervene faster, which in principle prevents you from getting worse." [Patient 111]

Useful Functions

Five different categories of functions an mHealth self-management intervention could contain were identified: "lifestyle," "education," "measurement devices," "psychological functions," and "other functions" (Textbox 3).

Within the lifestyle category, some patients wanted to record parameters related to their activity and weight. Other apps that could record food and calorie intake were also mentioned as useful tools that could be incorporated. Exercise was also considered a useful aspect to monitor. Pedometers provide more insights into the number of steps that are taken daily, and a function to set individual goals could motivate patients. However, some patients reported that tracking their physical conditions was not desirable for them. A patient also mentioned that tracking sleep quality would provide more insights into his asthma, as nocturnal symptoms are frequently present in patients with severe asthma. It was said that although it might be difficult to initiate lifestyle changes with just an app, it could support maintaining lifestyle changes.

Textbox 3. Useful functions with example quotes.

Lifestyle

- "But for us it [medication reminders] is not necessary, then again, I do not rule out there are people who certainly do need it." [Patient 111]
- "Since we often get prednisone or dexamethasone and then our weight after that increases, or some people lose weight. That you also can look at it [an app] and that it is motivating for others to lose weight, and to gain weight again." [Patient 0217]

Education

- "What is asthma actually, a piece of background information, but also if you do not take you medication in time, what happens to the lungs. I got booklets for this, it is something small, but still important." [HCP5]
- "I want to have it [education modules] as optional. If you think that you want it, that you can easily open it." [Patient 616]

Measurement devices

- "Yes, I think it is like you said, measuring is knowing, how are you doing at the moment. I think that is very important." [Patient 111]
- "With heart rate, that maybe people could insert a maximum heart rate and if you are too close to it you receive a notification. Often you can feel it already, but there are moments you do not realize it." [Patient 1116]

Psychological functions

- "Well if we are talking about an app, quick access for those who need it with the psychologist or with the care coordinator [from HACT] would be practical." [Patient 212]
- "Look in the end I think someone still needs a direct conversation. At the moment you fill in 'well I am feeling a little depressed today' or 'I am not really feeling well' if the app then says 'come on you can do it!' or 'look, the weather is nice outside' or 'the glass is half full instead of half empty'...I think you do not what to receive those kind of messages." [HCP3]

Other functions

- "Yes, you receive that [asthma action plan], but with me, it just hangs in my kitchen cupboard. If I open the door I see it and think 'oh yes, that is useful,' but I would personally find it convenient if I can find it on my phone." [Patient 212]
- "To return to the asthma action plan, everything is nicely written down on paper and to be very honest, things like that often ends up in a drawer, the drawer is closed and you never use it again." [Patient 0716]
- "I am missing the changes in weather for example. Pollen season, all indicators that are available that worsen asthma. It remains often a periodical thing, so yes, it would be nice to find it in there [in an app]." [Patient 516]

Patients reported that they wanted to have some form of education in an mHealth intervention. This varied from more general information about asthma to specific education about exacerbations and breathing techniques they learned during HACT. Patients reported that they also wanted information for their family and friends to help them understand the impact of severe asthma on daily life. HCPs reported that they wanted to have links to videos about correct inhaler use as a reminder for patients, such as the website [20], which has video tutorials about the use of inhalers. Patients from the focus groups reported that these videos could be presented in an app but thought that this might not be useful for themselves but rather for other patients. They also thought that medication reminders would be beneficial for these other patients because they reported that they mostly take their medication as prescribed. A 2-fold reminder system was suggested in a focus group: one reminder to take your daily inhalation medication as a push notification on your phone and another one after some time to check if the medication was indeed taken. With a single press, one should be able to dismiss these notifications.

Different devices were mentioned as useful additions to an effective mHealth self-management intervention. Heart rate and spirometry were often mentioned by patients as indicators that could be measured on a regular basis at home, with the addition of small home monitoring devices. Other parameters named in the focus groups were oxygen saturation, breathing frequency, and blood pressure. The HCPs thought that flooding the patients with numbers generated by these kinds of devices without proper medical guidance would not be helpful, as the patients also need to know what these values mean for their specific situation. Therefore, providing patients with data alone could create a false sense of security.

The patients also mentioned that they wanted a medication counter for their inhalers, indicating how many doses are left. Currently, only a few inhaler manufacturers include a counter on their casings. Patients are afraid that they will run out of (rescue) medication, and a counter would give them a feeling of security. They would like to receive a warning when they are running out of medication.

It was frequently mentioned by HCPs that severe asthma also has strong associations with psychological well-being. Severe asthma impacts the quality of life, and patients are provided with psychological consultations during their HACT revalidation if needed. Although both the patients and HCPs agreed that it would be difficult to incorporate adequate psychological support in mHealth, it is something that should not be overlooked when building an mHealth system. Although human interaction is key to psychological support, an mHealth system could provide a screening tool, for example, for anxiety and depression with

the help of a standardized questionnaire [21]. Opinions on this varied among patients; some thought it could be useful, whereas others definitely opposed it.

Other functions mentioned were having access to environmental data because many patients with asthma are hypersensitive to low air quality and allergens, including pollen. Patients and HCPs also wanted to monitor asthma-related parameters using questionnaires, including the ACQ and Asthma-Related Quality of Life Questionnaire. The frequency at which these questionnaires were suggested to be filled out varied between weekly and every 6 months or only when increased symptoms are experienced.

It was suggested that an asthma action plan be included in the app so that it would be available to patients when needed. Patients also wanted to have an option to share the data with their care team, including their pulmonologist, specialized asthma nurse practitioner, and, if necessary (varying per patient), general practitioner and physiotherapist. They would like to see treatment decisions based on the data collected, although this is not the primary aim of self-management. HCPs would also like to see the data of their patients if they help them make more informed decisions or gain insights into their patients' symptoms.

Perceived Barriers and Facilitators

As an intervention is only effective if it is actually used, patients and HCPs were asked about the barriers that prevent them from using an mHealth system (Textbox 4 provides example quotes). The most important barrier that the patients perceived was ineffectiveness. If they feel that the mHealth system does not support them in reaching their goals, they do not use it.

Textbox 4. Perceived barriers and facilitators with example quotes.

Effectiveness

- "It should not be a gadget that delivers more work to the patient and us and, like, in outcome in care is useless." [HCP1]

Time consumptions

- "The majority of thing should be automatic, so it will not cost us time, that it will not be a real burden. Look, if you tell me, like, I would think a minute a day is kind of acceptable to be actually actively involved with things." [Patient 116]

Too much information

- "I do not think it should be too full [of option], then it will not be attractive, it would seem to me." [Patient 0217]

Too complicated

- "But I think it will be very difficult to implement if it gets too complicated. So too many extra clicks on buttons, extra application and things." [HCP 6]

Privacy

- "I think I want it to be really secure [privacy protected], because it is really personal data. Because your things [data] from your pulmonologist are in it, from you psychologist are in it, a lot will be in." [Patient 212]

User-friendliness

- "It should be very user-friendly and I think it will be different in the future, that we can even have mHealth consultations." [HCP6]

Personalized

- "I think the action plan is practical, yes. But what would be even more practical is if you are able to adjust it. So if you return home, and things are changed in the Netherlands by your pulmonologist or whomever. So they can adjust it." [Patient 212]

Another important barrier is the time required for using the system. Patients would not use the system if entering data would take too much time. The use of automated data collection (with, eg, additional wearable devices) would partially address this problem. Patients were willing to wear additional devices, given that they were hypoallergenic, and if visible, it would be preferred if the devices were fashionable. The system must be part of their daily (or weekly) routine.

In addition to time consumption, the system should not contain too much information or be overly complicated. Patients were willing to use the system if the minimally required use to be effective was limited. The number of questionnaires that the patients were willing to fill varied, but they generally agreed that having a daily question to quantify their disease burden was too much work for insufficient disease gain. Implementing all previously named options into a single system could create an overly comprehensive system, which would prevent actual use because not all functions would be useful for all patients.

Privacy is also an important topic to consider, and the data need to be protected, as they are health data. Too many security measurements and log-in screens negatively affect user-friendliness.

Quantifying the Patient Opinions (Stage 2)

Stage 2 participants were predominantly female (37/52, 70%), and the vast majority had uncontrolled asthma (46/52, 89%; Table 1). The results of the questionnaire are provided in Table 2. Patients were positive with respect to all potential goals of

an mHealth system but were least positive about the possibility of preventing asthma exacerbation (31/52, 60%). With respect to potential use, the strongest agreement (>80%) was with the functionality to contact their care team; have information for important others; have their personal asthma action plan and outdoor air quality information available; and keep track of their rescue medication, lung function, and exercise behavior. However, only 69% (36/52) of the patients were willing to wear an additional wristwatch to track this. If a trade-off needed to be made between privacy and user-friendliness, 56% (29/52) of the patients preferred privacy.

In total, 65% (34/52) of the patients reported that they would like to have a smart inhaler that could record when they used their inhaler, and 58% (30/52) of the patients wanted medication reminders, although 78% (41/52) of the patients reported that they never forgot to take their medication (the remaining 11/52, 22% of the patients reported forgetting medication only sometimes).

In addition, 88% (46/52) of the patients found it important to customize an app according to their needs.

Table 2. Distribution of results from the stage 2 questionnaire study.

	Strongly disagree, n (%)	Unsure, n (%)	Strongly agree, n (%)
Potential goal of an mHealth[a] system			
Enhance the quality of life	4 (8)	10 (19)	38 (73)
Make me more aware of my asthma	6 (12)	4 (8)	42 (80)
Improve asthma control	4 (8)	3 (6)	45 (86)
Prevent an asthma exacerbation	7 (14)	14 (30)	31 (60)
Bring about a healthy lifestyle	9 (18)	7 (13)	36 (69)
Retain what I learned in Davos	3 (6)	0 (0)	49 (94)
Monitor my symptoms	2 (4)	3 (6)	47 (90)
I would like to			
Receive medication reminders	19 (37)	3 (6)	30 (58)
Digitally track how often I use my maintenance medication	13 (25)	5 (10)	34 (65)
Digitally track how often I use my rescue medication	5 (10)	4 (8)	43 (83)
Keep contact with the care team from Davos	2 (4)	7 (13)	43 (83)
Have routine outpatient visits to be digital rather than in person	18 (35)	8 (15)	26 (50)
Receive psychological assistance through an app	16 (31)	9 (17)	27 (52)
Receive tips on inhalation technique through an app	28 (54)	5 (10)	19 (36)
Receive asthma education through an app	11 (21)	11 (21)	30 (58)
Have information for family or friend through an app	9 (18)	1 (2)	42 (81)
Have the asthma action plan available on my smartphone	4 (8)	2 (4)	46 (88)
See the outdoor air quality on an app	6 (8)	3 (6)	45 (87)
Monitor my lung function at home with a spirometer	3 (6)	6 (12)	43 (83)
Keep track of my diet through an app	14 (27)	6 (12)	32 (62)
Keep track of my daily steps	7 (14)	2 (4)	43 (83)
Wear a heart rate monitor or pedometer as a wristband on daily basis	8 (16)	8 (15)	36 (69)
Have a device that monitors my inhalation technique	14 (27)	4 (8)	34 (65)
I think it is useful for my asthma to monitor my heart rate through a wristband	10 (20)	9 (17)	33 (63)
I think it is important that I could customize an app to my needs	1 (2)	5 (10)	46 (88)

[a]mHealth: mobile health.

Discussion

Principal Findings

This study showed that patients and HCPs expressed many different potential goals for an mHealth self-management support system and thought that numerous functions were useful to them. At the same time, having an overly comprehensive system would prevent them from using it. Different patients reported that they would benefit the most from different functions, and making all functions available to all patients could overwhelm the patient, possibly limiting the use of the

system. Therefore, creating a one-size-fits-all mHealth system is not the best strategy.

Personalization

As no 2 patients with asthma are the same, it is important to consider having different functions for different patients. In this group of patients with most severe asthma, some patients might benefit from reminders to take medication, whereas other patients might feel that these constant notifications interfere too much with their daily life and, because of these reminders, are more likely to ignore the app altogether. Many HCPs said that particular functions might work for some patients, but not for others, depending on the skills and willingness of the patient.

There is also a discrepancy between the views of HCPs and patients on the use of e-consultations. With e-consultations, patients could send messages to their HCP whenever they had a question regarding the management of their asthma. HCPs thought that new forms of HCP-patient communication could be more convenient for the patient. However, only 50% (26/52) of the patients from stage 2 reported that they wanted to replace their routine visits with e-consultations. HCPs reported that mHealth is unlikely to fully replace regular treatment; however, it could be used as an additional support tool.

Currently, many of the functions mentioned in the interviews and focus groups are already available in different apps or systems that focus on one specific function. However, it seems to be practical to integrate these numerous functionalities into a single system so that the patients do not have to familiarize themselves with all the different systems. Therefore, it is suggested that a single mHealth self-management support system that encompasses all these different functionalities be created. Patients in the focus groups also reported that they have more trust in a system developed and used in a care setting than in a system produced by a profit-based company.

According to international asthma guidelines, a (personalized) asthma action plan should be the backbone of asthma self-management. These guidelines describe that all patients with asthma should be provided with an asthma action plan to guide their self-management because it helps patients recognize and appropriately respond to the worsening of their asthma [6]. Although all patients receive such an action plan on paper during HACT, it is unknown to the HCPs if patients really use it. Although patients mentioned that written action plans (on paper) are often neglected, not received, or thrown in the trash, 88% (46/52) of them wanted the action plans to be available on their smartphone. By making it more accessible in an app, the action plan can be expanded and customized for the individual patient. Our suggestion is that the action plan can be the root of an app, whereas different additional functions mentioned as useful can be the possible branches. Although action plans are not often used in practice, they can help patients guide their self-management [6]. By adding an option to disable branches that are seen as irrelevant by or for a patient, we can prevent the patient from being overwhelmed. Personalization options are particularly important in this diverse group of patients with asthma, as indicated by the wide variety of potential goals and useful functions mentioned by patients and HCPs.

Strengths and Limitations

This study has several strengths and limitations. A strength of this study is the specific patient population studied. Patients with asthma undergoing HACT are among the patients with most severe asthma in the Netherlands. As conventional therapy is unsuccessful in accomplishing sufficient levels of asthma control, these patients must consider other forms of asthma management. Patients undergoing HACT spend 3 months in the desolated Swiss mountains, away from their friends and family, because HACT was their only solution left to improve the quality of life. Any tool able to assist them in increasing their asthma control at home would, therefore, be appreciated, and they are more likely to adhere to an (future) mHealth system.

The mixed methods study design allows for quantifying the attitudes expressed in qualitative stage 1 of the study. In the interviews and focus groups, the HCPs and patients were able to fully express their thoughts and ideas about mHealth in 30-minute or 2-hour sessions. Full train of thoughts could be formed and discussed, allowing for thoughtful considerations as to what is important and useful. Using the questionnaire, we tested whether these opinions were more widely supported by other patients.

A limitation of this study is the relatively modest number of respondents to the questionnaire. Every year, until 2019, approximately 80 patients in the Netherlands were admitted to HACT, and many of these patients went more often than once. This resulted in a small study population, and further research could help us better understand this population. The respondents agreed with most of the attitudes expressed in the focus groups, and the inclusion of the questionnaire was a useful addition to this study. As the first part of the study was qualitative, there could always be a selection bias. There were a limited number of patients and HCPs in the focus groups and interviews. This could have influenced the results. As we went to Davos at a random moment in time and invited all the patients present at that time to participate, we expected to have tackled the potential selection bias in these focus groups. We also tried to counteract bias by sending out the questionnaire to a larger group in stage 2.

Many different functionalities for the mHealth system were named. In our study, we did not ask patients to order the importance of different options. Therefore, we do not know whether every function is of equal importance or whether some can be left out without affecting user satisfaction. In this study, we gathered information on what was important to the end users but not on how important these options were to them. This could be explored further in future studies.

Comparison With Prior Work

In a previous study performed by Simpson et al [22], the perspectives of patients with general asthma and HCPs on mHealth were explored. Although most of the goals were similar to those of this study, the patients from the Simpson study reported that they wanted a system to assist them in emergency situations. The patients in this study did not mention this option. This could be attributed to the higher asthma severity of the

patients in our study because they are more used to having severe asthma symptoms.

In a 2017 review of the available asthma apps, 38 different apps were analyzed [23]. A total of 42 functions in 4 categories were identified: tracking, information, assessment, and notifications. Most of the tracking (monitoring) functions mentioned in the focus groups were available in these apps, except for the tracking of specific health data (eg, heart rate and oxygen saturation) and lifestyle parameters (eg, weight and exercise). An additional information (education) option mentioned in the focus groups was information for friends and family. "Assessment" provides the interpretation of recorded values, including the availability of the asthma action plan. Finally, all notification options implemented in these apps were reported in the focus groups, including medication reminders.

Van der Kleij et al [14] reported on 6 conditions that they regarded as vital for the development and implementation of safe eHealth apps in primary care [14]. The first condition they named was "together: stakeholder engagement and co-creation." An eHealth system must be improved in an iterative setting based on the attitudes of relevant stakeholders [24]. The second condition is called "blended: combining eHealth with regular care" in which it is outlined that an eHealth system is combined with regular face-to-face care. The third condition, "individualized and inclusive," poses that personalizing eHealth has the potential to be more effective than a one-size-fits-all app [25]. The next condition is "ethical: being attentive of ethical considerations, privacy, and patient safety." It is said, for example, that the privacy of patient data needs to be accounted for. These 4 conditions were named in our focus groups and interviews as the major topics. The condition "evidence-based: continuous research and educational guidance" will be met in the next phase of the Davos@home study when the effectiveness of the future mHealth system will be evaluated. The last condition "global: eHealth in primary care in high- and low-resource settings" is at the moment not applicable because Davos@home is entailing a specific patient population.

Future Perspective

In the next Davos@home project, an mHealth system will be built based on the results of this study. The effectiveness of the system will be evaluated by assessing its effect on asthma control after the discharge of patients from HACT compared with the regular aftercare process. The use of mHealth to support the self-management of asthma opens up a lot of possibilities. Many different aims can be targeted when developing an mHealth system, which can be a pitfall. Therefore, it must be clear to the developers of the system what the goals of the end users are to enable cocreation. Although it might be challenging to create a system with different functionalities, it is better to invest in a system that meets most of its users' demands than in a system that is unlikely to be used.

Conclusions

Different patients reported that they would benefit the most from different functionalities. Therefore, it is important to tailor functionalities to individual (treatment) goals. When developing an mHealth intervention, it is important to allow personalization to avoid overwhelming the users.

Conflicts of Interest

None declared.

References

1. Masoli M, Fabian D, Holt S, Beasley R, Global Initiative for Asthma (GINA) Program. The global burden of asthma: executive summary of the GINA Dissemination Committee report. Allergy 2004 May;59(5):469-478. [doi: 10.1111/j.1398-9995.2004.00526.x] [Medline: 15080825]
2. Rijssenbeek-Nouwens L, Bel E. High-altitude treatment: a therapeutic option for patients with severe, refractory asthma? Clin Exp Allergy 2011 Jun;41(6):775-782. [doi: 10.1111/j.1365-2222.2011.03733.x] [Medline: 21518039]
3. Bel EH, Sousa A, Fleming L, Bush A, Chung KF, Versnel J, Unbiased Biomarkers for the Prediction of Respiratory Disease Outcome (U-BIOPRED) Consortium, Consensus Generation. Diagnosis and definition of severe refractory asthma: an international consensus statement from the Innovative Medicine Initiative (IMI). Thorax 2011 Oct 23;66(10):910-917. [doi: 10.1136/thx.2010.153643] [Medline: 21106547]
4. Hashimoto S, Bel EH. Current treatment of severe asthma. Clin Exp Allergy 2012 Jan 18;42(5):693-705. [doi: 10.1111/j.1365-2222.2011.03936.x]
5. Fieten KB, Rijssenbeek-Nouwens LH, Hashimoto S, Bel EH, Weersink EJ. Less exacerbations and sustained asthma control 12 months after high altitude climate treatment for severe asthma. Allergy 2019 Mar 09;74(3):628-630. [doi: 10.1111/all.13664] [Medline: 30428132]
6. Global strategy for asthma management and prevention (2019 update). Global Initiative for Asthma. 2019. URL: https://ginasthma.org/wp-content/uploads/2019/06/GINA-2019-main-report-June-2019-wms.pdf [accessed 2022-09-19]
7. Rijssenbeek-Nouwens LH, Fieten KB, Bron AO, Hashimoto S, Bel EH, Weersink EJ. High-altitude treatment in atopic and nonatopic patients with severe asthma. Eur Respir J 2012 Dec 22;40(6):1374-1380 [FREE Full text] [doi: 10.1183/09031936.00195211] [Medline: 22441741]
8. Dutch association of physicians for pulmonary diseases and tuberculosis. Nederlandse Vereniging van Artsen voor Longziekten en Tuberculose. URL: https://www.nvalt.nl/ [accessed 2022-09-19]

9. Juniper E, O'Byrne PM, Guyatt G, Ferrie P, King D. Development and validation of a questionnaire to measure asthma control. Eur Respir J 1999 Oct;14(4):902-907 [FREE Full text] [doi: 10.1034/j.1399-3003.1999.14d29.x] [Medline: 10573240]
10. Wordt u behandeld voor astma en ervaart u nog veel beperkingen in uw dagelijks leven? Nederlands Astmacentrum Davos. URL: https://docplayer.nl/114523295-Wordt-u-behandeld-voor-astma-en-ervaart-u-nog-veel-beperkingen-in-uw-dagelijks-leven.html [accessed 2022-09-19]
11. Beerthuizen T, Rijssenbeek-Nouwens LH, van Koppen SM, Khusial RJ, Snoeck-Stroband JB, Sont JK. Internet-based self-management support after high-altitude climate treatment for severe asthma: randomized controlled trial. J Med Internet Res 2020 Jul 22;22(7):e13145 [FREE Full text] [doi: 10.2196/13145] [Medline: 32706692]
12. Gibson P, Powell H, Coughlan J, Wilson AJ, Abramson M, Haywood P, et al. Self-management education and regular practitioner review for adults with asthma. Cochrane Database Syst Rev 2003(1):CD001117. [doi: 10.1002/14651858.CD001117] [Medline: 12535399]
13. World Health Organization. MHealth New Horizons for Health Through Mobile Technologies. Geneva: World Health Organization; 2011.
14. van der Kleij RM, Kasteleyn MJ, Meijer E, Bonten TN, Houwink EJ, Teichert M, et al. SERIES: eHealth in primary care. Part 1: concepts, conditions and challenges. Eur J Gen Pract 2019 Oct 10;25(4):179-189 [FREE Full text] [doi: 10.1080/13814788.2019.1658190] [Medline: 31597502]
15. McCurdie T, Taneva S, Casselman M, Yeung M, McDaniel C, Ho W, et al. mHealth consumer apps: the case for user-centered design. Biomed Instrum Technol 2012 Sep;Suppl(s2):49-56. [doi: 10.2345/0899-8205-46.s2.49] [Medline: 23039777]
16. Thies K, Anderson D, Cramer B. Lack of adoption of a mobile app to support patient self-management of diabetes and hypertension in a federally qualified health center: interview analysis of staff and patients in a failed randomized trial. JMIR Hum Factors 2017 Oct 03;4(4):e24 [FREE Full text] [doi: 10.2196/humanfactors.7709] [Medline: 28974481]
17. P5 eHealth: An Agenda for the Health Technologies of the Future. Cham: Springer; 2020.
18. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
19. Hurt HT, Joseph K, Cook CD. Scales for the measurement of innovativeness. Human Comm Res 1977 Sep;4(1):58-65. [doi: 10.1111/j.1468-2958.1977.tb00597.x]
20. Welkom op inhalatorgebruik. Long Alliantie Nederland. URL: https://inhalatorgebruik.nl/nl/home [accessed 2022-09-20]
21. Stern AF. The hospital anxiety and depression scale. Occup Med (Lond) 2014 Jul 07;64(5):393-394. [doi: 10.1093/occmed/kqu024] [Medline: 25005549]
22. Simpson AJ, Honkoop PJ, Kennington E, Snoeck-Stroband J, Smith I, East J, et al. Perspectives of patients and healthcare professionals on mHealth for asthma self-management. Eur Respir J 2017 May 01;49(5):1601966 [FREE Full text] [doi: 10.1183/13993003.01966-2016] [Medline: 28461291]
23. Tinschert P, Jakob R, Barata F, Kramer J, Kowatsch T. The potential of mobile apps for improving asthma self-management: a review of publicly available and well-adopted asthma apps. JMIR Mhealth Uhealth 2017 Aug 02;5(8):e113 [FREE Full text] [doi: 10.2196/mhealth.7177] [Medline: 28768606]
24. van Limburg M, Wentzel J, Sanderman R, van Gemert-Pijnen L. Business modeling to implement an eHealth portal for infection control: a reflection on co-creation with stakeholders. JMIR Res Protoc 2015 Aug 13;4(3):e104 [FREE Full text] [doi: 10.2196/resprot.4519] [Medline: 26272510]
25. Kaptein M, Markopoulos P, de Ruyter B, Aarts E. Personalizing persuasive technologies: explicit and implicit personalization using persuasion profiles. Int J Human Comput Stud 2015 May;77:38-51. [doi: 10.1016/j.ijhcs.2015.01.004]

Abbreviations

ACQ: Asthma Control Questionnaire
HACT: high-altitude climate therapy
HCP: health care provider
mHealth: mobile health

Edited by A Mavragani; submitted 10.01.21; peer-reviewed by S Choi, K Fitzner, B Nievas Soriano; comments to author 11.05.21; revised version received 29.06.21; accepted 17.08.22; published 22.11.22.

Please cite as:
Khusial R, van Koppen S, Honkoop P, Rijssenbeek-Nouwens L, Fieten KB, Keij S, Drijver-Messelink M, Sont J
Patients' and Health Care Providers' Perceptions on mHealth Use After High-Altitude Climate Therapy for Severe Asthma: Mixed Methods Study
JMIR Form Res 2022;6(11):e26925
URL: https://formative.jmir.org/2022/11/e26925
doi:10.2196/26925
PMID:36413384

Original Paper

Chatbot-Delivered Cognitive Behavioral Therapy in Adolescents With Depression and Anxiety During the COVID-19 Pandemic: Feasibility and Acceptability Study

Ginger Nicol[1], MD; Ruoyun Wang[2], MS, MHA; Sharon Graham[2], BS; Sherry Dodd[2], BA; Jane Garbutt[2], MB, ChB

[1]Division of Child and Adolescent Psychiatry, Department of Psychiatry, Washington University School of Medicine, St Louis, MO, United States
[2]Division of Allergy, Immunology & Pulmonology, Department of Pediatrics, Washington University School of Medicine, St Louis, MO, United States

Corresponding Author:
Ginger Nicol, MD
Division of Child and Adolescent Psychiatry
Department of Psychiatry
Washington University School of Medicine
600 S Taylor Ave
Suite 121
St Louis, MO, 63110
United States
Phone: 1 3143625939
Fax: 1 3147471160
Email: nicolg@wustl.edu

Abstract

Background: Symptoms of depression and anxiety, suicidal ideation, and self-harm have escalated among adolescents to crisis levels during the COVID-19 pandemic. As a result, primary care providers (PCPs) are often called on to provide first-line care for these youth. Digital health interventions can extend mental health specialty care, but few are evidence based. We evaluated the feasibility of delivering an evidence-based mobile health (mHealth) app with an embedded conversational agent to deliver cognitive behavioral therapy (CBT) to symptomatic adolescents presenting in primary care settings during the pandemic.

Objective: In this 12-week pilot study, we evaluated the feasibility of delivering the app-based intervention to adolescents aged 13 to 17 years with moderate depressive symptoms who were treated in a practice-based research network (PBRN) of academically affiliated primary care clinics. We also obtained preliminary estimates of app acceptability, effectiveness, and usability.

Methods: This small, pilot randomized controlled trial (RCT) evaluated depressive symptom severity in adolescents randomized to the app or to a wait list control condition. The primary end point was depression severity at 4-weeks, measured by the 9-item Patient Health Questionnaire (PHQ-9). Data on acceptability, feasibility, and usability were collected from adolescents and their parent or legal guardian. Qualitative interviews were conducted with 13 PCPs from 11 PBRN clinics to identify facilitators and barriers to incorporating mental health apps in treatment planning for adolescents with depression and anxiety.

Results: The pilot randomized 18 participants to the app (n=10, 56%) or to a wait list control condition (n=8, 44%); 17 participants were included in the analysis, and 1 became ineligible upon chart review due to lack of eligibility based on documented diagnosis. The overall sample was predominantly female (15/17, 88%), White (15/17, 88%), and privately insured (15/17, 88%). Mean PHQ-9 scores at 4 weeks decreased by 3.3 points in the active treatment group (representing a shift in mean depression score from moderate to mild symptom severity categories) and 2 points in the wait list control group (no shift in symptom severity category). Teen- and parent-reported usability, feasibility, and acceptability of the app was high. PCPs reported preference for introducing mHealth interventions like the one in this study early in the course of care for individuals presenting with mild or moderate symptoms.

Conclusions: In this small study, we demonstrated the feasibility, acceptability, usability, and safety of using a CBT-based chatbot for adolescents presenting with moderate depressive symptoms in a network of PBRN-based primary care clinics. This pilot study could not establish effectiveness, but our results suggest that further study in a larger pediatric population is warranted. Future study inclusive of rural, socioeconomically disadvantaged, and underrepresented communities is needed to establish generalizability of effectiveness and identify implementation-related adaptations needed to promote broader uptake in pediatric primary care.

Trial Registration: ClinicalTrials.gov NCT04603053; https://clinicaltrials.gov/ct2/show/NCT04603053

(JMIR Form Res 2022;6(11):e40242) doi:10.2196/40242

KEYWORDS

COVID-19; adolescent depression; mobile health; cognitive behavioral therapy; chatbot; relational conversational agent; depression; anxiety; suicide; self-harm; pandemic; pediatric; youth; adolescent; adolescence; psychiatry; conversational agent; CBT; clinic; data; acceptability; feasibility; usability; primary care; intervention; mental health; digital health; technology mediated; computer mediated

Introduction

Depression and anxiety are common in adolescents. [1,2]. If left untreated, these common illnesses can result in significant sequelae including school dropout, substance abuse, and suicide [3]. At any one time, 18% of adolescents report depressive symptoms [1,2], and about 3% are diagnosed with a depressive disorder [2]. Depression and anxiety often co-occur; 1 in 3 teens will experience clinically significant anxiety symptoms [4]. These numbers have risen in recent years due to the isolating effects of quarantine and the stress of virtual learning during the COVID-19 pandemic [5,6]. Even before the pandemic, suicide was the second highest cause of death in this age group [7-9], but in 2020, emergency department (ED) visits for suspected suicide attempts increased by 30% compared to 2019 [10]. The effects of the pandemic on mental health in general has been unprecedented; prepandemic mental health workforce shortages have become more dire. Thus, there is an urgent need to provide prompt, effective treatments that can be readily integrated into the day-to-day lives of adolescents and disseminated remotely during periods of quarantine and ongoing workforce issues.

Evidence-based treatments for pediatric depression and anxiety include selective serotonin and serotonin-norepinephrine reuptake inhibitor (SSRI, SNRI) medications and cognitive behavioral therapy (CBT), used separately or in combination [11-13]. Yet, few teens receive treatment with these therapies [13]. CBT is typically provided through a series of face-to-face interactions with a trained therapist over several months and is often preferred by families over medication [3,14]. However, CBT is seldom used due to pervasive and persistent problems with access, cost, and stigma [15-18]. Internet and mobile health (mHealth) interventions reduce these barriers to utilization and are acceptable and effective alternatives to live CBT for the treatment of depression and anxiety in adults [19]. Conversational agents or “chatbots” that deliver CBT via a text-based, semiautomated algorithm are known to reduce mild-to-moderate depressive symptoms in nonclinical populations [17,20-23]. However, few have been evaluated in clinical populations, with scarce rigorous study in youth [21,24,25].

About 95% of teens in the United States use smartphones [26] or other web-enabled mobile devices, and nearly half report being online almost constantly. Over half of US teens use their devices for accessing social media platforms, and this number may be higher in adolescents with known mental health diagnoses [27]. Although internet use—social media use in particular—has known serious negative effects on adolescent mental health [28], mobile technology can also promote social support and even assistance to young people dealing with emotional challenges [29]. Adolescents are increasingly using their devices to search the internet for information about mental health [18]. Many teens want to be self-reliant when coping with emotional distress and prefer to manage difficult situations virtually rather than through in-person interactions [18]. However, few digital therapeutic interventions have been specifically developed for children and adolescents [26]. Using mHealth to deliver CBT to teens merits further evaluation, as it is inexpensive, easy to access and administer, and scalable, which are important attributes for widespread dissemination, adoption, and public health impact [30].

Because the management of adolescents with depression and anxiety is often done by primary care providers (PCPs), the goal of this study was to investigate the feasibility of implementing an existing mHealth intervention [20,31] as part of the treatment plan for adolescents diagnosed with depression and anxiety by their PCP. The intervention, Woebot for Adolesent Depression, or W-GenZ, is an mHealth application that uses a relational conversational agent known as Woebot to deliver CBT to adolescents presenting with mood and anxiety concerns. W-GenZ is one among a suite of emotional support Woebot-based applications that have been previously studied. Previous research demonstrated Woebot’s feasiblity and acceptability as well as efficacy to reduce (1) depression and anxiety among young adults [20], (2) depression among postpartum women [32,33], and (3) substance abuse among adults [34,35]. In addition, data indicate that Woebot establishes a therapeutic working alliance with users [33]. The W-GenZ app was developed specifically for use in adolescents, and while preliminary effectiveness data among this population are promising [36], it has not yet been rigorously tested for effectiveness. Moreover, the application has not yet been implemented in real-world clinical settings, where such interventions would ideally track and report progress to PCPs and parents, which is critical for maintaining safety and treatment adherence.

In this open-label, randomized pilot study, our objectives were 2-fold: (1) to establish that the app is usable, acceptable, and likely to provide benefit for adolescents, and (2) to determine the optimal way to integrate mHealth interventions such as the app into primary care management of adolescent depression. This study focused on adolescents newly diagnosed with depression and anxiety, as we anticipated immediate benefit from using the app while awaiting a therapeutic response from usual care, including antidepressant medication and/or referral for psychotherapy. We hypothesized that the app would be

acceptable and easy to use and that use of the app [16] would lead to greater improvement in symptoms for adolescents newly diagnosed with depression and anxiety compared to the wait list control group.

Methods

Methodology

Study activities were conducted within the Washington University Pediatric and Adolescent Ambulatory Research Consortium (WU PAARC), part of the WU Institute for Clinical and Translational Science (ICTS). The practice network includes over 100 pediatric care providers from approximately 60 practices in the greater metropolitan St Louis area and serves approximately 175,000 patients. Across the consortium, 25% (approximately 44,000) of the patients are publicly insured (eg, Medicaid) and 28% (approximately 50,000) are Black. Participation was restricted to patients and PCPs from 11 practices from the practice-based research network (PBRN) who had completed a 12-month quality improvement (QI) initiative targeted at improving care for adolescents with depression [37], specifically by establishing systems of care that follow the current national guidelines [12,13]. Study activities included a small pilot randomized trial to gather data from teens and their parent and semistructured interviews with PCPs. The study was conducted from October 2020 to April 2021, encompassing the second wave of COVID-19 infections in the region, which limited or at times completely shut down in-person clinical care delivery during the study period.

Ethics Approval

The study was approved by the Washington University Institutional Review Board (202005103). Parents or legal guardians provided informed consent, and adolescents provided informed assent to participate in the study. The consent/assent process was initiated by phone due to pandemic-related barriers to in-person participation. Formal consent/assent was conducted using a Research Electronic Data Capture (REDCap)–based electronic consent form. Following a prescreening phone call, the parents or legal guardians of potential adolescent participants received an email with a unique link to review the informed consent form online. After the research team explained the study and answered any questions, parents or legal guardians and potential participants clicked an “agree” button, which was accompanied by text indicating that by clicking the button, they were providing consent/assent to participate. They were then asked to type their full name, which served as their electronic signature confirming consent/assent. Upon completion of the form, participants were presented with the option to download an electronic copy of the executed form. A digital copy was also emailed to each participant’s parent or legal guardian. Participating parents or guardians and adolescents each received up to US $40 as reimbursement for their time to complete study assessments ($5 to $10 per survey).

Randomized Pilot Study

Participants were randomized 1:1 for 12 weeks either to a wait list control or to the app intervention. Time points for clinical assessments were based on clinical care guidelines (baseline, 4, 8, and 12 weeks) [13], with the primary analysis at 4 weeks. The primary outcome was depression symptom severity on the 9-item Patient Health Questionnaire (PHQ-9). Secondary outcomes included anxiety symptom severity assessed with the 7-item Generalized Anxiety Disorder questionnaire (GAD-7), which uses scoring similar to the PHQ-9 to indicate the frequency and severity of anxiety symptoms, and mental health self-efficacy assessed with the Mental Health Self-Efficacy Scale (MHSES). All patients in the intervention and wait list control groups received usual care for their depression and anxiety from their PCP. Participants in the control group were offered use of the app at the end of the 12-week study period.

Intervention

W-GenZ delivers evidence-based therapeutic elements via brief “conversations” with a fully automated, relational conversational agent named Woebot. Powered by natural language processing and machine learning techniques, the brief, self-guided intervention draws from CBT, interpersonal psychotherapy for adolescents (IPT-A), and some elements of dialectical behavior therapy (DBT), tailoring the conversation to the present situation to help the adolescent develop emotion regulation skills in the context of their everyday life for the problem at hand. The user experience is centered around mood tracking and goal-oriented, tailored conversations. Woebot checks in with the user, and depending on the user’s reported mood or desire to work on a problem or learn something new, Woebot will offer and guide the user through CBT-based psychoeducation and tools, tailored to the reported need in that moment. Daily push notifications prompt users to check in. Using proprietary neurolinguistic programing and artificial intelligence, the platform designs a personalized program to meet the user’s needs in real time. The app conforms with safety recommendations from the American Psychiatric Association (APA) and American Medical Association (AMA) [38,39]. As part of the app’s onboarding process, users are provided with the privacy policy, theoretical underpinning, targeted difficulties, expected results, research findings, and other information requested by potential young users [18]. Safety features include informing the user they are talking to a robot and not a real person, reminding them that the app is not a crisis intervention for suicidal ideation, encouraging them to seek additional support if they are feeling unwell, and providing helplines for assistance. The safety net protocol detects concerning language, confirms it with the user, and provides a local suicide crisis hotline number as needed.

Study Population

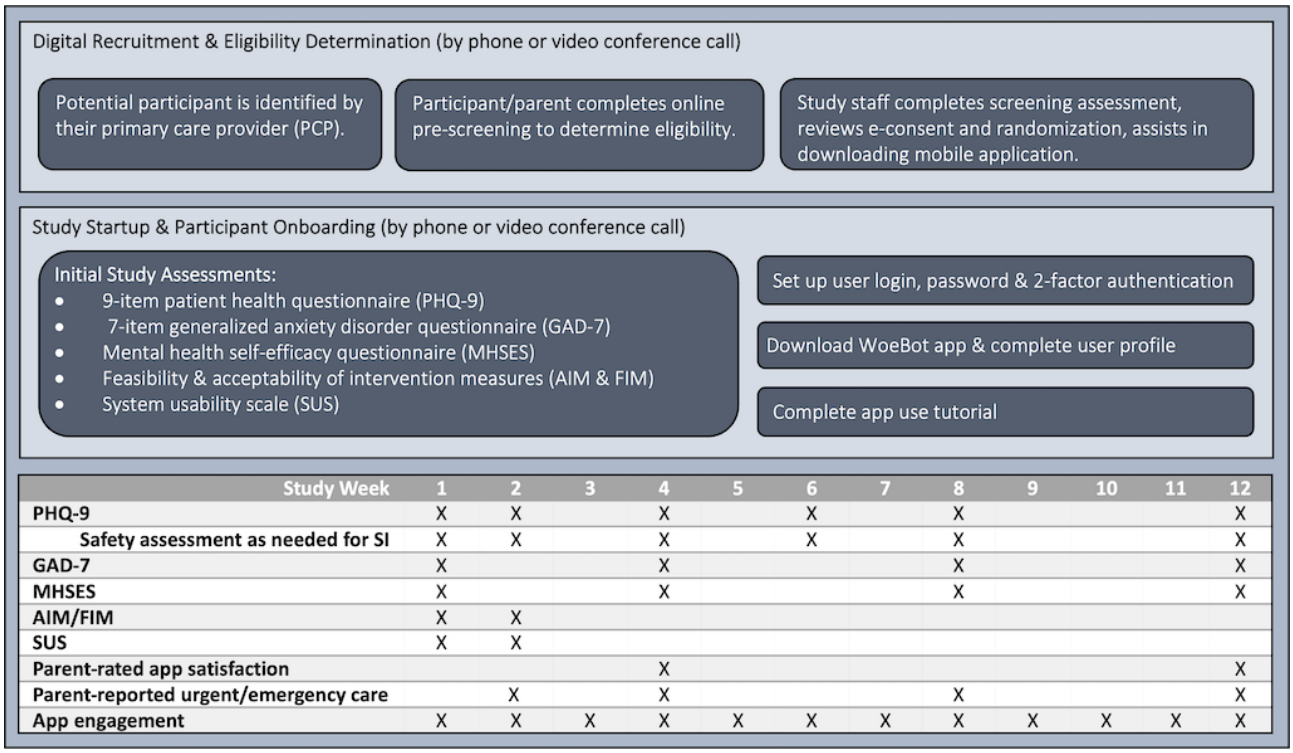
Adolescents were eligible for inclusion if they were 13 to 17 years old and had a new diagnosis of depression and anxiety in the past 3 months, as reported by their parent or legal guardian. Adolescents were excluded if their parent reported any lifetime history of severe depression, substance use disorder, psychotic illness, obsessive compulsive disorder, posttraumatic stress disorder, panic disorder, or specific phobias, as this could indicate that the current reported episode was a more complicated or severe presentation of depression or that the current episode was recurrent. Those who reported a psychiatric hospitalization in the previous month, were not accompanied by a guardian to the diagnostic visit, did not have access to a

mobile device (cell phone/tablet) for regular use, and were unable to read and write English were excluded.

Recruitment Procedures

Five of the 11 eligible practices referred patients to the randomized controlled trial (RCT). Potential participants were informed about the study via a flyer distributed by their PCP at the diagnostic visit or within the first 3 months after diagnosis. For those interested in participation, the parent completed the online eligibility screen. If the adolescent was eligible, the parent and adolescent reviewed the electronic consent document and had a telephone conversation with a member of the study team to review participation requirements and answer questions. If the family decided to participate, the adolescent and 1 parent signed the electronic consent and completed baseline web-based surveys via REDCap. Subsequently, the adolescent was randomized using REDCap [40]. Those randomized to the intervention received a link and a unique password to download the intervention app on their mobile device, proceed through the enrollment screens, and initiate the program. All participants were followed for 12 weeks with assessments, as indicated in Figure 1. After data collection was completed, participants' medical charts were reviewed to confirm any additional mental health diagnoses and record their initial treatment plan.

Figure 1. Flowchart indicating the study procedures and schedule of events. AIM: Acceptability of Intervention Measure; FIM: Feasibility of Intervention Measure; GAD-7: 7-item Generalized Anxiety Disorder questionnaire; MHSES: Mental Health Self-Efficacy Scale; PHQ-9: 9-item Patient Health Questionnaire; SUS: System Usability Scale.



Study Week	1	2	3	4	5	6	7	8	9	10	11	12
PHQ-9	X	X		X		X		X				X
Safety assessment as needed for SI	X	X		X		X		X				X
GAD-7	X			X				X				X
MHSES	X			X				X				X
AIM/FIM	X	X										
SUS	X	X										
Parent-rated app satisfaction				X								X
Parent-reported urgent/emergency care		X		X				X				X
App engagement	X	X	X	X	X	X	X	X	X	X	X	X

Measurement

In both study groups, study surveys were administered to adolescents and parents through a secure digital link to a REDCap survey. For each survey, participants were directed to a URL via email or text, with up to 2 push notification reminders. Time points for the primary (4 weeks) and safety (2 and 4 weeks) outcomes were selected based on the practice guidelines for PCPs managing adolescent depression [13]. If the measurement was not completed within 1 week, participants received a phone call from a study team member to assess study engagement. The surveys varied in duration and took 5 to 20 minutes to complete. Gift cards were provided to compensate participants for their time to complete data collection and were distributed immediately upon survey completion (up to US $40). Study design and assessments are indicated in Figure 1.

Depression symptoms were assessed using the PHQ-9 [41] modified for use in adolescents (PHQ-A) [42]. The PHQ-A also consists of 9 questions, but with wording changes to accommodate an adolescent reading level, and was used to screen for adolescent depression according to Diagnostic and Statistical Manual of Mental Disorders fifth edition (DSM-5) criteria. Although the unmodified version has good validity and reliability in adolescent populations, the PHQ-A has modified the wording of the questions so that it can be entirely self-administered in individuals with a grade 6 reading level. The overall PHQ-9 score ranges from 0 to 27, with individual item symptom severity scores of 0, 1, 2, and 3, which correspond to "not at all," "several days," "more than half the days," and "nearly every day," respectively. Overall cumulative scores of 5, 10, 15, and 20 represent cut points for mild, moderate, moderately severe, and severe depression, respectively.

Anxiety symptoms were assessed using GAD-7 [43]. This is a reliable, valid, and responsive screening instrument commonly used in primary care settings [44] and has been extensively studied and validated in diverse adolescent populations, including for screening in primary care settings in Finland [45],

China [46], Korea [47], Ghana [48], and Canada [49]. GAD-7 has also been validated in the United States in a clinical population of adolescents with generalized anxiety disorder [50]. Respondents use the same item scoring system as the PHQ-9 to indicate frequency and severity of anxiety symptoms, and a cut point of >10 is used to identify clinical cases of anxiety.

Confidence in self-management of mental health issues was measured using the MHSES [51], which has 6 items that address confidence in managing stress, depression, and anxiety and are scored on a 10-point scale (1: not at all confident, 10: very confident). Ratings are summed for an overall measure of self-efficacy (range 6 to 60), with higher scores indicating more self-efficacy.

Safety was assessed at 2, 4, 8, and 12 weeks. Parents were asked to report any hospitalizations or ED visits made by their child for depression/anxiety-related problems in the preceding 2 weeks up to 1 month.

Utilization, Acceptability, and Feasibility

Deidentified aggregate app usage data were obtained from the app developer. Metrics included aggregate descriptive summaries of engagement metrics, including total number of check-ins, days in app, and messages sent during the 4-week intervention period. Each conversation with the app is tailored to the users' reported needs at the moment and thus necessarily varies in length and composition. Nonetheless, these metrics are provided as examples to demonstrate various ways that users engaged with the app's offerings.

The acceptability and feasibility of using the intervention was assessed after 2 weeks from the perspective of the adolescents (n=10, 56%) and their parent or adult legal guardian (n=10, 44%) randomized to the intervention group of the RCT. Parents completed the 4-item Acceptability of Intervention Measure (AIM) and the 4-item Feasibility of Intervention Measure (FIM) developed by Weiner et al [52]. To aid understanding by the 13- to 17-year-old adolescent participants, only 1 item from each of these measures was used.

Adolescents assessed the usability of the app using 5 items from the 10-item System Usability Scale (SUS) [53] that includes statements about the effectiveness, efficiency, and satisfaction with use (score range 5 to 25). For the AIM, FIM, and SUS, respondents used a 5-item response scale (1: completely disagree, 5: completely agree) to indicate their agreement with item statements, and a summary score was created. For the AIM and FIM, higher scores indicated greater acceptability and feasibility [52,54].

Acceptability and feasibility for PCPs were assessed by semistructured interviews (described in the subsequent section).

Safety

For all participants, an additional layer of safety monitoring was provided by the study team through surveillance of the digital PHQ-9 assessment at baseline, 2, 4, 6, 8, and 12 weeks. Parents provided consent for communication between the study team and the PCP for this purpose. When the PHQ-9 flagged suicidal ideation, an email alert was sent to the study team, and the PI then contacted the participant's PCP to share information and discuss next steps regarding assessment of suicide risk. The principal investigator (PI; author JG) summarized the relevant information and alerted the psychiatrist. Management options included (1) the PCP electing to alert the family to the situation and complete additional suicide risk assessment themselves or (2) the PCP requesting further assessment be completed by the pediatric study psychiatrist, who then conducted additional assessment with patients and their legal guardians over the phone and provided a brief written assessment and recommendations back to the PCP. The PCP also had the option of further discussing the case with the study psychiatrist by phone. The PCP's preferred management approach was prioritized, and the psychiatrist provided further assessment by phone and offered treatment recommendations to the PCP when requested using an evidence-based telephone consultation model [55].

Statistical Analyses

Summary statistics are reported as percentages for categorical variables and mean and SD for continuous variables. We evaluated the Cohen *d* effect size and 95% CIs on the mean PHQ-9, GAD-7, and MHSES scores achieved by treatment group at the primary end point of 4 weeks. All statistical analyses were completed using SAS software version 9 (SAS Institute Inc).

PCP Interviews

We completed semistructured interviews with PCPs to better understand current attitudes toward and practices related to recommending CBT to pediatric patients as well as toward the use of mobile health applications for delivering behavioral mental health interventions to youth.

PCPs who participated in the PBRN quality improvement (QI) initiative were eligible to participate in the interviews, and all received an email invitation. Interviews were conducted concurrently with the pilot RCT portion of the study between October 2020 and April 2021. A total of 13 PCPs from the 11 participating clinics completed a 30-minute virtual video interview with the PI, who was also the PBRN director. She used a semistructured interview guide to ask PCP about their use of apps as part of a treatment plan to improve physical and mental health, how they found out about apps to recommend, and positive and negative features of the app considered when selecting an app to recommend. After these general questions, participating PCPs were provided with some promotional material about the intervention app that included a picture of the chatbot (without its name) and information about how it interacts with the user. They were asked about their first impressions, how likely they were to recommend it, and for which patients and when in the treatment course it would likely be most useful. Interviews were continued until thematic saturation was achieved. All interviews were digitally recorded and transcribed verbatim. Participants received a US $50 gift card for participation.

Qualitative Analysis of PCP Interviews

Transcripts were analyzed using an inductive coding approach based on pragmatic-variant grounded theory [56]. Authors JG

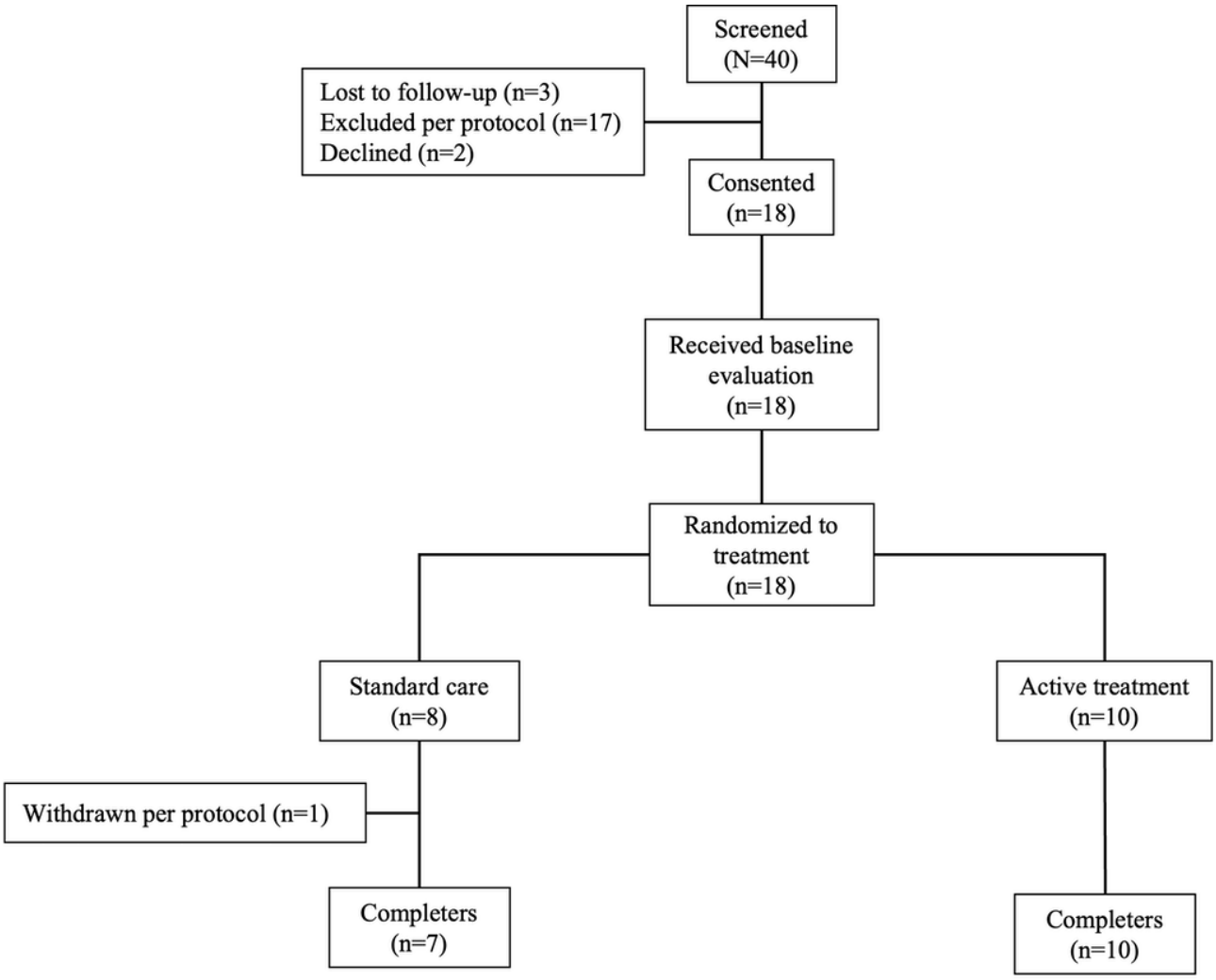
and GN independently reviewed interview transcripts for emergent themes in PCP perspectives on the use of mobile applications to augment physical and mental health care.

Results

Pilot RCT

Figure 2 shows participant flow through the pilot RCT. Of the 40 families that completed the eligibility screening for the RCT, 17 (43%) were ineligible, 2 (5%) declined to participate, 3 (8%) were lost to follow-up, and 18 (45%) were randomized to the intervention (10/40, 25%) or wait list control (8/40, 2%). Chart review revealed the following diagnoses: depression (8/18, 44%), depressive symptoms (2/18, 11%), anxiety (16/18, 89%), stress-related headaches (1/18, 6%), and functional abdominal pain and adjustment disorder (1/18, 6%). A review of the notes and baseline PHQ-9 and GAD-7 scores resulted in a sample of 17 for the analyses, with 10 (59%) that had both depression and anxiety 7 (41%) that had anxiety alone. One participant originally randomized to the wait list control group was excluded from analyses following chart review based on the chart notation "episodic tension headaches due to stress" and baseline measures (PHQ-9=5, GAD-7=6), as they were deemed to be ineligible (no diagnosis of depression or anxiety in their chart and no evidence of moderate disease).

Figure 2. Participant Disposition.



Participant Characteristics

The participating parent was typically the mother (16/17, 94%). Participating teens are described in Table 1 and did not differ by treatment group. They were predominantly female (15/17, 88%), White (15/17, 88%), and had family work-related health insurance (15/17, 88%). Initial treatment plans did not differ between study groups and included CBT (5/17, 29%), unspecified counseling, 15/17, 88%), and SSRI medication (11/17, 65%). By self-report, 7 (41%) teens had previously received at least 1 counselling session for depression or anxiety.

Table 1. Randomized Controlled Trial (RCT) participant baseline characteristics.

Characteristic	Total (N=17)	Intervention (N=10)	Wait list control (N=7)
Sex, n (%)			
Male	1 (5.9)	1 (10)	0 (0)
Female	15 (88.2)	9 (90)	6 (85.7)
Missing	1 (5.9)	0 (0)	1 (14.3)
Age in years, mean (SD)	14.7 (1.7)	14.7 (1.7)	14.8 (1.7)
Race, n (%)			
White	15 (88.2)	8 (80)	7 (100)
Mixed	2 (11.8)	2 (20)	0 (0)
Ethnicity, n (%)			
Non-Hispanic/Latino	17 (100)	10 (100)	7 (100)
Health Insurance, n (%)			
Private insurance	15 (88.2)	9 (90)	6 (85.7)
Medicaid	2 (11.8)	1 (10)	1 (14.3)
Living situation, n (%)			
Two parents	12 (70.6)	5 (50)	7 (100)
Diagnosis from chart review, n (%)			
Anxiety disorder	17 (100)	10 (100)	7 (100)
Depressive disorder	10 (58.8)	5 (50)	5 (71.4)
Initial treatment plan from chart review, n (%)			
SSRI[a]	11 (64.7)	6 (60)	5 (71.4)
Other medication(s)	0 (0)	0 (0)	0 (0)
CBT[b]	5 (29.4)	3 (30)	2 (28.6)
Unspecified counseling	15 (88.2)	8 (80)	6 (85.7)
Referral to psychiatry	2 (11.8)	1 (10)	1 (14.3)
Measures at Baseline, mean (SD)			
PHQ-9[c]	12.1 (5.3)	10.1 (3.9)	14.9 (6)
GAD-7[d]	12.8 (4.9)	11 (5)	15.3 (3.5)
MHSES[e]	30.9 (10.2)	32.4 (10.1)	28.9 (10.9)

[a]SSRI: selective serotonin reuptake inhibitor.

[b]CBT: cognitive behavioral therapy.

[c]PHQ-9: 9-item Patient Health Questionnaire.

[d]GAD-7: 7-item Generalized Anxiety Disorder questionnaire.

[e]MHSES: Mental Health Self-Efficacy Scale.

Pilot RCT Outcomes

Table 2 and Figure 3 show the overall study results. The small sample size precluded meaningful comparisons between study groups. However, large effect sizes were observed between groups on depression symptom severity score at 4 weeks, with the intervention group showing greater improvement on each scale than the wait list control group. Specifically, mean PHQ-9 scores at 4 weeks decreased by 3.3 units in the intervention group (representing a transition from moderate to mild symptom severity categories) and 2 units in the wait list control group (no shift in symptom severity category). In subgroup analyses among the 10 participants diagnosed with depression, PHQ-9 scores similarly decreased by 3 units in the intervention group (11.8 to 8.8; from moderate to mild severity categories) versus 1 unit in the wait list control group (15 to 14; no shift in symptom severity category).

Table 2. Primary and secondary study outcomes at 4 weeks.

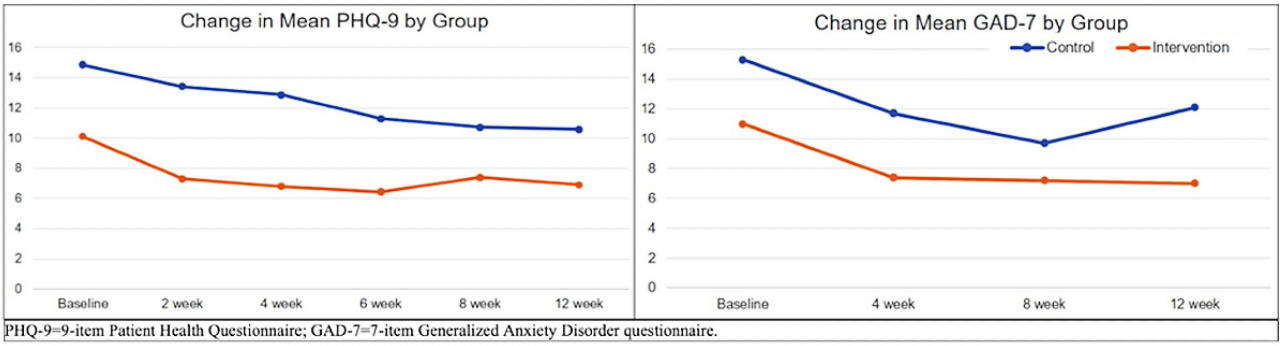
Outcome, mean (SD)	Intervention group	Wait list control group	95% CI	Cohen *d*
PHQ-9[a]	6.8 (6.9)	12.9 (5.4)	−12.9 to 0.7	.98
GAD-7[b]	7.4 (6.5)	11.7 (4.8)	−10.6 to 2	.75
MHSES[c]	38.7 (8.4)	31.7 (11)	−3.4 to 17.3	.71

[a]PHQ-9: 9-item Patient Health Questionnaire.

[b]GAD-7: 7-item Generalized Anxiety Disorder questionnaire.

[c]MHSES: Mental Health Self-Efficacy Scale.

Figure 3. Change in mean PHQ-9 and GAD-7 scores by treatment group over 12 weeks. GAD-7: 7-item Generalized Anxiety Disorder questionnaire; PHQ-9: 9-item Patient Health Questionnaire.



We also compared remission in depressive symptoms (PHQ-9 <5) at 4 and 12 weeks in patients with a baseline PHQ-9 >9 [57] in the subgroup diagnosed with depression. The percentage of participants achieving remission at both time points seemed to favor the active intervention, at 67% (2/3) and 0% (0/5) at 4 weeks and 50% (1/2) and 20% (1/5) at 12 weeks, respectively.

Safety and Satisfaction With Care

When completing the PHQ-9 on one of the 6 adolescent surveys, 10 (59%) of participants, of which 4 (40%) were in the intervention group and 6 (35%) were in the wait list control group, triggered at least 1 alarm to assess for suicidal ideation. During study participation, 4 (24%) of the participants had 1 alert, 4 (24%) had 3, and 2 (12%) had 6. These patients were managed by their PCP with support from a pediatric research psychiatrist, who interviewed all 4 (24%) participants by phone, discussed management with the legal guardian (who was the mother in all cases), and facilitated referral of 2 (12%) participants to a child psychiatrist for ongoing clinical management. One parent from the intervention group reported at week 12 that their teen was seen in an ED and discharged to home. Parental satisfaction with care did not differ between study groups.

Acceptability, Feasibility, and Usability for Adolescents and Parents

Adolescents found the app to be (1) acceptable, with a total of 8 (80%) agreeing or completely agreeing with the statement "I like using the app"; (2) feasible, with 7 (70%) agreeing with the statement "using the app in the treatment of depression seems possible"; and (3) usable (mean usability score 21.4, SD 1.7, possible range 5 to 25). The mean parental scores for acceptability and feasibility were 16.6 (SD 1.8) and 17 (SD 1.5), respectively (possible range 1 to 20).

Acceptability and Feasibility for Providers

Potential advantages for learning CBT skills via an app compared with face-to-face training identified by provider participants included ease/immediacy of access to evidence-based behavioral therapy, increased sense of anonymity or reduction in perception of stigma against mental health concerns, reduction in school absences to attend appointments, no need for transportation to and from a therapist's office, and reduction of out-of-pocket costs. Although providers acknowledged that an app might reduce common barriers to engagement in therapy, they voiced several concerns about using an app to augment clinical recommendations. Concerns commonly involved efficacy of the app, the potential safety risks of missing reports of suicidal ideation, overreliance on technology and reduced self-efficacy for adaptive help-seeking, increased screen time, and privacy and confidentiality of sensitive information.

Providers generally felt that an app to deliver CBT and positive psychology could be useful for teens with anxiety and mild-to-moderate depression, particularly early in the course of illness, and saw the app as being able to bridge the gap between diagnosis and obtaining an appointment with a therapist or the onset of therapeutic effects of prescribed antidepressant medications—both of which can take several weeks. Due to acute safety concerns, they recommended against using an app in teens with active suicidality or self-harm. They felt these patients needed closer clinical monitoring and were unlikely to use the app because of their low motivation to adopt new behaviors.

PCP participants suggested that the best time to introduce a CBT app to patients and families as part of the treatment plan would be at the time of diagnosis. Several suggested that usage

might be improved if the pediatrician introduced the app during an appointment and showed them how to download and use it (eg, by saying to the patient, "Here's what I'd like you to start doing…"). In summary, a CBT app was acceptable to these pediatric PCPs, and they suggested that it would be feasible to use in teens with mild/moderate depression or anxiety.

App Utilization

Participant app use during the intervention period (collected via the app) included days of app use, number of check-ins, and number of messages sent, lesson and tools use rates, lesson acceptability ratings indicated on a binary scale (ie, a thumbs up or thumbs down emoticon), and mood impact after tool utilization (ie, feeling same, better, or worse after completion). In-the-moment emotional state was reported through emoji selection with a default menu of 19 total moods, including options for negative (angry, sad, and anxious), positive (happy and content), and average mood (okay), with an additional ability to type in free text emotion words and self-selected emoji expressions.

Over the course of the intervention from baseline to 4 weeks, participants' app use averaged a mean of 6 (SD 6.9) days, with 55 (SD 7.14) mood check-ins, and 313.17 (SD 447.30) sent messages, with the use of 13.63 (SD 4.5) psychoeducational lessons and 9 (SD 14.91) tools. All completed psychoeducational lessons were rated positively (thumbs up). Over two-thirds (7/10, 69%) of participants reported feeling better after using the app, with 25% (3/10) and 1% (1/10) reporting feeling the same or worse, respectively.

Discussion

In this pragmatically designed, randomized pilot and feasibility study, we demonstrated the preliminary acceptability and feasibility of augmenting initial depression treatment in adolescents with CBT delivered via an evidence-based chatbot intervention. Although the sample size needed to detect between-group differences was not achieved, we observed a trend toward improvement in depressive symptoms with the app over a wait list control condition. Moreover, we observed that the detection and management of suicidal ideation in depressed adolescents was safely enhanced by use of the app in concert with PCP-based clinical care and telephonic consultation with a child psychiatrist. Finally, we found that using this app to address depression and anxiety as part of an overall treatment plan developed in collaboration with the treating PCP was acceptable to the adolescents, their parents, and their PCPs. These findings suggest that this app should be further evaluated for use in the primary care management of adolescents with depression and anxiety.

Adolescents visit their PCP 2 to 3 times per year on average and commonly report feeling most comfortable disclosing mental health concerns and obtaining mental health care in the primary care setting [14]. A recent review to assess the efficacy and acceptability of cell phone apps to support mental health management in adolescents found high levels of adherence, but few evaluative studies employed quantitative assessment of these process measures, which are key for guiding real-world implementation. This small pilot study provided additional lessons regarding real-world clinical implementation of mobile health interventions in pediatric primary care. For example, the frequency of suicidal ideation reporting (as assessed by question 9 of the PHQ-9) decreased over the course of study participation, even when the severity of other depressive symptoms remained in a clinically significant range. Although this study was not designed to detect differences between groups on suicidality, we observed a reduction in self-reported suicidal ideation that appeared to favor the active treatment group. This is particularly relevant in the delivery of mental health care to adolescent populations, who may feel more comfortable sharing such information anonymously through an app than with a trusted adult or health care provider.

Our results are subject to important limitations. First, despite using a fully remote study design and a well-established PBRN during a period of high mental health need, recruitment was challenged by the circumstances of the COVID-19 pandemic and the reduced access to and utilization of primary care resources, especially early on in the pandemic. Although the hybrid study design is important for piloting the intervention in the patient population of interest, the results must be interpreted with caution. Few conclusions can be drawn about intervention effectiveness due to the small sample size and the racial and geographic homogeneity and relative affluence of the enrolled participants, who were mostly female, White, and privately insured. The ability to detect differences between treatment groups may have been further limited by the concurrent use of antidepressant medications in some participants. The results regarding implementation effectiveness are subject to additional limitations, as the PBRN and PBRN providers in this study are not representative of the broader population of pediatric primary care clinics or providers in the United States [16]. Moreover, our safety approach of providing case review and rapid, individualized review and treatment planning with a child psychiatrist may also have contributed to the lack of separation between treatment groups, as everyone in the study received high-quality clinical care and close follow-up for suicidality reported on the PHQ-9.

The fully remote nature of the trial, which was conducted during critical periods in the COVID-19 pandemic, when child and adolescent mental health needs were most critical, was both a limitation and a strength. Additionally, this is the first study we are aware of that employed digital assessment and remote methods for managing suicidality in adolescents. The concern that suicidality will be exacerbated or missed with digital interventions has been a major barrier to offering these treatments to youth. The method we used to ensure safety (phone consultation with a child psychiatrist within 24 hours of reporting clinically significant suicidal ideation) could not have disguised symptom severity in the wait list control group, since we were responsive to all alerts. Another major strength of this pilot study was evaluating the feasibility of implementation in a real-world primary care setting. Although this study does not provide sufficient evidence to support the use of apps in primary care settings for teens with mental health problems [21], the results presented in this paper provide promising evidence that preliminarily support the feasibility of using a CBT-based

chatbot technology to supplement mental health treatment of adolescents with mild-to-moderate depression cared for in the primary care setting. Additional studies are needed to further evaluate the effectiveness of both the intervention and implementation in real-world primary care settings and in patient populations that include underrepresented populations and prioritize socioeconomically disadvantaged and rural communities, where access to high quality medical and mental health care are extremely limited.

Acknowledgments

Research reported in this publication was made possible through an academic partnership with the leadership of WoeBot, Allison Darcy, PhD (founder and president), Athena Robinson, PhD (chief clinical officer), and Sarah Parajito, MA (research scientist at WoeBot). This work was funded by the Washington University School of Medicine (WUSM) Clinical and Translational Funding Program (CTRFP) and the Institute of Clinical and Translational Sciences (ICTS) grant (UL1TR002345) from the National Center for Advancing Translational Sciences (NCATS) of the National Institutes of Health (NIH) and NIH/NCATS UL1 TR000445 (REDCap). The work reported in this manuscript was supported by the WUSM Pediatric and Adolescent Ambulatory Research Consortium (WU PAARC) and the WUSM Mobile Health Research Core (mHRC), both research cores of the WU ICTS. The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the official view of the NIH. Finally, we would like to thank the patients, families, and WU PAARC physicians for their participation, without which this study would not have been possible.

Conflicts of Interest

GN has received research support from Usona Institute (drug only), has served as principal or co-investigator on studies sponsored by Alkermes, Otsuka and LB Pharmaceuticals, and has received consulting fees from Alkermes, IngenioRx, Novartis and Sunovion. RW, SG, SD and JG have no conflicts of interest to disclose. AD, AR and SP are employees of WoeBot Health and were not involved in obtaining grant funding or IRB approval for this study, nor did they participate in data analyses. WoeBot Health provided app utilization data as reported in this manuscript.

References

1. Saluja G, Iachan R, Scheidt PC, Overpeck MD, Sun W, Giedd JN. Prevalence of and risk factors for depressive symptoms among young adolescents. Arch Pediatr Adolesc Med 2004 Aug;158(8):760-765. [doi: 10.1001/archpedi.158.8.760] [Medline: 15289248]
2. Perou R, Bitsko RH, Blumberg SJ, Pastor P, Ghandour RM, Gfroerer JC, Centers for Disease Control and Prevention (CDC). Mental health surveillance among children--United States, 2005-2011. MMWR Suppl 2013 May 17;62(2):1-35. [Medline: 23677130]
3. Asarnow JR, Jaycox LH, Duan N, LaBorde AP, Rea MM, Murray P, et al. Effectiveness of a quality improvement intervention for adolescent depression in primary care clinics: a randomized controlled trial. JAMA 2005 Jan 19;293(3):311-319. [doi: 10.1001/jama.293.3.311] [Medline: 15657324]
4. McCarthy C. Anxiety in teens is rising: What's going on? American Academy of Pediatrics. 2019. URL: https://www.healthychildren.org/English/health-issues/conditions/emotional-problems/Pages/Anxiety-Disorders.aspx [accessed 2020-09-10]
5. Leeb RT, Bitsko RH, Radhakrishnan L, Martinez P, Njai R, Holland KM. MMWR Morb Mortal Wkly Rep 2020 Nov 13;69(45):1675-1680 [FREE Full text] [doi: 10.15585/mmwr.mm6945a3] [Medline: 33180751]
6. Racine N, McArthur BA, Cooke JE, Eirich R, Zhu J, Madigan S. Global prevalence of depressive and anxiety symptoms in children and adolescents during COVID-19: a meta-analysis. JAMA Pediatr 2021 Nov 01;175(11):1142-1150 [FREE Full text] [doi: 10.1001/jamapediatrics.2021.2482] [Medline: 34369987]
7. Heron M. Deaths: leading causes for 2017. Natl Vital Stat Rep 2019 Jun;68(6):1-77 [FREE Full text] [Medline: 32501203]
8. Miron O, Yu K, Wilf-Miron R, Kohane IS. Suicide rates among adolescents and young adults in the United States, 2000-2017. JAMA 2019 Jun 18;321(23):2362-2364 [FREE Full text] [doi: 10.1001/jama.2019.5054] [Medline: 31211337]
9. Curtin SC. State suicide rates among adolescents and young adults aged 10-24: United States, 2000-2018. Natl Vital Stat Rep 2020 Sep;69(11):1-10. [Medline: 33054915]
10. Yard E, Radhakrishnan L, Ballesteros MF, Sheppard M, Gates A, Stein Z, et al. Emergency department visits for suspected suicide attempts among persons aged 12-25 years before and during the COVID-19 pandemic - United States, January 2019-May 2021. MMWR Morb Mortal Wkly Rep 2021 Jun 18;70(24):888-894 [FREE Full text] [doi: 10.15585/mmwr.mm7024e1] [Medline: 34138833]
11. March J, Silva S, Petrycki S, Curry J, Wells K, Fairbank J, et al. Fluoxetine, cognitive-behavioral therapy, and their combination for adolescents with depression: Treatment for Adolescents With Depression Study (TADS) randomized controlled trial. JAMA 2004 Aug 18;292(7):807-820. [doi: 10.1001/jama.292.7.807] [Medline: 15315995]

12. Cheung AH, Zuckerbrot RA, Jensen PS, Laraque D, Stein REK, GLAD-PC Steering Group. Guidelines for Adolescent Depression in Primary Care (GLAD-PC): Part II. Treatment and ongoing management. Pediatrics 2018 Mar;141(3) [FREE Full text] [doi: 10.1542/peds.2017-4082] [Medline: 29483201]
13. Zuckerbrot RA, Cheung A, Jensen PS, Stein REK, Laraque D, GLAD-PC Steering Group. Guidelines for Adolescent Depression in Primary Care (GLAD-PC): Part I. Practice preparation, identification, assessment, and initial management. Pediatrics 2018 Dec;141(3) [FREE Full text] [doi: 10.1542/peds.2017-4081] [Medline: 29483200]
14. Stein REK, Zitner LE, Jensen PS. Interventions for adolescent depression in primary care. Pediatrics 2006 Aug;118(2):669-682. [doi: 10.1542/peds.2005-2086] [Medline: 16882822]
15. Shafran R, Clark DM, Fairburn CG, Arntz A, Barlow DH, Ehlers A, et al. Mind the gap: Improving the dissemination of CBT. Behav Res Ther 2009 Nov;47(11):902-909. [doi: 10.1016/j.brat.2009.07.003] [Medline: 19664756]
16. Garbutt J, Sterkel R, Ruecker K, Dodd S, Smith E, Plax K. Ready for the challenge of depression care in the medical home. Clin Pediatr (Phila) 2019 Jun;58(7):816-818 [FREE Full text] [doi: 10.1177/0009922819834280] [Medline: 30866675]
17. 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 13;5(4):e64 [FREE Full text] [doi: 10.2196/mental.9782] [Medline: 30545815]
18. Kretzschmar K, Tyroll H, Pavarini G, Manzini A, Singh I, NeurOx Young People's Advisory Group. Can your phone be your therapist? Young people's ethical perspectives on the use of fully automated conversational agents (chatbots) in mental health support. Biomed Inform Insights 2019;11:1178222619829083 [FREE Full text] [doi: 10.1177/1178222619829083] [Medline: 30858710]
19. Andrews G, Basu A, Cuijpers P, Craske MG, McEvoy P, English CL, et al. Computer therapy for the anxiety and depression disorders is effective, acceptable and practical health care: An updated meta-analysis. J Anxiety Disord 2018 Apr;55:70-78 [FREE Full text] [doi: 10.1016/j.janxdis.2018.01.001] [Medline: 29422409]
20. 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 06;4(2):e19 [FREE Full text] [doi: 10.2196/mental.7785] [Medline: 28588005]
21. Merry SN, Stasiak K, Shepherd M, Frampton C, Fleming T, Lucassen MFG. The effectiveness of SPARX, a computerised self help intervention for adolescents seeking help for depression: randomised controlled non-inferiority trial. BMJ 2012 Apr 18;344:e2598 [FREE Full text] [doi: 10.1136/bmj.e2598] [Medline: 22517917]
22. 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]
23. Hoermann S, McCabe KL, Milne DN, Calvo RA. Application of synchronous text-based dialogue systems in mental health interventions: systematic review. J Med Internet Res 2017 Jul 21;19(8):e267 [FREE Full text] [doi: 10.2196/jmir.7023] [Medline: 28784594]
24. Ebert DD, Zarski A, Christensen H, Stikkelbroek Y, Cuijpers P, Berking M, et al. Internet and computer-based cognitive behavioral therapy for anxiety and depression in youth: a meta-analysis of randomized controlled outcome trials. PLoS One 2015;10(3):e0119895 [FREE Full text] [doi: 10.1371/journal.pone.0119895] [Medline: 25786025]
25. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. Cogn Behav Ther 2009;38(4):196-205. [doi: 10.1080/16506070903318960] [Medline: 20183695]
26. 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]
27. Aschbrenner KA, Naslund JA, Tomlinson EF, Kinney A, Pratt SI, Brunette MF. Adolescents' use of digital technologies and preferences for mobile health coaching in public mental health settings. Front Public Health 2019 Jul 2;7:178 [FREE Full text] [doi: 10.3389/fpubh.2019.00178] [Medline: 31312629]
28. Nick EA, Kilic Z, Nesi J, Telzer EH, Lindquist KA, Prinstein MJ. Adolescent digital stress: frequencies, correlates, and longitudinal association with depressive symptoms. J Adolesc Health 2022 Feb;70(2):336-339. [doi: 10.1016/j.jadohealth.2021.08.025] [Medline: 34666954]
29. Arias-de la Torre J, Puigdomenech E, García X, Valderas JM, Eiroa-Orosa FJ, Fernández-Villa T, et al. Relationship between depression and the use of mobile technologies and social media among adolescents: umbrella review. J Med Internet Res 2020 Aug 26;22(8):e16388 [FREE Full text] [doi: 10.2196/16388] [Medline: 32663157]
30. Glasgow RE, Fisher L, Strycker LA, Hessler D, Toobert DJ, King DK, et al. Minimal intervention needed for change: definition, use, and value for improving health and health research. Transl Behav Med 2014 Mar;4(1):26-33 [FREE Full text] [doi: 10.1007/s13142-013-0232-1] [Medline: 24653774]
31. Li LSE, Wong LL, Yap KY. Quality evaluation of stress, anxiety and depression apps for COVID-19. J Affect Disord Rep 2021 Dec;6:100255 [FREE Full text] [doi: 10.1016/j.jadr.2021.100255] [Medline: 34704086]
32. Ramachandran M, Suharwardy S, Leonard S, Gunaseelan A, Robinson A, Darcy A, et al. Acceptability of postnatal mood management through a smartphone-based automated conversational agent. Am J Obstet Gynecol 2020 Jan;222(1):S62 [FREE Full text] [doi: 10.1016/j.ajog.2019.11.090]

33. Suharwardy S, Ramachandran M, Leonard S, Gunaseelan A, Robinson A, Darcy A, et al. Effect of an automated conversational agent on postpartum mental health: A randomized, controlled trial. Am J Obstet Gynecol 2020 Jan;222(1):S91 [FREE Full text] [doi: 10.1016/j.ajog.2019.11.132]
34. Prochaska JJ, Vogel EA, Chieng A, Kendra M, Baiocchi M, Pajarito S, et al. A therapeutic relational agent for reducing problematic substance Use (Woebot): development and usability study. J Med Internet Res 2021 Mar 23;23(3):e24850 [FREE Full text] [doi: 10.2196/24850] [Medline: 33755028]
35. Prochaska JJ, Vogel EA, Chieng A, Baiocchi M, Maglalang DD, Pajarito S, et al. A randomized controlled trial of a therapeutic relational agent for reducing substance misuse during the COVID-19 pandemic. Drug Alcohol Depend 2021 Oct 01;227:108986 [FREE Full text] [doi: 10.1016/j.drugalcdep.2021.108986] [Medline: 34507061]
36. Darcy AP, Eaneff S, Parajito S, Robinson A. Feasibility, acceptability, and effectiveness of a digital therapeutic application vs psychoeducation content for mild-to-moderate depression among adolescents. 2020 Presented at: Association for Behavioral and Cognitive Therapies; November 19-22; Virtual p. 2022.
37. Garbutt J, Dodd S, Rook S, Graham S, Wang R, Sterkel R, et al. Improving follow-up for adolescents with depression in primary care. Pediatrics 2022 Jun 01;149(6):2021. [doi: 10.1542/peds.2021-051107] [Medline: 35641468]
38. Torous JB, Chan SR, Gipson SYT, Kim JW, Nguyen T, Luo J, et al. A hierarchical framework for evaluation and informed decision making regarding smartphone apps for clinical care. Psychiatr Serv 2018 Feb 15:appips201700423. [doi: 10.1176/appi.ps.201700423] [Medline: 29446337]
39. O'Reilly K. Proposed guidelines aim for safe, effective mobile health apps. American Medical Association. 2018. URL: https://www.ama-assn.org/practice-management/digital/proposed-guidelines-aim-safe-effective-mobile-health-apps [accessed 2022-09-10]
40. Kianersi S, Luetke M, Ludema C, Valenzuela A, Rosenberg M. Use of research electronic data capture (REDCap) in a COVID-19 randomized controlled trial: a practical example. BMC Med Res Methodol 2021 Aug 21;21(1):175 [FREE Full text] [doi: 10.1186/s12874-021-01362-2] [Medline: 34418958]
41. 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] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
42. Johnson JG, Harris ES, Spitzer RL, Williams JBW. The patient health questionnaire for adolescents: validation of an instrument for the assessment of mental disorders among adolescent primary care patients. J Adolesc Health 2002 Mar;30(3):196-204. [doi: 10.1016/s1054-139x(01)00333-0] [Medline: 11869927]
43. 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]
44. Dear BF, Titov N, Sunderland M, McMillan D, Anderson T, Lorian C, et al. Psychometric comparison of the generalized anxiety disorder scale-7 and the Penn State Worry Questionnaire for measuring response during treatment of generalised anxiety disorder. Cogn Behav Ther 2011;40(3):216-227. [doi: 10.1080/16506073.2011.582138] [Medline: 21770844]
45. Tiirikainen K, Haravuori H, Ranta K, Kaltiala-Heino R, Marttunen M. Psychometric properties of the 7-item Generalized Anxiety Disorder Scale (GAD-7) in a large representative sample of Finnish adolescents. Psychiatry Res 2019 Feb;272:30-35. [doi: 10.1016/j.psychres.2018.12.004] [Medline: 30579178]
46. Sun J, Liang K, Chi X, Chen S. Psychometric properties of the Generalized Anxiety Disorder Scale-7 Item (GAD-7) in a large sample of Chinese adolescents. Healthcare (Basel) 2021 Dec 09;9(12):1709 [FREE Full text] [doi: 10.3390/healthcare9121709] [Medline: 34946435]
47. Ryu HJ, Lee S, Chung JH. Association between asthma and generalized anxiety disorder-7 (GAD-7) in Korean adolescents. J Asthma 2022 Feb 02:1-7. [doi: 10.1080/02770903.2022.2032137] [Medline: 35060429]
48. Adjorlolo S. Generalised anxiety disorder in adolescents in Ghana: Examination of the psychometric properties of the Generalised Anxiety Disorder-7 scale. AJOPA 2019 Jul 18;1. [doi: 10.4102/ajopa.v1i0.10]
49. Romano I, Ferro M, Patte K, Leatherdale S. Measurement invariance of the GAD-7 and CESD-R-10 among adolescents in Canada. J Pediatr Psychol 2022 May 13;47(5):585-594 [FREE Full text] [doi: 10.1093/jpepsy/jsab119] [Medline: 35552429]
50. Mossman SA, Luft MJ, Schroeder HK, Varney ST, Fleck DE, Barzman DH, et al. The Generalized Anxiety Disorder 7-item scale in adolescents with generalized anxiety disorder: Signal detection and validation. Ann Clin Psychiatry 2017 Nov;29(4):227-234A [FREE Full text] [Medline: 29069107]
51. Clarke J, Proudfoot J, Birch M, 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;14:272 [FREE Full text] [doi: 10.1186/s12888-014-0272-1] [Medline: 25252853]
52. Weiner BJ, Lewis CC, Stanick C, Powell BJ, Dorsey CN, Clary AS, et al. Psychometric assessment of three newly developed implementation outcome measures. Implement Sci 2017 Aug 29;12(1):108 [FREE Full text] [doi: 10.1186/s13012-017-0635-3] [Medline: 28851459]
53. Lewis JR. The System Usability Scale: past, present, and future. Int J Hum-Comput Interact 2018 Mar 30;34(7):577-590 [FREE Full text] [doi: 10.1080/10447318.2018.1455307]

54. Brooke J. SUS: a "quick and dirty" usability scale. In: Usability Evaluation in Industry. London, UK: Taylor and Francis; 1996:189-194.
55. Ramtekkar U, Maras M, Ell W, Nicol G, Young-Walker L. Academic-community partnership to improve pediatric mental health access: Missouri Child Psychiatry Access Project. Psychiatr Serv 2022 May;73(5):588-591. [doi: 10.1176/appi.ps.202100074] [Medline: 34470505]
56. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
57. Schueller SM, Kwasny MJ, Dear BF, Titov N, Mohr DC. Cut points on the Patient Health Questionnaire (PHQ-9) that predict response to cognitive-behavioral treatments for depression. Gen Hosp Psychiatry 2015;37(5):470-475 [FREE Full text] [doi: 10.1016/j.genhosppsych.2015.05.009] [Medline: 26077754]

Abbreviations

AIM: Acceptability of Intervention Measure
AMA: American Medical Association
APA: American Psychiatric Association
CBT: cognitive behavioral therapy
DBT: dialectical behavior therapy
DSM-5: Diagnostic and Statistical Manual of Mental Disorders fifth edition
ED: emergency department
FIM: Feasibility of Intervention Measure
GAD-7: 7-item Generalized Anxiety Disorder questionnaire
ICTS: Institute for Clinical and Translational Science
IPT-A: interpersonal psychotherapy for adolescents
mHealth: mobile health
MHSES: Mental Health Self-Efficacy Scale
PBRN: practice-based research network
PCP: primary care provider
PHQ-9: 9-item Patient Health Questionnaire
PHQ-A: 9-item Patient Health Questionnaire modified for Adolescents
PI: principal investigator
QI: quality improvement
RCT: randomized controlled trial
REDCap: Research Electronic Data Capture
SNRI: serotonin-norepinephrine reuptake inhibitor
SSRI: selective serotonin reuptake inhibitor
SUS: System Usability Scale
WU PAARC: Washington University Pediatric and Adolescent Ambulatory Research Consortium

Edited by A Mavragani; submitted 12.06.22; peer-reviewed by R Benson, A Ecker; comments to author 14.09.22; revised version received 10.10.22; accepted 25.10.22; published 22.11.22.

Please cite as:
Nicol G, Wang R, Graham S, Dodd S, Garbutt J
Chatbot-Delivered Cognitive Behavioral Therapy in Adolescents With Depression and Anxiety During the COVID-19 Pandemic: Feasibility and Acceptability Study
JMIR Form Res 2022;6(11):e40242
URL: https://formative.jmir.org/2022/11/e40242
doi:10.2196/40242
PMID:36413390

Original Paper

Lessons Learned From the SoBeezy Program for Older Adults During the COVID-19 Pandemic: Experimentation and Evaluation

Marion Pech[1*], MSc; Antoine Gbessemehlan[1,2*], PhD; Lucile Dupuy[1], PhD; Hélène Sauzéon[1,3], PhD; Stéphane Lafitte[2,4,5,6], MD, PhD; Philippe Bachelet[2], MSc; Hélène Amieva[1], PhD; Karine Pérès[1,2*], PhD

[1]Bordeaux Population Health Research Center, Inserm, UMR 1219, University of Bordeaux, Bordeaux, France

[2]The SoBeezy Project, Pessac, France

[3]National Institute for Research in Digital Science and Technology, University of Bordeaux, Bordeaux, France

[4]Haut-Lévêque Cardiology Hospital, University Hôpital of Bordeaux, Pessac, France

[5]University of Bordeaux, Bordeaux, France

[6]Centre d'Investigation Clinique Plurithématique de Bordeaux, Pessac, France

[*]these authors contributed equally

Corresponding Author:
Marion Pech, MSc
Bordeaux Population Health Research Center
Inserm, UMR 1219
University of Bordeaux
146 rue Léo Saignat
Bordeaux, 33076
France
Phone: 33 667455145
Email: marion.pech@u-bordeaux.fr

Abstract

Background: The SoBeezy program is an innovative intervention aimed at promoting and fostering healthy aging and aging in place by proposing to older adults concrete solutions to face daily life, tackle loneliness, promote social participation, and reduce the digital divide, thanks to a specific, easy-to-use voice assistant (the BeeVA smart display).

Objective: This study aims to assess the acceptability of the SoBeezy program and its voice assistant and to identify potential areas of improvement.

Methods: A 12-month experimentation of the program was deployed in real-life conditions among older adults living in the community in 4 pilot cities of France. Launched during the first lockdown of the COVID-19 crisis, this multisite study aimed to assess acceptability using questionnaires and interviews conducted at baseline and at the end of the experimentation. In addition, a series of meetings were conducted with SoBeezy staff members to obtain direct feedback from the ground.

Results: In total, 109 older individuals were equipped with BeeVA to use the SoBeezy program; of these, 32 (29.4%) left the experimentation before its end and 69 (63.3%) completed the final questionnaires. In total, 335 interventions were conducted and 27 (39%) of the participants requested services, mainly for supportive calls and visits and assistance with shopping, transportation, and crafting-gardening. Of the whole sample, 52 (75%) considered BeeVA as a reassuring presence, and few persons (15/69, 22%) reported a negative opinion about the program. Among the participants, the voice assistant appeared easy to use (n=57, 82%) and useful (n=53, 77%). They also were positive about the BeeVA smart display and the SoBeezy intervention.

Conclusions: This multisite study conducted in real-life conditions among more than 100 older adults living in the community provides enlightening results of the reality from the ground of digital tools designed for the aging population. The COVID-19 context appeared both as an opportunity, given the massive needs of the older adults during this crisis, and as limiting due to sanitary constraints. Nevertheless, the experimentation showed overall good acceptability of the voice assistant and a high level of satisfaction of the participants among those who really used the system and could be a way of improving the autonomy and well-being of older adults and their families. However, the findings also highlighted resistance to change and difficulties for the users to ask for help. The experimentation also emphasized levers for next deployments and future research. The next step will be the experimentation of the activity-sharing component that could not be tested due to the COVID-19 context.

(JMIR Form Res 2022;6(11):e39185) doi:10.2196/39185

KEYWORDS

voice assistance; social isolation; healthy aging; living in place; acceptability; technologies; digital divide; older adults; aging; elderly population; voice assistant; COVID-19

Introduction

The increase in life expectancy along with the massive arrival of baby boomers at retirement age have led to a major transformation of the age population structure. In 2040, 1 in 4 inhabitants will be 65 years old or over compared to 18% in 2013, and by 2070, the population aged 75 years or more would be twice as numerous as in 2013 (+7.8 million) [1]. In this context, healthy aging has become a major challenge for societies, as suggested by various plans, programs, and policy orientations worldwide [2-6]. Beyond the obvious health factors central to the healthy aging process, personal factors (eg, resilience abilities, self-esteem, personality traits) and environmental ones (eg, social support, living environment) also play a crucial role [7]. For instance, loneliness and social isolation, accentuated during the COVID-19 crisis, are growing at an impressive pace, particularly in the older population [8-11]. Approximately half of the adults aged 60 years and over are at risk of social isolation [12], and one-third will experience some degree of loneliness in later life [13,14]. It is now well established that loneliness and social isolation compromise living in place [15,16] and are risk factors for an unhealthy lifestyle (eg, sedentary lifestyle, poor diet, tobacco, alcohol), morbidity (eg, chronic conditions, such as cardiovascular disease, stroke, dementia, depression, anxiety, and disability), and mortality [17-19]. Therefore, to tackle social isolation and loneliness among older adults, numerous interventions are being developed and deployed [16], many of them involving technological devices.

Technology plays an increasingly central role in the ways we communicate. Indeed, we are witnessing a paradigm shift in communication, where face-to-face exchanges are no longer the only way to maintain interpersonal connection [20], thanks to information communication technology (ICT) [21,22]. There are 2 types of online services [23,24], social network services (SNS) and social online services (SOS). SNS were created first. They are online environments where users create a personal profile, build a network of personal connections [23,25], and can thus stay in touch with friends, family, and acquaintances (eg, Facebook) [26]. For instance, Neves et al [21] developed an age-adapted app on iPad that allows nursing home residents to stay in touch with their families by providing easy access to photos, audio recordings, videos, and messages. This app increased residents' perceived levels of social interactions but only among people whose relatives were geographically distant and people who had a higher need to feel socially included, which drove them to adapt to technological constraints. Other SNS have been developed, such as Meeteetse [27], the ASTRA app [28], ShoddyPop, and PersonCard [29]. Following the advent of SNS, 2 types of social platforms arose, social commerce platforms (eg, eBay) and social solidarity platforms (eg, based on collaborative consumption platforms, such as Swaptree, Airbnb, Getaround, and Taskrabbit), both based on the trade and exchange of goods and services between people. Social solidarity platforms may be useful to older adults to help with daily activities and promote social participation [24]. A systematic review [30] identified social platforms dedicated to older persons that provide easy access to information sources and communication opportunities. These platforms enhance social connectivity and promote healthy lifestyles, including physical and cognitive activities (eg, games activities, such as the Wii console), a safer environment, and positive emotions. They include several functions, such as self-monitoring, calendars, photos, games, and online assistance [31-35]. Based on the sharing of personal services, which hold the potential to strengthen social integration and enable an independent lifestyle for older adults, Koene et al [24] developed a local, service-oriented collaborative consumption platform called "Bring Dich ein!". This platform aims to facilitate social interactions across generations and peer-to-peer services. The platform was implemented in a participatory development process. In the pilot phase, usability was good, but in the absence of subsequent publication, we cannot know whether the promise has been kept [24].

To summarize, SNS target more loneliness, enabling the creation or maintenance of social interactions [23], whereas SOS support accessibility of services to compensate and help people with health problems, disabilities (eg, walking difficulties), and social isolation issues [24,36]. With advancing age, older adults may need to stay in touch with their family or friends and exchange services within a secure community. For these reasons, a few studies have tried to combine SNS and SOS to increase the appropriation of such technologies. For instance, Personal Reminder Information and Social Management (PRISM) is a software application [31] designed to support social connectivity, memory, knowledge about various topics, leisure activities, and access to resources. Boll and Brune [36] proposed a prototype platform providing an integrated online environment, in particular to help bridge the gap between older adults, and services from professionals as well as from other users. However, we lack robust data to assess the benefit of this combination. On the one hand, grouping services and social networks in a single platform seems helpful. On the other hand, the use of the device could be harder to understand (longer menus, more services with more complicated pathways) [37], potentially leading to a negative user experience, especially in older users [38].

When focusing on older users, the conception and implementation of online services must address several ergonomic issues, such as age-appropriate design (ie, ease of use), compatibility with the user's needs (ie, usefulness), and technical issues, such as reliability [39] and privacy [23] concerns. Characteristics of the older users themselves, such as familiarity with technologies, adequate social support, cognitive

abilities, or health status, also need to be considered since these factors can also influence technology use [39]. For instance, interfaces are often too complex, have too many options, and are not appropriate for "nontypical" users who suffer from sensory or cognitive impairments or do not have a technological background. Furthermore, the user interface should be as simple as possible, for example, by grouping similar items and functionalities to help users who have no experience in using this type of interface, by providing the users with only essential information, by increasing header and content sizes, and by giving the option to zoom in and out for people with visual impairments. These solutions could be easily implemented and do not presume drastic changes in the standard user interaction of SNS, such as Facebook [23]. Multimodality (multiple modes of interacting with a system) is also recommended for intuitive use. It provides an opportunity to the user to choose the best-adapted mode regarding their skills, abilities, habits, and wishes. These functionalities, such as text-to-speech, text input, speech commands, or other augmentative alternatives, could have a positive impact on device appropriation [23]. Several requirements and design rationales were deemed essential by older adults, such as intuitive interaction and navigation, a closed community, strong privacy policies, and community consciousness [24].

Technology may deeply modify the ways we communicate and could be relevant to tackle loneliness. However, isolated people are also less likely to use these types of devices due to lower skills, greater reactance, and lack of motivation and support from family [40]. In the current intensive process of world digitalization, there is an urgent need to create accessible, adaptable, and easy-to-use tools for all to reduce the associated risk of social exclusion of the older population.

In this context, the SoBeezy program has been developed to foster healthy aging at home by facilitating and improving older adults' daily life [2]. The system proposes solutions to face the main difficulties encountered in daily life and fosters social participation by promoting community-based cooperation and the sharing of activities and experiences. The program relies on (1) an intelligent digital platform available on smartphones, tablets, and computers and also a voice assistant (BeeVA) specifically developed for people with a digital divide; (2) an extensive solidarity network that potentially relies on everyone's engagement through an intergenerational approach [41,42], where older people themselves are not only service receivers but also potential contributors; and (3) all the local partners and stakeholders available to cooperate (associations, social services of municipalities, health professionals, home care services, and all relevant local partners, such as artisans). The SoBeezy program is organized as a hub and connects all the territory's resources to provide the best solution to meet the user's needs. The program has been implemented for 12 months, specifically targeting older adults living alone or suffering from loneliness. The objectives of this study were (1) to assess the usage, service satisfaction, acceptability of BeeVA and, more generally, the SoBeezy program and (2) to identify the potential amendments that should be provided to improve the system.

Methods

A 12-Month Experimentation in Real-Life Conditions Among Older Adults Living in the Community

As previously published [2], the initial protocol of the SoBeezy program planned before-after analyses and a comparative approach with a control group to assess the impact and effectiveness on healthy aging, technical usage, mechanisms of intervention, and conditions of transferability and scalability. However, due to the particular context of the COVID-19 pandemic and given the massive needs of the older population at that time, we decided to anticipate the launch of the program and to prioritize the deployment of the device and the assistance given to older adults. Consequently, the evaluative research could not be implemented, as planned, and had to be adapted to this extraordinary context: removal of several services and activities, impossibility to recruit a control group in the pandemic context, and baseline data collection restricted to the bare minimum (as detailed later).

Initially scheduled in May 2020, the launch of the program was anticipated with a solidarity campaign during the first lockdown (supportive calls and assistance with shopping, transportation, dog walking without the SoBeezy technology) in April 2020 to respond to the massive needs generated by the COVID-19 crisis (Figure 1). Then from July, these services were extended to crafting-gardening, supportive visits, and at-home hairdressing and were made available on the SoBeezy platform and the BeeVA smart display (Hello 10 Archos). However, the program has only been partially deployed due to the restrictive barrier measures. Indeed, the activity-sharing component could not be analyzed and the assistance in daily living was only limited to the essentials (Figure 2). BeeVA also proposed several options, such as weather forecast, radio, a digital calendar, emergency numbers, games, and city news (Figure 3).

The experimentation started in July 2020 with the installation of BeeVA in the participants' homes as soon as it was allowed by sanitary measures (installations staggered over the first 6 months) and ended in June 2021. It took place in 4 pilot cities (2 urban cities, Pessac and Saint Jean de Luz, and 2 rural cities, St Yrieix la Perche and Langon) in southwest France, with a close partnership with the municipalities. The participants were recruited among older adults supported by the SoBeezy solidarity campaign but also with the support of local social services (municipalities), health professionals, local associations, a private social protection agency (AG2R La Mondiale), and communication campaigns (press, radio, social media). The eligibility criteria were being 50 years and over, living in 1 of the 4 pilot cities, living in an area with sufficient access to high-speed internet service, being free of severe visual or hearing impairments, being free of moderate-to-severe cognitive impairment, and being a French speaker (for better voice recognition by BeeVA). All participants were equipped (free of charge) with BeeVA and an internet connection (a 4G Wi-Fi device).

Each person potentially interested in participating received the first visit at home for a detailed presentation of the program and BeeVA. A second visit was scheduled to install the device and

propose the first user training if the person was still interested. A user manual provided instructions, including instructions in the case of technical problems as well as a specific hotline phone number. A few days later, a phone call aimed at ensuring the correct handling and use of the device. Other training/coaching visits were conducted as many times as necessary, and regular follow-up phone calls were made.

Figure 1. Representation of the SoBeezy experimentation in the general population.

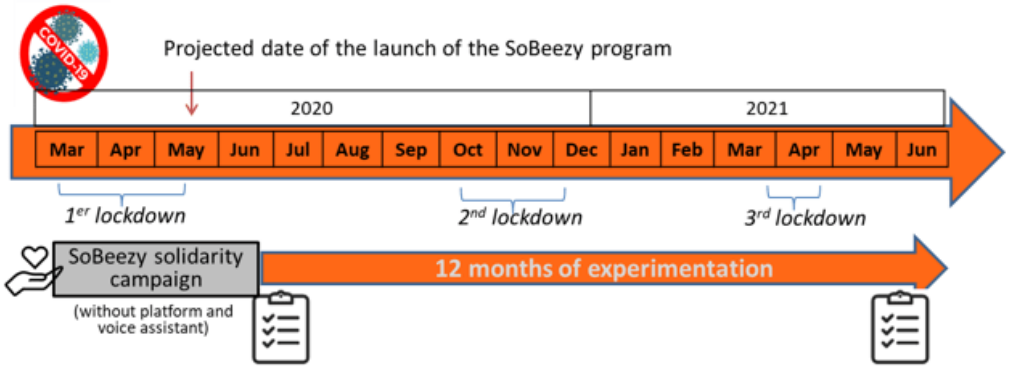


Figure 2. The BeeVA home page.

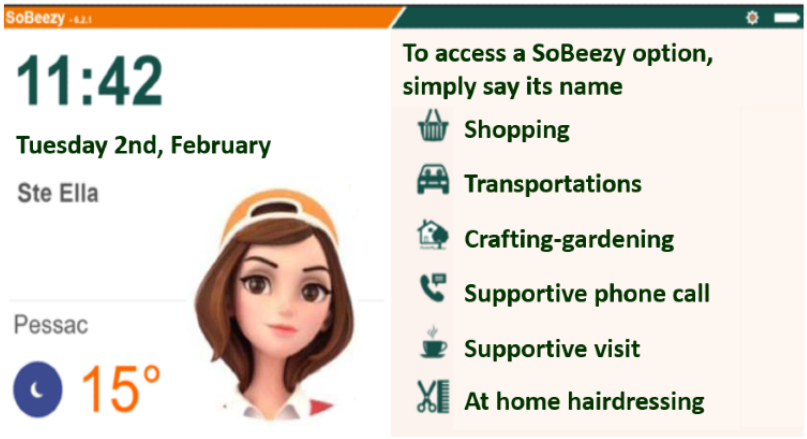
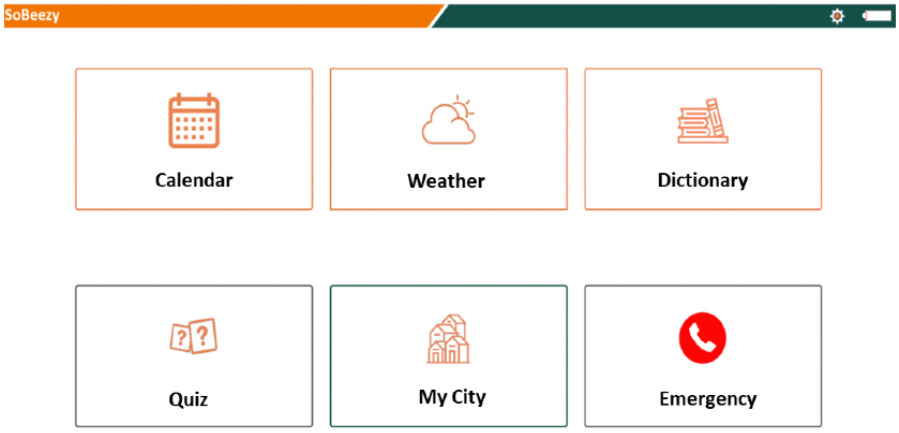


Figure 3. Options section.



Evaluation of the Experimentation

To assess the acceptability of the program and of BeeVA, questionnaires and interviews were administered at baseline and after 12 months (Figure 1). Due to the COVID-19 context, evaluations were conducted by phone to respect the barrier measures, as recommended.

General Individual Data

Information about age, gender, living status (living alone vs not living alone), city, beneficiary of home care services, member of associations, self-reported global health, and comfort level with online technologies was collected.

Acceptability of BeeVA

The number of requested services and options used by the participants was collected automatically by the system (log data). Usage frequency (number of interactions per month) was also recorded. At the final assessment, the participants were invited to answer 17 questions (Multimedia Appendix 1). As recommended by Chen and Chan [39], we distinguished the following 5 dimensions: perceived usefulness (4 items), usage behavior (4 items relative to the options used), ease of use (3 items), reliability (1 item), and appreciation of the design and ergonomics of the device (5 items).

Service Satisfaction of the SoBeezy Program

In total, 5 items were used (Multimedia Appendix 1): satisfaction regarding the services provided by SoBeezy, delay of answers and intervention, quality of the relationships with SoBeezy staff members, volunteers and professionals, and communication preferences with these 3 contributors, with 3 answer modalities (rather yes/rather no/never used).

Global Perception of the SoBeezy Intervention

We combined 1 item related to BeeVA ("Is the voice assistant a useful tool?") and 1 related to the services ("Are you satisfied with the services provided?"). We obtained 3 main opinions about SoBeezy: (1) 2 positive answers defined a positive opinion, (2) only 1 positive answer defined a mixed opinion, and (3) negative answers or no usage defined a negative opinion.

Improvement Tracks

In the perspective of improving the system, the participants were also invited to assess the usefulness/interest of 18 new possible options and features on a scale ranging from 1 (not interested at all) to 7 (very interested). These options included photos and messages sent by the family, music, radio, audiobooks, games or an e-calendar, and medication reminders (for a detailed description of the options, see Multimedia Appendix 2). Finally, the activity-sharing (leisure, physical and cultural activities) component that could not be experimented upon due to the COVID-19 crisis was also proposed as a future option.

In addition, to collect direct feedback from the SoBeezy staff members (comprising volunteers and employees), a series of meetings were conducted at the end of the experimentation, with 4 main topics: the health crisis context, older users, technological aspects, and organizational challenges.

Statistical Analysis

Statistical analyses were mainly descriptive and comparative. We reported means and SDs for continuous variables and frequencies for categorical variables. For comparative analyses, adequate statistical tests (chi-square and Fisher test) were performed. All analyses were performed using R version 4.1.2 (R Foundation for Statistical Computing).

Ethical Considerations

All participants provided written informed consent to participate in the study. Data protection complies with European and French data protection regulations (GDPR and CNIL). Privacy and confidentiality protection was ensured by systematically conducted statistical analyses on de-identified data. The protocol and informed consent and assent forms were approved by the Comité d'Evaluation Ethique de l'INSERM (CEEI; Institutional Review Board [IRB] approval 2020-16/05). Finally, as compensation for their participation in the research, the smart display was offered to each participant.

Results

Sample Description

In total, 256 persons were invited to participate in the study by telephone, of which 109 (43% participation rate) accepted. Among them, 77 (71%) completed the experimentation, and 69 (90%) of them completed the final assessment conducted at the end of the study in June 2021 (Figure 4). In total, 109 participants were equipped with BeeVA. The mean age was 81.2 years (SD 8.6), 86 (78.9%) were women, 66 (60.6%) lived in Pessac City, and 44 (55.7%) reported alteration in general health (eg, walking difficulty). See Table 1 for details.

Among the 109 participants, 32 (29.4%) requested the device to be uninstalled before the end of the study, after 2.8 months (SD 2.5) of use, on average. These participants were more likely to be older (mean 83.2 years, SD 7.9 years) compared to others (mean 80.4 years, SD 8.8 years), women (n=27, 84%, vs n=59, 77%), and tended to use less often the device and services.

Among the 109 participants, 47 (43.1%) used BeeVA for 8-12 months (the device installation being staggered over the first 6 months and 32 participants leaving the study prematurely). Of the 69 participants who completed the experimentation, 41 (59%) used it beyond 8 months.

Using the general information collected at baseline, we proposed a description of the characteristics of the completers and noncompleters in Table 1. Among the completers, almost 75% (n=52) were living alone, one-third (n=23, 33%) benefitted from home care services, and half (n=35, 51%) were members of an association (Table 1). Concerning the general use of technologies, the smartphone was the most frequently used device (n=39, 57%, using it regularly), followed by the computer (n=30, 21%), while only 9 (13%) participants used a digital tablet. In addition, 35 (56%) participants estimated their comfort level using technology as poor, with a higher frequency among the oldest participants, those living alone, and those with good global health/mobility (Table 1).

To identify the main reasons for ceasing participation, we conducted a qualitative analysis of the individual files (no information available for 5, 16%, of 32 participants). The 3 main reasons were technical problems (internet, breakdown, or difficulty using the BeeVA; n=8, 25%), followed by health problems (n=5, 16%), and no need of services (n=5, 16%).

Figure 4. Flowchart of the description of the sample.

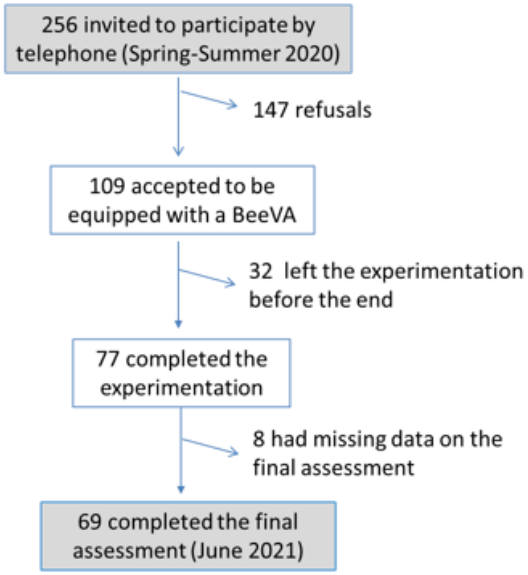


Table 1. Description and comparison of samples of completers (n=69) and noncompleters (n=40).

Characteristics	Completers (evaluation sample; n=69)	Noncompleters (n=40)
Age (years), mean (SD)	80.2 (8.9)	82.9 (7.8)
Women, n (%)	53 (77)	33 (83)
Pilot city, n (%)		
Pessac (urban area)	42 (61)	24 (60)
St Jean de Luz (urban area)	14 (20)	11 (28)
St Yrieix la Perche (rural area)	8 (12)	2 (5)
Langon (rural area)	5 (7)	3 (8)
Living alone[a], n (%)	49 (71)	N/A[b]
Recipient of home care services[a], n (%)	23 (33)	N/A
Member of an association[a], n (%)	32 (46)	N/A
Not comfortable with online technologies[a], n (%)	35 (56)	N/A
Global health, n (%)		
No self-reported problems	29 (42)	6 (60)
Walking difficulty	13 (19)	2 (20)
Other self-reported health problems	27 (39)	2 (20)

[a]Data only available for participants seen at the final visit (n=69).

[b]N/A: not applicable.

Usage and Satisfaction Regarding SoBeezy Services

Of the 69 participants who completed the experimentation, 27 (39%) used the services proposed by the platform at least once (Table 2). In total, 335 services were provided, 132 (39.4%) thanks to the solidarity campaign (before the technological deployment) and 203 (60.6%) thanks to the SoBeezy platform. Assistance with shopping and transportation, and home visits, were the 3 most frequently used services (n=54, 16.1%; n=51, 15.2%; and n=51, 15.2%, respectively), followed by homework/gardening (n=24, 7.2%), mobile hairdressing services (n=10, 3%), and other services (n=13, 3.9%). We then analyzed the characteristics of the service users compared with those who never used them (Table 2). The 2 groups did not differ in terms of age, living alone, number of interactions with BeeVA, and level of comfort with technologies. However, the SoBeezy service users all lived in Pessac. They were significantly more often women (n=23, 85%, vs n=4, 15%, men, $P<.001$), received home care services more often (n=13, 48%, vs n=14, 52%, $P=.04$), and tended to suffer more often from health problems ($P=.18$). See Table 2 for details. Among the 27 service beneficiaries, the level of satisfaction was very high (between 88% and 100%) for the following items: conditions of requests, quality of services, and delay of answers. It should be noted

that 18 (72%) of 25 beneficiaries preferred using the telephone rather than BeeVA to request services.

Table 2. Description and comparison of participants' characteristics according to the number of services received (n=69).

Characteristics	Service usage, n (%)	
	At least once (n=27)	Never (n=42)
Age (years)		
62-81	14 (52)	22 (52)
81-95	13 (48)	20 (48)
Gender		
Man	4 (15)	12 (29)
Woman	23 (85)	30 (71)
Pilot city		
Pessac (urban area)	27 (100)	15 (36)
St Jean de Luz (urban area)	0	14 (33)
St Yrieix la Perche (rural area)	0	8 (19)
Langon (rural area)	0	5 (12)
Marital status		
In a relationship	8 (30)	12 (29)
Living alone	19 (70)	30 (71)
Recipient of home care services		
Yes	13 (48)	10 (24)
No	14 (52)	32 (76)
Member of an association		
Yes	10 (37)	22 (52)
No	17 (63)	20 (48)
Global health		
No self-reported problems	8 (30)	21 (50)
Walking difficulty	5 (18)	8 (19)
Other self-reported health problems	14 (52)	13 (31)
Interaction with BeeVA (times/month)		
0-4	19 (70)	25 (60)
>4	6 (22)	9 (21)
Missing	2 (8)	8 (19)
Comfort level with online technologies		
Comfortable	9 (33)	19 (45)
Not comfortable	15 (56)	20 (48)
Missing	3 (11)	3 (7)

Usage and Acceptability Regarding BeeVA

In the whole sample (N=109), the median of monthly interactions with BeeVA was 1.5 (IQR 0.69-3.5), and for the sample of participants who completed the final experimentation (n=69), the median of monthly interaction was 1.6 (IQR 0.7-3.8).

Throughout the study, the participants interacted with BeeVA up to 23.7 times per month, with a median of 1.6 times per month. Those who interacted at least 4 times per month were considered as the highest users. Globally, the characteristics of the participants did not differ according to the level of interaction with BeeVA (Table 3). Nevertheless, we observed that the highest users were significantly younger (n=15, 75%, vs n=5, 25%), more often members of an association (n=12, 60%, vs n=8, 40%), and more likely to have walking and health problems (n=10, 25%, and n=10, 50%, respectively, vs n=5, 25%). The participants less comfortable with technology tended

to interact more often with BeeVA than the others (n=13, 65%, vs n=30%).

Globally, 41 (63%) users had a positive opinion toward BeeVA, and the most positive dimensions of acceptability concerned usefulness (n=50, 77%), ease of use (n=53, 82%), and ergonomics/design (n=55, 85%). Reliability and usage behavior showed poorer results, with 49% (n=32) and 45% (n=29) positive opinions, respectively. Note that the usage behavior dimension only concerns the use of BeeVA options and does not include the use of services. Regarding items related to the acceptability of BeeVA, 48 (77%) of 62 users had a good opinion of voice usage to request services and options. In addition, even though not everyone used the services, two-thirds (n=44, 69%) of the participants considered BeeVA as a reassuring presence, with a slightly higher proportion among nonusers of the services (29/40, 73%, vs 15/24, 63%).

Table 3. Description and comparison of participants' characteristics according to monthly interaction with BeeVA (n=59[a]).

Characteristics	Monthly interaction with BeeVA[b]		
	Lowest users (1st tertile; [0-1.05[times/month; n=20), n (%)	Middle users (1st-2nd tertile; 1.5-2.8 times/month; n=19), n (%)	Highest users (>2nd tertile; >2.8 times/month; n=20), n (%)
Age (years)			
62-81	7 (35)	9 (47)	15 (75)
81-95	13 (65)	10 (53)	5 (25)
Gender			
Man	3 (15)	5 (26)	5 (25)
Woman	17 (85)	14 (74)	15 (75)
Pilot city			
Pessac (urban area)	14 (70)	11 (58)	14 (70)
St Jean de Luz (urban area)	3 (15)	6 (31)	4 (20)
St Yrieix la Perche (rural area)	3 (15)	2 (11)	2 (10)
Langon (rural area)	0	0	0
Marital status			
In a relationship	7 (35)	7 (37)	5 (25)
Living alone	13 (65)	12 (63)	15 (75)
Recipient of home care services			
Yes	4 (20)	9 (47)	6 (30)
No	16 (80)	10 (53)	14 (70)
Member of an association			
Yes	8 (40)	8 (42)	12 (60)
No	12 (60)	11 (58)	8 (40)
Global health			
No self-reported problems	10 (50)	10 (53)	5 (25)
Walking difficulty	2 (10)	3 (16)	5 (25)
Other self-reported health problems	8 (40)	6 (32)	10 (50)
Comfort level with online technologies			
Comfortable	11 (55)	8 (42)	6 (30)
Not comfortable	8 (40)	7 (37)	13 (65)
Missing	1 (5)	4 (21)	1 (5)

[a]Data on 10 participants were missing.

[b]Percentages can add up to more than 100 because of rounding.

To highlight the influence of users' characteristics on acceptability dimensions, each dimension was described by age, gender, living status, city, global health, level of comfort with technology, and statistics of monthly interactions with BeeVA and services (Table 4). The main differences in age and gender mainly concerned usefulness and design. Men perceived BeeVA as more useful and were more positive about the design than women (n=13, 26%, vs n=37, 74%, and n=14, 25%, vs n=41,

75%, respectively). The youngest users were less positive regarding the design than the oldest ones (n=26, 47%, vs n=29, 53%) but reported more frequently the usefulness of BeeVA (n=28, 56%, vs n=22, 44%). Participants living alone used BeeVA more frequently than people living with someone (n=24, 53%, used all or almost all options vs n=5, 25%, of the others). The participants who were less comfortable with technology used BeeVA more often and considered it reliable more frequently than people who were more comfortable. However, no differences were observed in usefulness (n=27, 82%, vs n=20, 74%). Among users who benefitted from the services and answered the acceptability questionnaire (n=25, 39%), there were no differences in perceived usefulness, ease of use, and ergonomics/design. Nevertheless, the service users used BeeVA more often (n=13, 45%, vs n=16, 55%) and reported reliability problems more frequently (n=18, 78%, vs n=5, 22%) than their counterparts (n=40, 62%) who never used the services but answered the acceptability questionnaire.

Table 4. Description of participants' characteristics according to BeeVA acceptability dimensions (n=65[a]).[b]

Characteristics	Perceived usefulness, n (%)		Option usage behavior, n (%)			Ease of use, n (%)		Reliability, n (%)			Design and ergonomics, n (%)	
	Useful (n=50)	Mixed opinion/not useful (n=15)	0 (n=18)	1-2 (n=18)	3-4 (n=29)	Easy (n=53)	Mixed opinion/not easy (n=12)	Reliable (n=32)	Unreliable (n=23)	No opinion (n=10)	Positive (n=55)	Mixed opinion/negative (n=10)
Age (years)												
62-81	28 (56)	6 (40)	8 (44)	11 (61)	15 (52)	27 (51)	7 (58)	19 (59)	11 (48)	4 (40)	26 (47)	8 (80)
81-95	22 (44)	9 (60)	10 (56)	7 (39)	14 (48)	26 (49)	5 (42)	13 (41)	12 (52)	6 (60)	29 (53)	2 (20)
Gender												
Man	13 (26)	2 (13)	3 (17)	4 (22)	8 (28)	12 (23)	3 (25)	8 (25)	6 (26)	1 (10)	14 (25)	1 (10)
Woman	37 (74)	13 (87)	15 (83)	14 (78)	21 (72)	41 (77)	9 (75)	24 (75)	17 (74)	9 (90)	41 (75)	9 (90)
Pilot city												
Pessac (urban area)	32 (64)	8 (53)	10 (56)	9 (50)	21 (72)	32 (60)	8 (67)	21 (66)	11 (48)	8 (80)	34 (62)	6 (60)
St Jean de Luz (urban area)	8 (16)	5 (33)	3 (17)	8 (44)	2 (7)	10 (19)	3 (25)	4 (13)	8 (35)	1 (10)	10 (18)	3 (30)
St Yrieix la Perche (rural area)	7 (14)	1 (7)	4 (22)	1 (6)	3 (10)	7 (13)	1 (8)	4 (13)	3 (13)	1 (10)	7 (13)	1 (10)
Langon (rural area)	3 (6)	1 (7)	1 (6)	0	3 (10)	4 (8)	0	3 (9)	1 (4)	0	4 (7)	0
Living status												
Not living alone	15 (30)	5 (33)	6 (33)	9 (50)	5 (17)	16 (30)	4 (33)	10 (31)	10 (43)	0	19 (35)	1 (10)
Living alone	35 (70)	10 (67)	12 (67)	9 (50)	24 (83)	37 (70)	8 (67)	22 (69)	13 (57)	10 (100)	36 (65)	9 (90)
Comfort level with online technologies												
Very comfortable	20 (40)	7 (47)	6 (33)	10 (56)	11 (38)	21 (40)	6 (50)	12 (38)	11 (48)	4 (40)	23 (42)	4 (40)
Not comfortable	27 (54)	6 (40)	9 (50)	7 (39)	17 (59)	28 (53)	5 (42)	19 (59)	8 (35)	6 (60)	28 (51)	5 (50)
Missing	3 (6)	2 (13)	3 (17)	1 (6)	1 (3)	4 (8)	1 (8)	1 (3)	4 (17)	0	4 (7)	1 (10)
Number of services received												
At least 1	20 (40)	5 (33)	7 (39)	5 (28)	13 (45)	20 (38)	5 (42)	16 (50)	18 (78)	6 (60)	20 (36)	5 (50)
0	30 (60)	10 (67)	11 (61)	13 (72)	16 (55)	33 (62)	7 (58)	16 (50)	5 (22)	4 (40)	35 (64)	5 (50)
Global health												
No self-reported problems	19 (38)	10 (67)	8 (44)	6 (33)	15 (52)	23 (43)	6 (50)	14 (44)	12 (52)	3 (30)	26 (47)	3 (30)
Walking difficulty	10 (20)	1 (7)	4 (22)	4 (22)	3 (10)	10 (19)	1 (8)	7 (22)	3 (13)	1 (10)	9 (16)	2 (20)
Other self-reported health problems	21 (42)	4 (27)	6 (33)	8 (44)	11 (38)	20 (38)	5 (42)	11 (34)	8 (35)	6 (60)	20 (36)	5 (50)
Monthly interaction with BeeVA (times)												

Characteristics	Perceived usefulness, n (%)		Option usage behavior, n (%)			Ease of use, n (%)		Reliability, n (%)			Design and ergonomics, n (%)	
	Useful (n=50)	Mixed opinion/not useful (n=15)	0 (n=18)	1-2 (n=18)	3-4 (n=29)	Easy (n=53)	Mixed opinion/not easy (n=12)	Reliable (n=32)	Unreliable (n=23)	No opinion (n=10)	Positive (n=55)	Mixed opinion/negative (n=10)
0-4	30 (60)	13 (87)	15 (83)	11 (61)	17 (59)	33 (62)	10 (83)	18 (56)	19 (83)	6 (60)	36 (65)	7 (70)
>4	14 (28)	0	0	6 (33)	8 (28)	12 (23)	2 (17)	9 (28)	3 (13)	2 (20)	12 (22)	2 (20)
Missing	6 (12)	2 (13)	3 (17)	1 (6)	4 (14)	8 (15)	0	5 (16)	1 (4)	2 (20)	7 (13)	1 (10)

[a]Data on 4 participants were missing.

[b]Percentages can add up to more or less than 100 because of rounding.

Description of the Global Assessment of the Intervention

Of 65 participants, 14 (22%) had a positive opinion, 36 (55%) had a mixed opinion, and 15 (23%) had a negative one (Table 5).

Participants having a positive opinion were more often younger, were women, lived more often as couples, had poorer health (n=9, 64%, positive vs n=2, 14%, among those in good health), and were significantly more often users of the SoBeezy services (n=11, 79%, of them were positive vs only n=3, 21%, of the nonusers of the services, *P*=.002). However, the participants who found BeeVA adapted to older adults seemed to have a better opinion of the intervention (n=12, 86%, vs n=2, 14%). According to the experimental site, we observed that the participants of St Jean de Luz had the least positive opinion (n=6, 40%, were negative vs n=7, 47%, in Pessac and n=1, 7%, in St Yrieix la Perche). Finally, one-fourth of those who found the intervention useful had a positive perception of the intervention (vs none in the comparative group, *P*<.001).

Table 5. Description and comparison of participants' characteristics according to the overall perception of the intervention (n=65[a]).[b]

Characteristics	Positive (n=14), n (%)	Mixed (n=36), n (%)	Negative (n=15), n (%)
Age (years)			
62-81	11 (79)	17 (47)	6 (40)
81-95	3 (21)	19 (53)	9 (60)
Gender			
Man	2 (14)	11 (31)	2 (13)
Woman	12 (86)	25 (69)	13 (87)
Pilot city			
Pessac (urban area)	13 (93)	20 (56)	7 (47)
St Jean de Luz (urban area)	0	7 (19)	6 (40)
St Yrieix la Perche (rural area)	0	7 (19)	1 (7)
Langon (rural area)	1 (7)	2 (6)	1 (7)
Marital status			
In a relationship	6 (43)	10 (28)	4 (27)
Living alone	8 (57)	26 (72)	11 (73)
Global health			
No self-reported problems	2 (14)	19 (53)	8 (53)
Walking difficulty	3 (21)	6 (17)	2 (13)
Other self-reported health problems	9 (64)	11 (31)	5 (33)
Comfort level with online technology			
Comfortable	5 (36)	16 (44)	6 (40)
Not comfortable	8 (57)	18 (50)	7 (47)
Missing	1 (7)	2 (6)	2 (13)
Monthly interaction with BeeVA			
0-4	9 (64)	21 (58)	13 (87)
>4	4 (29)	9 (25)	1 (7)
Missing	1 (7)	6 (17)	1 (7)
Service usage			
At least once	11 (79)	11 (31)	3 (20)
Never	3 (21)	25 (69)	12 (80)
BeeVA was a presence			
Yes	9 (64)	27 (75)	8 (53)
No/no opinion	4 (29)	9 (25)	7 (47)
Missing	1 (7)		
BeeVA was adapted to older people			
Yes	12 (86)	25 (69)	7 (47)
No/no opinion	2 (14)	11 (31)	8 (53)
Perceived usefulness			
Useful	14 (100)	32 (89)	4 (27)
Mixed opinion/not useful	0	4 (11)	11 (73)
Reliability			
Reliable	10 (71)	18 (50)	4 (27)
Unreliable	2 (14)	14 (39)	7 (47)

Characteristics	Positive (n=14), n (%)	Mixed (n=36), n (%)	Negative (n=15), n (%)
No opinion	2 (14)	4 (11)	4 (27)
Usage behavior (times)			
0	4 (29)	6 (17)	8 (53)
1-2	3 (21)	11 (31)	4 (27)
3-4	7 (50)	19 (53)	3 (20)
Perceived ease of use			
Easy	11 (79)	31 (86)	11 (73)
Mixed opinion/not easy	3 (21)	5 (14)	4 (27)
Design and ergonomics			
Positive	12 (86)	31 (86)	12 (80)
Mixed/negative opinion	2 (14)	5 (14)	3 (20)

[a]Data on 4 participants were missing.

[b]Percentages can add up to more or less than 100 because of rounding.

Improvement Tracks

The 5 most popular options to be integrated into the future BeeVA were easy access to trusted professionals (50/63, 79%), communication about city events (42/65, 65%), late-night pharmacy (42/65, 65%), activity propositions tailored to their needs (40/65, 62%), and videoconferencing option (37/65, 57%); see Multimedia Appendix 2.

Regarding activity sharing, 28 (44%) of 63 participants were interested. The users expressed more interest in consulting the propositions (n=30, 48%) than to themselves propose an activity to the community (n=19, 30%). Among the users who expressed their motivations for shared activities (n=39, 57%), the 3 most frequent reasons were to meet people (n=29, 74%), to find a pastime (n=27, 69%), and to share a hobby with others (n=27, 69%). Sharing leisure activities interested 30 persons (81% of the sample), followed by physical activities (n=25, 68%) and cultural/touristic activities (n=24, 65%). In total, 29 (73%) of 40 participants were interested in using such an activity-sharing tool (Multimedia Appendix 3).

Feedback From the SoBeezy Operational Team

First, the team reported real satisfaction and gratitude from the older participants who were supported by SoBeezy throughout this health crisis period. This particular sanitary context clearly hindered the deployment of the program (impossibility to propose the activity-sharing component, yet particularly expected by the users) and its functioning. In this context, the team raised the difficulty of relying on volunteers for good functioning of the platform (insufficient number and lack of reactivity when solicited for help). Regarding the users, the team emphasized the resistance to change (“I’ve always used my paper calendar on my fridge, I will not change my functioning,” “I have my own radio, I don’t need a new one”) and the inflexibility and intransigency of some (“I want my shopping at 2:00 p.m., 6:00 is too late”) and mentioned the individual barriers to using BeeVA (older age, depression, cognitive impairment, poor health, and severe reluctance to technology). The team also insisted on the fact that in this generation, it appeared difficult to ask for help (“I don't want to bother anyone about this, I'll manage it as I can,” “I've always managed my life by myself, I don't want to rely on someone else”). Nevertheless, the participants usually accepted the assistance when it was proposed by the team.

Moreover, most participants succeeded in using BeeVA after a minimal training program, but most of them also called the SoBeezy team over the phone for confirmation, which induced an unplanned workload for the staff. The team identified the main obstacle regarding technological aspects: the lack of reliability of the internet connection and the hardware (with a series of failures). This issue seriously disturbed the users, especially as they appeared rapidly overtaken in dealing with technological problems, even mild ones (eg, switching on the device or restarting it). In addition, the team underlined that the first version of BeeVA proposed too many features and options on its home screen, which reduced the readability of the services proposed. A clearer and simplified version was developed and quickly replaced the first one, with a substantial increase in the comfort of use expressed by the users. In addition, an avatar was also added to the home screen (Figure 2) and was appreciated. Finally, the team highlighted the importance of an efficient network of local partners, with a central role of the municipality (for identifying persons to equip and local partners).

Discussion

Principal Findings

The SoBeezy program is an innovative intervention aiming at promoting and fostering healthy aging by proposing to older adults concrete solutions to face daily life, tackle loneliness, promote social participation, and reduce the digital divide, thanks to a specific voice assistant. The experimentation, conducted in real-life conditions among older adults living in the community, was launched during the COVID-19 crisis in 4 different sites for 12 months.

In total, 109 older persons were equipped with BeeVA to use the SoBeezy platform. Among them, 32 left the experimentation before its end. The 3 main reasons for discontinued technology use were concordant to the literature [43,44], with technical problems, health problems, and no need for services. In total, 335 interventions were conducted, and almost 40% of the participants requested services. Nevertheless, three-quarters of the whole sample considered BeeVA a reassuring presence, and few participants had a negative opinion about the program (15 of 69). Among the users, BeeVA appeared easy to use (82%) and useful (77%) for older participants. Alleviating social isolation and loneliness was an important goal of the program. However, the pandemic context did not allow experimentation with the main way to tackle loneliness (ie, activity sharing). Our conclusions on this issue are consequently more limited than expected, and further research is needed. Nevertheless, this experimentation (particularly thanks to feedback from the field) confirmed that loneliness is a complex status, often associated with difficult life paths, particular personality trait, isolation, depression, and poor health [13]. Fighting loneliness in the older population requires time and human resources to establish a relationship of trust to allow a person to recover from settled loneliness. Technology alone scarcely appears to be a solution in this vulnerable population. Therefore, we think that such a program (the initial one including sharing activities) could be more relevant to prevent the occurrence of loneliness among older adults at greater risk than to “treat” loneliness when settled. These results could suggest that such devices and services could be useful to deploy among older adults, particularly in persons in the digital divide, with mobility restrictions, limitations in activities of daily living, a small social network size, geographically distant relatives, or living in rural areas [16,45-47].

This study faced obstacles related to the targeted population and technological aspects. Indeed, the appropriation of a device requires new ways to perform some activities and to change one’s habits (eg, vocal communication, online order, listening to the radio), which is known to become more difficult with aging [48-51]. Second, to ensure the follow-up of their requests, the users progressively tended to use the telephone (not intended for this use), a well-established and reassuring habit. We thus faced a need for an immediate response to their request [52]. As previously reported in another study [21], this behavior is consistent with the preference of older adults to use synchronous communications (eg, phone calls and instant messaging). The SoBeezy team also faced the difficulty for an older adult to verbalize the need for help, which can be explained by the fear of disturbing or by the refusal to rely on someone else to perform activities of daily living that they have always done by themselves. For some people, requesting a service can be seen as a marker of old age, inducing the vision of vulnerable older people [53].

Nevertheless, the participants who used the services provided by the SoBeezy program were satisfied with the services' quality and the interactions with the contributors. Another lesson learned from this experimentation was the technological intransigence of the older users. Indeed, there is a common belief that technology must be doing better than other traditional existing things, otherwise technology could not be perceived as useful, nor easy to use [39]. Each technological incident was consequently difficult to accept by the users, negatively impacting usage behavior, acceptance, and possible long-term adoption [39,54]. Regarding specifically BeeVA, it gained rather positive acceptability, but it was penalized by reliability issues, such as instability of the internet connection, which strongly disrupted interactions with the voice assistant [55]. In addition, despite our recommendations, many participants switched off the voice assistant instead of leaving it on standby, with some consequences on the functioning of the devices (BeeVA and 4G Wi-Fi). We also faced a series of simultaneous unexplained device failures that required replacing the devices.

Our experimentation also emphasized interesting levers for actions for the next deployment and future research. First, the perception of ease of use, an ergonomic design, and good reliability appear to be facilitators of good acceptability of BeeVA, assuming that individual step-by-step training is conducted. A simple user manual and a hotline phone number are provided (with the risks of drift, as previously mentioned) [31]. Among the technological levers for action, we observed the importance of a user-centered approach, which is essential to understand users' needs in terms of technological skills and psychological characteristics (apprehension, reluctance, technophobia). A simplification of the steps to achieve the expected results is also essential for device appropriation, since older adults can be interested in technological devices. However, they can be discouraged when sophisticated computerized devices replace simpler ones, which are easier to use for them [40,56]. Moreover, the social environment may also play an important role in accepting the technology; the relatives could play the role of a mediator with a positive social pressure for the appropriation of new technology [21,46,53,57] and compensate for the prior lack of digital literacy [21]. As suggested by our study, the participants living with their partners tended to be more positive about the SoBeezy intervention and found BeeVA to be more reliable than people living alone. We also clearly identified important differences between the pilot sites, the program being more efficient when the local partnership was highly involved in the program, with a direct impact on the satisfaction of the users (46% had a negative opinion about the program in one site vs 13% in another). This result underlines the importance of solid partnerships with local actors, particularly for the diffusion and appropriation of the technology, the identification of available resources in a territory, and the identification of the “invisible” of the society (ie, isolated persons, often far from social and medical care systems, despite greater health or social problems) [58]. Taken together, our results converge with the conceptual framework developed by the Center for Research and Education on Aging and Technology Enhancement (CREATE). This framework highlights the importance of considering factors related to the microscopic level (eg, the capacity of the person, the services, and the interaction with the technology) and those related to the macroscopic level (eg, the social environment and support of the person) [45].

Finally, tackling loneliness is 1 of the main objectives of the SoBeezy program. Unfortunately, its deployment has been

greatly affected by the pandemic context, and we could not particularly experiment with the main component targeting loneliness (ie, the activity-sharing component). However, it is also well known that isolated people and those suffering from loneliness are more likely to refuse help and assistance, particularly when affected by the digital divide [34]. Nevertheless, the PRISM platform [31,32] reported a significant decrease in the feeling of loneliness and an increase in perceived social support and well-being among participants living alone and a good appropriation of the technology.

Strengths and Limitations

In experimenting on technologies for older people, the number of participants equipped with BeeVA (more than 100 older adults, including those in the digital divide) represents 1 of the strengths of the study. Second, the SoBeezy services, being free of charge, guaranteed wide and egalitarian access [59]. Moreover, the participants could choose the services that seemed appropriate to them at their convenience and according to their own needs. This choice ensured a good level of agentivity and self-determination in the appropriation of the SoBeezy platform [60]. Our results also showed the importance of the user-centered approach, which allows adapting a device and contributing to better learning and appropriation by the aging population [37,56].

However, our study also has some limitations. For even easier use of the device, natural language understanding may be improved using artificial intelligence methods. Due to the sanitary context, we could not experiment the sharing-activity component, whereas it was the first objective of this program to combat social isolation and loneliness. Moreover, the evaluation of the program was limited by a potential selection bias; the final assessment was not available for the participants who left the experimentation early—yet more likely to be unfavorable to the intervention. It is important to emphasize that our sample size was probably adequate to identify large significant effects but insufficient to detect medium or small effects. Therefore, we only reported the *P* values of significant relevant effects. Regarding the usage frequency of BeeVA, we did not have access to detailed individual data, such as time of usage of an option or an app, or accidental clicks, that represents a limitation of the data on the interactions with the device. We did not assess perceived loneliness but only the living status (living alone vs not living alone). Finally, the interviewers reported a risk of social desirability bias, the participants being particularly grateful for the support provided during the COVID-19 crisis.

Perspectives: Guidelines for Improved Deployment of SoBeezy

Several improvements were conducted during the experimentation. First, the final interviews allowed the older users to suggest new services or options that could interest them, such as easy access to trusted professionals or communication about city events. These services should be proposed in the next deployment of the SoBeezy program.

Second, our results also suggested the potential benefits of a close network of older users to improve the confidence that users have in services and activity sharing (see Multimedia Appendix 3 presenting the potential areas of improvement). Therefore, we suggest relying more on older adults already in a community, such as the seniors' clubs or independent living housing [36].

Regarding the personalization of the device, it could be interesting to propose to BeeVA users a mixed mode (voice and touch) and several versions of SoBeezy, for smartphones, tablets, or computers (ie, responsive design), and give them a choice to fit their preferences and habits.

Regarding the training phase, in addition to written instructions (hard-copy format), it could be interesting to provide interactive and personalized training sessions (eg, through games and videos of uses adapted to the level of the participants). These sessions would allow older people to better accept the technological solutions and help both the acceptance of asynchronous communications and a better use of the functionalities available on the device [31]. These individual and collective training sessions could not be implemented due to the COVID-19 context.

Finally, from a technical point of view, it could be judicious to record, for example, interaction errors under the use of some services to improve device reliability. For service providers on the platform, it could be also useful to give them feedback on the usage (number of users, frequency and type of use) to improve the proposed services.

Conclusion

This multisite study conducted in real-life conditions with more than 100 older adults living in the community provides enlightening results for the reality on the ground, specifically when we are interested in digital tools designed for the aging population. The context of the COVID-19 epidemic was, on the one hand, favorable in the light of the massive needs of the older adults during this crisis but, on the other hand, particularly limiting due to the sanitary measures that clearly affected the program. The experimentation overall showed a positive acceptability of the voice assistant (ie, perceived usefulness and ease of use) and a high level of satisfaction of the participants who used SoBeezy. Nevertheless, our findings also highlighted the issues of resistance to change, difficulties for the users in asking for help, and difficulties met to efficiently tackle chronic settled loneliness using ICT. The SoBeezy program could be a way to improve the autonomy and well-being of older adults and their families. The next step will be the experimentation with the activity-sharing component that could not be tested due to the COVID-19 context.

Acknowledgments

This work was supported by the Agence Régionale de Santé de Nouvelle Aquitaine, the Région Nouvelle Aquitaine, the Conférence des Financeurs de Gironde, des Pyrénées Atlantiques et de Haute-Vienne, 3 local municipalities (Pessac, St Jean de Luz, and St Yrieix la Perche), the Fondation John Bost, the AG2R La Mondiale, the Carsat Nouvelle-Aquitaine, the Fonds de dotation de la Caisse d'Epargne, the Laboratoire Bristol-Myers Squibb, and the Association Perfect Exploration CardioVasculaire.

Data Availability

The data sets generated and analyzed during the study are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Derivation of the variable measuring BeeVA acceptability.
[PNG File , 331 KB - formative_v6i11e39185_app1.png]

Multimedia Appendix 2
Distribution of opinions on possible functions to be integrated into BeeVA among participants surveyed.
[PNG File , 96 KB - formative_v6i11e39185_app2.png]

Multimedia Appendix 3
Repartition of the areas for improvement and reorientation of BeeVA, according to overall BeeVA acceptability.
[PDF File (Adobe PDF File), 502 KB - formative_v6i11e39185_app3.pdf]

References

1. Blanpain N, Buisson G. Projections de population à l'horizon 2070 - Insee Première - 1619. 2016. URL: https://www.insee.fr/fr/statistiques/2496228 [accessed 2022-03-14]
2. Pérès K, Zamudio-Rodriguez A, Dartigues J, Amieva H, Lafitte S. Prospective pragmatic quasi-experimental study to assess the impact and effectiveness of an innovative large-scale public health intervention to foster healthy ageing in place: the SoBeezy program protocol. BMJ Open 2021 Apr 29;11(4):e043082 [FREE Full text] [doi: 10.1136/bmjopen-2020-043082] [Medline: 33926977]
3. Pilotto A, Boi R, Petermans J. Technology in geriatrics. Age Ageing 2018 Nov 01;47(6):771-774. [doi: 10.1093/ageing/afy026] [Medline: 29546366]
4. O'Brien K, Liggett A, Ramirez-Zohfeld V, Sunkara P, Lindquist LA. Voice-controlled intelligent personal assistants to support aging in place. J Am Geriatr Soc 2020 Jan 16;68(1):176-179. [doi: 10.1111/jgs.16217] [Medline: 31617581]
5. Monahan PO, Kroenke K, Callahan CM, Bakas T, Harrawood A, Lofton P, et al. Reliability and validity of SymTrak, a multi-domain tool for monitoring symptoms of older adults with multiple chronic conditions. J Gen Intern Med 2019 Jun 18;34(6):908-914 [FREE Full text] [doi: 10.1007/s11606-018-4780-4] [Medline: 31001721]
6. Amini R, Chee KH, Mendieta M, Parker S. Online engagement and cognitive function among older adults. Geriatr Gerontol Int 2019 Sep 31;19(9):918-923. [doi: 10.1111/ggi.13749] [Medline: 31368165]
7. Pérès K, Verret C, Alioum A, Barberger-Gateau P. The disablement process: factors associated with progression of disability and recovery in French elderly people. Disabil Rehabil 2005 Mar 04;27(5):263-276. [doi: 10.1080/09638280400006515] [Medline: 16025753]
8. Sayin Kasar K, Karaman E. Life in lockdown: social isolation, loneliness and quality of life in the elderly during the COVID-19 pandemic: A scoping review. Geriatr Nurs 2021 Sep;42(5):1222-1229 [FREE Full text] [doi: 10.1016/j.gerinurse.2021.03.010] [Medline: 33824008]
9. van Tilburg TG, Steinmetz S, Stolte E, van der Roest H, de Vries DH. Loneliness and mental health during the COVID-19 pandemic: a study among Dutch older adults. J Gerontol B Psychol Sci Soc Sci 2021 Aug 13;76(7):e249-e255 [FREE Full text] [doi: 10.1093/geronb/gbaa111] [Medline: 32756931]
10. Stolz E, Mayerl H, Freidl W. The impact of COVID-19 restriction measures on loneliness among older adults in Austria. Eur J Public Health 2021 Feb 01;31(1):44-49 [FREE Full text] [doi: 10.1093/eurpub/ckaa238] [Medline: 33338225]
11. Macdonald B, Hülür G. Well-being and loneliness in Swiss older adults during the COVID-19 pandemic: the role of social relationships. Gerontologist 2021 Feb 23;61(2):240-250 [FREE Full text] [doi: 10.1093/geront/gnaa194] [Medline: 33258898]
12. Ibrahim R, Abolfathi Momtaz Y, Hamid TA. Social isolation in older Malaysians: prevalence and risk factors. Psychogeriatrics 2013 Jun 12;13(2):71-79 [FREE Full text] [doi: 10.1111/psyg.12000] [Medline: 23909963]

13. Grenade L, Boldy D. Social isolation and loneliness among older people: issues and future challenges in community and residential settings. Aust Health Rev 2008 Aug;32(3):468-478. [doi: 10.1071/ah080468] [Medline: 18666874]
14. Victor CR, Scambler SJ, Bowling A, Bond J. The prevalence of, and risk factors for, loneliness in later life: a survey of older people in Great Britain. Ageing Soc 2005 Apr 22;25(6):357-375. [doi: 10.1017/s0144686x04003332]
15. O Súilleabháin PS, Gallagher S, Steptoe A. Loneliness, living alone, and all-cause mortality: the role of emotional and social loneliness in the elderly during 19 years of follow-up. Psychosom Med 2019 May 14;81(6):521-526 [FREE Full text] [doi: 10.1097/PSY.0000000000000710] [Medline: 31094903]
16. Fakoya OA, McCorry NK, Donnelly M. Loneliness and social isolation interventions for older adults: a scoping review of reviews. BMC Public Health 2020 Feb 14;20(1):129 [FREE Full text] [doi: 10.1186/s12889-020-8251-6] [Medline: 32054474]
17. Landeiro F, Barrows P, Nuttall Musson E, Gray AM, Leal J. Reducing social isolation and loneliness in older people: a systematic review protocol. BMJ Open 2017 May 17;7(5):e013778 [FREE Full text] [doi: 10.1136/bmjopen-2016-013778] [Medline: 28515187]
18. Luanaigh CO, Lawlor BA. Loneliness and the health of older people. Int J Geriatr Psychiatry 2008 Dec;23(12):1213-1221. [doi: 10.1002/gps.2054] [Medline: 18537197]
19. Tabue Teguo M, Simo-Tabue N, Stoykova R, Meillon C, Cogne M, Amiéva H, et al. Feelings of loneliness and living alone as predictors of mortality in the elderly: the PAQUID study. Psychosom Med 2016;78(8):904-909. [doi: 10.1097/psy.0000000000000386]
20. van Bel DT, Smolders KCHJ, IJsselsteijn WA, de Kort Y. Social connectedness: concept and measurement. In: Callaghan V, Kameas A, Reyes A, Royo D, Weber M, editors. Proceedings of the 5th International Conference on Intelligent Environments. Amsterdam, the Netherlands: IOS Press; 2009:67-74.
21. Barbosa Neves B, Franz R, Judges R, Beermann C, Baecker R. Can digital technology enhance social connectedness among older adults? A feasibility study. J Appl Gerontol 2019 Jan 22;38(1):49-72. [doi: 10.1177/0733464817741369] [Medline: 29166818]
22. Hemberg J, Santamäki Fischer R. A window toward the world: older adults' experiences of becoming in health and developing as human beings through interacting with others using real video communication. Holist Nurs Pract 2018;32(2):90-97. [doi: 10.1097/hnp.0000000000000254]
23. Coelho J, Duarte C. A literature survey on older adults' use of social network services and social applications. Comput Hum Behav 2016 May;58:187-205. [doi: 10.1016/j.chb.2015.12.053]
24. Koene P, Köbler F, Esch S, Leimeister J, Krcmar H. Design and evaluation of a service-oriented collaborative consumption platform for the elderly. 2012 Presented at: CHI EA '12: CHI '12 Extended Abstracts on Human Factors in Computing Systems; May 2012; Austin, TX p. 2537-2542. [doi: 10.1145/2212776.2223832]
25. Boyd D, Ellison N. Social network sites: definition, history, and scholarship. J Comput-Mediat Commun 2007;13(1):210-230. [doi: 10.1111/j.1083-6101.2007.00393.x]
26. Burke M, Kraut R, Marlow C. Social capital on facebook: differentiating uses and users. 2011 Presented at: CHI '11: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; May 2011; Vancouver, BC p. 571-580. [doi: 10.1145/1978942.1979023]
27. Brunette K, Eisenstadt M, Pukinskis E, Ryan W. Meeteetse: social well-being through place attachment. 2005 Presented at: CHI EA '05: CHI '05 Extended Abstracts on Human Factors in Computing Systems; April 2005; Portland, OR p. 2065-2069. [doi: 10.1145/1056808.1057100]
28. Romero N, Markopoulos P, van Baren J, de Ruyter B, IJsselsteijn W, Farshchian B. Connecting the family with awareness systems. Pers Ubiquit Comput 2006 Aug 23;11(4):299-312. [doi: 10.1007/s00779-006-0089-0]
29. Lindley S, Harper R, Sellen A. Desiring to be in touch in a changing communications landscape: attitudes of older adults. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI ?09. Association for Computing Machinery; 2009 Presented at: CHI '09: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; April 4-9, 2009; Boston, MA p. 1693-1702. [doi: 10.1145/1518701.1518962]
30. Choi HK, Lee SH. Trends and effectiveness of ICT interventions for the elderly to reduce loneliness: a systematic review. Healthcare (Basel) 2021 Mar 07;9(3):293 [FREE Full text] [doi: 10.3390/healthcare9030293] [Medline: 33800099]
31. Czaja SJ, Boot WR, Charness N, A Rogers W, Sharit J, Fisk AD, et al. The personalized reminder information and social management system (PRISM) trial: rationale, methods and baseline characteristics. Contemp Clin Trials 2015 Jan;40:35-46 [FREE Full text] [doi: 10.1016/j.cct.2014.11.004] [Medline: 25460342]
32. Czaja SJ, Boot WR, Charness N, Rogers WA, Sharit J. Improving social support for older adults through technology: findings from the PRISM randomized controlled trial. Gerontologist 2018 May 08;58(3):467-477 [FREE Full text] [doi: 10.1093/geront/gnw249] [Medline: 28201730]
33. Gustafson DH, McTavish F, Gustafson DH, Mahoney JE, Johnson RA, Lee JD, et al. The effect of an information and communication technology (ICT) on older adults' quality of life: study protocol for a randomized control trial. Trials 2015 Apr 25;16(1):191. [doi: 10.1186/S13063-015-0713-2]

34. Jarvis MA, Padmanabhanunni A, Chipps J. An evaluation of a low-intensity cognitive behavioral therapy mHealth-supported intervention to reduce loneliness in older people. Int J Environ Res Public Health 2019 Apr 11;16(7):1305 [FREE Full text] [doi: 10.3390/ijerph16071305] [Medline: 30979042]
35. Lara J, O'Brien N, Godfrey A, Heaven B, Evans EH, Lloyd S, et al. Pilot randomised controlled trial of a web-based intervention to promote healthy eating, physical activity and meaningful social connections compared with usual care control in people of retirement age recruited from workplaces. PLoS One 2016 Jul 29;11(7):e0159703 [FREE Full text] [doi: 10.1371/journal.pone.0159703] [Medline: 27472560]
36. Boll F, Brune P. Online support for the elderly – why service and social network platforms should be integrated. Procedia Comput Sci 2016;98:395-400. [doi: 10.1016/j.procs.2016.09.060]
37. Czaja S, Boot W, Charness N, Rogers W. Designing for Older Adults: Principles and Creative Human Factors Approaches, Third Edition. Boca Raton, FL: CRC Press; 2019.
38. Consel C, Kabac M. Internet of things: a challenge for software engineering. Smart Cities 2014;98:20-21.
39. Chen K, Chan AHS. Gerontechnology acceptance by elderly Hong Kong Chinese: a senior technology acceptance model (STAM). Ergonomics 2014 Mar 24;57(5):635-652. [doi: 10.1080/00140139.2014.895855] [Medline: 24655221]
40. Czaja SJ, Charness N, Fisk AD, Hertzog C, Nair SN, Rogers WA, et al. Factors predicting the use of technology: findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE). Psychol Aging 2006 Jun;21(2):333-352 [FREE Full text] [doi: 10.1037/0882-7974.21.2.333] [Medline: 16768579]
41. Bagnasco A, Hayter M, Rossi S, Zanini MP, Pellegrini R, Aleo G, et al. Experiences of participating in intergenerational interventions in older people's care settings: a systematic review and meta-synthesis of qualitative literature. J Adv Nurs 2020 Jan 06;76(1):22-33. [doi: 10.1111/jan.14214] [Medline: 31566788]
42. Murayama Y, Murayama H, Hasebe M, Yamaguchi J, Fujiwara Y. The impact of intergenerational programs on social capital in Japan: a randomized population-based cross-sectional study. BMC Public Health 2019 Feb 06;19(1):156 [FREE Full text] [doi: 10.1186/s12889-019-6480-3] [Medline: 30727981]
43. Berkowsky R, Rikard R, Cotten S. Signing off: predicting discontinued ICT usage among older adults in assisted and independent living: a survival analysis. In: Zhou J, Salvendy G, editors. Human Aspects of IT for the Aged Population. Design for Everyday Life. Vol 9194. Lecture Notes in Computer Science. New York, NY: Springer International; 2015:389-398.
44. Hawley-Hague H, Boulton E, Hall A, Pfeiffer K, Todd C. Older adults' perceptions of technologies aimed at falls prevention, detection or monitoring: a systematic review. Int J Med Inform 2014 Jun;83(6):416-426. [doi: 10.1016/j.ijmedinf.2014.03.002] [Medline: 24798946]
45. Czaja S, Moxley J, Rogers W. Social support, isolation, loneliness, and health among older adults in the PRISM randomized controlled trial. Front Psychol 2021;12:728658 [FREE Full text] [doi: 10.3389/fpsyg.2021.728658] [Medline: 34675843]
46. Courtney K, Demiris G, Rantz M, Skubic M. Needing smart home technologies: the perspectives of older adults in continuing care retirement communities. Inform Prim Care 2008 Nov 01;16(3):195-201 [FREE Full text] [doi: 10.14236/jhi.v16i3.694] [Medline: 19094406]
47. Peek ST, Wouters EJ, van Hoof J, Luijkx KG, Boeije HR, Vrijhoef HJ. Factors influencing acceptance of technology for aging in place: a systematic review. Int J Med Inform 2014 Apr;83(4):235-248 [FREE Full text] [doi: 10.1016/j.ijmedinf.2014.01.004] [Medline: 24529817]
48. Bouisson J. Etude des stratégies de vieillissement. Psychol Clin Proj 2003;9(1):457-474. [doi: 10.3917/pcp.009.0457]
49. Yusif S, Soar J, Hafeez-Baig A. Older people, assistive technologies, and the barriers to adoption: a systematic review. Int J Med Inform 2016 Oct;94:112-116. [doi: 10.1016/j.ijmedinf.2016.07.004] [Medline: 27573318]
50. Durndell A, Haag Z. Computer self efficacy, computer anxiety, attitudes towards the internet and reported experience with the internet, by gender, in an East European sample. Comput Hum Behav 2002 Sep;18(5):521-535. [doi: 10.1016/S0747-5632(02)00006-7]
51. Tsai T, Lin W, Chang Y, Chang P, Lee M. Technology anxiety and resistance to change behavioral study of a wearable cardiac warming system using an extended TAM for older adults. PLoS One 2020 Jan 13;15(1):e0227270 [FREE Full text] [doi: 10.1371/journal.pone.0227270] [Medline: 31929560]
52. White H, McConnell E, Clipp E, Branch LG, Sloane R, Pieper C, et al. A randomized controlled trial of the psychosocial impact of providing internet training and access to older adults. Aging Ment Health 2002 Aug;6(3):213-221. [doi: 10.1080/13607860220142422] [Medline: 12217089]
53. Bobillier Chaumon M, Oprea Ciobanu R. Les nouvelles technologies au service des personnes âgées : entre promesses et interrogations – une revue de questions. Psychol Française 2009 Sep;54(3):271-285. [doi: 10.1016/j.psfr.2009.07.001]
54. Coughlin J, D'Ambrosio L, Reimer B, Pratt M. Annu Int Conf IEEE Eng Med Biol Soc 2007;2007:1810-1815. [doi: 10.1109/iembs.2007.4352665]
55. Baig MM, Afifi S, GholamHosseini H, Mirza F. A systematic review of wearable sensors and IoT-based monitoring applications for older adults - a focus on ageing population and independent living. J Med Syst 2019 Jun 15;43(8):233. [doi: 10.1007/s10916-019-1365-7] [Medline: 31203472]
56. Wandke H, Sengpiel M, Sönksen M. Myths about older people's use of information and communication technology. Gerontology 2012;58(6):564-570 [FREE Full text] [doi: 10.1159/000339104] [Medline: 22739502]

57. Pech M, Sauzeon H, Yebda T, Benois-Pineau J, Amieva H. Falls detection and prevention systems in home care for older adults: myth or reality? JMIR Aging 2021 Dec 09;4(4):e29744 [FREE Full text] [doi: 10.2196/29744] [Medline: 34889755]
58. Gorenko JA, Moran C, Flynn M, Dobson K, Konnert C. Social isolation and psychological distress among older adults related to COVID-19: a narrative review of remotely-delivered interventions and recommendations. J Appl Gerontol 2021 Jan 11;40(1):3-13. [doi: 10.1177/0733464820958550] [Medline: 32914668]
59. Mihailidis A, Cockburn A, Longley C, Boger J. The acceptability of home monitoring technology among community-dwelling older adults and baby boomers. Assist Technol 2008 Mar 31;20(1):1-12. [doi: 10.1080/10400435.2008.10131927] [Medline: 18751575]
60. Dupuy L, Consel C, Sauzéon H. Self determination-based design to achieve acceptance of assisted living technologies for older adults. Comput Hum Behav 2016 Dec;65:508-521. [doi: 10.1016/j.chb.2016.07.042]

Abbreviations

ICT: information communication technology
PRISM: Personal Reminder Information and Social Management
SNS: social network services
SOS: social online services

Edited by A Mavragani; submitted 02.05.22; peer-reviewed by N Charness, X Du; comments to author 14.07.22; revised version received 08.09.22; accepted 21.10.22; published 24.11.22.

Please cite as:
Pech M, Gbessemehlan A, Dupuy L, Sauzéon H, Lafitte S, Bachelet P, Amieva H, Pérès K
Lessons Learned From the SoBeezy Program for Older Adults During the COVID-19 Pandemic: Experimentation and Evaluation
JMIR Form Res 2022;6(11):e39185
URL: https://formative.jmir.org/2022/11/e39185
doi:10.2196/39185
PMID:36355629

Original Paper

Identifying Personality Characteristics and Indicators of Psychological Well-Being Associated With Attrition in the Motivation Makes the Move! Physical Activity Intervention: Randomized Technology-Supported Trial

Kaisa Kaseva[1,2], PhD; Mari Tervaniemi[2,3], PhD; Enni Heikura[2], MSc; Kaisamari Kostilainen[3], MSc; Maritta Pöyhönen-Alho[4], PhD; J Kevin Shoemaker[5], PhD; Robert J Petrella[6,7], PhD; Juha E Peltonen[1,8], PhD

[1]Department of Sport and Exercise Medicine, Clinicum, Faculty of Medicine, University of Helsinki, Helsinki, Finland
[2]Cicero Learning, Faculty of Educational Sciences, University of Helsinki, Helsinki, Finland
[3]Cognitive Brain Research Unit, Department of Psychology and Logopedics, University of Helsinki, Helsinki, Finland
[4]Obstetrics and Gynecology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland
[5]Neurovascular Research Laboratory, School of Kinesiology, Western University, London, ON, Canada
[6]Division of Sport and Exercise Medicine, Department of Family Practice, University of British Columbia, Vancouver, BC, Canada
[7]Center for Studies in Family Medicine, Schulich School of Medicine and Dentistry, Western University, London, ON, Canada
[8]Clinic for Sports and Exercise Medicine, Foundation for Sports and Exercise Medicine, Helsinki, Finland

Corresponding Author:
Kaisa Kaseva, PhD
Department of Sport and Exercise Medicine, Clinicum
Faculty of Medicine
University of Helsinki
Alppikatu 2
Helsinki, 00530
Finland
Phone: 358 443077737
Email: kaisa.kaseva@helsinki.fi

Abstract

Background: Data attrition has been a common problem in longitudinal lifestyle interventions. The contributors to attrition in technology-supported physical activity interventions have not been thoroughly studied.

Objective: The present study examined the roles of personality characteristics and indicators of psychological well-being in data attrition within a technology-supported, longitudinal intervention study with overweight adults.

Methods: Participants (N=89) were adults from the Motivation Makes the Move! intervention study. Data attrition was studied after a 3-month follow-up. Participants' personality characteristics were studied using the Short Five self-report questionnaire. Psychological well-being indicators were assessed with the RAND 36-item health survey, Positive and Negative Affect Schedule, and Beck Depression Inventory. Logistic regression analyses were conducted to assess the risk of discontinuing the study. The analyses were adjusted for sex, age, study group, and educational status.

Results: At the 3-month follow-up, 65 of 89 participants (73% of the initial sample) had continued in the study. Participants' personality characteristics and indicators of psychological well-being were not associated with the risk of dropping out of the study (all *P* values >.05). The results remained the same after covariate controls.

Conclusions: Participant attrition was not attributable to personality characteristics or psychological well-being in the Motivation Makes the Move! study conducted with overweight adults. As attrition remains a challenge within longitudinal, technology-supported lifestyle interventions, attention should be paid to the potentially dynamic natures of personality and psychological well-being, as well as other elements beyond these.

Trial Registration: ClinicalTrials.gov NCT02686502; https://clinicaltrials.gov/ct2/show/NCT02686502

(JMIR Form Res 2022;6(11):e30285) doi:10.2196/30285

KEYWORDS

randomized trial; physical activity; lifestyles; personality; psychological well-being; study attrition; mental health; lifestyle interventions

Introduction

Cardiovascular diseases are the leading cause of premature death worldwide. Overweight and obesity have been recognized as among the most severe risk factors for these diseases. According to the World Health Organization, more than 1 billion people worldwide are overweight (BMI ≥ 25), and more than 300 million are obese (BMI ≥ 30) [1]. Overweight and obesity may also be associated with comorbidities, referring to a person's vulnerability to other illnesses and diseases. Comorbidity may also refer to interactions between illnesses that can worsen the course of both. For instance, due to the possibility of unprecedented outbreaks of infectious diseases (eg, COVID-19), it is of utmost importance to pay attention to the prevention of risk factors, such as obesity, that are likely to contribute to the development of comorbidities.

Along with genes, unhealthy lifestyles are among the most prominent contributors to overweight and subsequent negative health outcomes. Physical inactivity has been recognized as one of the primary contributors to overweight and obesity [2,3]. Recently, it has been suggested that even small improvements in physical fitness reduce the risk of coronary heart disease and that any amount of physical activity is beneficial for overall health [4]. Despite this knowledge, the prevalence of obesity has increased during the last 3 decades, and therefore, there is a need to pay attention to, as well as target research resources toward, physical activity in preventing and hindering the epidemic of obesity [5].

Recently, many physical activity and exercise promotion actions and interventions have been designed to prevent the development of obesity and other health risk factors. Some physical activity interventions aiming to increase healthy physical activity habits among overweight people have demonstrated favorable results, but equivocal results regarding expected outcomes also exist [6]. Recently, technological innovations have been regarded as promising tools to optimize the effectiveness of interventions by improving intervention delivery and adherence to participation [7,8]. In addition, eHealth and mobile health (mHealth) technologies allow collection of reliable, comprehensive, and diverse information about the users' physical activity and can reach large populations quickly [9-11]. Many technology-based health interventions have proved to be efficient in increasing physical activity, contributing to weight loss, and improving overall health [12,13].

Along with the creation of valid study designs, the effectiveness of research is related to participants' commitment to the completion of studies. Attrition in longitudinal studies and interventions refers to participants dropping out of the study before its completion or stopping use of the program or application during participation [14]. High attrition is a common issue in longitudinal physical activity interventions and a major challenge to the validity of the research [15,16]. In addition to a variety of individual factors, psychological qualities have been suggested to be important contributors to health behaviors, as well as to participation in and commitment to interventions [17,18]. Further assessment of these factors could improve the effectiveness of interventions [17].

Personality characteristics have been regarded as among the most essential contributors to behavioral choices [19]. Personality reflects individual differences in thinking, feeling, and behaving, and it matures through age. In the widely acknowledged 5-factor model of personality (ie, the "Big Five" model), personality characteristics are categorized into 5 broader continuums [20,21]. Individuals scoring high in neuroticism tend to experience negative emotions, such as anger, fear, and stress. In contrast, individuals high in extraversion tend to exhibit positive emotions, including cheerfulness and enthusiasm. Conscientious individuals have been characterized as diligent, organized, and disciplined, and display planned, rather than spontaneous, behaviors. Individuals high in openness to experience are open-minded and creative, whereas those high in agreeableness tend to exhibit kindness, cooperativeness, and positivity in interpersonal relations [20]. The psychobiological theory of personality has introduced a 7-factor model, including 4 dimensions of temperament (novelty seeking, harm avoidance, reward dependence, and persistence) that are regarded as individuals' inherited, reactive tendencies, and 3 dimensions of character (self-directedness, cooperativeness, and self-transcendence) that mature through age [22,23]. Both the 5- and 7-factor paradigms of personality have also been applied within recent physical activity interventions.

High levels of openness to experience, as well as hedonistic orientation, have been shown to relate favorably to attitudes toward health behavior interventions [24,25]. Reward dependence, which correlates highly with extraversion and agreeableness [26], has been shown to play a role in determining high success rates in therapy-induced weight management [27]. Persons who were successful at achieving weight loss in a weight-loss program had a lower novelty-seeking trait, which correlates positively with conscientiousness [23], than those who did not reach their goals [28]. Personality seems to also play a role in becoming interested in participating in technology-based interventions [29], but more evidence on personality's role in the commitment to such interventions is needed.

Along with personality, the level of psychological well-being, referring to a person's experience of their own well-being, is associated with the commitment to a healthy lifestyle, as well as the commitment to lifestyle interventions [18,30]. The concept of health-related quality of life refers to a person's evaluation of their physical and mental health over time [31], and it can be regarded as one of the most essential contributors to psychological well-being. Past research has demonstrated that higher levels of health-related quality of life also contribute to increased physical activity [32]. Comparably, studies have

indicated that experiencing limitations related to physical and psychological functionality, negative thinking, and moods can become barriers to completing lifestyle interventions and managing obesity [30]. Furthermore, research has shown that experiencing depressive symptoms, which may result from prolonged focus on negative thinking and feelings, can contribute to dropout in weight-loss trials [33].

Along with the aforementioned psychological attributes, some studies have found that males seem to discontinue studies more often than females [30,34,35], and that older people tend to drop out of studies earlier than younger people [30]. Lower educational background has also been shown to contribute to dropout from obesity management studies, although there are also results that conflict with this [30]. The above-mentioned demographic factors, as well as personality and well-being related factors, have been addressed in examinations of attrition in previous longitudinal lifestyle interventions [36]. Regarding technology-based interventions, more evidence concerning the role of individual factors in committing to interventions is needed [14]. This research should target subgroups, in particular those with a severe risk of disease, such as people with overweight and obesity [30].

The present study examined the potential contribution of personality characteristics and psychological well-being to attrition within the technology-supported lifestyle intervention Motivation Makes the Move! (MoMaMo!) that was designed to reduce overweight and obesity and subsequent negative health outcomes. Based on evidence derived from previous weight management interventions, we hypothesized that scoring high in agreeableness, extraversion, and conscientiousness would reduce the risk of attrition in this technology-supported intervention. Furthermore, we hypothesized that persons who have experienced challenges with their psychological well-being would have a higher risk of dropping out from the study. Along with examining these hypotheses, we also test whether the potential findings are robust after controlling for participants' age, sex, study group (personalized intervention vs general guidelines), and educational status. Since we consider that program adherence (in contrast with attrition) is necessary to achieve successful outcomes in interventions, we consider this study as a highly important starting point for the MoMaMo! project. This study strives to provide evidence on whether personality characteristics and psychological well-being measured at baseline should be taken under consideration in future technology-assisted lifestyle interventions to support participants' commitment to the completion of such interventions.

Methods

Design of the Study

MoMaMo! was a part of the Bits of Health program, supported by Business Finland, the national governmental agency to support technological development and innovation [37]. MoMaMo! was registered at ClinicalTrials.gov (protocol record TYH2016215, NCT02686502). Recruitment for the study started in April 2016 and the last follow-ups were performed in April 2020.

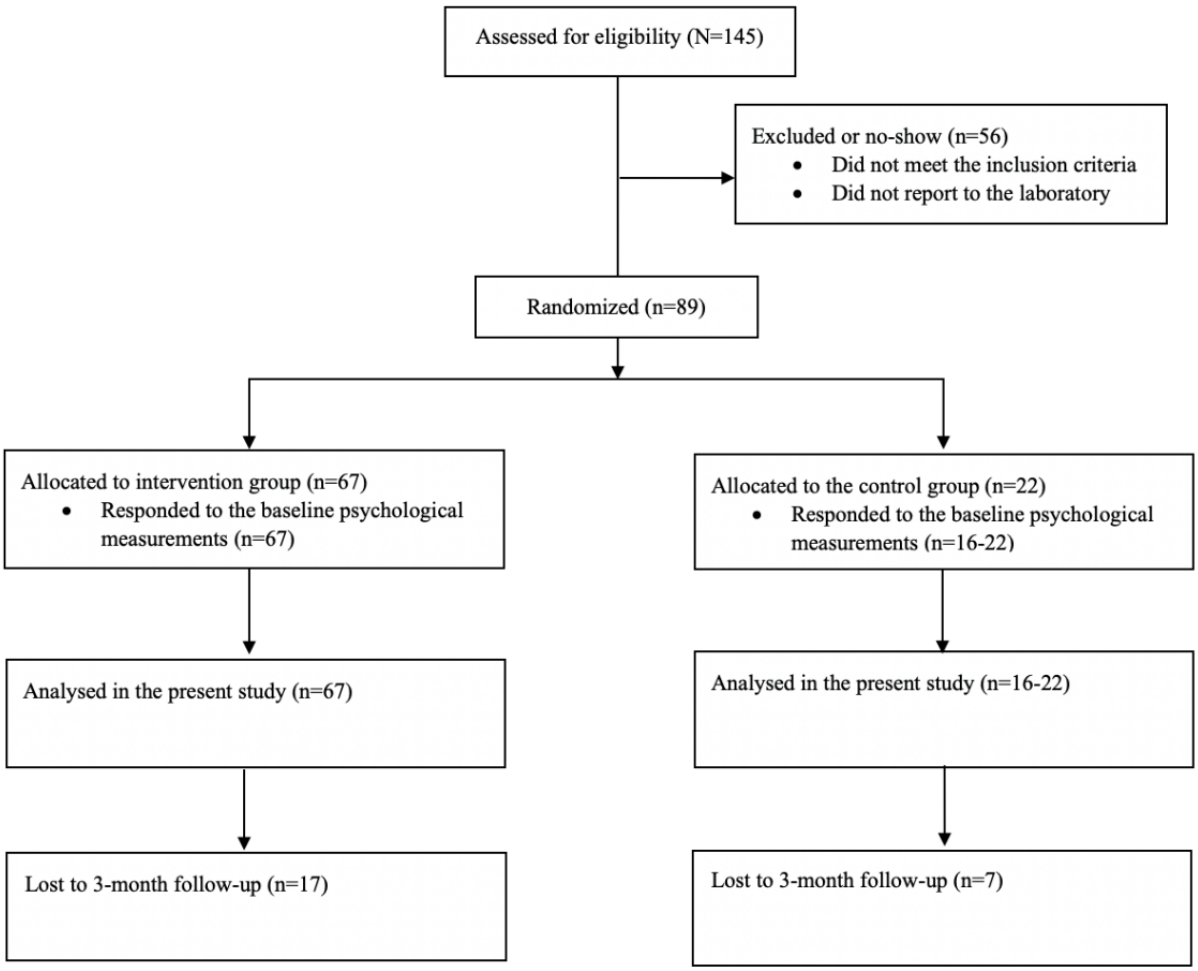
The overarching aim of the MoMaMo! study was to decrease and prevent sedentary lifestyles and obesity and subsequent health consequences among physically inactive and overweight or obese people. Furthermore, the study aimed to develop and validate lasting and individualized IT- and mHealth-assisted behavior change methodologies and practices for citizen engagement in health, well-being, and prevention of diseases. Specifically, the study aimed to quantify benefits and mechanisms of individualized exercise training in comparison with general guidelines. It also strived for identifying key elements to increase the adherence to and effectiveness of an individualized physical activity intervention and weight management program incorporating mHealth and other health technology.

The primary outcome measure in the MoMaMo! study was maximal oxygen uptake (VO_{2max}). VO_{2max} is a key parameter of cardiorespiratory fitness and an independent risk factor for several noncommunicable diseases and symptoms; it is associated with health-related quality of life [4]. VO_{2max} was assessed during a step-incremental cycle ergometer test (until volitional fatigue) with breath-by-breath alveolar gas exchange and pulmonary ventilation measurements. The highest 30-second moving average was calculated to obtain VO_{2max}.

Recruitment and Participants

Voluntary subjects were aged 18 to 40 years at entry, had BMI ≥ 27.5, had a referral from a physician for a consultation with a lifestyle clinic due to physical inactivity and overweight or obesity, and were deemed suitable for exercise testing and training. The subjects were recruited from different health care institutions in the Helsinki metropolitan area (Figure 1). In detail, participant recruitment was conducted by local public and private occupational health clinics by internet advertisement and recommendations by physicians. In addition, participants were recruited from the local University of Applied Sciences by internet advertisements. Exclusion criteria included the presence of a neurological or psychiatric disorder, use of medication influencing glucose homeostasis (except insulin) or autonomic nervous system function (eg, β-blockers or selective serotonin reuptake inhibitors), pregnancy, physical disability, substance abuse, significant co-operation difficulties, smoking, and severe anemia. The initial sample consisted of 89 subjects, of which 34 (38%) were men and 55 (62%) were women.

Figure 1. Flow diagram of the study (N=89).



MoMaMo! Intervention

The participants (N=89) were randomized into 3 groups. The principal investigator of the study generated the random allocation sequence. Research nurses, who were in charge of scheduling laboratory visit times, randomly allocated the participants to groups after the participants agreed to voluntarily participate in the study. A blinded draw of a paper containing the numbers 1, 2a, or 2b was performed to randomly allocate participant to group 1 (the general guidelines group), 2a (the individualized intervention group), or 2b (the highly individualized intervention group).

Group 1 visited the laboratory 3 times. During the first visit, the participants' height, weight, and body composition were assessed. The participants filled out self-report questionnaires on physical activity, psychology, and music use, as well as their work productivity and activity impairment. During the second visit, a physician examined each subject to ensure their suitability for exercise testing and training. Thereafter, each subject performed a step-incremental cardiopulmonary exercise test on a cycle ergometer (Monark Ergomedic 839 E, Monark Exercise AB) until voluntary fatigue. Data included recordings of pulmonary ventilation (Triple V, Jaeger Mijnhardt), alveolar gas exchange (Oxycon Pro), electrocardiography (ECG; PowerLab, AD Instruments) and the ratio of perceived exertion (RPE). On the third visit, blood tests were done, and the subjects received brief feedback on their results and measured health outcomes. Furthermore, group 1 subjects were provided general guidelines for a healthy diet and physical activity [38,39].

Groups 2a and 2b underwent the same measurements as group 1. Group 2a subjects also performed a submaximal step-incremental cycle ergometer test with ECG and RPE recordings during an additional visit. After completing all measurements, a personalized feedback meeting was organized for each subject in groups 2a and 2b. In this final meeting, each group 2a and group 2b subject received healthy lifestyle habit advice and an exercise prescription based on their results and their own preferences. For group 2a, a submaximal cycle ergometer test was used to personalize the exercise prescription. Thus, an estimate of maximal oxygen consumption (VO_{2max}), work-rate specific heart rate, and RPE were used to personalize the volume and intensity of the exercise prescription. For group 2b, measured VO_{2max}, determined ventilatory thresholds, heart rate response, and RPE during a maximal cycle ergometer test were used to provide a highly individualized exercise prescription. The purpose of using the simpler exercise test for the exercise prescriptions given to group 2a was to examine its usefulness in comparison with the more highly detailed cardiopulmonary exercise test given to group 2b.

To support the participants' own planning and follow-up of their physical activity, exercise training, and other health behaviors, subjects in groups 2a and 2b were instructed to use smartphone apps. Sports Tracker (Amer Sports Digital Services

Oy) with a heart rate belt (Suunto Oy) was used to guide and record exercise training, Argus (Azumio Inc) or its equivalent to count daily steps, Emotion Tracker (F8) to assess emotions, and Weight Diary (CurlyBrace Apps Ltd) or its equivalent to measure weight. The study web page provided further instructions and support for training. A 3-month Spotify gift card, along with example playlists, was provided to the subjects assigned to groups 2a and 2b to motivate their exercise training and support relaxation. The participants were instructed to fill out a web-based food diary (Nutri-Flow Oy) for ≥3 days at the beginning of intervention. The diary provided an analysis of nutritional habits, along with personalized feedback for suggested modifications. Subjects in groups 2a and 2b were also encouraged to report their training and weight loss on the study website (note the website was functional for the study duration and is presently not accessible) [40].

Sample size determination and power calculation (G*Power, Universität Kiel) were based on the expected increase in VO_{2max} (the primary outcome measure in the MoMaMo! study) after a 3-month exercise intervention in groups 1, 2a, and 2b. Our a priori hypothesis was that both the individualized (group 2a) and highly individualized (group 2b) interventions would induce a similar increase in VO_{2max} and be more beneficial than the intervention based on general guidelines (group 1). Therefore, groups 2a and 2b were considered as a single group (group 2) in comparisons with group 1. Based on the literature and our previous data, we assumed that there would be a similar initial VO_{2max} value of 25 (SD 2.5) mL/kg/min (representing a poor or very poor classification) in all groups [41,42]. Our hypothesis was that VO_{2max} would increase 15% in group 2 and 5% in group 1. To determine a priori differences between 2 independent means, a 2-tailed *t* test with α error probability = .05, power = 0.80, and effect size *d*=0.92, we needed 20 subjects in both groups. Due to an assumed attrition rate of 30% in the overweight or obese subjects and with a goal to also perform comparisons between groups 2a and 2b, we sought to recruit and randomize 33 subjects into each group, for a total of 99 subjects.

Final collected data came from 89 participants, including 22 (25%) participants in group 1, who received general guidelines on physical activity and nutrition, and 67 (75%) participants in group 2, who received combined individualized exercise and physical activity prescriptions, a food diary, and the health technology apps described above (Figure 1). The study procedure and its reporting followed the CONSORT (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) statement [43].

Psychological Questionnaires

Participants' personality characteristics were assessed with the Short Five (S5) personality assessment [21]. This measure is designed to assess personality traits, including neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness. Each dimension consists of 6 items, with higher scores reflecting higher levels of each trait. The scale ranges from –3 (the description is completely wrong/false) to 3 (the description is completely right or true). The sum scores of these 6 items were calculated. The reliability estimates (Cronbach α) were α=.84 for neuroticism, α=.75 for extraversion, α=.55 for openness to experience, α=.72 for agreeableness, and α=.65 for conscientiousness, indicating acceptable internal consistency for the items within each dimension. The major advantage of the S5 is that it allows for a detailed description of personality traits with a relatively small number of items compared to other instruments that assess the same components of personality [21]. This instrument has previously demonstrated good validity within general adult populations and correlates well with other measures of personality, including the NEO-Five-Factor Inventory [21].

Participants' psychological well-being was studied with the RAND 36-item health survey, consisting of a set of quality-of-life measures. This measure is designed to capture the following health concepts: physical functioning, bodily pain, role limitations due to physical health challenges, role limitations due to emotional problems, emotional well-being, social functioning, vitality (energy and fatigue), and perceptions of general health [44]. Higher scores within the scale reflect a more favorable health state. In the present study, the reliability of the subscales ranged from good to excellent (physical functioning: α=.72; role limitations due to physical health challenges: α=.68; role limitations due to personal or emotion-related problems: α=.78; vitality: α=.78; emotional well-being: α=.79; social functioning: α=.91; bodily pain: α=.76; and general health perceptions: α=.73). In previous studies, this instrument has been demonstrated to be valid in general adult populations [31,44].

Participants' mood was examined using the Positive and Negative Affect Schedule (PANAS). PANAS is a 20-item self-reported measure consisting of two 10-item scales, one for positive affect and one for negative affect. Participants were requested to rate single-word items describing positive or negative emotions. Higher scores reflected higher intensity of an emotion (1 indicated "not at all" and 5 indicated "extremely"). One missing value within each scale was allowed when calculating the scores. PANAS has been regarded as reliable and valid in previous studies [45]. Within the present study, Cronbach α for the scale assessing positive affectivity was α=.83, and the corresponding value for the scale for negative affectivity was α=.86.

Participants' depressive symptoms were assessed using the Beck Depression Inventory II (BDI-II) [46]. The BDI-II assesses 21 symptoms with a severity range from 0 (no symptoms) to 3 (severe level of depressive symptoms). A sum score of all items was computed for each participant (Table 1 shows descriptive statistics); no missing items were allowed. The reliability estimate (Cronbach α) for the depressive symptom scores was α=.90. The BDI-II has been demonstrated to be a valid instrument [46-48] and is an acknowledged standard in the measurement of depressive mood [46-48]. This instrument has also been demonstrated to be a useful screening tool for future risk for depression [46,47]. Each study group completed the above-mentioned self-administered questionnaires.

Table 1. Frequencies, percentages, and attrition rates of study variables.

Variables	Participants at month 0, n	Participants at month 3, n (%)	Attrition rate, n (%)
Study group			
Group 1	22	15 (68)	7 (32)
Group 2	67	50 (75)	17 (25)
Total	89	65 (73)	24 (27)
Sex			
Men	34	27 (79)	7 (21)
Women	55	38 (69)	17 (31)
Total	89	65 (73)	24 (27)
Educational background			
Vocational school	24	11 (46)	13 (54)
Lower academic degree	35	29 (83)	6 (17)
Academic degree	27	24 (89)	3 (11)
Total	86	64 (74)	22 (26)
Psychological measures			
Personality	88	64 (73)	24 (27)
RAND-36	89	65 (73)	24 (27)
Positive and Negative Affect Schedule	89	65 (73)	24 (27)
Beck Depression Inventory II	83	61 (73)	22 (27)

Statistical Analyses

Statistical analyses focused on assessing attrition with respect to the participants' self-reported personality characteristics and psychological well-being at the 3-month follow-up. The analyses were conducted using SPSS versions 25 and 27 (IBM Corp). Descriptive statistics of the data are presented as frequencies, percentages, means, and standard deviations in Tables 1-4. The minimum and maximum values are also reported in Tables 2-4.

Table 2. Mean values of personality characteristics in continuing (n=64) and dropout (n=24) groups of participants at the 3-month follow-up.

Variables	Continuing group (n=64)		Dropout group (n=24)	
	Minimum-maximum	Mean (SD)	Minimum-maximum	Mean (SD)
Neuroticism	–29.00 to 19.00	–8.28 (12.65)	–33.00 to 18.00	–6.13 (14.39)
Extraversion	19.00 to 30.00	7.31 (11.63)	–14.00 to 26.00	8.75 (11.40)
Openness	–2.00 to 35.00	14.34 (8.42)	–12.00 to 27.00	11.04 (9.56)
Agreeableness	–13.00 to 32.00	16.13 (10.17)	–7.00 to 32.00	14.92 (8.94)
Conscientiousness	–4.00 to 33.00	15.98 (9.02)	–8.00 to 33.00	15.25 (10.17)

Table 3. Mean values of psychological well-being indicators in continuing (n=61-65) and dropout (n=22-24) groups at the 3-month follow-up.

Variables	Continuing group (n=61-65)		Dropout group (n=22-24)	
	Minimum-maximum	Mean (SD)	Minimum-maximum	Mean (SD)
Physical functioning	50.00-100.00	91.23 (8.66)	60.00-100.00	91.25 (11.06)
Role limitations (physical)	0.00-100.00	89.62 (22.49)	50.00-100.00	93.75 (15.20)
Role limitations (psychological)	0.00-100.00	82.05 (31.22)	0.00-100.00	75.00 (37.11)
Vitality (energy/fatigue)	20.00-100.00	59.38 (16.83)	20.00-85.00	56.88 (21.51)
Emotional well-being	36.00-100.00	75.22 (13.68)	44.00-96.00	73.33 (15.72)
Social functioning	25.00-100.00	86.15 (16.55)	37.50-100.00	84.90 (21.17)
Bodily pain	22.50-100.00	78.35 (19.32)	32.50-100.00	74.38 (17.39)
General health perceptions	16.67-87.50	55.26 (16.49)	25.00-87.50	59.72 (17.71)
Positive mood	16.00-44.00	30.46(6.01)	21.00-40.00	30.29 (5.47)
Negative mood	10.00-30.00	13.74 (4.47)	10.00-26.00	14.46 (4.97)
Symptoms of depression	0.00-28.00	7.49 (6.37)	0.00-29.00	7.55 (7.32)

Table 4. Mean values of psychological well-being indicators with square root transformations in continuing (n=61-65) and dropout (n=22-24) groups at the 3-month follow-up.

Variables	Continuing group (n=61-65)		Dropout group (n=22-24)	
	Minimum-maximum	Mean (SD)	Minimum-maximum	Mean (SD)
Physical functioning	1.00-7.14	2.82 (1.37)	1.00-6.40	2.66 (1.66)
Role limitations (physical)	1.00-10.05	2.28 (2.50)	1.00-7.14	1.85 (2.00)
Symptoms of depression	0.00-5.29	2.47 (1.20)	0.00-5.39	2.36 (1.44)

The attrition analyses were conducted using hierarchical logistic regressions. Traditional statistical analysis methods have been used in previous longitudinal lifestyle interventions when examining attrition-related questions [36]. In the present study, the associations between the main predictors and the outcome, in other words, whether personality characteristics and indicators of psychological well-being contributed to attrition at the 3-month follow-up, were studied first (Table 5). Furthermore, the analyses were adjusted for the participants' sex, age, study group (group 1 vs group 2), and educational status (1 indicating vocational school or high school, 2 indicating a lower academic degree, and 3 indicating a higher academic degree) to test the robustness of the results (Table 6). The potential confounding factors were added to the models in the first step, and the main predictors in the second step (Table 6). The predictive power (Nagelkerke R^2) of the models is reported in Tables 5 and 6. For detailed estimates regarding the confounding factors, see Multimedia Appendix 1 (Tables S1-S16).

Table 5. Personality characteristics and indicators of psychological well-being predicting attrition at the 3-month follow-up.

Variables	B	SE	*P* value	OR	95% CI	Nagelkerke R^2
Personality characteristics						
Neuroticism	0.01	0.02	.49	1.01	0.98-1.05	0.01
Extraversion	0.01	0.02	.60	1.01	0.97-1.05	0.01
Openness	–0.04	0.03	.12	0.96	0.91-1.01	0.04
Agreeableness	–0.01	0.02	.61	0.99	0.94-1.04	0.00
Conscientiousness	–0.01	0.03	.74	0.99	0.94-1.04	0.00
Indicators of psychological well-being[a]						
Physical functioning	–0.07	0.17	.67	0.93	0.67-1.29	0.00
Role limitations (physical)	–0.09	0.11	.45	0.92	0.74-1.15	0.01
Role limitations (psychological)	–0.01	0.01	.37	0.99	0.98-1.01	0.01
Vitality (energy/fatigue)	–0.01	0.01	.56	0.99	0.97-1.02	0.01
Emotional well-being	–0.01	0.02	.58	0.99	0.96-1.02	0.01
Social functioning	–0.004	0.01	.77	1.00	0.97-1.02	0.00
Bodily pain	–0.01	0.01	.38	0.99	0.97-1.01	0.01
General health perceptions	0.02	0.02	.27	1.02	0.99-1.05	0.02
Positive mood	–0.01	0.04	.90	1.00	0.92-1.08	0.00
Negative mood	0.03	0.05	.51	1.03	0.94-1.14	0.01
Symptoms of depression	–0.07	0.20	.73	0.93	0.63-1.38	0.00

[a]Square root transformations were applied to predictors of physical functioning, role limitations (physical), and symptoms of depression.

Table 6. Personality characteristics and psychological well-being predicting attrition at the 3-month follow-up after adjusting for covariates.

Variables	B	SE	*P* value	OR	95% CI	Nagelkerke R^2
Personality characteristics						
Neuroticism	–0.03	0.03	.26	0.97	0.92-1.02	0.27
Extraversion	0.03	0.03	.26	1.03	0.98-1.09	0.28
Openness	0.02	0.03	.46	0.98	0.92-1.04	0.27
Agreeableness	–0.02	0.03	.53	0.98	0.93-1.04	0.27
Conscientiousness	0.02	0.03	.50	1.02	0.96-1.09	0.27
Indicators of psychological health[a]						
Physical functioning	–0.13	0.19	.50	0.88	0.61-1.27	0.27
Role limitations (physical)	–0.12	0.12	.35	0.89	0.70-1.13	0.28
Role limitations (psychological)	–0.003	0.01	.71	1.00	0.98-1.01	0.27
Vitality (energy/fatigue)	0.00	0.02	.95	1.00	0.97-1.03	0.27
Emotional well-being	0.01	0.02	.76	1.01	0.97-1.05	0.27
Social functioning	0.00	0.02	.93	1.00	0.97-1.04	0.27
Bodily pain	–0.001	0.01	.97	1.00	0.97-1.03	0.27
General health perceptions	0.03	0.02	.08	1.03	1.00-1.07	0.31
Positive mood	0.06	0.05	.29	1.06	0.95-1.18	0.28
Negative mood	0.01	0.06	.99	1.00	0.88-1.13	0.27
Symptoms of depression	–0.38	0.25	.13	0.68	0.42-1.12	0.31

[a]Square root transformations were applied to predictors of physical functioning, role limitations (physical), and symptoms of depression.

Ethics Approval

Subjects gave their written informed consent prior to participating in the study. The study conformed to the Declaration of Helsinki and was approved by the Ethics Committee (approval number 384/13/03/00/2015) of the Hospital District of Helsinki and Uusimaa, Helsinki, Finland.

Results

Of the 89 participants who were allocated to the intervention group (groups 2a and 2b, together comprising group 2), 65 (73%) participated in the 3-month follow-up. Tables 1-4 show further information on attrition rates and study variables. When the distributions of the variables were viewed within the whole sample, the distributions of personality factors were approximately symmetric. Regarding the psychological well-being indicators, the variable reflecting physical functioning was negatively skewed (skewness –1.75, kurtosis 4.47), as was the variable reflecting role limitations due to physical health challenges (skewness –2.49, kurtosis 5.95). The variable reflecting depressive symptoms was positively skewed (skewness 1.41, kurtosis 2.04). Skewness and kurtosis values exceeding the absolute value of 2 were considered as indicating severe nonnormality, and square root transformations were performed for these variables (Table 4). Considering other variables in RAND-36 and PANAS, no severe skewness was detected. A total of 77 of the 83 participants (93%) experienced only a minimal or small number of depressive symptoms (original scores 0-18), while 6 (7%) of the 83 participants experienced a moderate number of depressive symptoms (original scores 19-29). No information on severe depressive symptoms (original scores >30) was detected.

The examined personality characteristics and indicators of psychological well-being did not predict the risk for dropping out from the study at the 3-month follow-up (all *P* values >.05) (Table 5). The results remained the same when potential confounding factors, including age, sex, study group, and educational status, were adjusted for in the analyses (all *P* values >.05; Table 6). Perceptions of general health approached significance; in other words, health perceptions marginally predicted dropout from the study after covariate controls (*P*=.08; Table 6). Estimates for each confounder can be found in Multimedia Appendix 1 (Tables S1-S16).

Discussion

Principal Findings

The health-technology assisted MoMaMo! study focused on physically inactive overweight and obese men and women who were at risk for a permanent physically inactive lifestyle and lifestyle-related chronic diseases. We focused on assessing whether participant personality characteristics and indicators of psychological well-being contributed to attrition within the MoMaMo! study and whether these factors should be taken into consideration more carefully in future interventions to support adults' commitment to technology-based intervention programs.

Our results showed that 65 (73%) of the 89 initially recruited participants were still participating at the 3-month follow-up. Similar, and also lower, adherence rates have been demonstrated in other longitudinal weight management interventions [27,49]. Specifically, participant attrition rates from eHealth interventions have previously been shown to exceed 20% [50]. Here, the examined personality dimensions of neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness were not associated with the risk of dropping out from the study after the 3-month follow-up. Furthermore, the indicators of psychological well-being were not associated with the risk of attrition. All the results remained nonsignificant after adjusting for participant age, sex, study group, and educational status.

At present, studies on attrition in technology-based exercise intervention contexts are limited. Based on previous research [27], we carefully draw a hypothesis that scoring high in agreeableness and extraversion might reduce the risk of attrition. Our findings did not align with the evidence found by De Panfilis and colleagues [27] that characteristics related to agreeableness and extraversion (ie, reward dependence) may reduce the risk of dropping out in behavioral weight-loss interventions. Our study, however, is in the same direction as De Panfilis and colleague's [27] conclusion that personality as a whole seems to not be involved in whether subjects attend interventions associated with behavioral change (eg, weight loss). Some previous evidence has also indicated that scoring low in novelty seeking, which has been shown to correlate with conscientiousness [23], is associated with the completion of longitudinal weight management interventions [28]. Based on this, we expected that scoring high in conscientiousness might be associated with study adherence in MoMaMo! However, this characteristic was not related to the outcome. One possibility for our study's differences in relation to previous results could be that certain personality dimensions might be associated with later dropout [27]; in other words, personality-related dropout may occur after a longer period than 3 months. It has also been speculated that people who voluntarily join lifestyle programs have optimistic expectations for the program, and therefore their personality differences may not contribute to dropout at the beginning of the intervention [28]. Additionally, it has been suggested that there could be differences in lean and obese persons' personality characteristics, making the results from different samples vary [28]. In conclusion, our results suggest that commitment to the present technology-assisted physical activity intervention at the 3-month follow-up was not linked to any specific personality characteristics in the overweight adults.

Previous studies have indicated that a lower level of health-related quality of life (eg, experiences of health and physical limitations), negative thoughts and moods, and depressive symptoms could contribute to participant attrition [30,33]. Participants' expectations and emotions regarding technology-based interventions may also contribute to attrition [51]. Based on this evidence, it seems that people's personal resources related to well-being might be a key element in completing projects such as lifestyle interventions. In the present study, participants scored high in dimensions reflecting positive affect or mood. Positive mood has previously been shown to predict adherence to interventions [30]. Furthermore, almost all participants (77/83, 93%) experienced only minimal or mild

depressive symptoms, and no results on severe depressive symptoms were found. This might be the reason our results differed from those of Goode et al [33], who suggested that experiencing depression contributes to attrition. Taken together, we did not gain evidence for the contribution of the examined indicators of psychological well-being to dropout in the present study.

In addition to the above speculation, it is also possible that the designed intervention features in the technology-supported MoMaMo! trial were, in their current from, suitable for the recruited participants' characteristics, so that dropout could not be attributed to their personality or levels of psychological well-being. Our findings suggest that the psychological qualities do not contribute to discontinuation of the technology-based lifestyle intervention program among overweight adults, and that tailoring interventions more carefully to personality and well-being levels may not be needed in supporting participants' adherence to technology-assisted interventions during a 3-month period.

Potential causes of the relatively high attrition rate (27%) in our study remain, however, obscure. It has been stated that motivating users for behavior change before guiding them through action planning [35] could support their adherence to lifestyle interventions. The role of social support, such as integrating face-to-face intervention delivery strategies to technology-based intervention programs, may be a contributor to participants' motivation to commit to these studies [52]. Individuals may also be more likely to drop out of trials if they do not meet weight loss goals [53]. Use of positive reinforcement may also be beneficial, especially in cases where reaching outcomes requires time (eg, losing weight) [51]. Based on these perspectives, it seems possible that more attention should be paid to motivational factors, referring to the inner states that energize and direct behaviors. At present, examination of the effectiveness of the MoMaMo! study is ongoing, and future studies on attrition could focus on assessing motivational elements beyond the present research, such as on participants' meeting intervention goals.

Limitations

This study focused on assessing the potential role of personality and psychological well-being in study attrition. Other factors, such as physical or social circumstances (eg, sudden outbreaks of disease or lack of social support) and those related to meeting goals were not assessed. Furthermore, specific study variables, including psychological well-being, were studied only in the beginning of the intervention. As psychological well-being, along with many other factors relating to human experience, fluctuates over time, it will be important to address the dynamic nature of this phenomenon in relation to study attrition (eg, using survival analyses) in future studies. Additionally, we cannot be absolutely certain whether the participants who did not come to the laboratory follow-up tests quit only the study or also their intervention program at the individual level. That is, they may or may not have continued changing or improving their lifestyles after they discontinued their participation in the MoMaMo! study. Although all the instruments' subscales were considered acceptable according to general rules, openness to experience had poor reliability (Cronbach α=.55). Some studies have demonstrated similar α values when optimal subset items reflecting personality dimensions were researched [54]. Generally, the α coefficient reliability has been high in this instrument's validation studies [21]. Furthermore, due to the relatively high attrition rate, some of the analyses performed in the study may lack statistical power. Consequently, these study results can be regarded only as directional. Studies with larger sample sizes and longer follow-up periods are needed to confirm our results.

Strengths

We were able to study attrition with a well-validated set of psychological questionnaires. To our knowledge, this is one of the first studies assessing attrition in technology-supported, longitudinal lifestyle interventions conducted in overweight or obese adults. Our study is also one of the first to provide information on the role of participant psychological profiles in technology-supported interventions that aim at increasing physical activity and VO_{2max}.

Conclusions

The personality characteristics and indicators of psychological well-being examined in this study did not contribute to participant attrition in the technology-supported MoMaMo! lifestyle intervention conducted in overweight adults. As attrition remains a challenge within longitudinal, technology-supported lifestyle interventions, attention should be paid to the potentially dynamic nature of personality and psychological well-being, as well as to elements beyond these.

Acknowledgments

The study was funded by Business Finland (575/31/2015), University of Helsinki, Ministry of Education and Culture (Finland), Helsinki Deaconess Institute, Diaconia University of Applied Sciences, Terveystalo Oyj (former Diacor Terveyspalvelut Oy), HUR Oy, Firstbeat Technologies Oy, Nukute Oy, Suunto Oy, Amer Sports Digital Services Oy, Nutri-Flow Oy, F8 Helsinki Oy, Diamond (TYH2016215), and Amer Cultural Foundation. We gratefully acknowledge Sanni Malander and Laura Ginström for data acquisition and coding.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Detailed estimates for the confounding factors.
[DOCX File , 62 KB - formative_v6i11e30285_app1.docx]

Multimedia Appendix 2
CONSORT-eHEALTH checklist (V 1.6.2).
[PDF File (Adobe PDF File), 93 KB - formative_v6i11e30285_app2.pdf]

References

1. Van Camp G. Cardiovascular disease prevention. Acta Clinica Belgica 2014 Aug 31;69(6):407-411. [doi: 10.1179/2295333714y.0000000069]
2. Booth F, Roberts C, Laye M. Lack of exercise is a major cause of chronic diseases. Compr Physiol 2012 Apr;2(2):1143-1211 [FREE Full text] [doi: 10.1002/cphy.c110025] [Medline: 23798298]
3. Warburton DE, Nicol CW, Bredin SSD. Health benefits of physical activity: the evidence. CMAJ 2006 Mar 14;174(6):801-809 [FREE Full text] [doi: 10.1503/cmaj.051351] [Medline: 16534088]
4. Physical Activity Guidelines for Americans, 2nd ed. US Dept of Health and Human Services. 2018. URL: https://health.gov/sites/default/files/2019-09/Physical_Activity_Guidelines_2nd_edition.pdf [accessed 2022-08-16]
5. Wiklund P. The role of physical activity and exercise in obesity and weight management: Time for critical appraisal. J Sport Health Sci 2016 Jun;5(2):151-154 [FREE Full text] [doi: 10.1016/j.jshs.2016.04.001] [Medline: 30356545]
6. Bonomi AG, Westerterp KR. Advances in physical activity monitoring and lifestyle interventions in obesity: a review. Int J Obes (Lond) 2012 Feb 17;36(2):167-177. [doi: 10.1038/ijo.2011.99] [Medline: 21587199]
7. Zhao J, Freeman B, Li M. Can mobile phone apps influence peoples health behavior change? An evidence review. J Med Internet Res 2016 Oct 31;18(11):e287 [FREE Full text] [doi: 10.2196/jmir.5692] [Medline: 27806926]
8. Krukowski R, Tilford J, Harvey-Berino J, West D. Comparing behavioral weight loss modalities: incremental cost-effectiveness of an internet-based versus an in-person condition. Obesity (Silver Spring) 2011 Aug;19(8):1629-1635 [FREE Full text] [doi: 10.1038/oby.2010.341] [Medline: 21253001]
9. Eapen ZJ, Peterson ED. Can mobile health applications facilitate meaningful behavior change?: Time for answers. JAMA 2015 Sep 22;314(12):1236-1237. [doi: 10.1001/jama.2015.11067] [Medline: 26393844]
10. Denscombe M. Item non‐response rates: a comparison of online and paper questionnaires. Int J Soc Res Methodol 2009 Oct;12(4):281-291. [doi: 10.1080/13645570802054706]
11. Davies CA, Spence JC, Vandelanotte C, Caperchione CM, Mummery W. Meta-analysis of internet-delivered interventions to increase physical activity levels. Int J Behav Nutr Phys Act 2012 Apr 30;9(1):52 [FREE Full text] [doi: 10.1186/1479-5868-9-52] [Medline: 22546283]
12. Watson S, Woodside JV, Ware LJ, Hunter SJ, McGrath A, Cardwell CR, et al. Effect of a web-based behavior change program on weight loss and cardiovascular risk factors in overweight and obese adults at high risk of developing cardiovascular disease: Randomized controlled trial. J Med Internet Res 2015 Jul 16;17(7):e177 [FREE Full text] [doi: 10.2196/jmir.3828] [Medline: 26183659]
13. Rathbone AL, Prescott J. The use of mobile apps and SMS messaging as physical and mental health interventions: Systematic review. J Med Internet Res 2017 Aug 24;19(8):e295 [FREE Full text] [doi: 10.2196/jmir.7740] [Medline: 28838887]
14. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar 31;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
15. Geraghty AWA, Torres LD, Leykin Y, Pérez-Stable EJ, Muñoz RF. Understanding attrition from international Internet health interventions: a step towards global eHealth. Health Promot Int 2013 Sep 10;28(3):442-452 [FREE Full text] [doi: 10.1093/heapro/das029] [Medline: 22786673]
16. Dragana B, Vesna GJ, Veljko J, Dragan Z, Milica L. Baseline sociodemographic characteristics of participants and dropout in web-based weight reduction program in Serbia. Front Public Health 2016;4. [doi: 10.3389/conf.fpubh.2016.01.00113]
17. Rhodes R, Boudreau P. Physical activity and personality traits. In: Braddick O, editor. Oxford Research Encyclopedia of Psychology. Oxford, UK: Oxford University Press; 2017.
18. Biddle S, Fox K, Boutcher S. Physical Activity and Psychological Well-being. London, UK: Routledge; 2003.
19. Raynor DA, Levine H. Associations between the five-factor model of personality and health behaviors among college students. J Am Coll Health 2009 Jul;58(1):73-82. [doi: 10.3200/jach.58.1.73-82]
20. McCrae R, Costa P. The five-factor theory of personality. In: Handbook of Personality: Theory and Research (3rd ed). New York, NY: Guilford Press; 2008.
21. Konstabel K, Lönnqvist J, Walkowitz G, Konstabel K, Verkasalo M. The 'Short Five' (S5): Measuring personality traits using comprehensive single items. Eur J Pers 2012 Jan 01;26(1):13-29. [doi: 10.1002/per.813]
22. Cloninger CR, Svrakic DM, Przybeck TR. A psychobiological model of temperament and character. Arch Gen Psychiatry 1993 Dec 01;50(12):975-990. [doi: 10.1001/archpsyc.1993.01820240059008] [Medline: 8250684]
23. De Fruyt F, Van De Wiele L, Van Heeringen C. Cloninger's psychobiological model of temperament and character and the five-factor model of personality. Pers Individ Differ 2000 Sep;29(3):441-452. [doi: 10.1016/s0191-8869(99)00204-4]

24. Sundfør HB, Fyhri A. A push for public health: the effect of e-bikes on physical activity levels. BMC Public Health 2017 Oct 16;17(1):809 [FREE Full text] [doi: 10.1186/s12889-017-4817-3] [Medline: 29037235]
25. Wilson KE, Dishman RK. Personality and physical activity: A systematic review and meta-analysis. Pers Individ Differ 2015 Jan;72:230-242. [doi: 10.1016/j.paid.2014.08.023]
26. Jokela M, Hintsa T, Hintsanen M, Keltikangas‐Järvinen L. Adult temperament and childbearing over the life course. Eur J Pers 2020 Dec 02;24(2):151-166. [doi: 10.1002/per.749]
27. De Panfilis C, Torre M, Cero S, Salvatore P, Dall'Aglio E, Marchesi C, et al. Personality and attrition from behavioral weight-loss treatment for obesity. Gen Hosp Psychiatry 2008 Nov;30(6):515-520. [doi: 10.1016/j.genhosppsych.2008.06.003] [Medline: 19061677]
28. Sullivan S, Cloninger CR, Przybeck TR, Klein S. Personality characteristics in obesity and relationship with successful weight loss. Int J Obes (Lond) 2007 Apr 5;31(4):669-674 [FREE Full text] [doi: 10.1038/sj.ijo.0803464] [Medline: 16953251]
29. de Vries RAJ, Truong KP, Zaga C, Li J, Evers V. A word of advice: how to tailor motivational text messages based on behavior change theory to personality and gender. Pers Ubiquit Comput 2017 Jun 15;21(4):675-687. [doi: 10.1007/s00779-017-1025-1]
30. Burgess E, Hassmén P, Pumpa KL. Determinants of adherence to lifestyle intervention in adults with obesity: a systematic review. Clin Obes 2017 Jun 15;7(3):123-135. [doi: 10.1111/cob.12183] [Medline: 28296261]
31. Häkkinen A, Rinne M, Vasankari T, Santtila M, Häkkinen K, Kyröläinen H. Association of physical fitness with health-related quality of life in Finnish young men. Health Qual Life Outcomes 2010 Jan 29;8(1):15 [FREE Full text] [doi: 10.1186/1477-7525-8-15] [Medline: 20109241]
32. Bize R, Johnson JA, Plotnikoff RC. Physical activity level and health-related quality of life in the general adult population: a systematic review. Prev Med 2007 Dec;45(6):401-415. [doi: 10.1016/j.ypmed.2007.07.017] [Medline: 17707498]
33. Goode RW, Ye L, Sereika SM, Zheng Y, Mattos M, Acharya SD, et al. Socio-demographic, anthropometric, and psychosocial predictors of attrition across behavioral weight-loss trials. Eat Behav 2016 Jan;20:27-33 [FREE Full text] [doi: 10.1016/j.eatbeh.2015.11.009] [Medline: 26609668]
34. Van't Riet J, Crutzen R, De Vries H. Investigating predictors of visiting, using, and revisiting an online health-communication program: A longitudinal study. J Med Internet Res 2010 Sep 02;12(3):e37 [FREE Full text] [doi: 10.2196/jmir.1345] [Medline: 20813716]
35. Van der Mispel C, Poppe L, Crombez G, Verloigne M, De Bourdeaudhuij I. A self-regulation-based eHealth intervention to promote a healthy lifestyle: Investigating user and website characteristics related to attrition. J Med Internet Res 2017 Jul 11;19(7):e241 [FREE Full text] [doi: 10.2196/jmir.7277] [Medline: 28698168]
36. Kaseva K, Pulkki-Råback L, Elovainio M, Pahkala K, Keltikangas-Järvinen L, Hintsanen M, et al. Psychological wellbeing in 20-year-old adults receiving repeated lifestyle counselling since infancy. Acta Paediatr 2015 Aug 29;104(8):815-822. [doi: 10.1111/apa.13036] [Medline: 25939586]
37. Business Finland. URL: https://www.businessfinland.fi/en/for-finnish-customers/home [accessed 2022-07-27]
38. Garber C, Blissmer B, Deschenes M, Franklin B, Lamonte M, Lee IM, American College of Sports Medicine. American College of Sports Medicine position stand. Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise. Med Sci Sports Exerc 2011 Jul;43(7):1334-1359. [doi: 10.1249/MSS.0b013e318213fefb] [Medline: 21694556]
39. American College of Sports Medicine. ACSM's Resource Manual for Guidelines for Exercise Testing and Prescription (7th ed). Baltimore, MD: Wolters Kluwer Health, Lippincott Williams & Wilkins; 2014.
40. VTT Technical Research Centre of Finland Ltd. 2016. URL: https://www.vttresearch.com/en [accessed 2022-08-31]
41. Lunt H, Draper N, Marshall HC, Logan FJ, Hamlin MJ, Shearman JP, et al. High intensity interval training in a real world setting: a randomized controlled feasibility study in overweight inactive adults, measuring change in maximal oxygen uptake. PLoS One 2014 Jan 13;9(1):e83256 [FREE Full text] [doi: 10.1371/journal.pone.0083256] [Medline: 24454698]
42. Lamprecht M, Obermayer G, Steinbauer K, Cvirn G, Hofmann L, Ledinski G, et al. Supplementation with a juice powder concentrate and exercise decrease oxidation and inflammation, and improve the microcirculation in obese women: Randomised controlled trial data. Br J Nutr 2013 Apr 16;110(9):1685-1695. [doi: 10.1017/s00071145130001001]
43. Welcome to the CONSORT Website. CONSORT: Transparent Report of Trials. URL: http://www.consort-statement.org/ [accessed 2022-08-16]
44. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). Med Care 1992;30(6):473-483. [doi: 10.1097/00005650-199206000-00002]
45. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: The PANAS scales. J Pers Soc Psychol 1988;54(6):1063-1070. [doi: 10.1037/0022-3514.54.6.1063]
46. Beck A, Steer R, Brown G. Manual for the Beck Depression Inventory-II. San Antonio, TX: Psychological Corporation; 1996.
47. Storch EA, Roberti JW, Roth DA. Factor structure, concurrent validity, and internal consistency of the Beck Depression Inventory-Second Edition in a sample of college students. Depress Anxiety 2004;19(3):187-189. [doi: 10.1002/da.20002] [Medline: 15129421]

48. Wang Y, Gorenstein C. Psychometric properties of the Beck Depression Inventory-II: A comprehensive review. Braz J Psychiatry 2013 Dec;35(4):416-431. [doi: 10.1590/1516-4446-2012-1048] [Medline: 24402217]
49. Tobi P, Estacio EV, Yu G, Renton A, Foster N. Who stays, who drops out? Biosocial predictors of longer-term adherence in participants attending an exercise referral scheme in the UK. BMC Public Health 2012 May 11;12(1):347 [FREE Full text] [doi: 10.1186/1471-2458-12-347] [Medline: 22578185]
50. Neve MJ, Collins CE, Morgan PJ. Dropout, nonusage attrition, and pretreatment predictors of nonusage attrition in a commercial web-based weight loss program. J Med Internet Res 2010 Dec 14;12(4):e69 [FREE Full text] [doi: 10.2196/jmir.1640] [Medline: 21156470]
51. Short CE, Vandelanotte C, Dixon MW, Rosenkranz R, Caperchione C, Hooker C, et al. Examining participant engagement in an information technology-based physical activity and nutrition intervention for men: the manup randomized controlled trial. JMIR Res Protoc 2014 Jan 03;3(1):e2 [FREE Full text] [doi: 10.2196/resprot.2776] [Medline: 24389361]
52. Du H, Venkatakrishnan A, Youngblood GM, Ram A, Pirolli P. A group-based mobile application to increase adherence in exercise and nutrition programs: A factorial design feasibility study. JMIR Mhealth Uhealth 2016 Jan 15;4(1):e4 [FREE Full text] [doi: 10.2196/mhealth.4900] [Medline: 26772910]
53. Yackobovitch-Gavan M, Steinberg DM, Endevelt R, Benyamini Y. Factors associated with dropout in a group weight-loss programme: A longitudinal investigation. J Hum Nutr Diet 2015 Feb 14;28 Suppl 2:33-40. [doi: 10.1111/jhn.12220] [Medline: 24528102]
54. Konstabel K, Lönnqvist JE, Leikas S, García Velázquez R, Qin H, Verkasalo M, et al. Measuring single constructs by single items: Constructing an even shorter version of the "Short Five" personality inventory. PLoS One 2017 Aug 11;12(8):e0182714 [FREE Full text] [doi: 10.1371/journal.pone.0182714] [Medline: 28800630]

Abbreviations

BDI-II: Beck Depression Inventory II
CONSORT: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth
ECG: electrocardiography
mHealth: mobile health
MoMaMo!: Motivation Makes the Move!
PANAS: Positive and Negative Affect Schedule
RPE: ratio of perceived exertion
S5: Short Five
VO_{2max}: maximal oxygen uptake

Edited by A Mavragani; submitted 16.05.21; peer-reviewed by D Saleh, Y Zhou; comments to author 17.09.21; revised version received 01.03.22; accepted 02.06.22; published 25.11.22.

Please cite as:
Kaseva K, Tervaniemi M, Heikura E, Kostilainen K, Pöyhönen-Alho M, Shoemaker JK, Petrella RJ, Peltonen JE
Identifying Personality Characteristics and Indicators of Psychological Well-Being Associated With Attrition in the Motivation Makes the Move! Physical Activity Intervention: Randomized Technology-Supported Trial
JMIR Form Res 2022;6(11):e30285
URL: https://formative.jmir.org/2022/11/e30285
doi:10.2196/30285
PMID:36427239

Original Paper

Analyzing Person-Place Interactions During Walking Episodes: Innovative Ambulatory Assessment Approach of Walking-Triggered e-Diaries

Martina Kanning[1], Prof Dr; Lukas Bollenbach[1], MA; Julian Schmitz[2], MA; Christina Niermann[3], PhD; Stefan Fina[2], Prof Dr

[1]Department of Sport Science, University of Konstanz, Konstanz, Germany

[2]Faculty of Architecture and Civil Engineering, University of Applied Sciences Augsburg, Augsburg, Germany

[3]Institute of Interdisciplinary Exercise Science and Sports Medicine, Medical School Hamburg, Hamburg, Germany

Corresponding Author:
Martina Kanning, Prof Dr
Department of Sport Science
University of Konstanz
Universitätsstraße 10
Konstanz, 78464
Germany
Phone: 49 7531 883651
Email: martina.kanning@uni-konstanz.de

Abstract

Background: Walking behavior is positively associated with physiological and mental health as much evidence has already shown. Walking is also becoming a critical issue for health promotion in urban environments as it is the most often used form of active mobility and helps to replace carbon dioxide emissions from motorized forms of transport. It therefore contributes to mitigate the negative effects of climate change and heat islands within cities. However, to promote walking among urban dwellers and to utilize its health-enhancing potential, we need to know more about the way in which physical and social environments shape individual experiences during walking episodes. Such person-place interactions could not adequately be analyzed in former studies owing to methodological constraints.

Objective: This study introduces walking-triggered e-diaries as an innovative ambulatory assessment approach for time-varying associations, and investigates its accuracy with 2 different validation strategies.

Methods: The walking trigger consists of a combination of movement acceleration via an accelerometer and mobile positioning of the cellphone via GPS and transmission towers to track walking activities. The trigger starts an e-diary whenever a movement acceleration exceeds a predetermined threshold and participants' locations are identified as nonstationary outside a predefined place of residence. Every 420 (±300) seconds, repeated e-diaries were prompted as long as the trigger conditions were met. Data were assessed on 10 consecutive days. First, to investigate accuracy, we reconstructed walking routes and calculated a percentage score for all triggered prompts in relation to all walking routes where a prompt could have been triggered. Then, to provide data about its specificity, we used momentary self-reports and objectively assessed movement behavior to describe activity levels before the trigger prompted an e-diary.

Results: Data of 67 participants could be analyzed and the walking trigger led to 3283 e-diary prompts, from which 2258 (68.8%) were answered. Regarding accuracy, the walking trigger prompted an e-diary on 732 of 842 (86.9%) reconstructed walking routes. Further, in 838 of 1206 (69.5%) triggered e-diaries, participants self-reported that they were currently walking outdoors. Steps and acceleration movement was higher during these self-reported walking episodes than when participants denied walking outdoors (steps: 106 vs 32; acceleration>0.2 *g* in 58.4% vs 19% of these situations).

Conclusions: Accuracy analysis revealed that walking-triggered e-diaries are suitable to collect different data of individuals' current experiences in situations in which a person walks outdoors. Combined with environmental data, such an approach increases knowledge about person-place interactions and provides the possibility to gain knowledge about user preferences for health-enhancing urban environments. From a methodological viewpoint, however, specificity analysis showed how changes in trigger conditions (eg, increasing the threshold for movement acceleration) lead to changes in accuracy.

(JMIR Form Res 2022;6(11):e39322) doi:10.2196/39322

KEYWORDS

ecological momentary assessment; active transport; socio-ecological model; subjective well-being; mental health; urban health; GEMA; geographically explicit ecological momentary assessment; behaviour change; walking; experience; environment; monitoring; activity; tracking; e-diary; assessment

Introduction

Background

Walking with low to moderate intensity is associated with physiological health such as all-cause and cardiovascular mortality [1], obesity and diabetes [2], and mental health disorders such as stress [3], depression [4,5] and well-being [6-8]. Besides these health-enhancing associations, walking is the most often used type of active mobility [9]. Walking as well as active mobility replace carbon dioxide emissions from motorized transport and therefore contribute toward mitigating anthropogenic climate warming [10]. Such health and climate effects are of particular concern for urban residents who are exposed to higher levels of stress, isolation, and air pollutants and emissions [11-13]. Thus, increasing walking episodes and active mobility are important objectives for health promotion in urban environments [14,15]. However, to promote walking among urban dwellers in their everyday life and to utilize its health-enhancing potential, we need to know more about the way physical and social environments shape individual experiences during walking episodes. It is important to understand how a resident reacts to contextual circumstances (eg, street greenery and wide sidewalks) and which person or situation-specific moderators (eg, lifestyle, attitude, social interaction, and weather or noise condition) affect the association between walking outdoors and health [16,17]. Such knowledge would help to inform the development of urban designs, which are associated with positive effects while walking in daily urban life.

To analyze such person-place interactions between the individuum within its environment, we need time-varying associations, which previous studies mostly do not provide. The goal of this study is to introduce walking-triggered e-diaries as an innovative ambulatory assessment approach to overcome this methodological constraint.

Assessing and Analyzing the Health-Enhancing Potentials of Urban Environments

A plethora of studies and reviews investigated associations between the built or physical environment and physical activity (eg, walking). They indicated that high-quality walking infrastructure, access to recreational facilities, as well as parks, trails, or new infrastructure for walking are associated with increased levels of physical activity [18,19]. Regarding the impact on mental health, the findings are not as consistent as they are for physical activity but show that park-based and greenway interventions can promote improved mental health [20,21]. However, these findings are mostly based on cross-sectional or longitudinal data with longer time frames, and they rarely examine moderating effects of the environment in a way that allows investigation of person-place interactions. For this, an ambulatory assessment (AA) is a useful approach [22,23]; it allows researchers to examine time-varying associations of, for example, momentary affective states and physical activity and to what extent environmental features in that specific situation moderate this association. Hereinafter, we use the term “AA” instead of the “experience sampling method” or “ecological momentary assessment” (EMA), as it covers the daily assessment of self-reports (via paper-and-pencil or electronic diaries) in combination with the monitoring of physiological functions (eg, heart rate and electrodermal activity), behavior (eg, via accelerometers and mobile electrocardiograms), or environmental parameters (eg, via geolocation tracking) [24]. Several advantages, such as assessments in everyday life, in real time, and repeated measurements with a high sampling frequency led to the use of AA in a wide range of research areas [25]. Two general categories of data sampling are usually distinguished: (1) time-based designs and (2) event-based designs. To capture variability of a certain health or psychosocial outcome (eg, affective states and different kinds of social interactions), a time-based design is appropriate for recording data at fixed (eg, every hour) or random intervals. However, when the aim is to extract data in specific situations in everyday life (eg, walking outdoors), event-based designs are recommended. This design requires the participants to detect preselected situations on their own and to provide their self-reports during or after such situations. However, technical progress of accelerometers and sensors in recent years allow for triggers as a sophisticated feature within an event-based design. Sensors (eg, accelerometers and GPS) can discover in situ a predefined situation in everyday life and initiate an alarm to answer e-diaries. This triggered event–based approach minimizes barriers for participants and increases the objectivity of self-reporting when this predefined event occurs (for an overview, see Shiffman et al [26]).

Technical Progress in Assessing Person-Place Interactions

In recent years, some research groups in the field of urban development and city planning extended assessment ideas and collected momentary data of mental health or subjective environmental experiences during everyday life. These study designs combined mobile geographic location technologies (eg, GPS) to assess precisely the space where participants spend their time outdoors. Time-based assessments via daily diaries resulted in geographically explicit ecological momentary assessments (GEMA; for an overview, see Chaix [16] and Kirchner et al [27]). As an example, a GEMA-study assessed the association between urban green space and stress in 13- to 14-year-old adolescents living in urban surroundings [28]. Participants were recruited from an adolescent medicine outpatient clinic. Momentary experiences of stress were captured

3-6 times a day over a 4-day period every second month over a 2-year period. Participants were equipped with a GPS-enabled mobile phone. They received text messages including a link to a web-based questionnaire with a single item measuring momentary stress. Exposure to greenspace was calculated retrospectively for each participant using the normalized difference vegetation index of all image pixels within 100 m of each EMA location. This index measures leaf abundance in green vegetation. The analyses included 179 participants providing 9346 EMA responses and revealed significant associations between exposure to green space and lower psychological stress during active mobility (see also the feasibility study of GEMA by Boettner et al [29] with a representative sample of 1405 adolescents living in cities and suburban municipalities).

Electronic diaries within GEMA designs therefore allow capturing of subjective experiences in situ. The data can be linked retrospectively to participants' current position in time and space. However, not all GEMA studies explicitly assessed physical activity (eg, with accelerometers) but rather operationalized activity using time-space patterns of the GPS. In addition, most GEMA designs struggled with the disadvantages of "time-based" assessments because prompts to answer the web-based questionnaires were usually triggered at regular time intervals. With such a sampling scheme, self-reports during rare events (eg, walking episodes) are likely to be missed, which impedes the analysis of person-place interactions. A currently published study protocol of a GEMA study attempted to overcome this restriction (Fernandes et al 2020 [30]). This GEMA study included older residents of Paris (>60 years old) and aimed to assess whether the sequence of stressful environmental exposures in everyday life explained within-subject variability in stress and depression. They presented a novel methodology combining GEMA with an algorithm (based on the smartphone's built-in GPS receiver) that allowed triggering of EMA surveys only when subjects are outdoors. However, the trigger did not differentiate between active and inactive episodes and is therefore not able to identify walking episodes.

The purpose of this study is to introduce walking-triggered e-diaries as an innovative ambulatory assessment approach to overcome the methodological shortcomings mentioned above. It combines movement acceleration, time, and distance assessments (GPS) to trigger an e-diary. The aim of this study is to examine its accuracy through 2 different approaches. First, we reconstructed walking routes (using a GPS) and examined on how many walking routes the walking trigger prompted an e-diary (sensitivity). Second, we describe activity levels (movement acceleration and number of steps) immediately before the prompt, and used self-reports to identify "false positive" walking-triggered e-diaries; that is, a participant received a triggered e-diary, but denied walking outdoors (specificity).

Methods

Walking-Triggered e-Diaries: How Do They Work

Triggered e-diaries (or triggered EMAs) have previously been employed to assess data during physically active and inactive episodes [31], as well as during prolonged sedentary bouts [32] (see also the comparison of different GPS-based triggered e-diaries by Törnros et al [33]). To develop the walking-trigger, we used the following equipment: a hip-worn accelerometer (Move3, movisens.com), a smartphone (Motorola, running on the Android operating system) to show e-diaries (movisensXS), and location tracking via GPS and transmission towers, combined in a technical interface between the smartphone and the accelerometer (via Bluetooth low energy). Move3 captured the body position and movement acceleration within a range of ±16 *g* and with a sampling frequency of 64 Hz. Move3 has been validated for documenting movement acceleration and different body positions [31,34].

Taken together, the walking trigger consists of a combination of movement-acceleration and mobile positioning of the cellphone via GPS and transmission towers to identify walking episodes. The walking trigger prompts an e-diary between 6 AM and10 PM whenever an individual is walking outdoors. This is achieved by distinguishing between 2 states: stationary and nonstationary. The trigger state switches to nonstationary whenever a movement acceleration exceeds a predetermined threshold (>0.1 *g* for at least 1 minute) and the participants' location changes beyond a radius of 100 m. If these conditions are met, a walking-triggered e-diary (main) is prompted. Furthermore, participants receive follow-up walking-triggered e-diaries on their walking trip with a repeat interval of 420 seconds and a randomization time of 300 seconds for as long as the trigger conditions are met (repeated). For each e-dairy, participants need approximately 1 minute for responding.

To conduct the walking-triggered e-diary study, the trigger sampling scheme and the corresponding forms in the movisensXS browser app [35] have to be created. Next, the technical interface, consisting of a study smartphone (which depicts the e-diaries and enables tracking via GPS and transmission towers) and the Move3 accelerometer, was established via Bluetooth low energy. Finally, the study was prepared to be started by the participants once they received their smartphone and accelerometer.

Data Processing

The collected data were processed with software (UnisensViewer, DataAnalyzer, and DataMerger) available at the manufacturer's website home page [36]. All data forms from the smartphones were then uploaded to the movisensXS browser app. Corresponding data from the accelerometers were extracted with the manufacturer's data software, DataAnalyzer (versions 1.13.5 and 1.13.7). Next, physical activity was calculated in 1-minute intervals for better interpretation. All e-diary and smartphone data from the movisensXS browser app were then downloaded to synchronize and merge them with the accelerometer data, using the manufacturer's software DataMerger (version 1.8.0). After the successful combination of all relevant data, we applied predetermined validity criteria

of the data, regarding 8-hour wear-time, location accuracy, and answered prompts.

In a final step, spatial information and data on environmental features (eg, greenness and noise pollution) were added for further analyses. Data fusion of the e-diary and accelerometer data with GIS information require that GPS data are linked to the concomitant e-diary via a unique ID composed of the participant ID and the time stamp of the e-diary or accelerometer data.

Participant Recruitment and Study Procedure

For participant recruitment, 3000 letters were distributed to preselected urban households with an invitation to participate in a web-based, cross-sectional questionnaire. The data collected in the survey were used for purposes not emphasized in this study, except for the final step of the web-based questionnaire. Individuals, who want to and who fit the inclusion criteria could join the walking-triggered e-diary assessment. Inclusion criteria were for individuals to understand the German language (as the questionnaires were in German), to be older than 18 years, to live in an urban or suburban area, and to have no mental or physical health conditions that would restrain them from being physically active (eg, injuries or depression). Furthermore, participants were offered a profile of their activity pattern during the duration of the assessment and an incentive of €50 (US $51.86) if they provide data on at least 7 days. Potential participants received information regarding the study and the procedure and instructions on how to wear and use the sensors and smartphones before written consent was obtained. In addition, a video was provided with the same information for further clarification.

After a confirmatory phone call, participants were equipped with a smartphone and a sensor via a mailed package. In the cases where no questions came up, the participants were able to start the study by themselves. Participants were consecutively recruited from July to December 2020 (no lockdown conditions during this time owing to COVID-19 restrictions in Germany). The study period per participant lasted 10 consecutive days with an option to extend the study voluntarily. During the assessed time period, between 6 AM and 10 PM, participants received walking-triggered prompts, requesting them to answer short e-diaries that were shown on their smartphones. In addition, they randomly received 3 further e-diaries over the course of the day (between 10 AM and 10 PM, with a minimum 2.5-hour interval in between). After completing the study, the participants returned the smartphone-sensor combination via prepaid mail packages.

Ethical Considerations

The study received full ethical approval from the University of Konstanz (IRB18KN010-004; October 29, 2018).

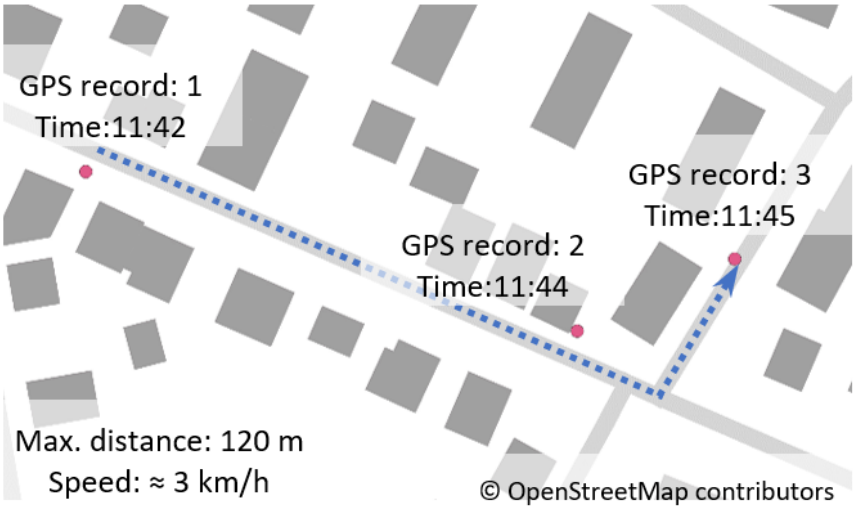
Statistical Analyses

Two different approaches were used to evaluate the accuracy of the walking trigger algorithm. First, we assessed accordance between walking routes and walking-triggered e-diaries (sensitivity). To reconstruct as many walking routes as possible, we used all recorded GPS signals, with the following conditions: GPS signals with a time span of 60-120 seconds, each GPS-signal has an accuracy of at least 30 m, mean speed of the total route is less than 10 km/hour, and a minimum walking distance of 100 m.

We modeled the walking routes with the statistical analysis software R and the OpenRouteService routing system (HeiGIT gGmbH). The reconstructed walking routes are calculated independently from the walking-triggered e-diaries and were used as reference data. We checked for how many walking routes were e-diaries triggered (Figure 1).

Furthermore, we compared how many trigger-based versus time-based prompts started an e-diary during a walking route to give an impression of the added value of using the walking-triggered approach instead of a time-based approach. In the next step, we focused on the specificity of the walking trigger approach and verified whether a participant was walking outside immediately before a walking-triggered e-diary was prompted (only main e-diaries were used). For this purpose, we used the self-reports of the main e-diaries. In addition, we compared steps and accelerometer data before main e-diaries were triggered between situations in which participants self-reported that they were or were not walking outdoors.

Figure 1. Survey locations along the walking routes (adapted from ©OpenStreetMap, licensed under the Open Data Commons Open Database License [37]).



Results

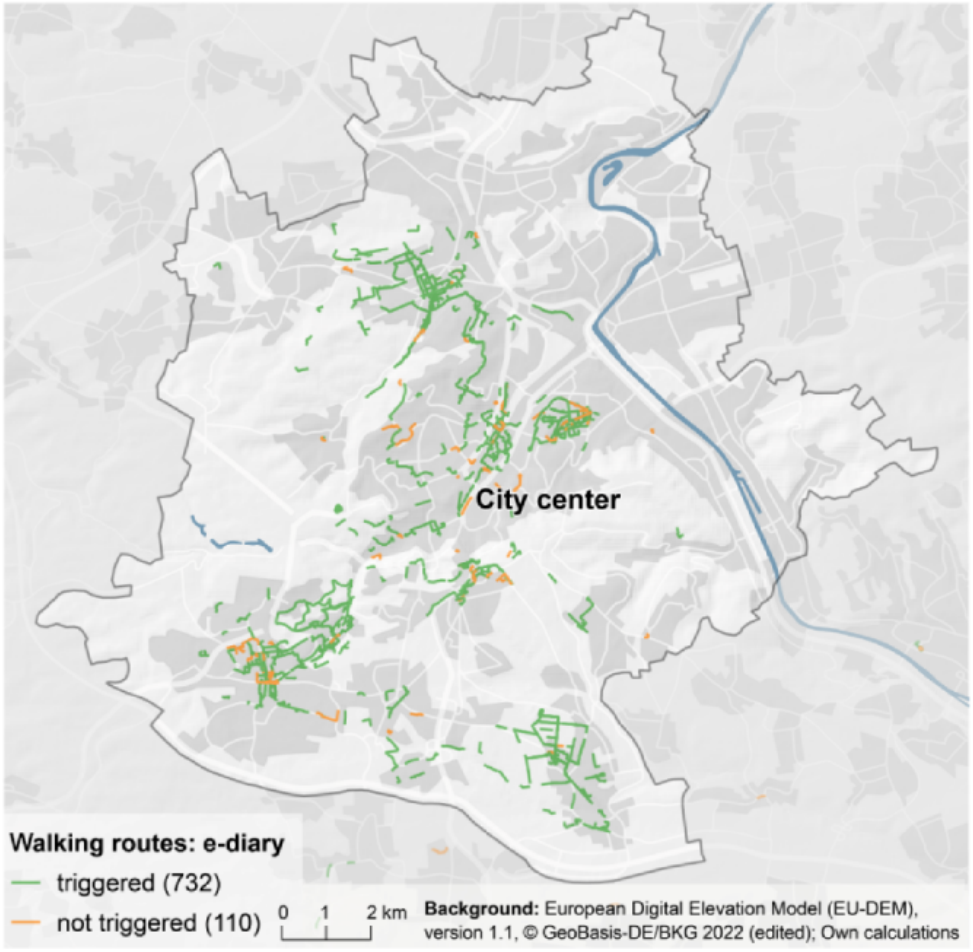
Descriptive Statistics

We could analyze data of 67 of 79 recruited participants (excluded participants, n=12; technical problems, n=4; wear time<8 hours, n=7; opted out of the study, n=1). The excluded participants (n=12) did not differ significantly in age (mean 40.4, SD 13 years) and sex (sex distribution of the excluded participants was 50% each for women and men) from the included participants. The included participants (N=67) had a mean age of 41.0 (SD 14.6) years, and 50.7% were female. 82.1% of participants had a high school certificate, and the average BMI was 23.9 (SD 3.8). Participants provided a total of 632 days of sensor data (mean per participant 9.4, SD 1.8, range 3-12 days; minimum wear time of 8 hours per day; mean wear time per participant per day of 11 hours 59 minutes, SD 2 hours 46 minutes). At least 1 e-diary was triggered on 452 days, which yielded an average of 6.7 (SD 2.9) diaries per day. The walking-trigger produced 3283 e-diary prompts within the study period (mean per participant 49, SD 44.3, range 1-179). Of a total of 3283 prompts, 2258 (68.8%) were answered, including 1206 of 1840 main e-diaries (mean per participant per day 2.8, SD 2.1, range 1-20) and 1052 of 1443 repeated e-diaries (mean per participant per day 2.5, SD 1.5, range 1-21). The resulting compliance rates were 65.5% (1206/1840; main e-diary) and 72.9% (1052/1443; repeated e-diary), respectively. In addition, participants received 1955 randomly prompted e-dairies (mean per participant per day 2.9, SD 0.1, range 1-3). Of a total of 1955 randomly prompted e-diaries, 1479 (75.6%) were completed (mean per participant 22.1, SD 7.1, range 1-35).

Did the Walking Trigger Algorithm Trigger an e-Diary During Walking Routes (Sensitivity)?

We reconstructed a total of 842 walking routes from recorded GPS signals. In 732 situations, at least 1 e-diary (*main* or *main and repeated*) was triggered during or immediately after a walking route (maximum 2 minutes). This yielded an overall accuracy of 86.9% (732/842; Figure 2). Using a time-based trigger (3 times randomly throughout the day), which was often used in previous GEMA studies, led to an e-diary during walking episodes in only 29 of 842 (3.4%) situations.

Figure 2. Walking routes in Stuttgart identified via GPS (adapted from ©GeoBasis-DE / BKG (2020) [38], licensed under Data licence Germany – attribution – Version 2.0 [39] and Copernicus Services that delivers data with funding by the European Union [40]). Note: the authors used their own color scheme and calculations.

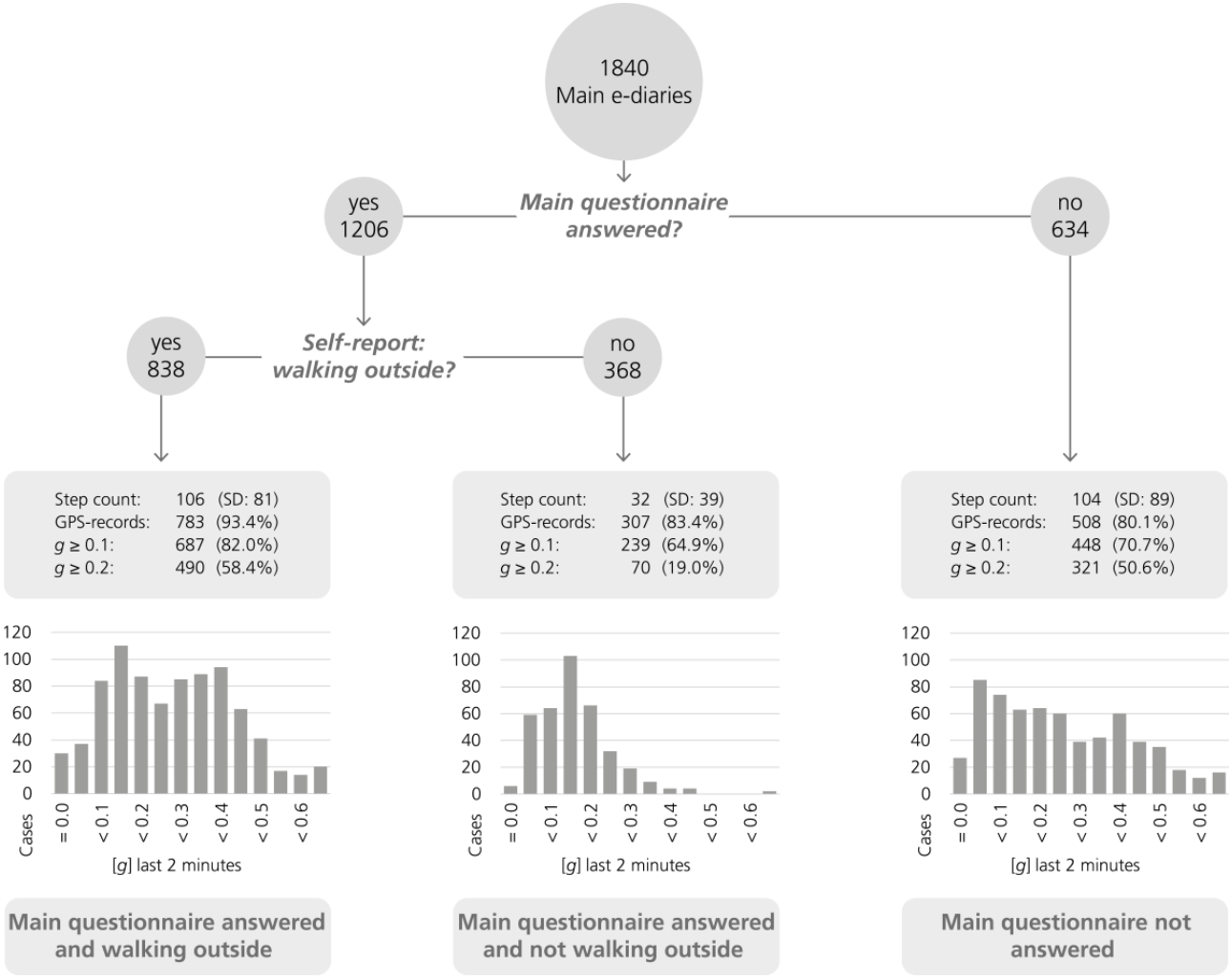


Did the Participants Walk Outdoors Immediately Before a Main e-Diary Was Triggered (Specificity)?

In total, 1840 main e-diaries were triggered, of which 1206 were answered. In 838 of the 1206 (69.5%) situations, the participants indicated via self-report that they were (currently) walking outdoors (Figure 3).

In these 838 situations, an average of 106 (SD 81) steps were taken in the 2 minutes before the e-dairy, and the acceleration was >0.1 *g* in 82.0% and ≥0.2 *g* in 58.4% of these situations. GPS records were present 783 out of 838 (93.4%) times. However, in 368 of the 1206 (30.5%) situations, participants indicated via self-report that they were not (currently) walking outdoors. In the 2 minutes before these triggered e-diaries, an average of 32 (SD 39) steps were taken, and in 64.9% and 19.0% of cases, the acceleration was >0.1 *g* and ≥0.2 *g*, respectively. GPS records were present 307 times (307/368, 83.4%). Unanswered main questionnaires amounted to a total of 634. The following movement data for these situations can be reported: an average of 104 (SD 89) steps were recorded in these situations, and in 70.7% and 50.6% of cases, the acceleration was at least 0.1 *g* and ≥0.2 *g*, respectively. GPS records were present 508 out of 634 (80.1%) times.

Figure 3. Comparison of activity level (step count and movement acceleration in *g*) between situations in which participants self-reported walking or not walking outdoors and unanswered situations.



Discussion

Principal Findings

This study introduces walking-triggered e-diaries as an innovative measurement approach for walking behavior and evaluates its accuracy for this purpose. This new method allows for the analysis of time-varying associations to deepen our understanding of person-place interactions. It can potentially assess data of the person (eg, mental health criteria including momentary affective states) as well as the social and physical environment (eg, interacting with others and experiencing greenness) during the very situations in which a person walks outdoors.

According to the results of the reconstructed walking routes about sensitivity, in 86.9% of these walking episodes, participants received a triggered e-diary. This finding shows that walking-triggered e-diaries had high accuracy in capturing data during episodes of walking outdoors. Compared to a time-based assessment, which was usually used in previous studies (eg, GEMA studies), the walking trigger is more sensitive and leads to a relevant higher number of e-diaries during episodes of walking outdoors (732 vs 29, respectively). Nevertheless, there are some reconstructed walking routes where

participants did not receive a triggered e-diary. As the trigger depends on movement acceleration and not explicitly on gait velocity, it is possible that the participant's movement acceleration had been lower than the set threshold value of 0.1 *g*. Detection of gait velocity or inclusion of steps in the algorithm of the walking trigger potentially mitigate this inaccuracy.

The assessment approach of walking-triggered e-diaries extends former studies, which have already shown that, for instance, more green space during active episodes is associated with lower stress levels [28]. However, to improve our knowledge about environmental effects on participants' behavioral outcomes and mental health, relevant constructs and variables should be assessed in situ, precisely in the situation during which a person experienced his or her environment [17,41]. With a time-based assessment (eg, e-diaries were prompted 4 times a day), it is difficult to explicitly capture episodes of walking outdoors. For instance, former GEMA studies showed that almost more than two-thirds of their GEMA responses occurred at home and not during episodes of walking outdoors (eg, 72% of home responses) [28]. This high number of home responses instead of assessments during outdoor walking episodes was to be expected in time-based assessments, as physically active episodes are rare during everyday life; this is because most people worldwide are insufficiently physically active and do not meet the recommended activity level of 150 minutes of moderate activity per week [42]. Calculated at 12 hours for 1 day of assessment, most people only use 3% of their time for physical activity. It is therefore highly probable that time-based assessments miss some episodes of walking outdoors. Walking-triggered e-diaries enhance the possibility to assess data during episodes of walking outdoors.

The second analysis was related to the specificity of the walking trigger assessment. It used participants' self-reports to provide an impression of false positive triggered e-diaries and used movement acceleration and step counts to estimate activity levels immediately before the e-diary was triggered. However, participants reported, in 30.5% of the walking-triggered e-diaries, that they were currently not walking outdoors. According to the objectively measured activity data during these situations, there had been sufficient movement acceleration, and the accelerometer identified on average 32 steps per person during the last 2 minutes immediately before the prompt. Thus, owing to the objectively measured activity levels, the walking trigger algorithm's prompt was accurate. Participants, however, did not actively indicate these movements as outdoor walking behavior. A possible explanation is that a person was strolling around his or her home or performed some leisure time activities such as slow roller-skating or slow bicycling. Such activities can lead to a nonstationary movement acceleration greater than 0.1 *g* (as GPS and acceleration condition is true by value), but the person did not interpret the movement as walking outdoors. Overall, our findings emphasize the importance of assessing participants' interpretations and behavior patterns apart from objective measures of physical activity. Furthermore, the histograms in Figure 3 provide an impression of how the movement acceleration of physical activity is related to self-reported (non-) stationary active episodes. When participants denied that they were engaged in walking outdoors, acceleration was mostly lower (19%; >0.2 *g*; 32 steps) than that in situations where participants confirmed activity (58.4%; >0.2 *g*; 106 steps). This observation leads to the conclusion that an increase in the activity threshold to 0.2 *g* of movement acceleration would help reduce false positive e-diaries but would also reduce sensitivity (ie, fewer walking episodes will be identified).

Our study design with walking-triggered e-diaries reveals acceptable compliance rates of 65% for the triggered main e-diaries and 72% for the repeated ones. These numbers are in line with those of other studies assessing walking triggers [43]. However, the compliance rate is difficult to compare because it depends heavily on the study design and the burden participants have with providing self-reports several times per day for several consecutive days [44].

Although the data set included 632 days of sensor data, at least 1 walking-triggered e-diary was triggered on only 452 days. We therefore have 180 days of no data on outdoor walking activity (an average of 2.7 days per participant during a 10-day assessment period). This number seems quite high. Part of the explanation is, however, that data assessments have been conducted during a time where COVID-19 restrictions imposed heavy constraints on participants to meet other people, and working from home was recommended or even mandatory for some employees. Several studies showed that physical activity levels and walking activities were reduced during this time [45].

Strengths, Limitations, and Implications for Future Studies

The strength of our new measurement approach of walking-triggered e-diaries is 2-fold: first, it incorporates accelerometers to assess physical activity directly and combines activity level and mobile positioning of the cellphone via GPS and transmission towers to identify walking outdoors. Second, it uses these identified outdoor walking episodes to start an e-diary. Such a design expanded existing activity-triggered assessments because it can differentiate between indoor activities or stationary outdoor activities and walking behavior. Thus, person-place interactions with the physical and social environment, participants' walking behavior, and how these constructs might moderate health-enhancing effects can be examined exactly in those situations in which a person is walking outdoors. This design also minimizes the burden for participants, since they receive questionnaires predominantly in situations in which a relevant behavior occurs. In addition, this study highlights the importance of assessing participants' interpretation of their current behavior when assessing and analyzing person-place interactions.

Nevertheless, walking-triggered e-diaries do face some challenges. First, the topography of cities limits the accuracy of GPS signals because buildings restrict their transmission to satellites [46]. Second, participants had to wear an extra sensor (accelerometer) and were equipped with a smartphone for the study. To increase the compliance rate and reduce missing data, technical solutions in which participants install an app on their own smartphone would be preferable. Third, expertise in data handling is an essential precondition for processing (eg,

movement acceleration, GPS signals, and merging of different data sets) and analyzing (eg, multilevel modeling) the results. It requires interdisciplinary competencies (eg, spatial science, sport and movement science, and psychology) to facilitate such studies in the field of urban health.

In future studies, ambulatory assessments with walking-triggered e-diaries could be enriched with further sensor-based data (eg, momentary noise or heart rate and electrodermal activity) to enhance knowledge about different momentary exposures, experiences, and behavior during active mobility [17]. Furthermore, walking-triggered e-diaries can be used to develop just-in-time adaptive intervention by delivering support (eg, feedback prompts) when users need them most (eg, encourage to walk for a longer distance or start to walk after a longer episode of inactivity) [35].

Conclusions

This study presents an accuracy analysis for a new method in ambulatory assessments. Our results reveal that walking-triggered e-diaries are suitable to collect different data of individuals' current experiences (eg, affective states, social interactions, and environmental features) during episodes of walking outdoors. Analyzing time-varying data during these predefined situations provides new insights into person-place interactions; that is, how physical and social environments shape individual experiences when walking outdoors. Based on the increasing interest in the design of health-promoting urban environments in recent years, such a design can offer substantial methodological progress. However, analyzing person-place interactions requires expert knowledge on data handling (eg, GPS and walking routes, movement acceleration, and e-diaries) and the management of intensive longitudinal studies in interdisciplinary teams.

Acknowledgments

The study was funded by the Deutsche Forschungsgemeinschaft (KA 3279/4-1).

Authors' Contributions

MK conceptualized and designed the study, formulated the research question, interpreted the data, and drafted the manuscript. LB conducted the study and drafted the Methods and Results sections of the manuscript. JS analyzed the data and drafted the Results section of the manuscript. CN conceptualized and designed the study and substantially revised the manuscript. SF substantially revised the manuscript.

Conflicts of Interest

None declared.

References

1. Dutheil F, Pélangeon S, Duclos M, Vorilhon P, Mermillod M, Baker JS, et al. Correction to: protective effect on mortality of active commuting to work: a systematic review and meta-analysis. Sports Med 2020 Dec;50(12):2251. [doi: 10.1007/s40279-020-01371-z] [Medline: 33104996]
2. Pucher J, Buehler R, Bassett DR, Dannenberg AL. Walking and cycling to health: a comparative analysis of city, state, and international data. Am J Public Health 2010 Oct;100(10):1986-1992. [doi: 10.2105/AJPH.2009.189324] [Medline: 20724675]
3. Avila-Palencia I, Cole-Hunter T, de Nazelle A, Donaire D, Jerrett M, Nieuwenhuijsen M. A43 The relationship between bicycle commuting and levels of self-perceived stress. J Transp Health 2015 Jun;2(2):S27. [doi: 10.1016/j.jth.2015.04.531]
4. Smith TL, Masaki KH, Fong K, Abbott RD, Ross GW, Petrovitch H, et al. Effect of walking distance on 8-year incident depressive symptoms in elderly men with and without chronic disease: the Honolulu-Asia Aging Study. J Am Geriatr Soc 2010 Aug;58(8):1447-1452 [FREE Full text] [doi: 10.1111/j.1532-5415.2010.02981.x] [Medline: 20670378]
5. Marques A, Peralta M, Henriques-Neto D, Frasquilho D, Rubio Gouveira É, Gomez-Baya D. Active Commuting and Depression Symptoms in Adults: A Systematic Review. Int J Environ Res Public Health 2020 Feb 06;17(3) [FREE Full text] [doi: 10.3390/ijerph17031041] [Medline: 32041331]
6. Dinu M, Pagliai G, Macchi C, Sofi F. Active commuting and multiple health outcomes: a systematic review and meta-analysis. Sports Med 2019 Mar;49(3):437-452. [doi: 10.1007/s40279-018-1023-0] [Medline: 30446905]
7. Martin A, Goryakin Y, Suhrcke M. Does active commuting improve psychological wellbeing? Longitudinal evidence from eighteen waves of the British Household Panel Survey. Prev Med 2014 Dec;69:296-303 [FREE Full text] [doi: 10.1016/j.ypmed.2014.08.023] [Medline: 25152507]
8. Humphreys DK, Goodman A, Ogilvie D. Associations between active commuting and physical and mental wellbeing. Prev Med 2013 Aug;57(2):135-139 [FREE Full text] [doi: 10.1016/j.ypmed.2013.04.008] [Medline: 23618913]
9. Nobis C, Kuhnimhof T. Mobilität in Deutschland − MiD Ergebnisbericht. 2018. URL: https://elib.dlr.de/125879/ [accessed 2022-11-11]
10. Keall MD, Shaw C, Chapman R, Howden-Chapman P. Reductions in carbon dioxide emissions from an intervention to promote cycling and walking: a case study from New Zealand. Transp Res D Transp Environ 2018 Dec;65:687-696. [doi: 10.1016/j.trd.2018.10.004]

11. Lederbogen F, Kirsch P, Haddad L, Streit F, Tost H, Schuch P, et al. City living and urban upbringing affect neural social stress processing in humans. Nature 2011 Jun 22;474(7352):498-501. [doi: 10.1038/nature10190] [Medline: 21697947]
12. Ventriglio A, Torales J, Castaldelli-Maia JM, De Berardis D, Bhugra D. Urbanization and emerging mental health issues. CNS Spectr 2021 Feb;26(1):43-50. [doi: 10.1017/S1092852920001236] [Medline: 32248860]
13. Abbott A. Stress and the city: Urban decay. Nature 2012 Oct 11;490(7419):162-164. [doi: 10.1038/490162a] [Medline: 23060170]
14. Saelens BE, Sallis JF, Frank LD. Environmental correlates of walking and cycling: findings from the transportation, urban design, and planning literatures. Ann Behav Med 2003;25(2):80-91. [doi: 10.1207/S15324796ABM2502_03] [Medline: 12704009]
15. de Nazelle A, Nieuwenhuijsen MJ, Antó JM, Brauer M, Briggs D, Braun-Fahrlander C, et al. Improving health through policies that promote active travel: a review of evidence to support integrated health impact assessment. Environ Int 2011 May;37(4):766-777. [doi: 10.1016/j.envint.2011.02.003] [Medline: 21419493]
16. Chaix B. Mobile sensing in environmental health and neighborhood research. Annu Rev Public Health 2018 Apr 01;39:367-384 [FREE Full text] [doi: 10.1146/annurev-publhealth-040617-013731] [Medline: 29608869]
17. Chaix B. How daily environments and situations shape behaviors and health: momentary studies of mobile sensing and smartphone survey data. Health Place 2020 Jan;61:102241. [doi: 10.1016/j.healthplace.2019.102241] [Medline: 31784331]
18. Dixon BN, Ugwoaba UA, Brockmann AN, Ross KM. Associations between the built environment and dietary intake, physical activity, and obesity: a scoping review of reviews. Obes Rev 2021 Apr;22(4):e13171 [FREE Full text] [doi: 10.1111/obr.13171] [Medline: 33369097]
19. Tcymbal A, Demetriou Y, Kelso A, Wolbring L, Wunsch K, Wäsche H, et al. Effects of the built environment on physical activity: a systematic review of longitudinal studies taking sex/gender into account. Environ Health Prev Med 2020 Nov 27;25(1):75 [FREE Full text] [doi: 10.1186/s12199-020-00915-z] [Medline: 33246405]
20. Hunter RF, Cleland C, Cleary A, Droomers M, Wheeler BW, Sinnett D, et al. Environmental, health, wellbeing, social and equity effects of urban green space interventions: a meta-narrative evidence synthesis. Environ Int 2019 Sep;130:104923 [FREE Full text] [doi: 10.1016/j.envint.2019.104923] [Medline: 31228780]
21. Núñez-González S, Delgado-Ron JA, Gault C, Lara-Vinueza A, Calle-Celi D, Porreca R, et al. Overview of "Systematic Reviews" of the built environment's effects on mental health. J Environ Public Health 2020;2020:9523127 [FREE Full text] [doi: 10.1155/2020/9523127] [Medline: 32256618]
22. Ebner-Priemer UW, Trull TJ. Ambulatory assessment. Eur Psychol 2009 Jan;14(2):109-119. [doi: 10.1027/1016-9040.14.2.109]
23. Kanning MK, Ebner-Priemer UW, Schlicht WM. How to investigate within-subject associations between physical activity and momentary affective states in everyday life: a position statement based on a Literature overview. Front Psychol 2013;4:187 [FREE Full text] [doi: 10.3389/fpsyg.2013.00187] [Medline: 23641221]
24. Reichert M, Giurgiu M, Koch E, Wieland LM, Lautenbach S, Neubauer AB, et al. Ambulatory assessment for physical activity research: state of the science, best practices and future directions. Psychol Sport Exerc 2020 Sep;50 [FREE Full text] [doi: 10.1016/j.psychsport.2020.101742] [Medline: 32831643]
25. Bolger N, Laurenceau JP. Intensive longitudinal methods: an introduction to diary and experience sampling research. New York, NY: The Guilford Press; 2013:A.
26. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4:1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]
27. Kirchner TR, Shiffman S. Spatio-temporal determinants of mental health and well-being: advances in geographically-explicit ecological momentary assessment (GEMA). Soc Psychiatry Psychiatr Epidemiol 2016 Sep;51(9):1211-1223 [FREE Full text] [doi: 10.1007/s00127-016-1277-5] [Medline: 27558710]
28. Mennis J, Mason M, Ambrus A. Urban greenspace is associated with reduced psychological stress among adolescents: a geographic ecological momentary assessment (GEMA) analysis of activity space. Landsc Urban Plan 2018 Jun;174:1-9 [FREE Full text] [doi: 10.1016/j.landurbplan.2018.02.008] [Medline: 29576668]
29. Boettner B, Browning CR, Calder CA. Feasibility and validity of geographically explicit ecological momentary assessment with recall-aided space-time budgets. J Res Adolesc 2019 Sep;29(3):627-645 [FREE Full text] [doi: 10.1111/jora.12474] [Medline: 31573764]
30. Fernandes A, Van Lenthe FJ, Vallée J, Sueur C, Chaix B. Linking physical and social environments with mental health in old age: a multisensor approach for continuous real-life ecological and emotional assessment. J Epidemiol Community Health 2021 May;75(5):477-483 [FREE Full text] [doi: 10.1136/jech-2020-214274] [Medline: 33148684]
31. Ebner-Priemer UW, Koudela S, Mutz G, Kanning M. Interactive multimodal ambulatory monitoring to investigate the association between physical activity and affect. Front Psychol 2012;3:596 [FREE Full text] [doi: 10.3389/fpsyg.2012.00596] [Medline: 23355828]
32. Giurgiu M, Niermann C, Ebner-Priemer U, Kanning M. Accuracy of sedentary behavior-triggered ecological momentary assessment for collecting contextual information: development and feasibility study. JMIR Mhealth Uhealth 2020 Sep 15;8(9):e17852 [FREE Full text] [doi: 10.2196/17852] [Medline: 32930668]

33. Törnros T, Dorn H, Reichert M, Ebner-Priemer U, Salize H, Tost H, et al. A comparison of temporal and location-based sampling strategies for global positioning system-triggered electronic diaries. Geospat Health 2016 Nov 21;11(3):473 [FREE Full text] [doi: 10.4081/gh.2016.473] [Medline: 27903053]
34. Anastasopoulou P, Tubic M, Schmidt S, Neumann R, Woll A, Härtel S. Validation and comparison of two methods to assess human energy expenditure during free-living activities. PLoS One 2014;9(2):e90606 [FREE Full text] [doi: 10.1371/journal.pone.0090606] [Medline: 24587401]
35. movisens GmbH. URL: https://xs.movisens.com [accessed 2022-11-11]
36. Products. movisens. URL: https://www.movisens.com/en/products/ [accessed 2022-11-14]
37. Copyright and License. OpenStreetMap. URL: openstreetmap.org/copyright [WebCite Cache ID openstreetmap.org/copyright]
38. Bundesamt für Kartographie und Geodäsie. BKG. URL: https://gdz.bkg.bund.de/index.php/default/digitale-geodaten/digitale-landschaftsmodelle/digitales-landschaftsmodell-1-250-000-ebenen-dlm250-ebenen.html
39. Datenlizenz Deutschland – Namensnennung – Version 2.0. GovData. URL: www.govdata.de/dl-de/by-2-0
40. European Digital Elevation Model (EU-DEM), version 1.1. Copernicus. URL: https://land.copernicus.eu/imagery-in-situ/eu-dem/eu-dem-v1.1?tab=metadata
41. Yi L, Wilson JP, Mason TB, Habre R, Wang S, Dunton GF. Methodologies for assessing contextual exposure to the built environment in physical activity studies: a systematic review. Health Place 2019 Nov;60:102226 [FREE Full text] [doi: 10.1016/j.healthplace.2019.102226] [Medline: 31797771]
42. Guthold R, Stevens GA, Riley LM, Bull FC. Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1·9 million participants. Lancet Global Health 2018 Oct;6(10):e1077-e1086. [doi: 10.1016/s2214-109x(18)30357-7]
43. Kanning M, Niermann C, Ebner-Primer U, Giurgiu M. The context matters - not all prolonged sitting bouts are equally related to momentary affective states: an ambulatory assessment with sedentary-triggered E-diaries. Int J Behav Nutr Phys Act 2021 Aug 14;18(1):106 [FREE Full text] [doi: 10.1186/s12966-021-01170-3] [Medline: 34391442]
44. Stone AA, Broderick JE, Schwartz JE, Shiffman S, Litcher-Kelly L, Calvanese P. Intensive momentary reporting of pain with an electronic diary: reactivity, compliance, and patient satisfaction. Pain 2003 Jul;104(1-2):343-351. [doi: 10.1016/s0304-3959(03)00040-x] [Medline: 12855344]
45. Christensen A, Bond S, McKenna J. The COVID-19 conundrum: keeping safe while becoming inactive. A rapid review of physical activity, sedentary behaviour, and exercise in adults by gender and age. PLoS One 2022;17(1):e0263053 [FREE Full text] [doi: 10.1371/journal.pone.0263053] [Medline: 35085330]
46. Schipperijn J, Kerr J, Duncan S, Madsen T, Klinker CD, Troelsen J. Dynamic accuracy of GPS receivers for use in health research: a novel method to assess GPS accuracy in real-world settings. Front Public Health 2014;2:21 [FREE Full text] [doi: 10.3389/fpubh.2014.00021] [Medline: 24653984]

Abbreviations

AA: ambulatory assessment
EMA: ecological momentary assessment
GEMA: geographically-explicit ecological momentary assessment

Edited by G Eysenbach; submitted 06.05.22; peer-reviewed by M Schneider, J de Brito; comments to author 18.08.22; revised version received 01.09.22; accepted 08.10.22; published 25.11.22.

Please cite as:
Kanning M, Bollenbach L, Schmitz J, Niermann C, Fina S
Analyzing Person-Place Interactions During Walking Episodes: Innovative Ambulatory Assessment Approach of Walking-Triggered e-Diaries
JMIR Form Res 2022;6(11):e39322
URL: https://formative.jmir.org/2022/11/e39322
doi:10.2196/39322
PMID:36427231

Original Paper

Feasibility of a Reinforcement Learning–Enabled Digital Health Intervention to Promote Mammograms: Retrospective, Single-Arm, Observational Study

Amy Bucher[1], PhD; E Susanne Blazek[1], PhD; Ashley B West[1], PhD

Lirio, Knoxville, TN, United States

Corresponding Author:
Amy Bucher, PhD
Lirio
320 Corporate Drive, NW
Knoxville, TN, 37923
United States
Phone: 1 865 839 6158
Email: abucher@lirio.com

Abstract

Background: Preventive screenings such as mammograms promote health and detect disease. However, mammogram attendance lags clinical guidelines, with roughly one-quarter of women not completing their recommended mammograms. A scalable digital health intervention leveraging behavioral science and reinforcement learning and delivered via email was implemented in a US health system to promote uptake of recommended mammograms among patients who were 1 or more years overdue for the screening (ie, 2 or more years from last mammogram).

Objective: The aim of this study was to establish the feasibility of a reinforcement learning–enabled mammography digital health intervention delivered via email. The research aims included understanding the intervention's reach and ability to elicit behavioral outcomes of scheduling and attending mammograms, as well as understanding reach and behavioral outcomes for women of different ages, races, educational attainment levels, and household incomes.

Methods: The digital health intervention was implemented in a large Catholic health system in the Midwestern United States and targeted the system's existing patients who had not received a recommended mammogram in 2 or more years. From August 2020 to July 2022, 139,164 eligible women received behavioral science–based email messages assembled and delivered by a reinforcement learning model to encourage clinically recommended mammograms. Target outcome behaviors included scheduling and ultimately attending the mammogram appointment.

Results: In total, 139,164 women received at least one intervention email during the study period, and 81.52% engaged with at least one email. Deliverability of emails exceeded 98%. Among message recipients, 24.99% scheduled mammograms and 22.02% attended mammograms (88.08% attendance rate among women who scheduled appointments). Results indicate no practical differences in the frequency at which people engage with the intervention or take action following a message based on their age, race, educational attainment, or household income, suggesting the intervention may equitably drive mammography across diverse populations.

Conclusions: The reinforcement learning–enabled email intervention is feasible to implement in a health system to engage patients who are overdue for their mammograms to schedule and attend a recommended screening. In this feasibility study, the intervention was associated with scheduling and attending mammograms for patients who were significantly overdue for recommended screening. Moreover, the intervention showed proportionate reach across demographic subpopulations. This suggests that the intervention may be effective at engaging patients of many different backgrounds who are overdue for screening. Future research will establish the effectiveness of this type of intervention compared to typical health system outreach to patients who have not had recommended screenings as well as identify ways to enhance its reach and impact.

(JMIR Form Res 2022;6(11):e42343) doi:10.2196/42343

KEYWORDS

artificial intelligence; reinforcement learning; feasibility studies; mammograms; nudging; behavioral intervention; digital health; email; health equity; cancer screening

Introduction

In the United States, breast cancer is the second most common type of cancer among women [1]. Globally, it represented 11.7% of all new cancer cases in 2020, accounting for 1 in 4 cancer diagnoses among women [2]. Mammograms are a valuable tool in detecting breast cancer early, when treatment options may be less invasive, intensive [3], and costly [4], and are associated with lower frequencies of advanced and fatal breast cancer [5]. Many patients do not generally adhere to national guidelines on the frequency and timing of preventative care [6]. Adherence to recommended mammogram screenings is no different. According to the Behavioral Risk Factor Surveillance System's 2020 US data, only 78.2% of women between the ages of 50 and 74 years had a mammogram in the past 2 years [7]. The COVID-19 pandemic further exacerbated the gap between recommended and attended screenings, with mammogram rates falling dramatically in April 2020 compared to April 2019 [8]. While rates have since rebounded to close to prepandemic levels, it is estimated that it could take as long as 22 weeks to clear the backlog of delayed mammograms [9]. While there is some debate over appropriate mammogram usage given drawbacks associated with overdiagnosis [10], the gap between recommended and actual screening behavior is likely to persist regardless of adjustments to the recommendations, suggesting the need for behavioral interventions targeting those who remain overdue for mammograms.

Evidence suggests that such behavioral interventions are most effective when they address a comprehensive set of barriers to performing a health behavior [11-13], apply different behavior change ingredients to overcome each barrier [12-14], and personalize these ingredients to each person as their barriers dynamically change over time [15,16]. Tailoring digital health messages overcomes person-specific barriers [15,17], facilitates the behavior [15,17], and, most notably, improves health outcomes [14-16]. However, personalizing behavior change ingredients to people's changing barriers over time involves reassessing barriers [18], which is labor intensive, costly, and not scalable [13]. Sophisticated technologies, such as artificial intelligence (AI), offer promise to overcome some of these limitations [19], but these are not yet widely used.

It is important to note that while digital health presents a promising means to deliver interventions at scale, not all people are equally likely to use technology for health-related purposes. For example, Black and Hispanic people engage less in digital health than their White counterparts [20], and many people in both rural and urban areas lack broadband internet access to support more data-intensive applications [21]. Email remains a relatively accessible modality to deliver behavioral interventions, being widely used across racial and ethnic groups [20] and requiring much less data than an app to access on a computer or mobile device. Emails are also a typical method for health systems to communicate with patients, meaning little additional technological support is required to use them in that environment, and patients are familiar with this type of digital interaction with their health system.

The purpose of this study is to assess the feasibility of developing and implementing a digital health intervention incorporating reinforcement learning, a type of AI, to personalize email content in order to increase mammography scheduling and attendance among patients of a large health system who are significantly overdue for their recommended screenings. This retrospective, single-arm, observational study explores the reach of the intervention within the patient population, outcomes related to engagement, and outcomes related to the target behaviors of scheduling and attending mammograms.

Methods

Background

This study had several purposes. First, we sought to establish the feasibility of implementing a behavioral science–informed, reinforcement learning–powered digital health intervention intended to increase the scheduling of and attendance at mammograms within a health system. Next, we wanted to understand the reach of such an intervention within the patient population. Finally, the study was intended to identify the behavioral outcomes of mammogram scheduling and attendance associated with use of the intervention.

Ethical Considerations

Solutions IRB, a private institutional review board accredited by the Association for the Accreditation of Human Research Protections Programs, approved analyses of deidentified, aggregated derived data with a waiver of informed consent (study ID: 2021/05/28). Study data were deidentified and anonymous.

Setting and Participants

The intervention was implemented in a large Catholic health system in the Midwestern United States. The implementation focused on patients who were overdue for, and eligible to schedule, a recommended mammogram. Patients were eligible for the intervention if they were female, between 49.5 and 74 years of age, had not had a mammogram in the past 24 months, were subscribed to health system communications, and had a valid email address on file. Patients were excluded if they had a future mammogram scheduled, had a history within the last 12 months of a breast cancer diagnosis or associated surgery, had health maintenance modifiers excluding them from outreach, or indicated participation in hospice, palliative care, or long-term nursing home care.

Data Collection and Rolling Eligibility

At intervention launch, the health system provided a population-level historical data file of all patients from their Epic electronic medical record system to facilitate the establishment of eligibility criteria and set up data integration. This data file included the email addresses for which the system has permission to contact patients about health-related matters. Then, during the study period, the system sent daily data file updates with information on whether patients had scheduled or attended a mammogram (behavioral outcome), as well as changes to age, health status, or other variables affecting

eligibility for communication. Patients whose data rendered them newly eligible or ineligible for the intervention were added or removed to the distribution list accordingly. Eligible individuals who did not schedule and attend a mammogram and who did not unsubscribe continued to receive communications for the duration of the study period. Eligible individuals received up to 40 emails during the 2-year study period.

Data for this study were collected between August 27, 2020, and July 12, 2022.

Intervention

Precision Nudging for mammography is an English-language messaging intervention designed to influence the target behaviors of scheduling and attending a mammogram. The messages are designed to address specific determinants of completing a mammogram, identified through a combination of literature review and primary research with health systems. A sample of these determinants can be found in Textbox 1. Those determinants are then organized using an intervention mapping process [22] that links barriers and facilitators to evidence-based behavior change techniques (BCTs) [23]. The BCTs are operationalized into a set of message components, such as subject lines, body copy paragraphs, and visual illustrations, that form a content library [24]. Interrater agreement of the content (ie, subject lines and body content) was assessed by two trained coders [25] to ensure that each component accurately operationalized the intended BCT; agreement exceeded the acceptability threshold of κ=0.80.

A behavioral reinforcement learning (BRL) algorithm [26-29] then selected components to compile into a complete message based on recipient characteristics. A total of 468 email combinations were possible using the components used in this implementation. Over time, the BRL algorithm optimized message composition based on recipients' past behavioral responses (ie, opening messages, clicking calls to action, and scheduling and attending mammograms) by selecting components that maximize the probability that the recipient will complete the target behaviors. All emails were white labeled so that they appeared to come from the health system. Figure 1 shows some sample assembled messages.

Approximately 30 days prior to launch, we conducted an IP warming exercise intended to establish a reputation for the IP address used to send intervention emails. This minimizes the likelihood that intervention emails will be flagged as spam by the most common email providers, including Gmail and Yahoo.

In order to avoid creating excess demand on the health system, eligible women were randomly assigned to cohorts of approximately 2000 people; the intervention start date was staggered across cohorts. Intervention communications were sent out once per week on Tuesday mornings via a third-party email vendor. Each cohort received one message per week for 5 weeks, with an 8-week pause, and then another pulse of one message per week for 5 weeks. This message patterning was designed to balance intervention exposure and potential notification fatigue. This pattern continued until women scheduled a mammogram, unsubscribed from the intervention, or otherwise became ineligible for continued communication. Figure 2 shows the communication patterns incorporating both cohorts and message timing.

The calls to action to schedule a mammogram were based on the location where each eligible patient received care according to the eligibility data file provided by the health system. For all care locations, patients were provided the appropriate scheduling phone number. Patients with an established patient portal account who received care at a location with online scheduling enabled also received a link to schedule in the portal.

Textbox 1. A sample of the determinants for mammogram scheduling and attendance incorporated into the development of the Precision Nudging intervention.

Intrapersonal barriers

- Low perceived risk
- Fear of diagnosis

Social context barriers

- Lack of social support
- Social norms around mammograms

Environmental context barriers

- Cost
- Scheduling and wait times

Figure 1. Sample assembled messages from the Precision Nudging mammography intervention.

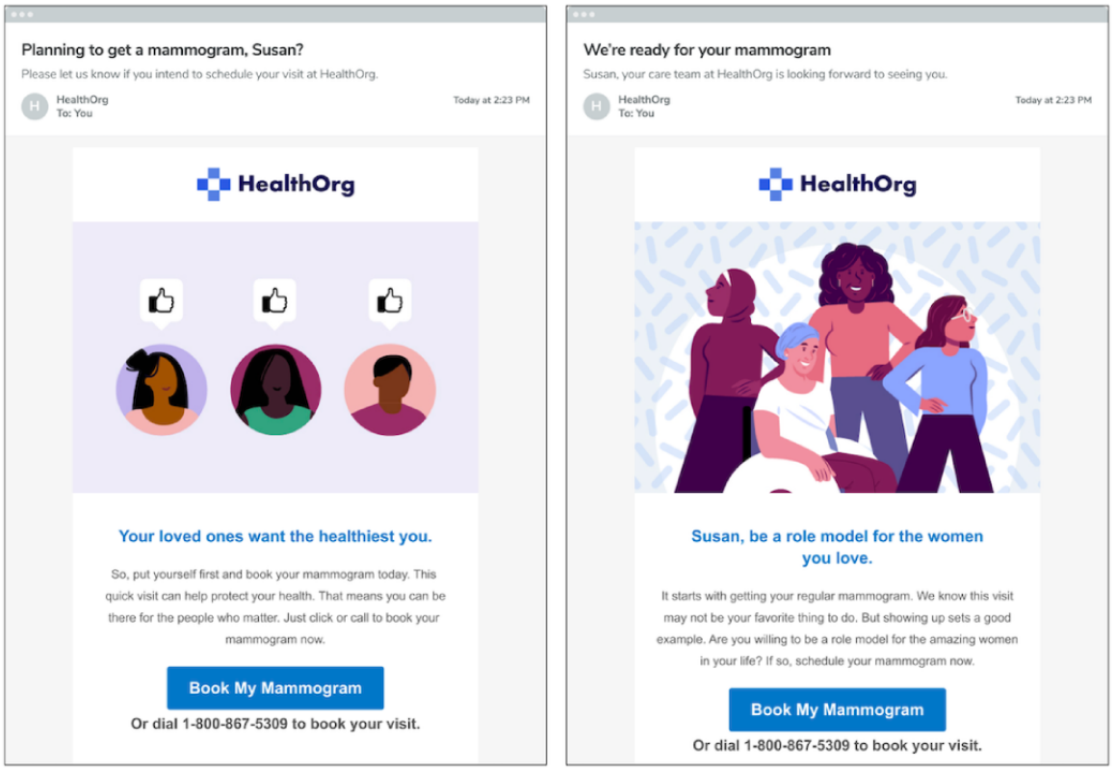
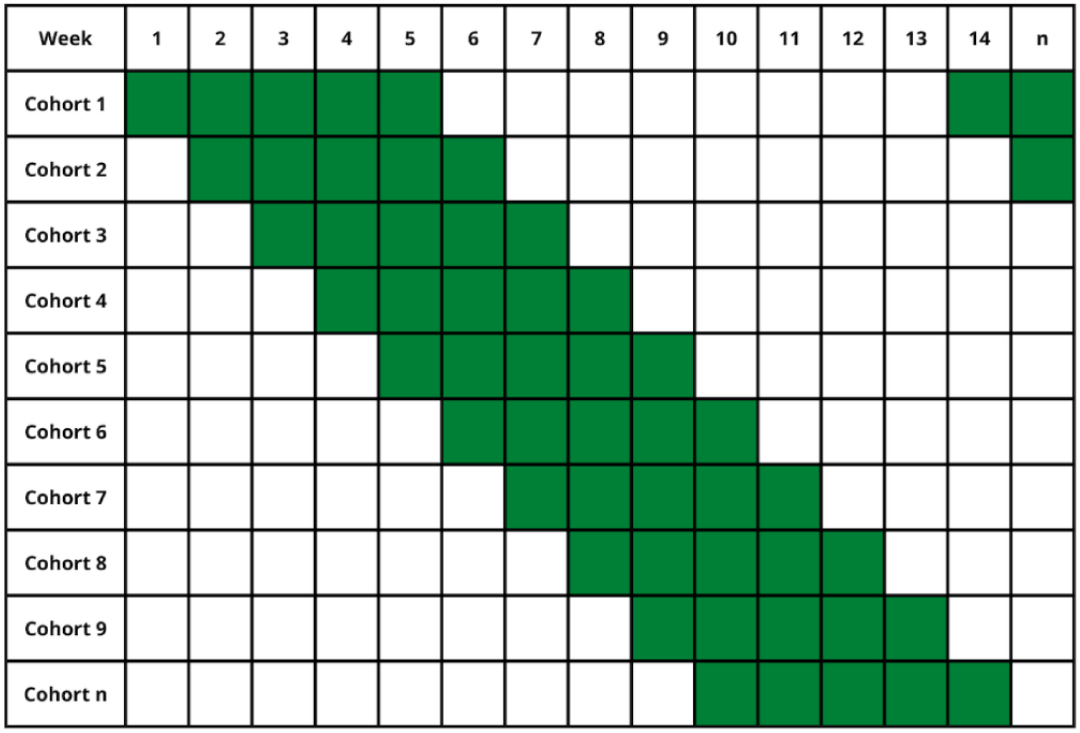


Figure 2. An example of the messaging cadence by cohort showing alternating pulses of one message per week for 5 weeks and an 8-week pause.



Outcomes

We assessed the feasibility of the Precision Nudging intervention by investigating engagement with emails, measured via open rates and clicks on the call to action, and behavioral outcomes, measured as mammograms scheduled and attended. Clicking the call to action was not a prerequisite for successfully scheduling or attending a mammogram, as patients had options to either call for an appointment or self-navigate to their patient portal for online scheduling. We also examined the demographic characteristics of the women who received, engaged with, and took action following a Precision Nudging communication.

Data Analysis

Data were analyzed using the Python programming language. Univariate statistics were used to understand who was reached with the intervention and who responded by scheduling or attending mammograms. Chi-square goodness-of-fit tests were used to understand whether any demographic groups were more likely than others to engage or take action following an intervention message.

Results

Email Deliverability

During the study period, a total of 2,761,270 messages were sent. Overall, 98.91% of emails sent were successfully delivered to a total of 139,164 women (ie, reached). A total of 32.35% of emails were opened at least once. Over the study period, a total of 14,625 women (10.51%) unsubscribed from the intervention messaging.

Demographic Reach

Of the 139,164 women reached with the intervention, the majority (n=121,909, 87.60%) were Caucasian, with the next largest racial group being Black (n=11,879, 8.54%). The most common level of educational attainment was completion of high school (n=61,001, 43.83%), and the majority of message recipients had a household income level under US $100,000 (n=101,164, 72.69%). The mean age of message recipients was 62.13 (SD 7.23) years. One person older than 80 years received an intervention message; this was in error, as eligibility to receive mammography outreach requires patients be 74 years of age or younger. The sample characteristics are summarized in Table 1.

Table 1. Summary of the demographic characteristics of the women receiving intervention messages, with engagement and behavioral responses.

Characteristics	Reached (N=139,164), n (%)	Opened message (n=113,452), n (%)[a]	Clicked call to action (n=15,636), n (%)[a]	Scheduled mammogram (n=34,780), n (%)[a]	Attended mammogram (n=30,637), n (%)[a]
Age (years)					
50-59	58,439 (41.99)	48,240 (82.55)	7825 (13.39)	15,242 (26.08)	13,368 (22.88)
60-69	56,047 (40.27)	45,671 (81.49)	5938 (10.59)	14,489 (25.85)	12,801 (22.84)
70-79	24,677 (17.73)	19,541 (79.19)	1873 (7.59)	5049 (20.46)	4468 (18.11)
80-89	1 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Race					
Caucasian	121,909 (87.60)	99,649 (81.74)	13,272 (10.89)	30,104 (24.69)	26,670 (21.88)
Black	11,879 (8.54)	9370 (78.88)	1773 (14.92)	3431 (28.89)	2848 (23.98)
Asian	659 (0.47)	554 (84.07)	90 (13.66)	230 (34.90)	208 (31.56)
Two or more races	187 (0.13)	137 (73.26)	17 (9.09)	40 (21.39)	36 (19.25)
Other	4530 (3.25)	3742 (82.60)	484 (10.68)	975 (21.52)	875 (19.32)
Educational attainment					
Completed high school	56,767 (40.79)	45,665 (80.44)	6158 (10.85)	15,889 (26.99)	13,920 (24.52)
Completed vocational or technical training	1585 (1.14)	1281 (80.82)	177 (11.17)	425 (26.81)	378 (23.85)
Completed college	44,054 (31.66)	36,389 (82.60)	5232 (11.88)	12,137 (27.55)	10,776 (24.46)
Completed graduate school	17,747 (12.75)	15,217 (85.74)	2115 (11.92)	4397 (24.78)	3924 (22.11)
Unknown	18,901 (13.58)	14,838 (78.50)	1974 (10.44)	5218 (27.60)	4551 (24.08)
Household income level (US $)					
<40,000	35,306 (25.37)	27,129 (76.84)	3564 (10.09)	9176 (25.99)	7966 (22.56)
40,000-69,999	30,374 (21.83)	24,312 (80.04)	3293 (10.84)	7870 (25.91)	6962 (22.92)
70,000-99,999	35,964 (25.84)	29,944 (83.26)	4232 (11.77)	8989 (24.99)	7907 (21.99)
≥100,000	32,608 (23.43)	28,275 (86.71)	4032 (12.37)	7467 (22.90)	6678 (20.48)
Unknown	4912 (3.53)	3792 (77.19)	515 (10.48)	1278 (26.02)	1124 (22.88)

[a]Percentages in these columns are based on values in the “Reached” column.

Engagement With Messages

Overall, of the 139,164 people who received an intervention email, 113,452 (81.52%) opened at least one message. A total of 15,636 people (11.24%) clicked the call to action in at least one of the messages (ie, "clicks"). Engagement by sample characteristics can be found in Table 1.

For those who opened an email, it took an average of 4.46 (SD 5.87) emails for them to do so. For those who clicked an email, it took opening of an average of 4.26 (SD 4.49) emails first; the data do not indicate how many emails were opened prior to people booking a mammogram by phone or via the patient portal without clicking a call to action. Time delays in the scheduling and attendance data prevent calculating the average number of messages prior to behavioral engagement.

Behavioral Outcomes

Among the 139,164 people messaged, 34,780 people (24.99%) scheduled an appointment for a mammogram. At the time of data analysis, 30,637 people (22.02% of the total; 88.09% of those who scheduled) had attended a mammogram. Behavioral response by sample characteristics can be found in Table 1.

Proportionate Engagement and Behavioral Outcomes Across Demographic Subgroups

An important goal in digital health intervention development and research is the achievement of health equity, reached when every person has the opportunity to "attain his or her full health potential" [30]. Statistical methods for analyzing health equity—or equivalent outcomes between groups—largely stem from clinical trial research and tend to focus on comparisons between two groups (ie, two means, two proportions, etc) [31].

One approach to showing equivalence is to carry out a chi-square goodness-of-fit test based on the null hypothesis of no treatment difference [32]. Chi-square goodness-of-fit tests were used to analyze whether any patient subsamples were more likely than others to engage with the intervention emails (ie, opened and clicked). Due to giant sample sizes, all the chi-square goodness-of-fit tests were significant at $P<.001$—with the exception of the chi-square goodness-of-fit test comparing email engagement between race subgroups, which was significant at $P=.008$—obscuring the fact that the expected engagement resembled the observed engagement across demographic subgroups. In samples of this size, the *P* values quickly approach zero [33]. Unlike in clinical trials, where insufficient sample sizes, insensitive outcome measures, or insensitive analyses unduly threaten nonsignificant results [32], this research is challenged by the giant sample size. Basing conclusions on the *P* values of the chi-square goodness-of-fit tests might suggest that there are important statistical differences in engagement between population subgroups, when in fact the results indicate little to no practical difference.

For example, women across age groups demonstrated practically similar levels of engagement with the emails and close-to-expected values from the baseline established by the percentage of women in each group reached. In total, 82.55% of the 58,439 women reached who were aged 50 to 59 years (n=48,240) opened an email, compared to the expected baseline of 81.60%. Even closer to the baseline, 81.49% of the 56,047 women reached who were aged 60 to 69 years (n=45,671) opened an email. Slightly further from the baseline were women aged 70 to 79 years, of whom 79.19% out of 24,677 women reached (n=19,541) opened an email.

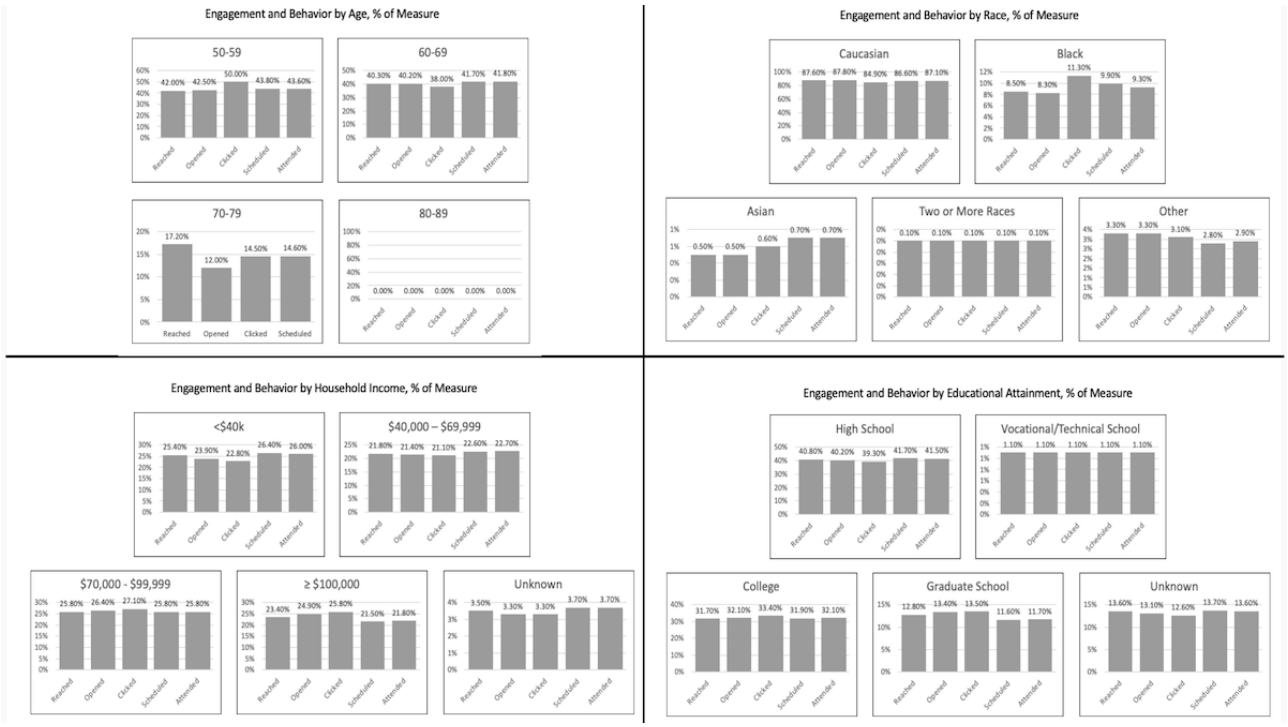
Chi-square goodness-of-fit tests were also used to analyze whether any patient subsamples were more likely than others to schedule or attend a mammogram following receipt of the intervention emails. Similar to the tests conducted for email engagement, due to giant sample sizes, all the chi-square goodness-of-fit tests for behavioral outcomes were significant at $P<.001$. Again, the large sample sizes obscure the fact that the expected engagement resembled the observed engagement across demographic subgroups.

For example, women across income levels demonstrated practically similar levels of behavioral outcomes and close-to-expected values based on the baseline. Out of 35,306 women reached who were making less than US $40,000 per year, 22.56% (n=7966) attended a mammogram, compared to the expected baseline of 25.37%. Out of 30,374 women reached who were making US $40,000 to US $69,999 per year, 22.92% (n=6962) attended a mammogram; also, 21.99% of the 35,964 women reached making US $70,000 to US $99,999 per year (n=7907) attended a mammogram. Slightly further from the baseline were the 32,608 women making more than US $100,000 per year, of which 20.48% (n=6678) attended a mammogram. Out of 4912 women with unknown income, 22.88% (n=1124) attended a mammogram.

An alternative approach to a chi-square goodness-of-fit test based on the null hypothesis of no treatment difference is to preselect a value for the treatment difference that is of practical importance [32,34]. This value should be chosen a priori such that proportions can be considered equivalent if their observed differences do not exceed it. Figure 3 shows the proportion of each demographic subgroup for each measure along the behavioral funnel. We did not choose an a priori value for the treatment difference because of the novel nature of the reinforcement learning–enabled digital health intervention under study. Instead, these exploratory analyses seek to complement the chi-square goodness-of-fit tests and to qualitatively ascertain equitable reach by demonstrating that the proportion of each demographic subgroup remains close to the same for each measure along the behavioral funnel.

For example, 87.60% of the reached population consisted of Caucasian people, so equivalent outcomes would require that close to 87.60% each of the populations who opened at least one message, clicked the call to action, scheduled a mammogram, and attended a mammogram should consist of Caucasian people. For the population who opened at least one message, the proportion of Caucasian people was slightly higher than the expected 87.60%. For the populations who clicked the call to action, scheduled a mammogram, and attended a mammogram, the proportion of Caucasian people was slightly lower than the expected 87.60%.

Figure 3. Bar graphs showing the proportion of each demographic subgroup who engaged and took action after receiving an intervention. Relatively small changes down the funnel from reached to opened to clicked to scheduled to attended suggest, from a practical perspective, proportional response to the intervention for people in that group.



Discussion

Reach and Behavioral Outcomes

This study explored the feasibility of developing and implementing Precision Nudging (ie, a tailored messaging intervention delivered through a BRL algorithm) to promote mammograms to eligible patients of a large health system. Overall, of the 139,164 people who received an intervention email, 113,452 (81.52%) opened at least one message and 15,636 (11.24%) clicked a call to action. A total of 34,780 people (24.99%) scheduled a mammogram, and 30,637 people (22.02% of the total; 88.08% of those who scheduled a mammogram) attended a mammogram. The results support similar health-related pilot studies [35-39] and demonstrate that a reinforcement learning–enabled digital health intervention is capable of reaching women overdue for recommended screenings and prompting behavioral responses, such as scheduling and attending mammograms. The results also demonstrate that engagement and behavioral response are proportional within demographic subgroups of race, age, educational attainment, and household income level.

The intervention used in this study was specifically designed to address a broad range of behavioral determinants, including those more common among underserved groups, like racial and ethnic minorities, and those experiencing poor social determinants of health [24]. A test of whether that approach was successful is assessing whether engagement and behavioral responses from members of those groups are at levels equal to or greater than responses from people who are Caucasian or of higher socioeconomic status. Within the population who was eligible to receive intervention messaging, we saw no practical differences [40] in the frequency at which people engaged with the intervention or took action following a message based on their age, race, educational attainment, or household income. This has implications for the ability of this type of intervention to support health equity in being able to communicate with, and overcome barriers to, preventive care across population subgroups, including many that are historically underserved by digital health, such as non-White people and people of lower educational and income levels. Importantly, this study suggests that such equitable outreach can be done at scale by leveraging email, reinforcement learning, and behavioral science.

That said, we do see differences in the baselines by which people of demographic subgroups were reached by the intervention. For example, 17.73% of the sample were women aged 70 to 79 years, compared to 41.99% who were women aged 50 to 59 years and 40.27% who were women aged 60 to 69 years. While it is likely this is partly due to the fact that eligibility criteria for the intervention was cut off at 74 years of age, curtailing the number of potentially eligible women, it is worth investigating alternative channels to ensure that people who are not frequent email users receive prompts about recommended health behaviors.

It is important to note that the patients included in the feasibility pilot were considered less engaged with their health care by nature of being overdue for their recommended mammograms without a future appointment scheduled. Highly activated patients tend to be compliant with health recommendations, including breast cancer screening [41]. Relatedly, in this sample, 88.08% of women who scheduled a mammogram went on to attend it, suggesting a no-show rate as high as 11.92%. This is higher than the no-show rates found in other research studies looking at a general population (ie, 6.20%) [42]. It seems likely that expanding this intervention to a more heterogeneously

engaged sample (ie, women within 6 to 24 months of their last mammogram) may yield higher mammogram scheduling and completion rates.

Implications and Future Directions

Having established the feasibility of this digital health intervention to improve uptake of mammograms in a health system, the most obvious and urgent next steps are to understand causal effects. It is important to understand whether this intervention improves mammogram uptake compared to standard of care or alternatives, such as a static reminder message. This research may be accomplished via a randomized controlled trial or quasi-experimental implementation (eg, comparing synchronous mammogram rates between two similar health systems or markets where one uses the intervention and the other does not). It also may be fruitful to look at historical screening behaviors among the patients eligible for the intervention—or a yoked sample of similar patients—to establish the incremental influence of Precision Nudging on mammography behaviors.

Future research should investigate the economics of a behavioral intervention such as this one to increase mammography uptake in a health system. Given the potential of mammography to detect breast cancers at an earlier stage where treatment is less costly, widespread implementation of this sort of intervention may yield observable return on investment at the health system level over time, especially if used with unengaged patients who may have historically skipped recommended screenings and checkups. Understanding the economic impact of mammography interventions will help health systems determine whether and when to implement such programs as part of their prevention and disease management portfolios. As this research will necessarily account for cost savings associated with early detection as well as expenses associated with false positives, it may also help to clarify the costs and benefits of annual mammograms, in general, and for specific demographic groups.

Another promising area of future study is the use of behavioral interventions to improve operational efficiency in health systems. A potential drawback to patient-directed behavioral interventions is that they may increase provider workload. The Precision Nudging intervention was designed to have limited impact on clinician workflow. Patient data were automatically captured from the medical record without additional steps from providers, and communication timing and frequency were considered in terms of provider capacity. The intervention also accommodates message throttling to help mitigate excess demand on screening centers. We believe that interventions that help close the loop within the health system so that patients complete recommended behaviors in a timely manner may actually create operational efficiencies by smoothing demand for mammograms and other screenings and make productive use of existing patient data to support engagement with recommended care. It may also have the benefit of making it clear to providers which patients do not have or use email and may require high-touch outreach, so that those channels can be used appropriately. It would be worth quantifying whether monitoring and adjustment of outreach smooths mammography schedules, maximizing throughput without creating additional stress on providers. This could include both increasing mammogram appointments at slow times of the day, week, month, or year, as well as shifting mammography demand subsequent to campaigns such as Breast Cancer Awareness Month in October [43] to times where capacity is greater.

Another lens to understand how the intervention impacts patient behaviors is through patient experience research. Especially given that the population reached in this feasibility study was not proactively engaged in scheduling their recommended mammograms, there is value in a qualitative understanding of their response to the email communications and whether they perceived them as different or more compelling than typical health system communications. We hope to study patient perception of Precision Nudging as well as perception of the communications to better understand the intervention's effects and continually improve its acceptability and effectiveness.

Finally, there is opportunity beyond patient experience research to investigate improvements to the intervention itself. This may include advances to the reinforcement learning platform that assembles messages based on patient behavioral responses, enhancements to the library content to address barriers more effectively or to accommodate emerging barriers, or expansion to other channels, such as text message or chatbot, to better engage the full patient population. Although the feasibility study shows promising reach and engagement across patient subgroups, future research should focus on ensuring equitable access and support for preventive care among groups with historical experience of structural inequalities [44]. This will require engaging members of those groups to understand their barriers to action and partnering to co-design solutions [24]. Ensuring equity will also require that the data used to train interventions like the one under study, which is driven by reinforcement learning, are representative of the populations at large and that the benefits conferred are available to all [45].

Limitations

This feasibility study offers real-world pilot results, while laying the groundwork for further investigation. First, and most obviously, while this study demonstrates that a BRL-powered email-based behavioral intervention is feasible to deliver a behavioral intervention for mammograms, its efficacy in achieving behavioral results can be better understood through a randomized controlled trial or other experimental methods in future work. The hypothesis that the personalization enabled by BRL enhances outcomes relative to a standard nonpersonalized health system messaging campaign should be rigorously tested.

In terms of better demonstrating equitable outcomes, a major limitation to this study came from the lack of an a priori value for the treatment difference that would be of practical importance. Given the giant sample sizes in studies like this, alternative methods for establishing equivalence need to be employed. Future research should suggest and test a priori values of true treatment differences between population subgroups.

This work was also confined to a single health system whose patients are geographically concentrated in the Midwestern United States. Moreover, the system is a mission-driven Catholic

health care organization. Although this does not have bearing on recommendations around breast cancer screening, the health system's Catholic identity may attract a patient population who differs from the general US patient population. Future research should examine the success of this mammography intervention in other health systems to establish the generalizability of the results.

The unsubscribe rate for intervention emails over the 2 years of study was 10.5%. Despite being generally much higher than the 2022 average for health care services marketing emails [46], it is difficult to draw direct comparisons between this novel reinforcement learning–driven digital health intervention and other digital communications interventions. In this case, it is reasonable to assume that a large proportion of those who unsubscribed were women who had attended their mammograms, as they were not informed that they would no longer receive intervention emails once they had scheduled or attended their mammograms. Other women might have been induced to feel annoyed or guilty by the ongoing messages. Thus, a limitation involved not exploring the demographics, engagement, and behavioral outcomes of those who unsubscribed, to better understand their motivations for doing so, and, ultimately, improve the intervention to reduce the unsubscribe rate.

Finally, the time period during which the feasibility study was conducted coincided with the resumption of preventive care, such as mammograms, during the COVID-19 pandemic, which may have artificially influenced mammography behaviors among the patient population. Continued monitoring of the intervention's outcomes over time should provide clarity as to its performance in times of reduced demand for screenings.

Conclusions

This retrospective, single-arm, observational study suggests that a reinforcement learning–enabled email intervention can be used in a health system to engage patients who are significantly overdue for their mammograms to schedule and attend a recommended screening. In this feasibility investigation, the intervention was associated with scheduling and attending mammograms for patients who were significantly overdue for recommended screening. Moreover, the intervention showed proportionate reach across demographic subpopulations. This suggests that the intervention may be effective at engaging overdue patients of many different backgrounds. In a time where many patients are behind on preventative screenings, with potentially life-altering results, and where many health care organizations are eager to manage costs and deliver quality care, interventions that engage the most disengaged patients are a vital tool to improve outcomes. These interventions will be successful to the extent they can be delivered in a low-cost and scalable fashion, offer flexibility to systems and providers to support established workflows, and concretely help patients overcome the barriers that have kept them from recommended care.

Acknowledgments

The authors wish to thank Kelsey Ford, DrPH, for foundational input to the reporting of this research; Erin Hurlburt, MD, for guidance and leadership on intervention design and implementation; and Geoff Alday for support in creating figures. The authors also wish to thank Kathleen Benitez; Folasade Phillips, MPH; and especially Darrin Weber, PhD, for analytics support. Lirio LLC financially supported this research.

Data Availability

The data sets generated and analyzed during this study are not publicly available due to contractual agreements between Lirio and the health system where the intervention was implemented. Derived data used in the analyses are available from the corresponding author upon reasonable request.

Conflicts of Interest

AB, ESB, and ABW are paid employees of Lirio, LLC, which developed the intervention studied.

References

1. Breast cancer statistics. Centers for Disease Control and Prevention. 2022. URL: https://www.cdc.gov/cancer/breast/statistics/index.htm [accessed 2022-06-30]
2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021 May;71(3):209-249 [FREE Full text] [doi: 10.3322/caac.21660] [Medline: 33538338]
3. Treatment of breast cancer stages I-III. American Cancer Society. 2022. URL: https://www.cancer.org/cancer/breast-cancer/treatment/treatment-of-breast-cancer-by-stage/treatment-of-breast-cancer-stages-i-iii.html [accessed 2022-06-30]
4. Blumen H, Fitch K, Polkus V. Comparison of treatment costs for breast cancer, by tumor stage and type of service. Am Health Drug Benefits 2016 Feb;9(1):23-32 [FREE Full text] [Medline: 27066193]
5. Duffy SW, Tabár L, Yen AM, Dean PB, Smith RA, Jonsson H, et al. Mammography screening reduces rates of advanced and fatal breast cancers: Results in 549,091 women. Cancer 2020 Jul 01;126(13):2971-2979 [FREE Full text] [doi: 10.1002/cncr.32859] [Medline: 32390151]

6. Song S, White A, Kucik JE. Use of selected recommended clinical preventive services - Behavioral Risk Factor Surveillance System, United States, 2018. MMWR Morb Mortal Wkly Rep 2021 Apr 02;70(13):461-466 [FREE Full text] [doi: 10.15585/mmwr.mm7013a1] [Medline: 33793461]
7. United States Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics. National Health Interview Survey, 1992: Cancer Control Supplement. Ann Arbor, MI: Inter-university Consortium for Political and Social Research; 1994. URL: https://www.icpsr.umich.edu/web/NACDA/studies/6344/versions/V1 [accessed 2022-07-25]
8. Sprague B, Lowry K, Miglioretti D, Alsheik N, Bowles E, Tosteson A, et al. Changes in mammography use by women's characteristics during the first 5 months of the COVID-19 pandemic. J Natl Cancer Inst 2021 Sep 04;113(9):1161-1167 [FREE Full text] [doi: 10.1093/jnci/djab045] [Medline: 33778894]
9. Song H, Bergman A, Chen AT, Ellis D, David G, Friedman AB, et al. Disruptions in preventive care: Mammograms during the COVID-19 pandemic. Health Serv Res 2021 Feb;56(1):95-101 [FREE Full text] [doi: 10.1111/1475-6773.13596] [Medline: 33146429]
10. Ryser MD, Lange J, Inoue LYT, O'Meara ES, Gard C, Miglioretti DL, et al. Estimation of breast cancer overdiagnosis in a U.S. breast screening cohort. Ann Intern Med 2022 Apr;175(4):471-478 [FREE Full text] [doi: 10.7326/M21-3577] [Medline: 35226520]
11. Bucher A. Engaged: Designing for Behavior Change. New York, NY: Rosenfeld Media; 2020.
12. Michie S, van Stralen MM, West R. The behaviour change wheel: A new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
13. 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 08;3(2):e23 [FREE Full text] [doi: 10.2196/humanfactors.6029] [Medline: 27609738]
14. Baker R, Camosso-Stefinovic J, Gillies C, Shaw EJ, Cheater F, Flottorp S, et al. Tailored interventions to overcome identified barriers to change: Effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2010 Mar 17(3):CD005470 [FREE Full text] [doi: 10.1002/14651858.CD005470.pub2] [Medline: 20238340]
15. Nelson L, Greevy R, Spieker A, Wallston K, Elasy T, Kripalani S, et al. Effects of a tailored text messaging intervention among diverse adults with type 2 diabetes: Evidence from the 15-month reach randomized controlled trial. Diabetes Care 2021 Jan;44(1):26-34 [FREE Full text] [doi: 10.2337/dc20-0961] [Medline: 33154039]
16. Mayberry L, Nelson L, Greevy R, Wallston K, Kripalani S, Osborn C. Personalized texts improve adherence and A1c over 6 months among high-risk adults with type 2 diabetes. Diabetes 2019;68(Supplement_1):819-P. [doi: 10.2337/db19-819-p]
17. Mayberry LS, Mulvaney SA, Johnson KB, Osborn CY. The MEssaging for Diabetes intervention reduced barriers to medication adherence among low-income, diverse adults with type 2. J Diabetes Sci Technol 2017 Jan;11(1):92-99 [FREE Full text] [doi: 10.1177/1932296816668374] [Medline: 27595710]
18. Nelson LA, Wallston KA, Kripalani S, Greevy RA, Elasy TA, Bergner EM, et al. Mobile phone support for diabetes self-care among diverse adults: Protocol for a three-arm randomized controlled trial. JMIR Res Protoc 2018 Apr 10;7(4):e92 [FREE Full text] [doi: 10.2196/resprot.9443] [Medline: 29636319]
19. Mac Aonghusa P, Michie S. Artificial intelligence and behavioral science through the looking glass: Challenges for real-world application. Ann Behav Med 2020 Dec 01;54(12):942-947 [FREE Full text] [doi: 10.1093/abm/kaaa095] [Medline: 33416835]
20. Mitchell U, Chebli P, Ruggiero L, Muramatsu N. The digital divide in health-related technology use: The significance of race/ethnicity. Gerontologist 2019 Jan 09;59(1):6-14. [doi: 10.1093/geront/gny138] [Medline: 30452660]
21. Eruchalu CN, Pichardo MS, Bharadwaj M, Rodriguez CB, Rodriguez JA, Bergmark RW, et al. The expanding digital divide: Digital health access inequities during the COVID-19 pandemic in New York City. J Urban Health 2021 Apr;98(2):183-186 [FREE Full text] [doi: 10.1007/s11524-020-00508-9] [Medline: 33471281]
22. Fernandez ME, Ruiter RAC, Markham CM, Kok G. Intervention mapping: Theory- and evidence-based health promotion program planning: Perspective and examples. Front Public Health 2019;7:209 [FREE Full text] [doi: 10.3389/fpubh.2019.00209] [Medline: 31475126]
23. 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]
24. Ford KL, West AB, Bucher A, Osborn CY. Personalized digital health communications to increase COVID-19 vaccination in underserved populations: A double diamond approach to behavioral design. Front Digit Health 2022;4:831093 [FREE Full text] [doi: 10.3389/fdgth.2022.831093] [Medline: 35493533]
25. Wood CE, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. Applying the behaviour change technique (BCT) taxonomy v1: A study of coder training. Transl Behav Med 2015 Jun;5(2):134-148 [FREE Full text] [doi: 10.1007/s13142-014-0290-z] [Medline: 26029276]
26. Mnih V, Kavukcuoglu K, Silver D, Rusu A, Veness J, Bellemare M, et al. Human-level control through deep reinforcement learning. Nature 2015 Feb 26;518(7540):529-533. [doi: 10.1038/nature14236] [Medline: 25719670]

27. Sutton R, Precup D, Singh S. Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning. Artif Intell 1999 Aug;112(1-2):181-211. [doi: 10.1016/s0004-3702(99)00052-1]
28. Dulac-Arnold G, Levine N, Mankowitz DJ, Li J, Paduraru C, Gowal S, et al. Challenges of real-world reinforcement learning: Definitions, benchmarks and analysis. Mach Learn 2021 Apr 22;110(9):2419-2468. [doi: 10.1007/s10994-021-05961-4]
29. Dulac-Arnold G, Mankowitz D, Hester T. Challenges of real-world reinforcement learning. In: Proceedings of the 36th International Conference on Machine Learning. 2019 Presented at: The 36th International Conference on Machine Learning; June 9-15, 2019; Long Beach, CA URL: https://arxiv.org/pdf/1904.12901.pdf
30. Health equity. Centers for Disease Control and Prevention. 2022. URL: https://www.cdc.gov/chronicdisease/healthequity/index.htm [accessed 2022-07-29]
31. Tunes da Silva G, Logan BR, Klein JP. Methods for equivalence and noninferiority testing. Biol Blood Marrow Transplant 2009 Jan;15(1 Suppl):120-127 [FREE Full text] [doi: 10.1016/j.bbmt.2008.10.004] [Medline: 19147090]
32. Betensky RA. Optimally selected chi square statistics for equivalence testing. J Stat Plan Inference 2001 Feb;93(1-2):247-257. [doi: 10.1016/s0378-3758(00)00204-4]
33. Lin M, Lucas HC, Shmueli G. Too big to fail: Large samples and the p-value problem. Inf Syst Res 2013 Dec;24(4):906-917. [doi: 10.1287/isre.2013.0480]
34. Dunnett CW, Gent M. Significance testing to establish equivalence between treatments, with special reference to data in the form of 2 x 2 tables. Biometrics 1977 Dec;33(4):593. [doi: 10.2307/2529457]
35. Park C, Otobo E, Ullman J, Rogers J, Fasihuddin F, Garg S, et al. Impact on readmission reduction among heart failure patients using digital health monitoring: Feasibility and adoptability study. JMIR Med Inform 2019 Nov 15;7(4):e13353 [FREE Full text] [doi: 10.2196/13353] [Medline: 31730039]
36. Luong P, Glorioso TJ, Grunwald GK, Peterson P, Allen LA, Khanna A, et al. Text message medication adherence reminders automated and delivered at scale across two institutions: Testing the nudge system: Pilot study. Circ Cardiovasc Qual Outcomes 2021 May;14(5):e007015 [FREE Full text] [doi: 10.1161/CIRCOUTCOMES.120.007015] [Medline: 33993727]
37. Rouyard T, Leal J, Baskerville R, Velardo C, Salvi D, Gray A. Nudging people with type 2 diabetes towards better self-management through personalized risk communication: A pilot randomized controlled trial in primary care. Endocrinol Diabetes Metab 2018 Jul;1(3):e00022 [FREE Full text] [doi: 10.1002/edm2.22] [Medline: 30815556]
38. Brar Prayaga R, Jeong EW, Feger E, Noble HK, Kmiec M, Prayaga RS. Improving refill adherence in Medicare patients with tailored and interactive mobile text messaging: Pilot study. JMIR Mhealth Uhealth 2018 Jan 30;6(1):e30 [FREE Full text] [doi: 10.2196/mhealth.8930] [Medline: 29382623]
39. Schoenthaler A, Leon M, Butler M, Steinhaeuser K, Wardzinski W. Development and evaluation of a tailored mobile health intervention to improve medication adherence in Black patients with uncontrolled hypertension and type 2 diabetes: Pilot randomized feasibility trial. JMIR Mhealth Uhealth 2020 Sep 23;8(9):e17135 [FREE Full text] [doi: 10.2196/17135] [Medline: 32965230]
40. Wasserstein RL, Lazar NA. The ASA statement on p-values: Context, process, and purpose. Am Stat 2016 Jun 09;70(2):129-133. [doi: 10.1080/00031305.2016.1154108]
41. Greene J, Hibbard JH, Sacks R, Overton V, Parrotta CD. When patient activation levels change, health outcomes and costs change, too. Health Aff (Millwood) 2015 Mar;34(3):431-437. [doi: 10.1377/hlthaff.2014.0452] [Medline: 25732493]
42. Hensing WL, Poplack SP, Herman CR, Sutcliffe S, Colditz GA, Ademuyiwa FO. Racial differences in no-show rates for screening mammography. Cancer 2021 Jun 01;127(11):1857-1863 [FREE Full text] [doi: 10.1002/cncr.33435] [Medline: 33792894]
43. Barlow W. Stat bite: Effect of breast cancer awareness month on mammography use. J Natl Cancer Inst 2005 Oct 19;97(20):1493. [doi: 10.1093/jnci/dji368] [Medline: 16234559]
44. Bailey ZD, Krieger N, Agénor M, Graves J, Linos N, Bassett MT. Structural racism and health inequities in the USA: Evidence and interventions. Lancet 2017 Apr 08;389(10077):1453-1463. [doi: 10.1016/S0140-6736(17)30569-X] [Medline: 28402827]
45. Ibrahim SA, Charlson ME, Neill DB. Big data analytics and the struggle for equity in health care: The promise and perils. Health Equity 2020;4(1):99-101 [FREE Full text] [doi: 10.1089/heq.2019.0112] [Medline: 32258961]
46. Ultimate email marketing benchmarks for 2022: By industry and day. Campaign Monitor. 2022. URL: https://www.campaignmonitor.com/resources/guides/email-marketing-benchmarks/ [accessed 2022-07-29]

Abbreviations

AI: artificial intelligence
BCT: behavior change technique
BRL: behavioral reinforcement learning

Edited by A Mavragani; submitted 31.08.22; peer-reviewed by M Kapsetaki, C Castro Sweet; comments to author 18.10.22; revised version received 03.11.22; accepted 04.11.22; published 28.11.22.

Please cite as:
Bucher A, Blazek ES, West AB
Feasibility of a Reinforcement Learning–Enabled Digital Health Intervention to Promote Mammograms: Retrospective, Single-Arm, Observational Study
JMIR Form Res 2022;6(11):e42343
URL: https://formative.jmir.org/2022/11/e42343
doi:10.2196/42343
PMID:36441579

Original Paper

Fitbits for Monitoring Depressive Symptoms in Older Aged Persons: Qualitative Feasibility Study

Fiza Mughal[1], PhD; William Raffe[1], PhD; Peter Stubbs[2], PhD; Ian Kneebone[3], MPsych, PsychD; Jaime Garcia[1], PhD

[1]Faculty of Engineering and IT, University of Technology Sydney, Sydney, Australia

[2]Discipline of Physiotherapy, University of Technology Sydney, Sydney, Australia

[3]Discipline of Clinical Psychology, University of Technology Sydney, Sydney, Australia

Corresponding Author:
Fiza Mughal, PhD
Faculty of Engineering and IT
University of Technology Sydney
15 Broadway, Ultimo
Sydney, 2007
Australia
Phone: 61 4 5262 7824
Email: fiza.mughal@uts.edu.au

Abstract

Background: In 2022, an estimated 1.105 billion people used smart wearables and 31 million used Fitbit devices worldwide. Although there is growing evidence for the use of smart wearables to benefit physical health, more research is required on the feasibility of using these devices for mental health and well-being. In studies focusing on emotion recognition, emotions are often inferred and dependent on external cues, which may not be representative of true emotional states.

Objective: The aim of this study was to evaluate the feasibility and acceptability of using consumer-grade activity trackers for apps in the remote mental health monitoring of older aged people.

Methods: Older adults were recruited using criterion sampling. Participants were provided an activity tracker (Fitbit Alta HR) and completed weekly online questionnaires, including the Geriatric Depression Scale, for 4 weeks. Before and after the study period, semistructured qualitative interviews were conducted to provide insight into the acceptance and feasibility of performing the protocol over a 4-week period. Interview transcripts were analyzed using a hybrid inductive-deductive thematic analysis.

Results: In total, 12 participants enrolled in the study, and 9 returned for interviews after the study period. Participants had positive attitudes toward being remotely monitored, with 78% (7/9) of participants experiencing no inconvenience throughout the study period. Moreover, 67% (6/9) were interested in trialing our prototype when it is implemented. Participants stated they would feel more comfortable if mental well-being was being monitored by carers remotely.

Conclusions: Fitbit-like devices were an unobtrusive and convenient tool to collect physiological user data. Future research should integrate physiological user inputs to differentiate and predict depressive tendencies in users.

(JMIR Form Res 2022;6(11):e33952) doi:10.2196/33952

KEYWORDS

digital mental health; Fitbit; smartwatch; smart wearable; geriatric; aging; health informatics; feasibility; usability; older aged

Introduction

Background

In 2021, there were approximately 280 million people diagnosed with depression worldwide [1]. Although the majority of older adults (65+ years old) are not depressed, older adults are at a higher risk of depression [2]. Eighty percent of older adults have at least one chronic health condition, which can contribute to mental illness [3,4]. Thirty percent of older adults in residential care are at risk of depression [5]. Mental illness in older adults is sometimes viewed as an inevitable reaction to changes in socioeconomic standing or increased age and has been deemed untreatable for some people [2,3]. Older adults are also more likely to be concerned about the stigma of seeking treatment for mental illness [6]. Consequently, older adults are less likely to seek help when affected by mental illness, such as depression [3].

Cost-effective smart wearables are increasingly used in the general population [7]. In 2022, there were 1.105 billion smart wearable users worldwide [7]. Of these, approximately 31 million were Fitbit users [8]. Therefore, more studies are incorporating smart wearables, particularly Fitbit devices, in health research [9-11]. A recent scoping review investigated the effectiveness and efficiency of mobile health procedures for improving physical health [12]. From 2012 to 2022, including 148 studies, there was no "one-size-fits-all approach for monitoring physical health. However, the authors found that mobile health interventions are promising for facilitating behavioral change. A similar review [13] concluded that more research was required on the effectiveness and feasibility of smart wearables for changing or assessing physical health. Using only Fitbit devices, Ringeval et al [11] assessed the effectiveness of using Fitbit devices in interventions to promote healthy lifestyle habits. Evaluating 41 studies [11], the authors concluded that Fitbit devices improved lifestyles in users. There was increased daily step counts, increased physical activity, and reduced mass in participants [11]. A further review [14] investigated the applicability of wearable devices for vital sign monitoring in outpatients and found that early detection of physiological deterioration via wearable devices likely improves patient outcomes. Although positive, on-body, potentially obtrusive sensors, such as heart rate monitors, patches, and arm bands were used, these may not be feasible for prolonged use.

The aforementioned reviews demonstrate that wearable devices can result in positive physical activity changes. Despite this, few studies have investigated how mental health might change from using smart wearables. Furthermore, very few studies have investigated older adults. This is problematic, as older adults are more likely to be apprehensive toward newer technologies [15,16]. With most studies focusing on laboratory-based settings, we proposed the use of consumer-grade smart wearables for mental health monitoring in order to bridge the gap between older people, modern technologies, and mental health [17].

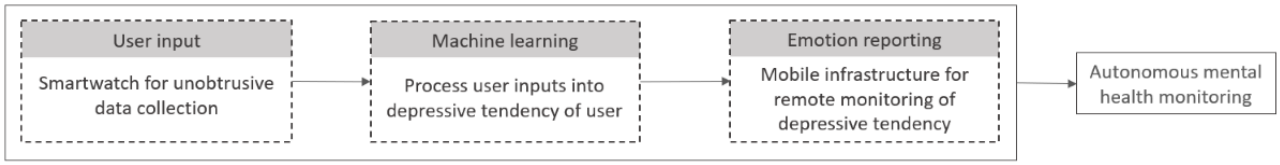
Prior Work

Smart wearables, particularly Fitbit-like devices, can be used to assist in the treatment of mental health conditions. A recent study investigated the feasibility (frequency of use) and acceptability (through follow-up interviews) of mobile health technologies to increase physical activity among users with severe mental illness [18]. Participants reported high satisfaction and increased motivation through goal setting and self-monitoring, facilitated by the mobile health technology. Another qualitative study integrated Fitbit devices with behavioral activation therapy [10]. People had more positive self-awareness, improved peer-based motivation, and improved motivation to set goals. Negative feedback included inconvenience, disinterest in participating in the study, and inaccuracy of the measurements. Furthermore, a 12-week randomized trial incorporated a non-Fitbit device (tablet) and telephone counseling to increase physical activity in older adults [19]. The intervention was feasible and accepted by the participants. It also reduced mass and increased physical activity time in older adults.

Factors such as heart rate and gait patterns can indicate depression which can be monitored using smart wearables [20,21]. Research has also validated the use of smart wearables for emotion recognition [22-25]. Smartwatches or bracelets were used and paired with external emotion-eliciting stimuli to evaluate the reliability and accuracy for emotion recognition. In laboratory-based settings, 3 studies provided participants with stimuli, including video and music clips, which were used to elicit happy, neutral, or sad emotions [23-25]. Poststimuli activities, conducted within 2 hours of emotion elicitation, included walking with a chest-mounted heart rate monitor [23]. Results showed a higher accuracy (74%) for detecting happy emotions as opposed to sad emotions. An earlier study observed walking changes 1 minute after participants received a visual stimulus to elicit an emotional response [25]. In that study, walking patterns could predict the expected emotional response 81.2% of the time. The studies show that walking patterns and smart wearables can be used to recognize user emotions. Adopting a different approach for emotion elicitation, using electrodermal activity, skin temperature, heart rate, and a Self-Assessment Manikin form, another study [22] developed an algorithm to predict emotion based on electrodermal activity signals. The algorithm accurately tagged emotions 57% of the time.

In a previous study [17], we proposed an autonomous mental health–monitoring system for older aged people (AutoMAP), which provided an emotion recognition framework using minimal to no explicit user input or interaction (Figure 1). The aim of AutoMAP is to monitor the mental health of the user, without explicit user interaction (ie, not requiring the user to manually log their moods at a specific time). This framework used physiological data collected using smart wearables with the intention of predicting user mood using mood measures such as the Geriatric Depression Scale (scored from of 0 to 15, with 0 indicating no depressive tendencies and 15 indicating high depressive tendencies; scores > 9 would indicate very high depressive tendencies). The intention, when finalized, is to report depressive tendency scores to caregivers via a mobile app. In this paper, we focus on validating the feasibility (Are we able to do it?) and acceptability (Are participants willing to do it?) of our framework (Figure 1). Therefore, this paper presents the findings of a feasibility study. Although our study consists of qualitative and quantitative components, this paper describes the qualitative outcomes, centered around the feasibility and acceptability of AutoMAP. Quantitative Fitbit data, efficacy, and accuracy of the emotion recognition techniques in this framework will be presented in future papers.

Figure 1. The autonomous mental health–monitoring system for older aged people (AutoMAP) framework.



Methods

Overview

Our study investigated the feasibility (Are the researchers able to perform the study, with the numbers required, and using the framework proposed?) and participant acceptability (How accepting are participants of the daily monitoring and questionnaires, the mobile app, and use of the mobile app?) of AutoMAP for older adults (65+ years old). Participants wore a Fitbit smartwatch for a 4-week period, completed a validated depressive symptom survey weekly, and provided a self-reported mood score daily.

Ethical Considerations

The study was approved by the Human Research Ethics Committee (HREC) of the University of Technology Sydney (HREC reference #ETH20-4912).

Procedures

To identify the feasibility and acceptability of our framework, we performed a hybrid inductive-deductive thematic analysis [26,27] on pre-post semistructured interviews with participants. The study protocol is accessible under the Open Science Frameworks [28].

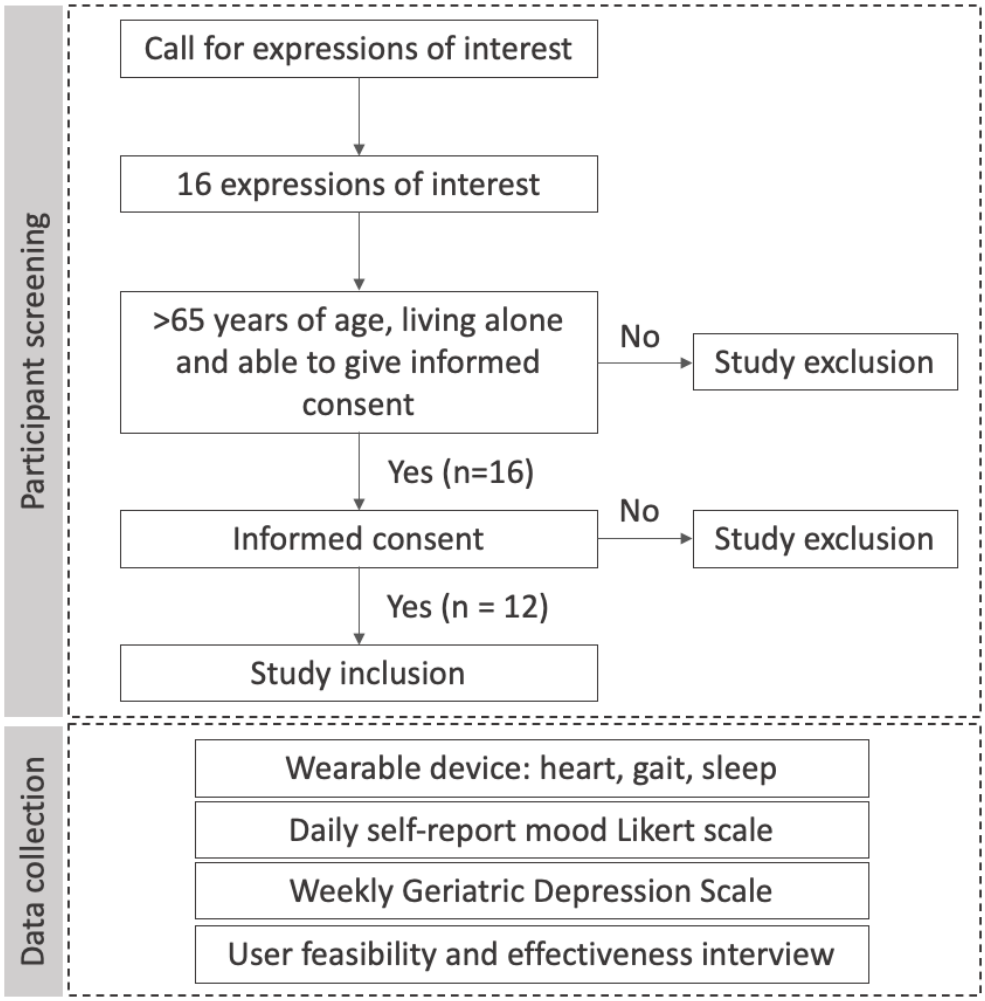
Participant Recruitment and Onboarding

Participants were recruited by word of mouth and calls for expressions of interest among senior communities and aged care facilities. There were six main inclusion criteria for the study: (1) being able to speak and understand English, (2) being aged over 65 years, (3) willing to wear a Fitbit for 4 weeks, (4) living alone with no external assistance, (5) being able to provide informed consent, and (6) having access to the internet through a computer or mobile phone. The participant recruitment flowchart is shown Figure 2. Individuals living independently were chosen for this study, as this is the target cohort for our proposed AutoMAP framework.

Similarly, there were 8 exclusion criteria for the study. These exclusion criteria included pre-existing conditions affecting sleep, gait, or heart recordings; travel plans within the duration of the 4-week period; foreseen out of the ordinary plans within the 4-week period; an inability to walk; an inability to access online surveys within the time requirement; having a history of skin rashes around the wrist area; a nickel allergy that may cause a reaction to the smartwatch charging probe; and any pre-existing diagnosis of mental health conditions. At this preliminary stage, we explicitly excluded participants with pre-existing mental health issues, as we are in the feasibility and participant acceptability stage of this research. Due to difficulties recruiting participants, 2 participants aged 64 years were also included; however they did not return for the poststudy interviews. Their data were excluded from study.

After providing an expression of interest (Figure 2), participants were provided an information sheet detailing the inclusion criteria and requirements during the 4-week period. Twelve participants were deemed eligible and were sent a participant pack. Participant packs consisted of five components: (1) a Fitbit Alta HR device, (2) a participation confirmation and information sheet, (3) a distress management resource list, (4) a simplified Fitbit user guide, and (5) a summarized task checklist. To avoid subconscious response bias, participants were blinded to the study objectives until the end of the 4-week study period. Of the 12 participants, 3 did not return for the final interviews. Figure 2 illustrates the initial screening and onboarding process at the start of the 4-week study period.

Figure 2. Participant recruitment flow.



Wearable Device

For this study, all participants were provided a Fitbit Alta HR smartwatch. The Fitbit Alta HR smartwatch was used due to its cost-effectiveness for the eventual end user, as well as its ease of setup and usage [4,22]. Studies have shown fair performance accuracy for similar Fitbit devices (Fitbit Charge 2), with a heart rate estimation error of 14% [2]. A review of Fitbit-centered sleep studies also showed sensitivity values of 0.95-0.96 and specificity values of 0.58-0.69 for detecting sleep stages [29].

Participants were required to wear the device at all times during the 4-week period, except during showering. As the study required sleep data, participants wore the device when sleeping. The Fitbit Alta HR takes approximately 2 hours to fully charge. To ensure minimal data loss, we recommended that participants charge the Fitbit before they showered. To limit further inconvenience, we did not set fixed times for recharging.

Participant Protocols

Participants were provided a list of resources in case of emotional distress during the 4-week study period. Although introductory interviews were basic questions about the everyday lives of the participants, there were protocols to terminate the interview session and stop participation if a participant became distressed. Participants were informed that they were allowed to withdraw from the study at any point.

Researcher Interventions

Researchers could advise participants to seek assistance from a resource list or medical professionals. Participants were provided with contact information of the research team.

COVID-19 Protocols

The study took place in zero-contact settings to accommodate COVID-19. Participant document packs and devices were sent via post, and all interviews were held via phone.

Data Acquisition and Analysis

Questionnaires

Participants completed an online daily questionnaire consisting of a self-report mood rating Likert scale, open questions on activity and food preferences (diversion questions away from depressive symptoms), and an optional section to add details of any out-of-the-ordinary events in the preceding 24-hour period. For the mood scale, participants were asked “How would you rate your mood in the past 24 hours,” rated on a scale from 1 to 10, once per day. Diversion questions, such as “What do you feel like eating?” and “What do you feel like doing?” were added to the daily mood reported questionnaire to reduce response bias.

Participants also completed an online 15-item Geriatric Depression Scale (GDS-15) once per week [30]. This is a short

form of the GDS and consists of 15 yes or no questions with a single-point score for each response, which can indicate depression [31]. The participants also completed the GDS-15 questionnaire online during the introductory briefing. These responses were used as the baseline GDS scores to exclude participants (if GDS scores were >10) and for data labeling to train and test a machine learning algorithm that is part of the AutoMAP framework. As this stage is part of the quantitative phase of the research, it is not detailed further in this paper. Any participants scoring >6 on the GDS would be advised to attend a general practitioner consultation prior to participating in the study. Higher-scoring participants were not excluded immediately, as the GDS is a screening rather than a diagnostic tool for depression [31]. Both sets of questionnaires were administered through an online platform (Google Forms) dedicated to conducting and circulating online surveys.

Interviews

Two sets of interviews (Textbox 1) were performed at different times: (1) baseline and (2) on completion of the 4-week study period. Interviews were transcribed verbatim. The mean interview time was 11 minutes (range 4 minutes to 16 minutes) and were administered via phone by the primary researcher and coder (FM). Anonymized interviews were coded by a team of 3 additional coders (JG, PS, and WR) that had no interaction with the participants. Data were analyzed using hybrid inductive-deductive thematic analysis [26].

Prestudy interviews (Textbox 1) provided insight into the everyday lives of participants in the 4 weeks preceding the study and into the impact of COVID-19 on regular interactions (if any). We added the COVID-19 component to account for anomalies in participant data if they in fact did feel impacted by the pandemic. Poststudy interviews gained feedback on the study period itself. Poststudy interviews were held after the 4-week study period (Textbox 1). Participants were provided further details on the study and its objectives. To determine feasibility and acceptability, we asked the participants about their experience during the 4-week period, whether they would be interested in the autonomous mental health–monitoring system we have proposed, and what they would or would not want to see in the mobile app. Such an app would report their emotional state, for example, using clean, straightforward notifications versus a detailed dashboard view of vital signs.

To provide insight on whether participants felt any inconvenience or discomfort during the 4-week study period, we asked participants a set of questions about the Fitbit device. Questions 1 to 3 (Textbox 1) discussed the user experience. Questions 4 to 6 and 10 were designed to obtain feedback for improvement on the study design should the study be replicated in the future. Questions 7 and 8 asked participants about the likelihood of being more relaxed if their caregivers were informed of their mental health. Finally, question 9 was designed to validate the potential impact that the framework, as a whole, may have on older adults living alone.

As the aim of this study is to evaluate the feasibility and acceptability of AutoMAP, we conducted a qualitative analysis of both sets of interviews. We assessed levels of agreement through study retention and qualitative outcomes. We performed a hybrid inductive-deductive thematic analysis [26] with 4 coders (FM, JG, PS, and WR). Our deductive thematic analysis was initially directed toward usability; however, we identified commonly mentioned subthemes to provide more insight on the interviews. In a workshop discussion, deidentified interview responses were assessed with four aims: (1) generate initial codes, (2) search for themes, (3) review themes, and (4) finally define and name the themes. Figure 3 shows an overview of how interview responses were coded and analyzed.

Textbox 1. Questions asked during the pre-post interviews.

Preintervention (introductory) interview

1. How have you generally felt over the past four weeks?
2. How often do your caregivers visit you/you visit them (family, friends, or medical professionals)?
3. Was this the same/less/more prior to the COVID-19 pandemic?
4. Do you feel that your lifestyle or your daily life has been impacted by COVID-19 in anyway?
5. Do you feel more dependent on others now as opposed to a few years ago?
6. Geriatric Depression Scale (GDS) Questionnaire
7. Do you have any questions for us?

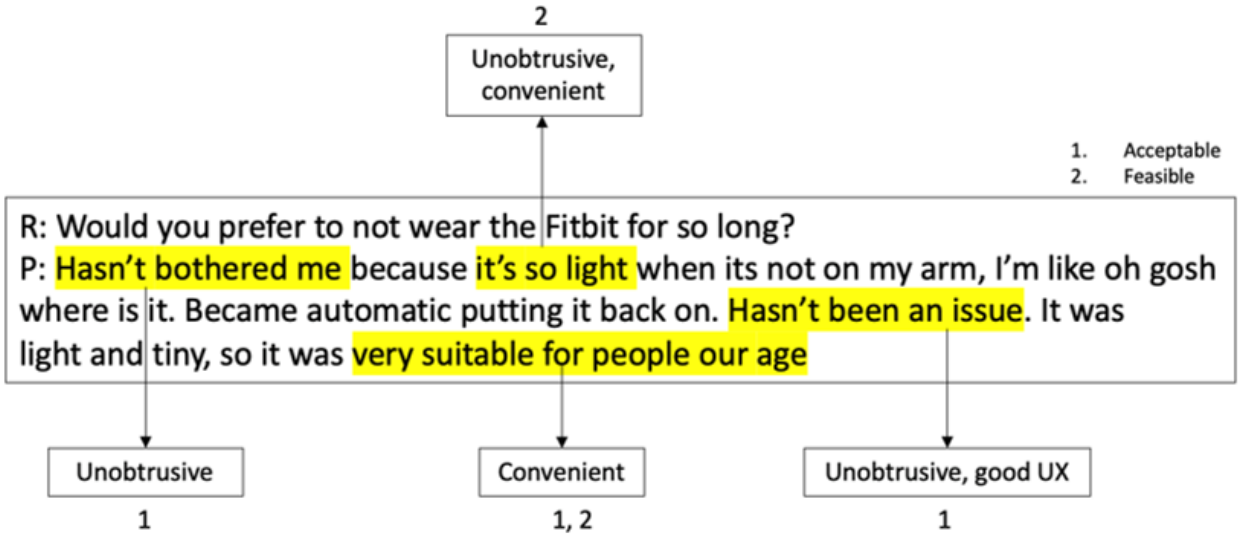
Follow-up questions will be asked based on their responses to the above questions.

Postintervention (closing) interview

1. How was your experience over the past four weeks?
2. Did anything cause you discomfort over this time?
3. How was wearing the Fitbit for 4 weeks
4. Was the daily survey inconvenient?
5. Was there any part in the survey that made you feel uncomfortable or distressed?
6. Do you have any feedback or suggestions for how we can improve the study in the future?
7. Let's say we build a mobile app to send notifications to you or your families based on our findings, in real-time. Would that make you feel more relaxed while living independently?
8. What would you want in such an app?
9. Do you feel more at comfort when you know someone is looking after you or concerned about your inner wellbeing, regardless of whether they are with you all the time?
10. Do you have any further questions from us about our study?
11. Would your experience or answers with this study have been different prior to the COVID-19 pandemic?

Follow-up questions will be asked based on their responses to the above questions.

Figure 3. Sample excerpt analysis between the researcher and participant. P: participant; R: researcher; UX: user experience.



Results

General Summary

A summary of participant demographics is presented in Table 1. In all, 12 participants (mean age 68.58 years, SD 3.475 years) were interviewed prior to the 4 weeks with 9 out of 12 participants (mean age 69.11 years, SD 2.998 years) returning for a poststudy interview (75% retention). The 3 participants that did not attend the closing interviews were uncontactable, and the Fitbit data and prestudy interview responses were excluded from the study. Closing interviews were held with the

remaining participants to identify the feasibility and acceptability of the framework as well as potential improvements to the proposed framework (Figure 1). Figures 3 and 4 provide an example of how interviews were analyzed. Themes were formed based on key words and the context of the overall response.

Table 1. Participant demographics.

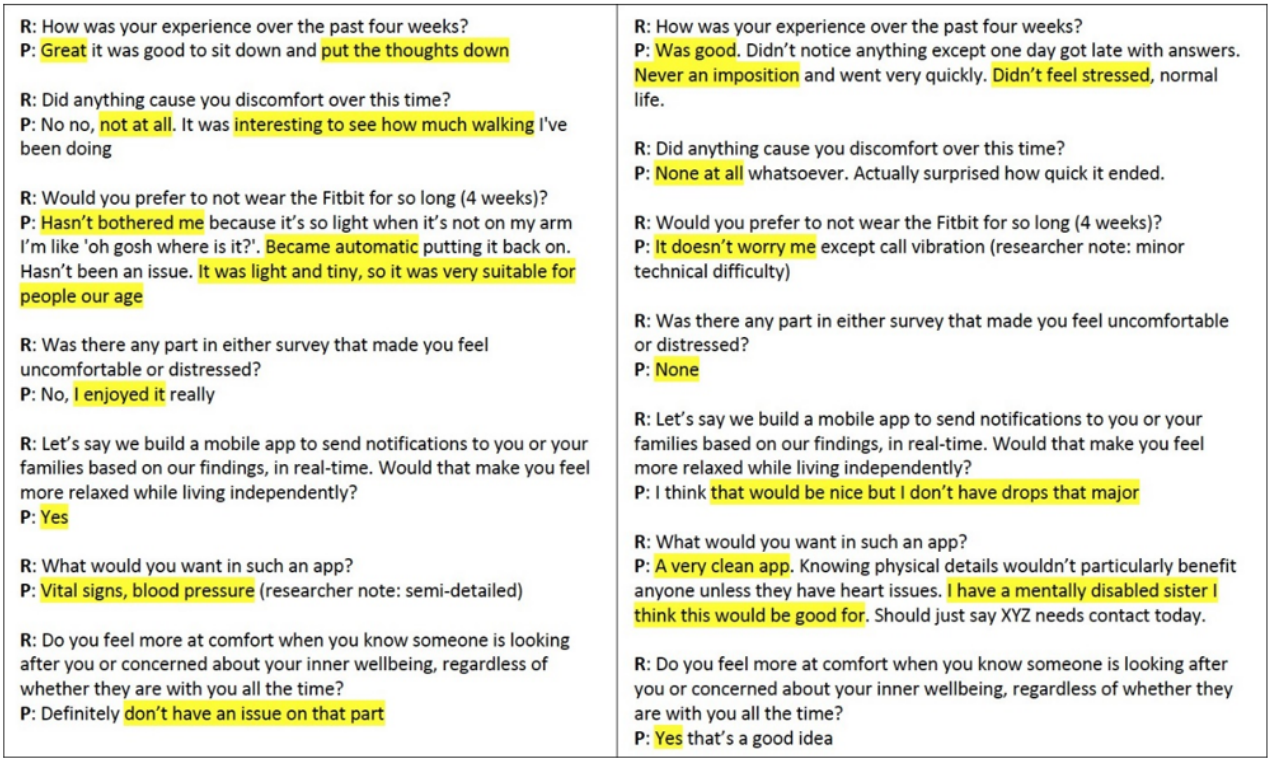
ID	Sex	Age (years)	GDS[a], mean (baseline)
1[b]	F	64	3.406 (3)
2	F	68	0.464 (0)
3	F	72	0 (0)[c]
4[b]	F	64	7.071 (6)
5	M	67	0.316 (0)
6	M	70	0 (0)[c]
7	F	68	0.974 (1)
8	F	75	6.310 (9)
9	M	71	0.4 (0)
10	F	65	0 (0)[c]
11[b]	F	73	2 (2)
12	M	66	0.448 (1)

[a]GDS: Geriatric Depression Scale.

[b]Did not return for postintervention interview and was excluded from the study.

[c]Had no change in GDS scores throughout the 4-week period.

Figure 4. Excerpts from two researcher-participant interviews with key words highlighted by one coder. P: participant; R: researcher.

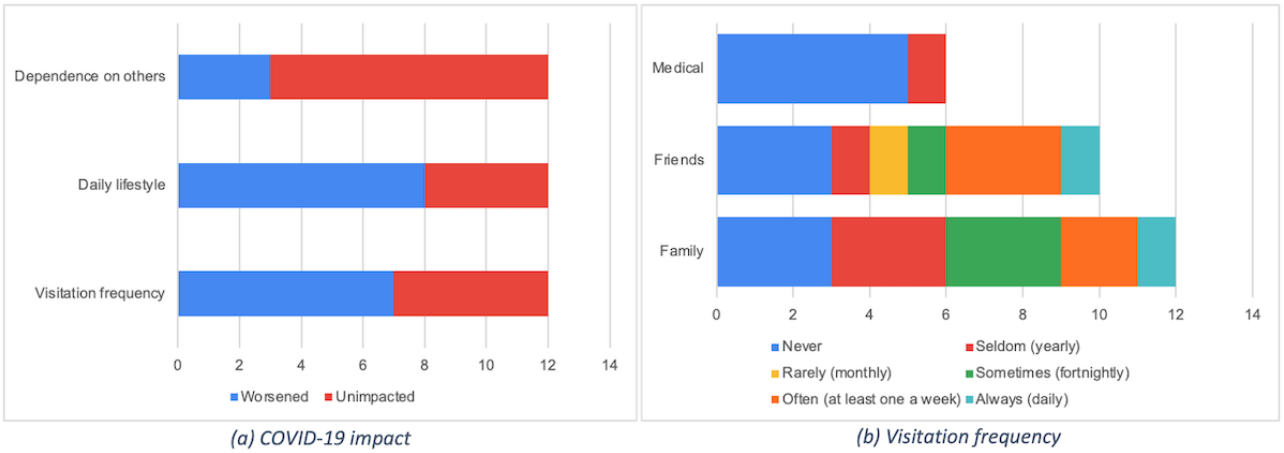


Preprocedure Analysis

Emergent themes from our analyses showed that participants had a positive life outlook over the preceding 4-week period. Specifically, 75% (9/12) of the participants had mostly positive emotions, with 41% (5/12) experiencing some negative emotions, 60% (3/5) of whom experienced significant negative life events. Moreover, 50% (6/12) of interviewees expressed a range of health concerns, with 83% (5/6) showing minor concerns and 17% (1/6) having major health concerns.

All (12/12) participants had family members visiting them occasionally, while 83.3% (10/12) had more frequent visitation with friends. Only 50% (6/12) had appointments with medical practitioners, usually once per year. Visitation frequency responses (n=28) exceeded the interviewee count (n=12), as categories were not mutually exclusive; that is, participants could state visitations with any or all categories (friends, family, or medical professionals). Categorically, 57% (16/28) of the responses showed visits ranging between family, friends, or medical practitioners, once or twice per year. However, among participants that had more frequent interactions, 32% (9/28) indicated weekly interaction with friends. Of the 12 participants, 7 had less interaction since the start of the COVID-19 pandemic (Figure 5b). Additionally, 67% (8/12) of the participants felt that their lifestyle and daily lives were negatively impacted by COVID-19 (Figure 5a).

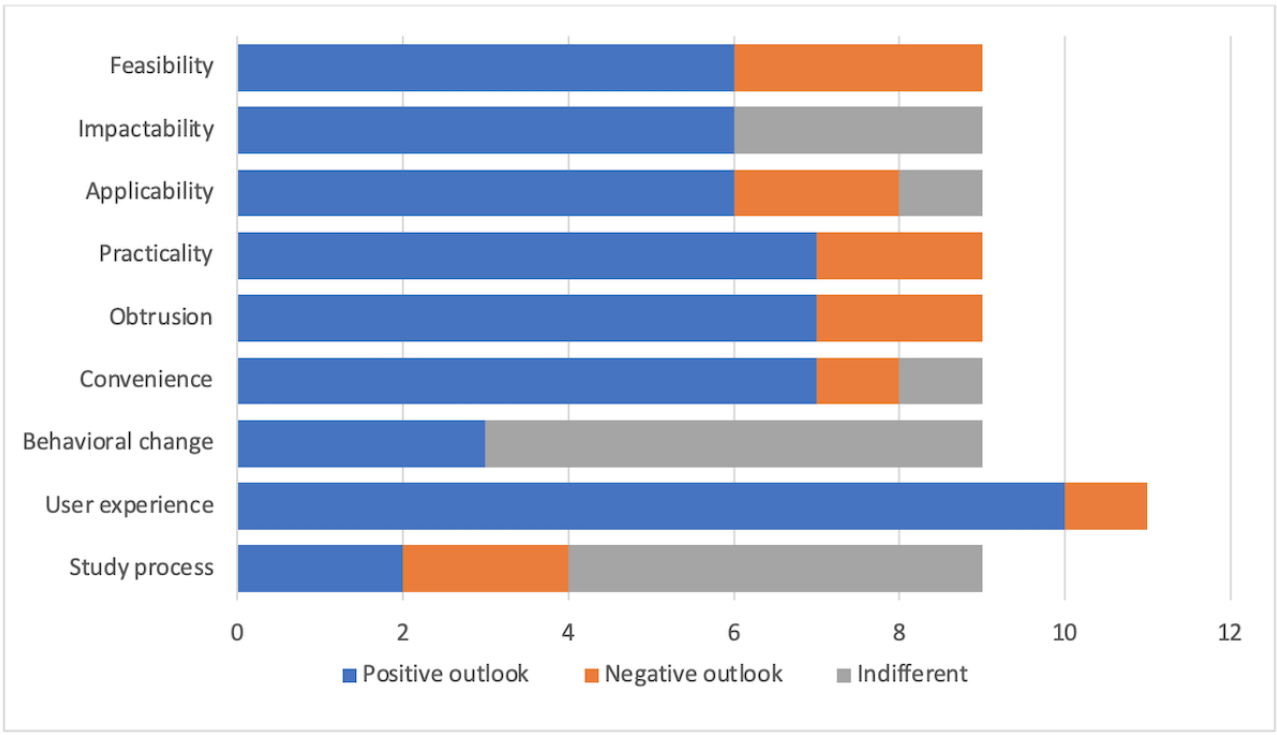
Figure 5. Prestudy themes: (a) COVID-19 impact on users in preceding 4-week timeframe; (b) frequency of visitation with friends, family, or medical practitioners. Bars are different lengths, as not all patients mentioned each of these items during the interviews.



Postprocedure Analysis

Responses after the 4-week study period were generally positive: 7 out of the 9 participants were pleased with the convenience and ease of study participation as well as the limited hands-on time commitment that was required. Figure 6 illustrates a summary of the thematic analysis for the postprocedure interviews. The user experience theme exceeds 9 participants, as 2 of the 9 participants mentioned this more than once through the interview. Moreover, 33% (3/9) of participants also experienced positive behavioral changes over the 4-week study period, particularly increased awareness and motivation to exercise more.

Figure 6. Dominant postprocedure themes for overall autonomous mental health–monitoring system for older aged people (AutoMAP) implementation and overall procedure period. User experience exceeds interview count due to repeat emphasis in responses.



Study Feasibility

Of the 9 participants, 6 (67%) felt no discomfort during the study, while 1 participant felt mild discomfort with having to wear the watch during warmer summer days. Of the 9 participants, 2 stated feeling slight frustration with having to wear the device overnight. Additionally, 7 of 9 (78%) interviewees found no inconvenience during the procedure, with the most common concerns being with the ability to report daily mood on time. Of the 9 participants, 2 felt worried about responding correctly even though there were no right or wrong answers. One participant found the diversion question "What do you feel like eating right now?" to be irrelevant, which was the purpose of the question. The procedure also resulted in behavioral change in 33% (3/9) of the participants particularly through exercise awareness and journaling during daily self-reports.

The device setup process and usage were favored by most participants, with only 1 participant having difficulty with the initial setup. Despite this, during data extraction, we found that 2 additional participants incorrectly linked their devices to the Fitbit app, which was resolved by requesting the participants to pair their devices again.

Overall, our analysis showed generally positive outcomes during the 4-week procedure. We achieved a retention rate of 75% (9/12). Some participants were seemingly aware of the study objectives although not explicitly disclosed, while others agreed it to be a good aim once they were provided more detail on our study goals.

Participant Usability

Further evaluating the usefulness and relevance of AutoMAP, we assessed the usability of the framework. Six of nine participants (67%) were interested in the full implementation of our working prototype and stated that they would also feel more relaxed and at comfort knowing that their mental well-being was being monitored, while the remainder felt it would be more useful for those people with major depression or diagnosed ailments. However, there were emergent themes pertaining to privacy, false positives (ie, the end user is alerted when there is no problem), and false negatives (ie, the end user is not alerted when there is a problem). These will be addressed under the mobile app requirements and constraints. Furthermore, 22% (2/9) of the participants suggested that the prototype would be more beneficial for users with specialized needs or diagnosed conditions such as Alzheimer disease, dementia, autism, Down syndrome, Asperger syndrome, or other mental health issues.

For 1 of the 9 participants, the pre-post study interviews showed a positive outlook, while the GDS scores were relatively high (mean 7). One participant that did not return for the postprocedure interviews had higher GDS scores (score ≥8) and an indifferent-to-positive outlook on life in the prestudy interview. The same participants also rated their own moods very highly (mean 8.31 and 7.8, respectively), showing that self-perceptions of emotion are not always accurate.

Of the 9 participants, 7 (67%) were pleased with the concept and potential of the prototype but raised concerns on its practicality. They suggested that the prototype may be a better fit for the special needs of persons of all ages or older adults with more serious problems, aside from general older adult populations. The envisioned end users of AutoMAP are the caregivers of older adults. However, one participant presented circumstances where the user does not have close friends or family, and the end user may not follow up on alerts.

Mobile App

The poststudy interviews were designed to collect suggestions and preferences for our proposed mobile app. The app would notify caregivers of the users' depressive tendencies. Participants were provided 2 suggestions for app design: one being a neat, minimal app with only the bare essential information, such as their depressive score; and the other having more detailed information. Expanding on their preferences, 56% (5/9) of the participants preferred a clean, minimal app that would only send plain-text alerts to the app end user (caregiver) when the user showed depressive tendencies. For 2 of the 9 participants, a semidetailed app interface was preferred to a text-only interface. Another 2 participants did not specify their interface preferences but raised concerns on potential issues such as misinterpretation and alarm in case of false positives, as well as reduced privacy. If these concerns were addressed, participants would be allowed to choose the amount of information shared with their caregivers, with depressive scores being the minimum amount of information.

Discussion

Principal Results

Despite some participants facing minor technical difficulties during the device setup and syncing process, participants were positive about the framework. Two participants had issues with syncing the device to their Fitbit accounts through the app. Of the 2, 1 participant successfully resynced the device upon viewing an instructional video on the syncing process, while the other was unreachable despite our attempts to engage with the participant. This participant was one of the excluded participants. The participant that resynced the device was not concerned or particularly inconvenienced by the issues. The procedure resulted in positive behavioral change and improvements in physical and mental well-being. Participants commented that participation made them more aware of their physical activity, with some using the daily survey as a means of journaling, leading to mental relaxation. For people with no caregivers, we recommend a volunteer function within the AutoMAP mobile app that could allow other nominated people to check on users. This could also benefit the volunteers living through the COVID-19 pandemic through added purpose or interaction.

Participants were concerned about the possibility of false positives sent to caregivers or false negatives not being sent to caregivers. To mitigate this, we will need to train and test our algorithm to achieve high performance when determining emotion recognition from the smart watch data. When privacy is a concern, device users might choose what extra information is visible to the caregiver. This could include vital signs, depressive symptom range history, and movement patterns. We are unsure if device users should see their own emotion-level

information, as it could cause subconscious bias or emotion alteration.

Although most participants favored a minimalist visual-based notification app, we will also provide options for semidetailed views for those that prefer this level of information in the app. We propose a mobile app that notifies caregivers when a user's scores are indicative of depressive tendencies, facilitates autonomy and privacy, and allows for interface selection. Users' scores will be determined through their physiological data and through machine learning techniques.

Strengths and Limitations

The procedure was convenient and easy to implement from a user experience perspective. Although some participants felt mildly concerned about the device setup and compliance with weekly and daily surveys, all participants that attended the follow-up interviews were positive about the general study design and ease of participation. We blinded participants to our study aims during our preprocedure interviews and informed participants of the study aims after the 4-week procedure. This type of blinding prevented subconscious response bias. The overall impression of the study was positive, with participants reporting that the 4-week study was well conducted, easy to follow, and had no significant inconvenience to their everyday lives. Generally, the daily survey was perceived to be clear and concise although in future iterations, efforts should be made to increase the range of the mood-rating components in the survey. This could potentially allow for finer, more-detailed self-reported emotion mapping.

Our study exclusively recruited older adults with no diagnosed mental health issues. A more diverse sample could improve the future application of the AutoMAP, as well as its performance and accuracy. Although participants generally found the study favorable, some preferred to not wear the device overnight. There may also be future issues regarding the sharing of private medical user data. The latter will be assessed and detailed in future research. Despite our efforts to maintain engagement with all participants, we were unable to interview 3 of the 12 participants for postprocedure interviews. Including the perspectives of these dropouts in our analysis might have provided a different perspective on the feasibility and acceptability of the procedure and framework.

Comparison With Prior Work

Our findings align with previous research on smart wearables for emotion recognition [23,24] or mental well-being [10,19]. These findings have shown that wearable sensors can be used for emotion recognition. Some of these studies [23,24] were in controlled settings, with deliberate emotion stimulation, where participants were not emotionally invested in the stimulus. These studies also observed participants for short time periods, meaning it is uncertain whether the stimuli used for emotion elicitation had the intended effect. By contrast, our procedure was performed in real-world settings, with participants wearing the Fitbit throughout their daily lives. Therefore, collected sensor data reflect emotions in real day-to-day life.

Promoting behavioral change through telephone counseling and smart wearables has shown promising outcomes [19]. Investigators also highlighted that competing wearable products may have different applications and accuracy, which should be accounted for in future studies using smart wearables. This study used a device from the Jawbone company, which is no longer manufacturing devices. This could render the implementation of procedures using Jawbone devices potentially invalid for long-term implementation. On the other hand, our study found feasibility for Fitbit devices in detecting preemptive depressive tendency rather than in being implemented as devices for behavioral therapy in people with diagnosed depression.

Conclusions

We will make four modifications for the future replication of our study: (1) predefined weekly and daily survey times for all participants, (2) fewer diversion questions, (3) assisted device setup and walk-throughs, and (4) more check-ins with participants during the study period.

This study is a vital step toward validating our framework and identifying requirements for the development of our proposed mobile app. Moving forward, we will perform a quantitative analysis on participant Fitbit data, develop the mobile app, and train and test our machine learning algorithm to detect depressive tendencies using physiological inputs. The machine learning component will be integrated with the mobile app to notify caregivers if the user's score is indicative of potential depression.

Acknowledgments

The authors would like to thank associate professors Phoebe Bailey and Kris Rodgers for their support and guidance.

Authors' Contributions

FM, WR, and JG conceptualized the study. IK consulted on the study design. FM, WR, PS, and JG contributed to the interview guideline, interpretation of results, and drafting and revision of the manuscript. FM conducted the interviews, performed data analysis, and wrote the first draft of the manuscript. WR, PS, and JG provided guidance on data analysis and critical feedback on the manuscript. All the authors read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

References

1. World Health Organization. Depression. URL: https://www.who.int/news-room/fact-sheets/detail/depression [accessed 2022-10-26]
2. Mental health of older adults Internet. WHO. 2017. URL: https://www.who.int/news-room/fact-sheets/detail/mental-health-of-older-adults [accessed 2022-10-26]
3. Depression is not a normal part of growing older. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/aging/depression/index.html [accessed 2021-09-08]
4. NIA. Supporting older patients with chronic conditions internet. National Institute on Aging. URL: http://www.nia.nih.gov/health/supporting-older-patients-chronic-conditions [accessed 2021-09-20]
5. Statistics. Beyond Blue. URL: https://www.beyondblue.org.au/media/statistics [accessed 2021-09-08]
6. Segal DL, Coolidge FL, Mincic MS, O'Riley A. Beliefs about mental illness and willingness to seek help: a cross-sectional study. Aging Ment Health 2005 Jul;9(4):363-367. [doi: 10.1080/13607860500131047] [Medline: 16019293]
7. Statista. Global connected wearable devices 2016-2022 Internet. URL: https://www.statista.com/statistics/487291/global-connected-wearable-devices/ [accessed 2021-09-08]
8. FitbitR revenue usage statistics (2022). Business of Apps. 2020. URL: https://www.businessofapps.com/data/fitbit-statistics/ [accessed 2021-09-08]
9. Bai Y, Tompkins C, Gell N, Dione D, Zhang T, Byun W. Comprehensive comparison of Apple Watch and Fitbit monitors in a free-living setting. PLoS One 2021;16(5):e0251975 [FREE Full text] [doi: 10.1371/journal.pone.0251975] [Medline: 34038458]
10. Chum J, Kim MS, Zielinski L, Bhatt M, Chung D, Yeung S, et al. Acceptability of the Fitbit in behavioural activation therapy for depression: a qualitative study. Evid Based Ment Health 2017 Nov;20(4):128-133 [FREE Full text] [doi: 10.1136/eb-2017-102763] [Medline: 29056608]
11. Ringeval M, Wagner G, Denford J, Paré G, Kitsiou S. Fitbit-based interventions for healthy lifestyle outcomes: systematic review and meta-analysis. J Med Internet Res 2020 Oct 12;22(10):e23954 [FREE Full text] [doi: 10.2196/23954] [Medline: 33044175]
12. Domin A, Spruijt-Metz D, Theisen D, Ouzzahra Y, Vögele C. Smartphone-Based Interventions for Physical Activity Promotion: Scoping Review of the Evidence Over the Last 10 Years. JMIR Mhealth Uhealth 2021 Jul 21;9(7):e24308 [FREE Full text] [doi: 10.2196/24308] [Medline: 34287209]
13. Ridgers ND, McNarry MA, Mackintosh KA. Feasibility and Effectiveness of Using Wearable Activity Trackers in Youth: A Systematic Review. JMIR Mhealth Uhealth 2016 Nov 23;4(4):e129 [FREE Full text] [doi: 10.2196/mhealth.6540] [Medline: 27881359]
14. Soon S, Svavarsdottir H, Downey C, Jayne DG. Wearable devices for remote vital signs monitoring in the outpatient setting: an overview of the field. BMJ Innov 2020 Jan 14;6(2):55-71. [doi: 10.1136/bmjinnov-2019-000354]
15. Vaportzis E, Clausen MG, Gow AJ. Older Adults Perceptions of Technology and Barriers to Interacting with Tablet Computers: A Focus Group Study. Front Psychol 2017 Oct 04;8:1687 [FREE Full text] [doi: 10.3389/fpsyg.2017.01687] [Medline: 29071004]
16. No I. Why do many seniors have trouble using technology? No Isolation. URL: https://www.noisolation.com/research/why-do-many-seniors-have-trouble-using-technology [accessed 2022-02-21]
17. Mughal F, Raffe W, Garcia J. Emotion recognition techniques for geriatric users: a snapshot. 2020 Aug 12 Presented at: IEEE 8th International Conference on Serious Games and Applications for Health (SeGAH); Aug 12-14, 2020; Vancouver, Canada. [doi: 10.1109/segah49190.2020.9201749]
18. Naslund JA, Aschbrenner KA, Barre LK, Bartels SJ. Feasibility of popular m-health technologies for activity tracking among individuals with serious mental illness. Telemed J E Health 2015 Mar;21(3):213-216 [FREE Full text] [doi: 10.1089/tmj.2014.0105] [Medline: 25536190]
19. Lyons EJ, Swartz MC, Lewis ZH, Martinez E, Jennings K. Feasibility and Acceptability of a Wearable Technology Physical Activity Intervention With Telephone Counseling for Mid-Aged and Older Adults: A Randomized Controlled Pilot Trial. JMIR Mhealth Uhealth 2017 Mar 06;5(3):e28 [FREE Full text] [doi: 10.2196/mhealth.6967] [Medline: 28264796]
20. Hartmann R, Schmidt FM, Sander C, Hegerl U. Heart Rate Variability as Indicator of Clinical State in Depression. Front Psychiatry 2018;9:735 [FREE Full text] [doi: 10.3389/fpsyt.2018.00735] [Medline: 30705641]
21. Trivedi MH. The link between depression and physical symptoms. Prim Care Companion J Clin Psychiatry 2004;6(Suppl 1):12-16 [FREE Full text] [Medline: 16001092]
22. Pollreisz D, TaheriNejad N. A simple algorithm for emotion recognition, using physiological signals of a smart watch. 2017 Presented at: 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); July11-15, 2017; Jeju Island, Korea p. 2353. [doi: 10.1109/embc.2017.8037328]
23. Quiroz JC, Geangu E, Yong MH. Emotion Recognition Using Smart Watch Sensor Data: Mixed-Design Study. JMIR Ment Health 2018 Aug 08;5(3):e10153 [FREE Full text] [doi: 10.2196/10153] [Medline: 30089610]
24. Shu L, Yu Y, Chen W, Hua H, Li Q, Jin J, et al. Wearable Emotion Recognition Using Heart Rate Data from a Smart Bracelet. Sensors (Basel) 2020 Jan 28;20(3):718 [FREE Full text] [doi: 10.3390/s20030718] [Medline: 32012920]
25. Zhang Z, Song Y, Cui L, Liu X, Zhu T. Emotion recognition based on customized smart bracelet with built-in accelerometer. PeerJ 2016;4:e2258 [FREE Full text] [doi: 10.7717/peerj.2258] [Medline: 27547564]

26. Xu W, Zammit K. Applying Thematic Analysis to Education: A Hybrid Approach to Interpreting Data in Practitioner Research. International Journal of Qualitative Methods 2020 Apr 14;19:160940692091881 [FREE Full text] [doi: 10.1177/1609406920918810]
27. Fereday J, Muir-Cochrane E. Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. International Journal of Qualitative Methods 2016 Nov 29;5(1):80-92. [doi: 10.1177/160940690600500107]
28. Mughal F, Garcia J, Raffe W, Stubbs P. Autonomous Monitoring for Older Aged Persons. OSF 2020 Dec 30:A. [doi: 10.17605/OSF.IO/GCYRM]
29. Haghayegh S, Khoshnevis S, Smolensky MH, Diller KR, Castriotta RJ. Accuracy of Wristband Fitbit Models in Assessing Sleep: Systematic Review and Meta-Analysis. J Med Internet Res 2019 Nov 28;21(11):e16273 [FREE Full text] [doi: 10.2196/16273] [Medline: 31778122]
30. Yesavage JA, Sheikh JI. 9/Geriatric Depression Scale (GDS). Clinical Gerontologist 2008 Oct 25;5(1-2):165-173. [doi: 10.1300/J018v05n01_09]
31. Greenberg S. How to try this: the Geriatric Depression Scale: Short Form. Am J Nurs 2007 Oct;107(10):60-69. [doi: 10.1097/01.NAJ.0000292204.52313.f3] [Medline: 17895733]

Abbreviations

AutoMAP: autonomous mental health monitoring for older aged people
GDS: Geriatric Depression Scale
HREC: Human Research Ethics Committee

Edited by T Leung, A Mavragani; submitted 30.09.21; peer-reviewed by M Bulgheroni, J Ye, W Zhang, Y Yu; comments to author 11.01.22; revised version received 06.04.22; accepted 09.09.22; published 29.11.22.

Please cite as:
Mughal F, Raffe W, Stubbs P, Kneebone I, Garcia J
Fitbits for Monitoring Depressive Symptoms in Older Aged Persons: Qualitative Feasibility Study
JMIR Form Res 2022;6(11):e33952
URL: https://formative.jmir.org/2022/11/e33952
doi:10.2196/33952
PMID:36268552

Original Paper

Using Magnetic Resonance Imaging During Childbirth to Demonstrate Fetal Head Moldability and Brain Compression: Prospective Cohort Study

Olivier Ami[1,2*], MD; Jean-Christophe Maran[2*], BE; Dominique Musset[3*], MD, PhD; Claude Dubray[4*], MD, PhD; Gérard Mage[5*], MD, PhD; Louis Boyer[2,6*], MD, PhD

[1]Clinique de la Muette - Ramsay Santé, Paris, France

[2]Image Guided Therapies, Institut Pascal, UMR 6602, Université Clermont Auvergne and CNRS, Paris, France

[3]Université Paris Sud 11, Paris, France

[4]Centre d'investigations cliniques, Université Clermont Auvergne, Clermont Ferrand, France

[5]Service de gynécologie-obstétrique, Université Clermont Auvergne, Clermont Ferrand, France

[6]Pole radiologie, CHU Clermont Ferrand Gabriel Montpied, Clermont-Ferrand, France

[*]all authors contributed equally

Corresponding Author:
Olivier Ami, MD
Clinique de la Muette - Ramsay Santé
46 rue Nicolo
Paris, 75116
France
Phone: 33 661143924
Fax: 33 178767852
Email: olivier.ami@birthimaging.com

Abstract

Background: Childbirth is a physiological process with significant medical risk, given that neurological impairment due to the birthing process can occur at any time. Improvements in risk assessment and anticipatory interventions are constantly needed; however, the birthing process is difficult to assess using simple imaging technology because the maternal bony pelvis and fetal skeleton interfere with visualizing the soft tissues. Magnetic resonance imaging (MRI) is a noninvasive technique with no ionizing radiation that can monitor the biomechanics of the birthing process. However, the effective use of this modality requires teamwork and the implementation of the appropriate safeguards to achieve appropriate safety levels.

Objective: This study describes a clinically effective and safe method to perform real-time MRI during the birthing process. We reported the experience of our team as part of the IMAGINAITRE study protocol (France), which aimed to better understand the biomechanics of childbirth.

Methods: A total of 27 pregnant women were examined with 3D MRI sequences before going into labor using a 1-Tesla open-field MRI. Of these 27 patients, 7 (26%) subsequently had another set of 3D MRI sequences during the second stage of labor. Volumes of 2D images were transformed into finite element 3D reconstructions. Polygonal meshes for each part of the fetal body were used to study fetal head moldability and brain compression.

Results: All 7 observed babies showed a sugarloaf skull deformity and brain compression at the middle strait. The fetus showing the greatest degree of molding and brain shape deformation weighed 4525 g and was born spontaneously but also presented with a low Apgar score. In this case, observable brain shape deformation demonstrated that brain compression had occurred, and it was not necessarily well tolerated by the fetus. Depending on fetal head moldability, these observations suggest that cephalopelvic disproportion can result in either obstructed labor or major fetal head molding with brain compression.

Conclusions: This study suggests the presence of skull moldability as a confounding factor explaining why MRI, even with the best precision to measure radiological landmarks, fails to accurately predict the modality of childbirth. This introduces the fetal head compliance criterion as a way to better understand cephalopelvic disproportion mechanisms in obstetrics. MRI might be the best imaging technology by which to explore all combined aspects of cephalopelvic disproportion and achieve a better understanding of the underlying mechanisms of fetal head molding and moldability.

(JMIR Form Res 2022;6(11):e27421) doi:10.2196/27421

KEYWORDS

parturition; magnetic resonance imaging; obstetrics; fetus; cephalopelvic disproportion

Introduction

Background

Events occurring within the first few hours of birth can cause fetal cerebral palsy or maternal perineum trauma, determining a human's life trajectory. There is an urgent need for predictive and preventive tools to prevent avoidable birth traumas. Magnetic resonance imaging (MRI) is a noninvasive and nonirradiating tool suitable for exploring the biomechanics of the birthing process and evaluating fetal well-being. A clinical protocol to achieve this goal requires numerous safeguards to ensure the same level of safety in the MRI suite as in the delivery room.

In 1948, Mengert [1] described five components of cephalopelvic disproportion: (1) size and shape of the bony pelvis, (2) size of the fetal head, (3) force exerted by the uterine contractions, (4) moldability of the fetal head, and (5) presentation and position of the fetal head. A sixth component with intrapelvic organs and structures shaping the birth canal might also be considered [2]. Exploring these factors with imaging during birth is crucial to the prediction and prevention of cephalopelvic disproportion.

MRI is routinely used to explore fetal malformations antenatally and perform obstetric pelvimetry [3]. This technique, which is neither invasive nor irradiating, is safe for the birthing person and the fetus [4-9]. Parturients near term usually require an open-field MRI or a large oval bore, given their increased abdominal circumference. However, despite the potential advantages of MRI when evaluating the birthing process in real time, we identified only a few sites worldwide with an open-field MRI near a maternity ward that would enable the exploration of childbirth in all positions.

Prior Work

In 2010, Dr Christian Bamberg (Charité Hospital, Berlin) demonstrated the feasibility of MRI in childbirth [10,11] in response to a clear need to observe images of this process, which had only previously been described by obstetricians. Since then, MRI has made an enormous progress, given that it offers high-resolution 3D imaging and can evaluate the properties and composition of various tissues. Magnetic resonance (MR) elastography techniques are now available to measure tissue elasticity [12], and diffusion MRI explores water molecule movement in cells and tracks fibers on fetal cerebral nerve cells, uterine cells, and perineal muscle cells [13,14].

MR spectroscopy can explore tissue composition while acquiring qualitative and quantitative information on the presence of various molecules such as lactates [15-17], their tissue distribution, and changes in their concentrations over time. Multinuclear MR spectroscopy can explore fetal energy reserves in models of chronic fetal distress [18]; it can also study the maturity of fetal organs [19] as part of prematurity studies or in the exploration of certain metabolic diseases or birthing person-to-fetus drug transmission. Functional brain MRI with blood oxygen level–dependent imaging [20-22] may also be used, whereas proton spectroscopy can be used to assess lactate levels to examine the health of fetal tissues and the placenta [23].

In the short term, information available from 3D MRI during the birthing process can simulate childbirth and facilitate virtual reality birthing tests that are closer to reality than pelvimetry [24,25]. Moreover, a better understanding of fetal head compression could explain spontaneously observed cerebral hemorrhages and fetal heart rate abnormalities that can occur during the birthing process [26-28]. Thus, MRI can reveal important information that could be used at this crucial time.

Goal

We report our team's experience with the IMAGINAITRE research protocol (France), which was approved by French authorities and a French National Ethics Committee. We aimed to observe the biomechanical changes of the fetal head and maternal perineum during human birth [2,29]. This paper describes a clinically effective and safe method for both the birthing person and child to perform real-time MRI during the birthing process.

Methods

Research Protocol

This research protocol was developed by the University Hospital of Clermont-Ferrand (Auvergne University) to meet specific research objectives and identify suitable locations to conduct this research in France. In our study, a public or private partnership emerged because of the location of and the equipment available at the private establishment that hosted this protocol. This private institution, located in Evry (Hopital Privé de l'Essonne, F91000, France), had clinical permits to operate both a maternity ward and a radiology department with open-field MRI in the same building.

Ethical Considerations

This prospective biomedical interventional study (IMAGINAITRE) was approved by the French ethical institutional review board "Ile de France II" (ID-RCB 2012-A01469-34) and the French National Agency for Drug and Medical Product Safety (*Agence Nationale de Sécurité du Médicament et des Produits de Santé*) and promoted by the University of Clermont-Ferrand Medical Center. All the women agreed to participate and signed an informed consent.

Preliminary Organization of Clinical and Radiological Teams

Performing an MRI during childbirth requires close collaboration among the clinical teams that usually support the patient in the delivery room. The birthing process is

unpredictable in onset and duration; pregnant patients can deliver at any time, which does not correspond well with the scheduled studies of most radiology departments. It is important that both the obstetric and radiology teams are comfortable with performing the required MRI protocol safely and that patient management remains focused on the birth and not on obtaining images. It is necessary that the team dedicated to obtaining images be available 24 hours a day, 7 days a week, which requires close cooperation between the imaging and maternity departments; an on-call list of staff should be dedicated to this activity. All members of the research protocol team should be trained and ready to make themselves MR safe (ie, by removing any metal or devices from their body and selecting only nonmagnetic, MR-safe equipment), should they be required to enter the magnetic field to perform any emergency procedure.

The delivery suite should be located near the imaging suite to allow staff to carefully monitor the patient during transfer and throughout the examination. A formal protocol should exist within each department outlining the equipment needed for patient transfer between the departments, including a dedicated stretcher, a portable epidural kit with a syringe pump and remote bolus, a sterile protection equipment and aseptic products, obstetric instruments, and a nonmagnetic delivery kit.

The staff dedicated to the protocol must remain with the patient from the time they leave the delivery room. The obstetrician and anesthesiologist must always remain at the patient's side. The other parent's presence during transfer is often reassuring for the parturient in labor. The transfer time between the departments must not exceed a few minutes; fetal monitoring should be performed continuously. If battery monitoring tools are used, monaural stethoscopes must be available should the electronic devices fail. The transfer of the patient to the MRI suite should not occur if there are signs of fetal distress, an abnormal heart rhythm, clear amniotic fluid, or any pain or other abnormalities noted during labor. An obstetrician or midwife should perform an examination before departure, according to the strict protocol.

A dedicated operating room must be available before the patient's departure from the obstetric suite and throughout the imaging procedure; staff should be trained and provided with the appropriate equipment to perform the protocol and the requisite emergency procedures. These research activities must have ethics and institutional approval, and they must also be covered by insurance.

Patient Selection

The inclusion criteria are as follows: parturients who are in the third trimester; who are aged 21 to 39 years; who are primipara, secundipara, or tertipara; who are with no known risk factors that can affect delivery; and who have provided their informed consent to participate. The exclusion criteria include fetal breech presentation or presentation that is not strictly cephalic at the time of delivery, a scarred uterus, multiple pregnancy, known maternal or fetal medical conditions or those requiring urgent care, and a fetal heart rate requiring care every 30 minutes. The contraindications for open-field MRI are as follows: a ferromagnetic foreign body in the patient's body, claustrophobia, parturients aged <21 years, a vulnerable adult, an abnormal fetal heart rate requiring treatment within 30 minutes, contraindications to MRI, or no prior obstetrical disease during pregnancy.

The patients are selected in advance during pregnancy monitoring and are asked about contraindications to MRI (ie, claustrophobia and ferromagnetic foreign bodies). The physician then discusses the research protocol with the couple and arranges a visit to the MRI, maternity ward, and imaging services. The patients review, sign, and return the consent forms if they agree to participate.

The patient is then invited to a prepartum MRI examination on the date closest to the start of the term period to establish a baseline examination; ideally, the other parent is present so that the patient can experience the imaging protocol as it will take place during labor, which usually lasts only 3 to 5 minutes. Any remaining questions are addressed, and the signed consent forms are obtained. Participants' charts are indicated with a label unique to them.

Clinical Support When Leaving the Delivery Room

An MRI-specific delivery procedure is displayed in the delivery room. When the enrolled patient goes into spontaneous labor and enters the delivery room, the on-duty midwife calls the protocol coordinator. The progress of the labor can be monitored by the local on-call team until dilation is achieved. If imaging is scheduled to take place during the second phase of labor, the entire team should arrive before the patient reaches cervical dilation no greater than 8 cm. Once the entire team is on site, the obstetrician examines the patient to confirm presentation height and cervical dilation.

Once the second stage of labor is reached, an anesthetic epidural bolus is administered, and the patients are asked not to push; they are then transferred to the MRI table. In our hospital, the average transfer time between the bed in the delivery room and the MRI table is 2 minutes and 36 seconds. Once the patient is in the MRI suite, the nonmagnetic transport stretcher is positioned parallel to the MRI table outside the range of the 5-Gauss line. Sealed and sterile protections are placed on the MRI table (Figure 1).

Figure 1. (A) Open magnetic resonance imaging (MRI) room with sterile drapes and a delivering patient. (B) Patient in the open-field MRI after antenna installation and being positioned inside the magnetic field.

Imaging Process

The patient is settled on the nonmagnetic stretcher for transfer to the MRI machine. After the administration of a last epidural bolus, the epidural catheter can be closed by a plastic cap. The patient is then positioned in the MRI machine with a surface or bridge antenna in the open-field MRI, enabling mobility. The bearing areas should be protected by padded reinforcements; the patient should be comfortable (Figure 2). The patient wears earplugs and ear protectors to reduce the scanner acoustic noise. The antenna is centered on the pelvic region, which is wide enough to get the signal below the perineum and back to the top of the uterus. The patient is provided with preliminary instructions; a first localization examination is performed. The sequences are then implemented with centering in the space.

The imaging protocol may be adjusted as needed. The MRI sequences should be optimized beforehand with an application engineer to obtain fine contiguous images that balance resolution and acquisition speed. The region of interest should be broad, and the number of excitations is limited to the acquisition of a correct signal. The use of the rapid-filling technique in the Fourier plane and a superconducting antenna significantly reduces acquisition times. The fetus moves very little in the pelvis; however, diaphragmatic movements are more significant because of maternal breathing. Therefore, it is advisable to ask the birthing person to maintain apnea when acquiring images of the subcostal region. Once static acquisitions are obtained in the 3 planes in the decubitus dorsal position in T1 gradient echo throughout the uterine contents and just below the perineum, rapid-sequence, contiguous T2 images centered on the fetal brain and the birthing person's perineum are performed (Figure 3). A dynamic sequence can then be launched in true fast imaging with steady state precession in which 1 image per second is obtained for 60 seconds along the sagittal and frontal planes at the base of the iliococcygeal head of the levator ani; this informs the movement of various muscle groups or other organs during pushing (Figure 4). 3D reconstructions can be adjusted, as the contours observed in 2D can be altered when a change is observed in a region of interest. Posture changes can then be made with a new 3D, T1 gradient echo in the 3 planes; these scans can observe the modification of the trajectory of the fetal head as a consequence of postural change. Furthermore, the delivery process can be observed by dynamic imaging or the acquisition of successive static volumes.

Figure 2. Patient installed in the open-field magnetic resonance imaging, with a gynecologist by her side.

Figure 3. (A) T2 gradient echo sequence showing an axial slice centered on the fetal brain at the superior inlet level. (B) The same sequence at the iliac bone level. (C) T1 gradient echo sequence in the sagittal plane showing bony pelvis and fetal head before entering into labor. The empty bladder is in retropubic position. (D) Same T1 gradient echo sequence showing the fetal head molding in the middle brim during second phase of labor with a full bladder. The fetal head is rotated in occipito-pubic position, and the full bladder is ascended above the upper limit of the pubic bone.

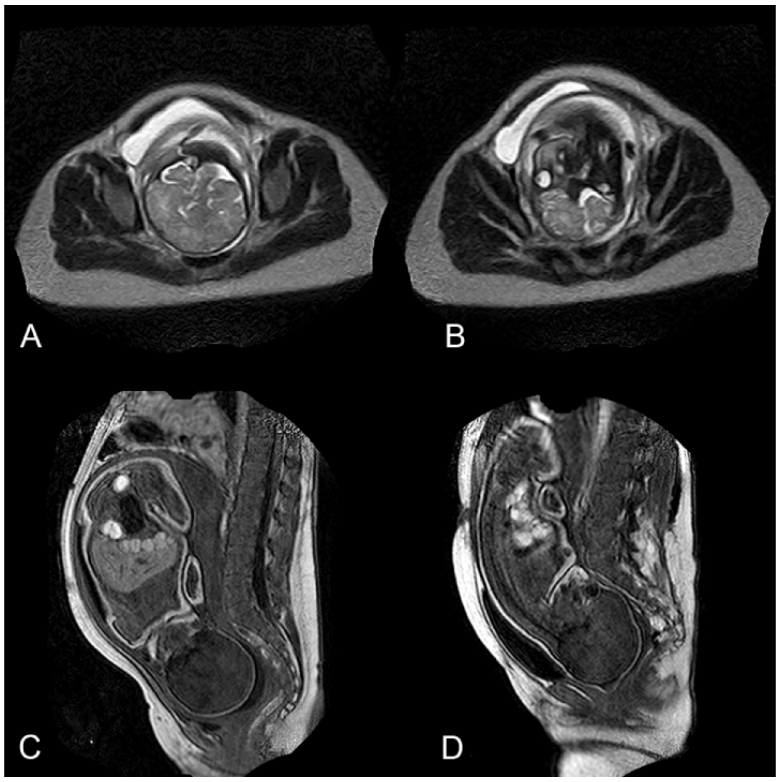
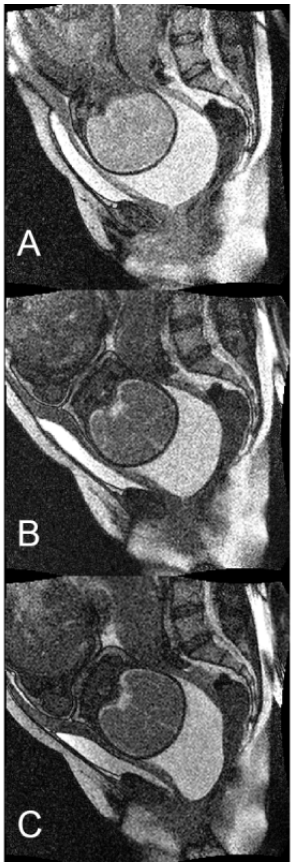


Figure 4. T2 dynamic imaging during maternal effort (successive images) at complete cervix dilation and unruptured membranes (dynamic acquisition). (A) Before maternal efforts: fetal head is above the sacro-pubic line. (B) Beginning of maternal efforts: the membranes start to change form to underline the vaginal route. (C) The intact membranes are approaching the superficial perineum level inside the vagina, the coccygeal bone is pushed posteriorly, the fetal head is under the superior limit of the pubic bone, and the bladder and the urinary sphincter are descending with the fetal head.



Foreseeable Risks and Emergency Protocols

Overview

According to the legal provisions in France, maternal and fetal exposure to magnetic fields must not exceed 2 Tesla for temporary, whole-body exposure. No adverse effects have been reported since MRI exists and has been used worldwide during pregnancy [30]. Given the potential adverse effects of gadolinium on fetuses, the use of MR contrast media is not recommended for parturients [31].

We used an open-field MRI machine low-magnetic field (1T; Phillips Panorama). The measurement of the cumulative maternal specific absorption rate should remain below the maximum exposure recommendations. For comfort and tolerance, we do not recommend using MRI in parturients for more than 90 minutes [32-34]. In our protocols, we never exceeded more than 15 minutes of cumulative examination time.

There are 2 types of foreseeable risks during such protocols: those related to the use of MRI and those unrelated to the use of MRI. Risks associated with the use of an MRI machine in parturients during labor are detailed below.

Risks Related to the Use of MRI

Claustrophobia

Anxiety and agitation related to the phobia of enclosed spaces is possible, although this type of anxiety is usually mild and quickly resolved. These incidents are typically mitigated by using an open-field device and by short imaging sequences. The patient can be quickly and easily removed from the MRI room with staff standing nearby. To anticipate and mitigate this risk, pelvic MRI should be scheduled during the eighth or ninth month of pregnancy; this allows patients to become familiar with the MRI procedure and meet the protocol team. Any questions or concerns can be addressed to enhance the patients' understanding of the study.

Scanner Acoustic Noise to the Birthing Person and Fetus

The intensity of the sound wave generated by the MRI machine depends on the amplitude of the vibrations transmitted to the coils by the geometry of the sequence, which means that the amount of noise depends on the sequence type. The noisiest sequence with the Philips Panorama 1T MRI machine is the *real-time balanced TFE* at 108 dB, whereas the gradient echo T1 and T2 sequences proposed to image the birthing process generate approximately 90 to 95 dB. Temporary exposure of the public in France is limited to an average intensity of 105 to 120 dB peak sound pressure to prevent hearing damage.

For the birthing person, the sound level related to MRI sequence acquisition is reduced by using earplugs or ear protection, which are put in place before entering the machine. For the fetus, scanner acoustic noise can be perceived with intact or ruptured membranes.

Intact Membranes

The fetus is surrounded by amniotic fluid; sound waves reaching here initially propagate in the air. If a planar sound wave propagates without weakening in medium 1 and passes through medium 2, the sound intensity generally decreases. According to Snell law, with the passage of sound waves from air to water, the fetus is protected from scanner acoustic noise in its amniotic fluid, given that the sound intensity transmitted to the fetus is on the order of 0.12% of the incident intensity.

Ruptured Membranes

Ruptured membranes occur in cases where the cervix starts to expand and the patient is in labor. Here, the lower segment is extended, and the myometrium in the lower segment is molded on the fetal head. Owing to the proximity of the impedance of the biological tissue to water, we are only interested in reducing scanner acoustic noise by blocking the fetus's ears in the lower segment, where sound is directly transmitted to the eardrum through fetal aerial vibration. In this case, the geometry of the acoustic wave is highly attenuated by the fact that the fetus is in a sealed enclosure: a little amniotic fluid is often trapped over the fetus's head, as the lower segment molds the cephalic pole.

Acoustic attenuation depends on the thickness of the maternal tissues that protect the fetal eardrum from scanner acoustic noise. Although this has never been measured in utero, the calculation of noise reduction was at least 30 dB, which would reduce the noise from 95 to 65 dB for about 10 minutes, which is less than the noise made by a domestic vacuum cleaner. There are no adverse consequences to the unborn child's hearing; however, when the child is born, the absence of tympanic protection requires the sequences can be continued after the use of specific hearing protection when the ears are viewed at the level of the vulva.

Risks Not Related to the Use of MRI

In the event of unexpected maternal discomfort, the imaging procedure is stopped immediately, and the anesthesiologist can provide immediate care. All equipment required for resuscitation is available throughout the procedure, as is an operating room, if required.

If membrane rupture occurs during the examination, it can cause the sudden emission of a stream of amniotic fluid. Therefore, the exposed MRI surfaces are protected by disposable, sterile, and waterproof drapes. Furthermore, if delivery occurs during the MRI scan, a dedicated midwife and obstetrician (in addition to the clinical care team) are present in the MRI room. All equipment necessary to manage childbirth in the MRI room is readily available (plastic disposable material, nonmagnetic forceps and scissors, drapes, and compresses), including a nonferromagnetic plastic suction cup if an instrumental delivery is required.

Finally, there could be instances in which maternal or fetal emergency or both occur during the examination. Three situations can occur: cord prolapse, abnormal fetal heart rate with the risk of acidosis, and sudden bleeding. To detect these risks, the entire obstetric team will remain with the patient at every stage of the imaging procedure. An operating room to perform a cesarean section and emergency extraction is available for the duration of the procedure. This operating room is located such that the child's extraction can be guaranteed within 10 minutes of problem detection. Furthermore, fetal heart rate and uterine contractions are monitored during patient transfer and throughout the procedure using a portable, low-current

measurement device. If a suitable fetal heart rate monitoring device is not available in the MRI suite, intermittent auscultation can be used for a cumulative period not exceeding 10 minutes.

Patients participating in the protocol will be transferred to the MRI room only if an effective epidural is administered and dosed correctly. An anesthesiologist dedicated to the protocol will be on site to administer the epidural and maintain optimal levels of pain relief.

Image Postprocessing Method

We recommend using a 3D finite element postprocessing station. This can aid in the manual or automated reconstruction of the vector contours of the fetal organ and birth canal surfaces (Figure 5). Finite elements meshing thus permits the removal of most imaging-related artifacts. Vector meshing can be printed in 3D or exported for use in downstream imaging simulation software. A 3D fetal mesh can then be used for biomechanical studies and to retro-evaluate forces applied to the fetal head (Figure 6).

Figure 5. 3D reconstruction and finite element meshing. (A) 2D image with mesh contouring. (B) Artifacted 3D mesh. (C) 3D volumes after optimization process.

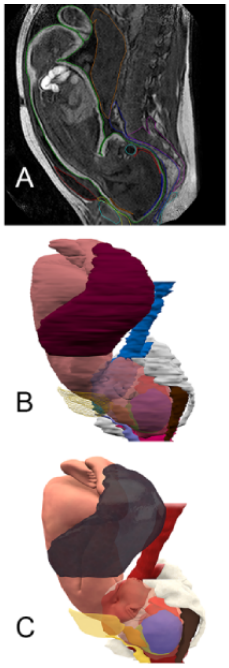
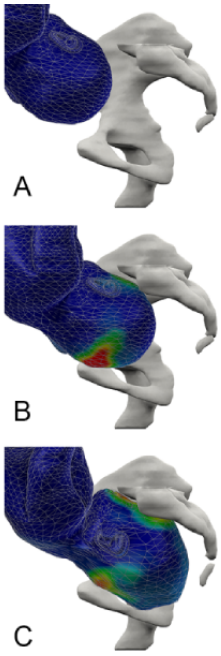


Figure 6. 3D fetal mesh in false colors (blue) showing the forces applied on the fetal head and their gradient (the greatest forces are displayed in hot colors) from left to right. (A) Head is above the superior inlet. (B) The fetal head is engaging at the superior brim with a contact of the occiput on the pubic bone. (C) The fetal head is engaging at the inferior brim and shows an important fetal head molding.



Results

Outcomes Description

Of the 27 patients who agreed to participate in the protocol, 7 (26%) presented with conditions that allowed for an MRI during the second phase of labor. The 7 patients were, on average, aged 28 (range 23-34) years ; 3 (43%) were primipara, 3 (43%) were secundipara, and 1 (14%) was tertipara. Moreover, of these 7 patients, 3 (43%) were beyond 39 weeks of gestation, and 4 (57%) were past their due date. The membranes were ruptured in 86% (6/7) of the patients, and the patients were at rest during the MRI scans, just before pushing.

The Magnin index was calculated by adding the obstetric conjugate diameter to the median transverse diameter.

The median Magnin score of the 7 patients with imaging during the second stage of labor was 26 (IQR 25.1-27.5). None of the patients in our series showed any modification of pelvic diameters at the superior brim during the second phase of labor, in comparison with the MRI acquired before entering into labor.

The newborns had a median weight of 3755 (IQR 3095-4525) g, and all 10-minute Apgar scores were 10.

In total, 71% (5/7) of babies were born by natural delivery without instrumental extraction, and 29% (2/7) were born by emergency cesarean section because of stagnation when fully dilated, of which 1 (14%) was after forceps failure. In both cases, the pregnancies were overdue.

No patient had pelvic dystocia predictable from the size of their pelvis, given that all Magnin scores were above 25.

Rotation During Labor

Of the 7 observed babies, 6 (86%) rotated to face their placentas during labor, driven by contractions, with the convexity of the baby's back responding to the smooth concavity of the uterus.

Bladder Behavior

The bladder of all birthing persons who retained a postmicturition residue before delivery reascended above the pubic bone plane at the time of expulsive efforts. When emptied, the bladder fell below the pubic bone plane and appeared at risk of being dragged by rubbing between the immobile bony wall of the pelvis and the fetal head (Figure 3).

Fetal Skull Compliance and Brain Compression

All 7 observed babies showed a sugarloaf skull deformity and a brain compression at the middle strait. This deformation remained visible in only 29% (2/7) of babies immediately after delivery. During this deformation, a sequence involving the skull bones, fontanel closure, tentorium cerebelli, and falx cerebri was initiated, resulting in movements of cerebrospinal fluid from the lateral ventricles and brain periphery toward the skull base and large cistern. All 7 observed babies showed cerebrospinal fluid distributed around the cerebellum during the fetal skull molding at the middle brim.

The fetus showing the greatest degree of molding and brain shape deformation weighed 4525 g and was born spontaneously but also presented with a low Apgar score.

Discussion

Principal Findings

In the case of the fetus with the greatest degree of fetal skull molding, observable brain shape deformation demonstrated that brain compression occurred, and it was not necessarily well tolerated by the fetus, which illustrates the moldability criteria described by Mengert [1]. This observation, far from ruling out a cephalopelvic disproportion, suggests the presence of skull moldability as a confounding factor explaining why MRI, even with the best precision to measure radiological landmarks, fails to accurately predict the modality of childbirth. This introduces the fetal head compliance criterion as a way to better understand cephalopelvic disproportion mechanisms. This observation is probably the most remarkable, given that it allows the rediscovery of the wealth of scientific work performed on the phenomenon of static brain compression of the child during the labor and delivery processes [35-39], and revives the notion of fetal head compliance in the interpretation of cephalopelvic disproportion situations at birth.

The fact that the cerebrospinal fluid moves to the posterior fossa during the shaping of the fetal skull is another interesting finding on the shock-absorber role that this fluid can play, and the circulation channels that are responsible for these hydraulic movements. During this process, the subarachnoid spaces widen in the posterior fossa.

The rotation of infants facing their placenta during labor was also an interesting finding in this work. It appears that the placenta is a major determinant of the presentation of the child to the upper strait engagement, whose position can be determined well in advance. Whether the risk of occipito-sacral presentation can be effectively predicted when the placenta is anterior, for example, should be investigated in the future. This would allow for more complete maternal information about the upcoming delivery to be given early in the second trimester.

Finally, the important logistics realized by our MRI during patients' labor allowed us to observe the behavior of the bladder, which had filled by the time the imaging was performed, when the head of the child was at the middle strait. For organizational reasons, we could not empty the bladder in the MRI room, so we had to do so before leaving the delivery room. Most of our patients' bladders filled during the successive examinations, and we could see that full bladders moved up above the pubic bone, whereas when they were empty, they were literally stuck between the infant's head and the pubic bone. This is important because most teams empty the bladder just before expulsive efforts. However, this work suggests that emptying the bladder at least half an hour before beginning maternal expulsive efforts might be more protective and prevent the bladder from being dragged by the fetal head below the pubic plane.

Comparison With Prior Work

Fetal head molding and brain shape changes during the second stage of labor were previously reported [29], but the criterion of fetal head compliance as a confounding factor in the interpretation of the presence or absence of cephalopelvic disproportion has not been sufficiently made explicit.

Regarding cephalopelvic disproportion, 2 biomechanical situations can occur during labor: the fetal head is perfectly adapted and fits the maternal birth canal without the need of molding, which occurs in eutocia; or the fetal head is of a size or geometry unsuited to pass through the birth canal and the fetal skull needs to mold to fit the maternal birth canal, which is a case of cephalopelvic disproportion. In the second case, depending on the compliance degree of the fetal head, we then observe 2 possible scenarios.

First, in deliveries with low fetal skull compliance, cephalopelvic disproportion will not allow the fetal skull to mold sufficiently to pass through the birth canal. Given that the skull bones will not overlap accordingly, uterine cervix dilatation will stop, and the labor will be defined as obstructed. The fetal head will remain high in the pelvis, and fetal skull molding will be very slow. If not resolved by a cesarean section in time, the prolonged ischemia of the fetal tissues can lead to craniotabes, caput succedaneum, and cephalohematoma; the compression of the maternal tissues can lead to necrosis or fistulas, and a postpartum hemorrhage due to uterine atony can occur. Currently, these situations can fortunately be resolved by most obstetrical teams by performing a cesarean section within 3 hours of consecutive cervix dilation stagnation, which is generally well tolerated by the fetus, except in cases of relentlessness and cesarean section performed too late or if an attempt at instrumental extraction fails. Nevertheless, traumatic delivery for the mother is frequent in these situations (red-code cesarean section, perineal tears, difficult extraction, and postpartum hemorrhage).

The deliveries presenting a cephalopelvic disproportion with high compliance are among the most puzzling because the fetal skull will adapt itself in a sugarloaf shape, thus reducing its lateral caliber in favor of an elongation in height in the occipito-bregmatic direction. The brain shape will follow this deformation. The consequences observed here are the migration of the cerebrospinal fluid from the ventricular and pericerebral spaces to the posterior fossa as well as a global folding of the cerebral mass. Thus, there are shear forces and pressure forces present, which are propelled by uterine contractions. These births can appear to be “normal deliveries” according to the usual obstetrical criteria because they will occur vaginally, most often without any identified mechanical difficulty, and they can even be rapid, with an apparent asymptomatic child. Therefore, these deliveries could appear normal for the birthing parent but be genuinely traumatic for the child and might not be identified by the current standard neonatal evaluation criteria.

The presence of these forces [28], which are exerted directly on the brain and whose effect is directly visible in imaging in our study through the morphological deformation of the skull and the brain that they induce [29], are crucial to recognize and investigate.

The cerebral handicap linked to birth is based on 2 possible events: a brain hemorrhage and a brain ischemia. Ischemia is easily detected by noting increased lactate levels and by the presence of metabolic acidosis according to pH and blood gases. Cerebral and retinal hemorrhage, on the other hand, is often minimally symptomatic in the newborn and does not alter pH or lactate. Some of these hemorrhages have radiological similarities to shaken baby syndrome, which is known to often progress to disability when occurring in a developing brain.

The cerebral hemorrhages described in the literature can reportedly occur in up to 43% of asymptomatic newborns [26,27,40] and retinal hemorrhages in up to 50% of all asymptomatic newborns [41-44]. Most of the affected newborns were born vaginally or by late cesarean section. Very few were born by scheduled cesarean section; however, the possible presence of uterine contractions even in the cesarean sections described as scheduled does not mean that they did not occur during labor. Prospective studies with a thorough analysis of these elements are needed to increase the knowledge on this subject. It is important to remember that these newborns are typically asymptomatic with the usual evaluation criteria, that is, the children do not show loss of consciousness, motricity is preserved, and there are no visible seizures, so the occurrence of these abnormalities is never routinely investigated and is, therefore, mostly ignored. The detection, study, and prevention of these hemorrhages could soon be the object of investment by national and international public health organizations.

The detection of neonatal concussion due to high-compliance cephalopelvic disproportion [28,38,39] could be improved by the identification of indirect signs of such dystocia by the obstetrical team [45], such as abnormalities of the fetal heart rate with a vagal pattern during labor [46], the presence of significant dislocation or overlap of the skull bones on examination during labor, the presence of significant sugarloaf molding of the fetal head after delivery, and the presence of a significant caput succedaneum or a cephalhematoma, sometimes with softening of the skull vault bones on touch.

It would also be necessary to reinforce the evaluation of the neurological state of the child at birth, whose behavior alone and the conservation of motricity will not alert on the presence of a brain concussion, such as after a “boxing match without knockout.” This could be achieved by the systematic application of the Sarnat score [47,48] by midwives at birth; the integration of retinoscopy without pupillary dilation in the neonatal examination by pediatricians [49]; the umbilical cord sampling by obstetricians for brain concussion markers in the cord blood, such as the S100B protein [50]; and the performance of brain imaging in the event of an abnormal value of these markers.

Our MRI observations during delivery were all performed on patients with a normal or rather large pelvis according to the usual obstetrical management criteria, which allowed us to better understand, by having been able to observe these different cases of high or low fetal head compliance dystocia, why radiological pelvic examinations alone typically fail to detect these situations effectively. It is necessary to compare the geometry and size of the child’s skull with the geometry and the dimensions of the maternal pelvis, and not to forget that the compliance of the fetal head remains for the moment an unknown criterion; therefore, the focus of interpretation should not be on the simple modality of the delivery route but on the existence or not of a traumatic delivery situation for the mother or child or both and the development of the resources required to detect these traumas.

Limitations

Providing imaging evidence of what happens during the physiological process of birth is paramount. However, the small sample size of the populations studied with MRI during labor and delivery causes the requirement for many more prospective imaging studies (with all available modalities of medical imaging) to further explore the complexity of the biomechanical events occurring during birth.

Conclusions

Our observations with MRI during childbirth suggest that, depending on fetal head moldability, cephalopelvic disproportion can result in either obstructed labor or apparent normal vaginal delivery with major fetal head molding and brain compression. MRI might be the best imaging technology by which to explore all the combined aspects of cephalopelvic disproportion already described by Mengert [1] and to achieve a better understanding of the underlying mechanisms of fetal head molding and moldability.

Prospective randomized cohort studies should establish a policy of routine bladder emptying earlier than at the time of expulsive efforts. Placental location could also be useful information for anticipating the presentation of the child at delivery.

The use of MRI in pregnancy and delivery creates new avenues through which to understand the physiology of the biomechanical process of childbirth and fetal adaptation during the passage to extrauterine life. The ability to share experiences related to the safety protocols that are currently being undertaken or planned using real-time MRI of the human birthing process could accelerate research and knowledge acquisition on this topic. Thus, a new era of "prenatological imaging" is emerging in a world of predictive and preventive medicine.

Conflicts of Interest

None declared.

References

1. Mengert WF. Estimation of pelvic capacity. J Am Med Assoc 1948 Sep 18;138(3):169-174. [doi: 10.1001/jama.1948.02900030001001] [Medline: 18877482]
2. Maran JC, Cassagnes L, Delmas V, Musset D, Frydman R, Mage G, et al. Comparative anatomy on 3-D MRI of the urogenital sinus and the periurethral area before and during the second stage of labor during childbirth. Surg Radiol Anat 2018 Apr;40(4):371-380. [doi: 10.1007/s00276-017-1925-9] [Medline: 28948372]
3. Nagayama M, Watanabe Y, Okumura A, Amoh Y, Nakashita S, Dodo Y. Fast MR imaging in obstetrics. Radiographics 2002;22(3):563-582. [doi: 10.1148/radiographics.22.3.g02ma03563] [Medline: 12006687]
4. Shaikh MS, Lombay B. Fetal MRI : reviewing the history, indications, technique, safety and drawbacks. J Coll Physicians Surg Pak 2004 Sep;14(9):576-579. [Medline: 15353148]
5. Jackson HA, Panigrahy A. Fetal magnetic resonance imaging: the basics. Pediatr Ann 2008 Jun;37(6):388-393. [doi: 10.3928/00904481-20080601-08] [Medline: 18616192]
6. De Wilde JP, Rivers AW, Price DL. A review of the current use of magnetic resonance imaging in pregnancy and safety implications for the fetus. Prog Biophys Mol Biol 2005;87(2-3):335-353 [FREE Full text] [doi: 10.1016/j.pbiomolbio.2004.08.010] [Medline: 15556670]
7. Tsai LL, Grant AK, Mortele KJ, Kung JW, Smith MP. A practical guide to MR imaging safety: what radiologists need to know. Radiographics 2015 Oct;35(6):1722-1737. [doi: 10.1148/rg.2015150108] [Medline: 26466181]
8. Tirada N, Dreizin D, Khati NJ, Akin EA, Zeman RK. Imaging pregnant and lactating patients. Radiographics 2015 Oct;35(6):1751-1765. [doi: 10.1148/rg.2015150031] [Medline: 26466183]
9. Ray JG, Vermeulen MJ, Bharatha A, Montanera WJ, Park AL. Association between MRI exposure during pregnancy and fetal and childhood outcomes. JAMA 2016 Sep 06;316(9):952-961. [doi: 10.1001/jama.2016.12126] [Medline: 27599330]
10. Bamberg C, Rademacher G, Güttler F, Teichgräber U, Cremer M, Bührer C, et al. Human birth observed in real-time open magnetic resonance imaging. Am J Obstet Gynecol 2012 Jun;206(6):505.e1-505.e6. [doi: 10.1016/j.ajog.2012.01.011] [Medline: 22425409]
11. Güttler FV, Heinrich A, Rump J, de Bucourt M, Schnackenburg B, Bamberg C, et al. Magnetic resonance imaging of the active second stage of labour: proof of principle. Eur Radiol 2012 Sep;22(9):2020-2026. [doi: 10.1007/s00330-012-2455-9] [Medline: 22549105]
12. Jiang X, Asbach P, Streitberger KJ, Thomas A, Hamm B, Braun J, et al. In vivo high-resolution magnetic resonance elastography of the uterine corpus and cervix. Eur Radiol 2014 Dec;24(12):3025-3033. [doi: 10.1007/s00330-014-3305-8] [Medline: 25038856]
13. Scola E, Sirgiovanni I, Avignone S, Cinnante CM, Biffi R, Fumagalli M, et al. Fetal development of the corpus callosum: insights from a 3T DTI and tractography study in a patient with segmental callosal agenesis. Neuroradiol J 2016 Oct;29(5):323-325 [FREE Full text] [doi: 10.1177/1971400916665390] [Medline: 27549148]
14. Huang H, Xue R, Zhang J, Ren T, Richards LJ, Yarowsky P, et al. Anatomical characterization of human fetal brain development with diffusion tensor magnetic resonance imaging. J Neurosci 2009 Apr 01;29(13):4263-4273 [FREE Full text] [doi: 10.1523/JNEUROSCI.2769-08.2009] [Medline: 19339620]

15. Roelants-van Rijn AM, Groenendaal F, Stoutenbeek P, van der Grond J. Lactate in the foetal brain: detection and implications. Acta Paediatr 2004 Jul;93(7):937-940. [Medline: 15303809]
16. van Cappellen van Walsum AM, Jongsma HW, Wevers RA, Nijhuis JG, Crevels J, Engelke UF, et al. Hypoxia in fetal lambs: a study with (1)H-MNR spectroscopy of cerebrospinal fluid. Pediatr Res 2001 May;49(5):698-704. [doi: 10.1203/00006450-200105000-00015] [Medline: 11328955]
17. Story L, Damodaram MS, Allsop JM, McGuinness A, Patel A, Wylezinska M, et al. Brain metabolism in fetal intrauterine growth restriction: a proton magnetic resonance spectroscopy study. Am J Obstet Gynecol 2011 Nov;205(5):483.e1-483.e8. [doi: 10.1016/j.ajog.2011.06.032] [Medline: 21861969]
18. Simões RV, Muñoz-Moreno E, Carbajo RJ, González-Tendero A, Illa M, Sanz-Cortés M, et al. In vivo detection of perinatal brain metabolite changes in a rabbit model of intrauterine growth restriction (IUGR). PLoS One 2015 Jul 24;10(7):e0131310 [FREE Full text] [doi: 10.1371/journal.pone.0131310] [Medline: 26208165]
19. Kim DH, Vahidi K, Caughey AB, Coakley FV, Vigneron DB, Kurhanewicz J, et al. In vivo (1)H magnetic resonance spectroscopy of amniotic fluid and fetal lung at 1.5 T: technical challenges. J Magn Reson Imaging 2008 Oct;28(4):1033-1038. [doi: 10.1002/jmri.21528] [Medline: 18821604]
20. Wedegärtner U, Tchirikov M, Koch M, Adam G, Schröder H. Functional magnetic resonance imaging (fMRI) for fetal oxygenation during maternal hypoxia: initial results. Rofo 2002 Jun;174(6):700-703. [doi: 10.1055/s-2002-32222] [Medline: 12063598]
21. Wedegärtner U, Tchirikov M, Schäfer S, Priest AN, Walther M, Adam G, et al. Fetal sheep brains: findings at functional blood oxygen level-dependent 3-T MR imaging--relationship to maternal oxygen saturation during hypoxia. Radiology 2005 Dec;237(3):919-926. [doi: 10.1148/radiol.2373041625] [Medline: 16304112]
22. Wedegärtner U, Kooijman H, Andreas T, Beindorff N, Hecher K, Adam G. T2 and T2* measurements of fetal brain oxygenation during hypoxia with MRI at 3T: correlation with fetal arterial blood oxygen saturation. Eur Radiol 2010 Jan;20(1):121-127. [doi: 10.1007/s00330-009-1513-4] [Medline: 19618188]
23. Siauve N, Chalouhi GE, Deloison B, Alison M, Clement O, Ville Y, et al. Functional imaging of the human placenta with magnetic resonance. Am J Obstet Gynecol 2015 Oct;213(4 Suppl):S103-S114. [doi: 10.1016/j.ajog.2015.06.045] [Medline: 26428488]
24. Ami O, Chabrot P, Jardon K, Rocas D, Delmas V, Boyer L, et al. [Detection of cephalopelvic disproportion using a virtual reality model: a feasibility study of three cases]. J Radiol 2011 Jan;92(1):40-45. [doi: 10.1016/j.jradio.2009.05.001] [Medline: 21352725]
25. Ami O, Chabrot P, Rabischong B, Rocas D, Delmas V, Boyer L, et al. [Tridimensional vector animation from fetal MRI as a simulation of delivery]. J Radiol 2010 Apr;91(4):515-517. [doi: 10.1016/s0221-0363(10)70070-6] [Medline: 20514011]
26. Looney CB, Smith JK, Merck LH, Wolfe HM, Chescheir NC, Hamer RM, et al. Intracranial hemorrhage in asymptomatic neonates: prevalence on MR images and relationship to obstetric and neonatal risk factors. Radiology 2007 Feb;242(2):535-541. [doi: 10.1148/radiol.2422060133] [Medline: 17179400]
27. Rooks VJ, Eaton JP, Ruess L, Petermann GW, Keck-Wherley J, Pedersen RC. Prevalence and evolution of intracranial hemorrhage in asymptomatic term infants. AJNR Am J Neuroradiol 2008 Jun;29(6):1082-1089 [FREE Full text] [doi: 10.3174/ajnr.A1004] [Medline: 18388219]
28. Amiel-Tison C, Sureau C, Shnider SM. Cerebral handicap in full-term neonates related to the mechanical forces of labour. Baillieres Clin Obstet Gynaecol 1988 Mar;2(1):145-165. [doi: 10.1016/s0950-3552(88)80069-5] [Medline: 3046797]
29. Ami O, Maran JC, Gabor P, Whitacre EB, Musset D, Dubray C, et al. Three-dimensional magnetic resonance imaging of fetal head molding and brain shape changes during the second stage of labor. PLoS One 2019 May 15;14(5):e0215721 [FREE Full text] [doi: 10.1371/journal.pone.0215721] [Medline: 31091263]
30. Bulas D, Egloff A. Benefits and risks of MRI in pregnancy. Semin Perinatol 2013 Oct;37(5):301-304. [doi: 10.1053/j.semperi.2013.06.005] [Medline: 24176150]
31. Victoria T, Jaramillo D, Roberts TP, Zarnow D, Johnson AM, Delgado J, et al. Fetal magnetic resonance imaging: jumping from 1.5 to 3 tesla (preliminary experience). Pediatr Radiol 2014 Apr;44(4):376-386. [doi: 10.1007/s00247-013-2857-0] [Medline: 24671739]
32. Chung SM. Safety issues in magnetic resonance imaging. J Neuroophthalmol 2002 Mar;22(1):35-39. [doi: 10.1097/00041327-200203000-00009] [Medline: 11937904]
33. Jagannathan NR. Magnetic resonance imaging: bioeffects and safety concerns. Indian J Biochem Biophys 1999 Oct;36(5):341-347. [Medline: 10844986]
34. Brugger PC, Prayer D. Actual imaging time in fetal MRI. Eur J Radiol 2012 Mar;81(3):e194-e196. [doi: 10.1016/j.ejrad.2011.01.115] [Medline: 21345632]
35. Schifrin BS, Deymier P, Cohen WR. Cranial compression ischemic encephalopathy: fetal neurological injury related to the mechanical forces of labor and delivery. In: Zhang L, Longo LD, editors. Stress and Developmental Programming of Health and Disease: Beyond Phenomenology. Hauppauge, NY, USA: Nova Science Publishers; 2014:651-688.
36. Kelly JV. Compression of the fetal brain. Am J Obstet Gynecol 1963 Mar 01;85:687-694. [doi: 10.1016/0002-9378(63)90317-x] [Medline: 14031815]

37. Carlan SJ, Wyble L, Lense J, Mastrogiannis DS, Parsons MT. Fetal head molding. Diagnosis by ultrasound and a review of the literature. J Perinatol 1991 Jun;11(2):105-111. [Medline: 1890466]
38. Vlasiuk VV, Zirakadze AN. [Compression of the fetal head during labor as a cause of disorders of cerebral circulation and intrauterine asphyxia]. Akush Ginekol (Mosk) 1987 Oct(10):60-61. [Medline: 3434690]
39. Vlasyuk VV. [Compression-hypoxic birth injuries and skull configuration]. Arkh Patol 2021;83(3):70-74. [doi: 10.17116/patol20218303170] [Medline: 34041900]
40. Bartoli M, Mannes I, Aikem N, Rambaud C, de Boissieu P, Adamsbaum C. Is bridging vein rupture/thrombosis associated with subdural hematoma at birth? Pediatr Radiol 2022 May;52(5):932-940. [doi: 10.1007/s00247-021-05255-w] [Medline: 34988598]
41. Emerson MV, Pieramici DJ, Stoessel KM, Berreen JP, Gariano RF. Incidence and rate of disappearance of retinal hemorrhage in newborns. Ophthalmology 2001 Jan;108(1):36-39. [doi: 10.1016/s0161-6420(00)00474-7] [Medline: 11150261]
42. Watts P, Maguire S, Kwok T, Talabani B, Mann M, Wiener J, et al. Newborn retinal hemorrhages: a systematic review. J AAPOS 2013 Feb;17(1):70-78. [doi: 10.1016/j.jaapos.2012.07.012] [Medline: 23363882]
43. Hughes LA, May K, Talbot JF, Parsons MA. Incidence, distribution, and duration of birth-related retinal hemorrhages: a prospective study. J AAPOS 2006 Apr;10(2):102-106. [doi: 10.1016/j.jaapos.2005.12.005] [Medline: 16678742]
44. Laghmari M, Skiker H, Handor H, Mansouri B, Ouazzani Chahdi K, Lachkar R, et al. [Birth-related retinal hemorrhages in the newborn: incidence and relationship with maternal, obstetric and neonatal factors. Prospective study of 2,031 cases]. J Fr Ophtalmol 2014 Apr;37(4):313-319. [doi: 10.1016/j.jfo.2013.06.005] [Medline: 24576566]
45. Schifrin BS. Dystocia. J Reprod Med 1975 Jun;14(06):266-267. [Medline: 1151938]
46. Mocsáry P, Gaál J, Komáromy B, Mihály G, Pohánka O, Surányi S. Relationship between fetal intracranial pressure and fetal heart rate during labor. Am J Obstet Gynecol 1970 Feb 01;106(3):407-411. [doi: 10.1016/0002-9378(70)90367-4] [Medline: 5410877]
47. Chalak LF, Adams-Huet B, Sant'Anna G. A total sarnat score in mild hypoxic-ischemic encephalopathy can detect infants at higher risk of disability. J Pediatr 2019 Nov;214:217-21.e1. [doi: 10.1016/j.jpeds.2019.06.026] [Medline: 31301853]
48. Chansarn P, Torgalkar R, Wilson D, Fan CP, Widjaja E, Whyte H, et al. Correlation of Thompson and modified Sarnat scores in neonatal hypoxic ischemic encephalopathy. J Perinatol 2021 Jun;41(6):1522-1523. [doi: 10.1038/s41372-021-00987-x] [Medline: 33649434]
49. Li LH, Li N, Zhao JY, Fei P, Zhang GM, Mao JB, et al. Findings of perinatal ocular examination performed on 3573, healthy full-term newborns. Br J Ophthalmol 2013 May;97(5):588-591 [FREE Full text] [doi: 10.1136/bjophthalmol-2012-302539] [Medline: 23426739]
50. Bouvier D, Giguère Y, Pereira B, Bernard N, Marc I, Sapin V, et al. Cord blood S100B: reference ranges and interest for early identification of newborns with brain injury. Clin Chem Lab Med 2020 Jan 28;58(2):285-293. [doi: 10.1515/cclm-2019-0737] [Medline: 31622243]

Abbreviations

MR: magnetic resonance
MRI: magnetic resonance imaging

Edited by A Mavragani; submitted 31.01.21; peer-reviewed by R Dewey; comments to author 27.02.21; revised version received 14.03.21; accepted 02.11.22; published 30.11.22.

Please cite as:
Ami O, Maran JC, Musset D, Dubray C, Mage G, Boyer L
Using Magnetic Resonance Imaging During Childbirth to Demonstrate Fetal Head Moldability and Brain Compression: Prospective Cohort Study
JMIR Form Res 2022;6(11):e27421
URL: https://formative.jmir.org/2022/11/e27421
doi:10.2196/27421
PMID:36322921

Original Paper

End Users' and Other Stakeholders' Needs and Requirements in the Development of a Personalized Integrated Care Platform (PROCare4Life) for Older People With Dementia or Parkinson Disease: Mixed Methods Study

Mona Ahmed[1], MA; Mayca Marín[2], MSc; Daniella How[1], MSc; Elda Judica[3], MD; Peppino Tropea[3], PhD; Ellen Bentlage[1], MSc; Joaquim J Ferreira[4], MD, PhD; Raquel Bouça-Machado[4], PhD; Michael Brach[1], PD

[1]Institute of Sport and Exercise Sciences, Münster University, Münster, Germany

[2]Asociación Parkinson Madrid (APM), Madrid, Spain

[3]Department of Neurorehabilitation Sciences, Casa di Cura del Policlinico, Milan, Italy

[4]Campus Neurológico Sénior (CNS), Torres Vedras, Portugal

Corresponding Author:
Mona Ahmed, MA
Institute of Sport and Exercise Sciences
Münster University
Horstmarer Landweg 62b
Münster, 48149
Germany
Phone: 49 25183 ext 34812
Email: mona.ahmad@uni-muenster.de

Abstract

Background: With what has been known as the "*triple-win effect*", introducing information and communication technologies (ICTs) in the health care of neurodegenerative diseases is beneficial in delaying the need for institutional care, reducing the associated health care costs, reducing the caregiving burden, and improving individuals' quality of life. Nevertheless, the mismatch between the users' expectations and their actual needs remains one of the main challenges that can reduce the usability of technology solutions. Therefore, the European project Personalized Integrated Care Promoting Quality of Life for Older People (PROCare4Life), which aimed to develop an ICT-based platform for all parties involved in the health care of neurodegenerative diseases, adopted a user-centered design approach, where all users are involved from the inception and throughout the platform development and implementation to integrate their needs and requirements in the proposed platform.

Objective: This paper presents the results of a study on the needs and requirements of the potential end users (older people with neurodegenerative diseases, caregivers, and health care professionals) and other key stakeholders in the development of the PROCare4Life platform.

Methods: A mixed qualitative and quantitative study design was used, including 2 web-based surveys, 40 interviews, and 4 workshops. The study was conducted between April and September 2020 in 5 European countries: Germany, Italy, Portugal, Romania, and Spain. Both data types were analyzed separately and then merged and interpreted, with greater priority placed on qualitative research.

Results: A total of 217 participants were recruited; 157 (72.4%) of them completed the web-based surveys (n=85, 54.1% patients and n=72, 45.9% caregivers), and 60 (27.6%) individuals participated in the qualitative research (20/60, 33% health care professionals; 5/60, 8% patients; 5/60, 8% caregivers; and 30/60, 50% key stakeholders). We identified 3 main themes (T): (T1) experiences associated with illness, (T2) thoughts about the platform technology, and (T3) desired properties. Alerts for adverse events, communication tools, reminders, and monitoring are constantly needed functionalities, whereas ease of use, personalization, and user-friendliness are foreseen as necessary features.

Conclusions: This paper identified the key personal, social, and health factors that influence the daily lives of the potential end users and reflected on their needs and expectations regarding the design of the proposed PROCare4Life platform. The collected data were useful for the development of the PROCare4Life platform. Although the combination and collection of features for diverse user groups are typical for integrated care platforms, it results in exponential complexity for designers, developers, and

users. Contradicting opinions and several concerns in this study demonstrate that an ICT-integrated care platform should not promise too much for too many. Instead, selection, focus, and, sometimes, restriction to essentials are necessary. Users and other stakeholders should be involved in these decisions.

International Registered Report Identifier (IRRID): RR2-10.2196/22463

(JMIR Form Res 2022;6(11):e39199) doi:10.2196/39199

KEYWORDS

neurodegenerative; Parkinson disease; dementia; chronic diseases; health care technologies; integrated care; information and communication technology; ICT; user-centered design; mobile phone

Introduction

Background

Neurodegenerative diseases (NDDs), including dementia and Parkinson disease (PD), are among the most common chronic diseases associated with aging [1]. Characterized by a continuous decline in motor and cognitive abilities [2], difficulties in performing daily activities, and altered behavior [3], NDDs are mostly disabling diseases that negatively impact the quality of life of older populations and their families [1,4]. With the increased prevalence of NDDs, an enormous burden is placed on health care systems in terms of both resources and costs [2,5-9]. Therefore, implementing alternative health care solutions is needed [10].

Integrated care, which coordinates and brings together different health services, has the advantages of optimizing health care resources and being able to respond to the needs of older populations with chronic diseases [11,12]. In Europe, as a part of the eHealth action plan in supporting active aging, the introduction of the integrated care jointly with information and communication technology (ICT) supportive tools has contributed to improving patient experience and providing more efficient health care services at lower costs [13,14]. Divided into wearable, nonwearable, and hybrid-based categories, ICTs offer a wide variety of technological solutions with the purpose of either monitoring or managing the users' health [15].

In the care of NDDs, with what has been known as the "*triple-win effect*," ICTs are beneficial in "(1) delaying the need for institutional care, and reduction of the associated health care costs, (2) reducing the caregiving burden and (3) improving individuals' QoL by helping to keep an independent lifestyle, autonomy and social interaction" [16]. In addition, the integration of health-related data of patients in an interactive web interface enables health care professionals (HCPs) to better monitor and support their patients [17,18].

Although older people, their families, and HCPs are positive about using ICTs [19], the mismatch between the users' expectations and their actual needs remains one of the main challenges that can reduce the usability of technology solutions targeting older people with mild cognitive decline [20]. In fact, developing an ill-fitting ICT for this target group can be a burden instead of being a supportive tool [21,22]. Therefore, it is crucial to first identify the needs of the users when implementing ICTs in health care to better develop a suitable solution [23].

In this paper, we present the results of a study on the needs and requirements of older people with NDDs, their caregivers, HCPs, and other key stakeholders in the development of a personalized integrated ICT-based, Personalized Integrated Care Promoting Quality of Life for Older People Platform (PROCare4Life). These results ought to drive implications on the design and the properties of the platform.

PROCare4Life

PROCare4Life is an ICT-based, integrated, scalable, and interactive health care platform. The intended end users include older people with NDDs, caregivers, and HCPs involved in the care process. The PROCare4Life platform plans to collect disease, cognitive, and behavior related data about the patients via wearables, stationary devices, medical records, and other sensors. In a highly secured and protected cloud environment, algorithms analyze and process these data to create a profile for each patient. On the basis of this profile, personalized information and recommendations are provided to those involved in the care plan. The end users will be able to interact with a wide range of services via various digital devices such as smartphone, tablet, or smart television. More details about the aims and the technology of PROCare4Life are reported elsewhere [24].

Throughout the entire development process of the PROCare4Life platform, the project adopted a user-centered design (UCD) approach [25]. This has been recommended for decision makers and leaders in the process of developing ICT in integrated care [26] to ensure active engagement and incorporation of the intended users' feedback. In line with this approach, the following steps were incorporated: (1) study of the user needs and requirements; (2) iterative design throughout the pilot phases; (3) iterative user evaluation, refining the design throughout the pilots; and (4) a final product that is developed based on the iterative cycles and evaluation or a developed final product based on the iterative cycles and evaluation. This study focused on the first step, understanding the users' needs and requirements.

Research on Users' Needs

In general, ICT solutions need to be easy to use, private, secure, and affordable in terms of costs [27]. However, older people tend to have heterogeneous needs [28], with possible conflicts among patients, caregivers, and HCPs [29]. In dementia, previous studies have summarized the main need areas as information, company, memory and daily activity support, and reduction of psychological stress [30,31]. A systematic review by Lauriks et al [32] that aimed to identify the unmet needs of

patients and caregivers stated that ICTs need to be personalized according to the users' needs and abilities. Boman et al [33] studied the needs of people with cognitive impairment in the design of an ICT-based device. The study reported that the participants were positive about including calendars as memory support, whereas HCPs valued a feature that allows them to view the previous and current care plans. However, low participant numbers were reported as one of the study limitations. In PD, clinical symptoms have high daily fluctuations, meaning that nonmotor- and motor-related symptoms vary within and between days [34]. Therefore, the ICT solution needs to be able to monitor and identify all relevant changes and develop personalized strategies to counteract them [35]. In addition, social support positively affects the ability of patients with PD to cope with the difficulties in daily living and reduces the risk of developing nonmotor symptoms, such as depression [36].

The critique of previous research was that it did not consider the users in the early stages of development. Both patients and caregivers were included at later stages, which resulted in the lost value of their experiences [37].

Following the multidisciplinary principle of the UCD approach, in addition to the intended end users identified as patients, caregivers, and HCPs, the PROCare4Life study on users' needs includes other key stakeholders from different related health care disciplines. These stakeholders are academic researchers, decision makers, markets, and media actors. The overall objectives of this study were as follows:

1. Collecting detailed information on the opinions, thoughts, experiences, and feelings of the end users (patients, caregivers, and HCPs) and other key stakeholders regarding NDDs, health care processes, and digital health care solutions to identify those aspects where the PROCare4Life platform would best suit and support them.

2. Identifying the aspects that the PROCare4Life platform should consider to achieve success in its acceptance, development, and marketing (eg, strengths and weaknesses, factors that influence the digital health care market, and communication channels through which to adequately diffuse the product).

This paper presents the results related to the first aim of the study on users' needs and requirements.

Methods

Study Design

In a mixed methods study design, we followed the formative and summative research methodologies to identify and analyze the end user needs (identified as patients, caregivers, and HCPs), key stakeholder perspectives (identified as academic researchers, media actors, policy makers, and market actors), and context. In the 2-step approach, we applied both qualitative and quantitative research methods. The quantitative data included 2 web-based surveys involving patients and caregivers. The qualitative data included semistructured interviews and workshops involving the end users and other key stakeholders. This study placed greater priority on qualitative research, with quantitative research playing a supportive role [38].

Study Procedure and Eligibility Criteria

This study was conducted between April and September 2020 in 5 European countries: Germany, Italy, Portugal, Romania, and Spain. The web-based surveys were launched first, followed by the semistructured interviews. Finally, 4 workshops that involved HCPs were conducted, in which the preliminary results from the surveys were presented and discussed.

In this study, patients were included if they were aged ≥65 years and diagnosed with PD or dementia, including Alzheimer disease and other dementias (OD). Patients with substantial cognitive impairment, intellectual disability, or other serious psychiatric conditions that affect their ability to use mobile phones or computers were excluded. Caregivers were referred to as those who care for patients diagnosed with PD or dementia as formal (ie, paid) or informal (ie, unpaid) caregivers. HCPs included those who worked in the medical or social care of patients diagnosed with PD or dementia (eg, physiotherapists, physicians, and occupational therapists). Further details regarding eligibility criteria and recruitment are reported in the study protocol [24].

Quantitative Data Collection and Analysis

In total, 2 anonymous web-based surveys were created in English through EUSurvey tool and were translated into the other 5 project languages (German, Italian, Portuguese, Romanian, and Spanish). Both surveys were available on the PROCare4Life official website [39], in the period between May 27, 2020, and July 31, 2020, along with a short explanation of the purpose of the surveys. In addition, the surveys were disseminated through consortium member databases, networks, and national patient associations. The questions were developed in collaboration with clinical partners and aimed to gather answers regarding the topics listed in Textbox 1.

A descriptive analysis of the quantitative data was applied, including descriptive statistics and frequencies. The 7-item abbreviated Zarit scale was analyzed using SPSS statistical software (version 27; IBM Corp).

Textbox 1. Topics covered by the web-based surveys.

1. Demographic data
2. Disease-related symptoms, assessed through a list of formulated questions regularly used by one of the pilot centers in this study—Asociación Parkinson Madrid—in evaluating the disease symptoms and medication side effects. These questions were approved for their content by the Movement Disorders Study Group of the Spanish Society of Neurology [40]
3. Difficulties in activities of daily living, assessed through domains derived from the self-reported Barthel index [41], in addition to other domains related to difficulties moving around and accessing health care centers
4. Caregiver family burden, assessed through the 7-items abbreviated Zarit scale [42,43]
5. Assistive health technology experiences and acceptance, assessed through developed Likert-items questionnaires assessing the frequency of use and the importance of technology devices intended to be included in the PROCare4Life platform (eg, tablet or mobile app)
6. Expected benefits and desired features in the proposed PROCare4Life platform. On a Likert-items questionnaires, the key performance indicators to be achieved within the project (eg, feelings of safety, feelings of autonomy, perception of empowerment, improvement of social participation, mental condition, and physical condition) and the main functionalities to be included were assessed

Qualitative Data Collection and Analysis

Owing to the explorative nature of the study, individual interviews and workshops with open-ended questions were conducted in the period between June and July 2020. Each interview lasted between 30 and 60 minutes, and the following three types of interviews were conducted:

1. Interview for patients covering the same topics as in the quantitative study and allowing more exploratory answers through open-ended questions
2. Interview for caregivers covering their working experience, in addition to the topics mentioned in the web-based survey
3. Interview for key stakeholders covering their opinions, experiences, and ideas about using integrated digital health care platforms in the management of older people with NDDs, in addition to strengths, weaknesses, and the possible ways to promote the proposed PROCare4Life platform from their point of view

Additionally, 4 workshops were conducted between July and August 2020, with a duration of approximately 2 hours each. They involved HCPs and covered topics related to their experiences with using integrated digital health care technologies in the care of people with NDDs, expected benefits, barriers, and their requirements regarding the properties of the proposed PROCare4Life platform. The workshops also included discussions about preliminary results from the web-based surveys, which allowed more interaction between the participants, aiming to enrich the data collected.

Owing to the pandemic situation in most of the European countries at the time of the study, this qualitative study was conducted on the web except for a few interviews, where participants requested a face-to-face interaction; in those cases, COVID-19 social distancing and safety measures were all considered.

Gathered data were recorded, transcribed verbatim, and translated into English. Thematic analysis [44,45] was applied, using MAXQDA software (version 20; VERBI GmbH). Following a deductive-inductive approach, a framework containing the key topics covered in the quantitative study was developed. In total, 2 researchers worked independently and performed a first round of identifying the relevant text and coding (the four-eyes principle). The developed framework was applied to the entire data set but considering that the qualitative study included all the target groups, unlike the quantitative study, the researchers conducted a second round of open coding to identify additional topics and subtopics. The identified initial codes were discussed in a workshop involving the aforementioned researchers along with another researcher; significant data overarching the key topics were also discussed and validated in an iterative manner (discussion<->modification) to develop themes and subthemes (STs). The results and illustrative quotes were discussed with a researcher from Asociación Parkinson Madrid until a consensus was reached and themes were finalized.

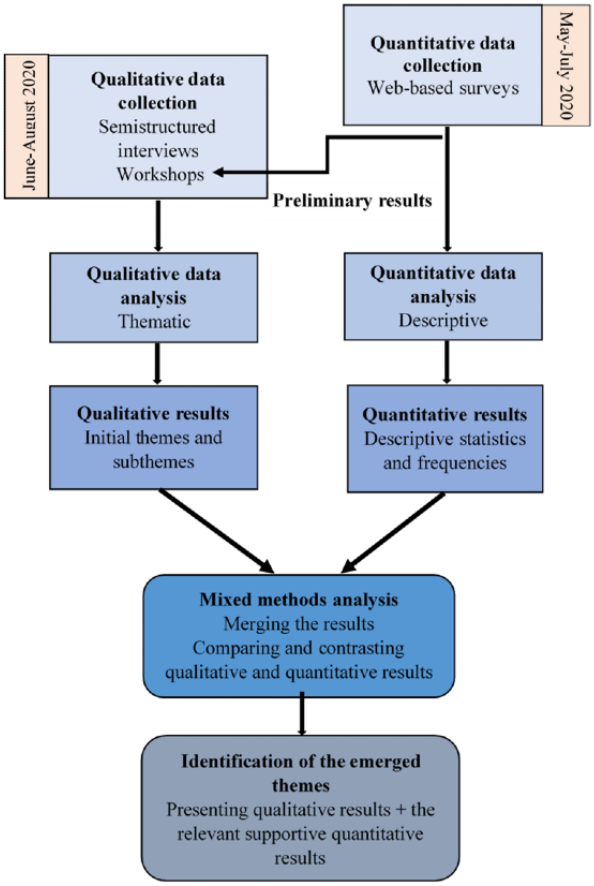
Mixed Methods Analysis

Both data types were analyzed separately. The identified initial qualitative themes and the main quantitative results were merged, aiming to combine the results and present them as the final emerged themes and STs (Figure 1). Although quantitative results provided numerical conclusions from our research and predicted outcomes about each theme, qualitative results allowed more comprehensive insights and in-depth knowledge from our participants about the same themes.

Because the topics in the surveys were also discussed during the interviews, the results were consistent, in particular those related to patients and caregivers. However, as the qualitative data were broader and involved all the participants, it included exclusive STs that had no corresponding supportive quantitative data. As these additional STs were considered important and provided a comprehensive reflection of the participants' point of view, we included them in the mixed methods analysis as well.

Finally, 3 main emerged themes along with several STs were identified, which are presented in the *Results* section.

Figure 1. Mixed methods flowchart.



Ethics Approval

The study protocol was approved by local ethical committees in Germany (number 020-37-MB), Italy (number 493-2020), Portugal (number 10-20), Romania (number 7/10.06.2020), and Spain (number 20/453-E). The organizations conducting this study established procedures for data protection management before the start of any processing of personal data, according to legal regulations and following good practices in research.

According to Good Clinical Practice and International Conference on Harmonization standards, once the study was fully explained, a written or digital informed consent was obtained from each participant before any study-related procedures.

There were no direct physical risks to the participants. Participation was entirely voluntary, and the participants had the right to withdraw from the study at any time, without giving reasons or experiencing any disadvantage. In case of withdrawal, no replacement was considered.

Results

Participants

Table 1 shows an overview of the participants in this study. Across study methods, countries, and target groups, a total of 217 participants were recruited. Overall, 72.4% (157/217) completed the 2 web-based surveys, distributed as 85 (54.1%) patients and 72 (45.9%) caregivers. The remaining 27.6% (60/157) of participants took part in the qualitative research: 20 (33.3%) HCPs were included in the workshops and 5 (8.3%) patients, 5 (8.3%) caregivers, and 30 (50%) key stakeholders took part in the semistructured interviews.

The characteristics of the end users are presented in Table 2. A total of 187 end users participated in the qualitative and quantitative strands of the study; most of them (71/187, 37.9%) were aged between 61 and 75 years, with more than half (112/187, 59.9%) being female. Most of the patients (64/90, 71%) who took part in this study were diagnosed with PD, and 96% (86/90) lived at home. More than one-third of them (32/90, 36%) rated their general health status in the past 4 weeks as fair. Most caregivers (68/77, 88%) were informal, and around half of them (38/77, 49%) lived with the patients they care for. In the workshops, HCPs from different specialties participated; however, 25% (5/20) were physiotherapists.

Table 1. Overview of the participants across study methods, countries, and target groups (N=217).

Country	Study method, n (%)									Total, n (%)
	Quantitative (n=157)		Qualitative (n=60)							
	Surveys		Interviews (n=40)						Workshops (n=20)	
	Patients (n=85)	CGs[a] (n=72)	Patients (n=5)	CGs (n=5)	Media actors (n=6)	Academia (n=8)	Policy makers (n=8)	Market actors (n=8)	HCPs[b] (n=20)	
Germany	3 (3.5)	13 (18)	1 (20)	1 (20)	2 (33.3)	2 (25)	3 (37.5)	4 (50)	5 (25)	34 (15.7)
Italy	13 (15.3)	11 (15.3)	1 (20)	1 (20)	1 (16.6)	N/A[c]	2 (25)	2 (25)	3 (15)	34 (15.7)
Portugal	27 (31.8)	8 (11.1)	1 (20)	1 (20)	2 (33.3)	4 (50)	2 (25)	2 (25)	7 (35)	54 (24.9)
Romania	27 (31.8)	23 (32)	N/A	1 (20)	N/A	2 (25)	N/A	N/A	N/A	53 (24.4)
Spain	14 (16.5)	17 (23.6)	2 (40)	1 (20)	1 (16.6)	N/A	1 (12.5)	N/A	5 (25)	41 (18.9)
Others[d]	1 (1.2)	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	1 (0.4)

[a]CG: caregiver.

[b]HCP: health care professional.

[c]N/A: not applicable.

[d]"Others" was one of the country choices listed in the web-based surveys, and the participants who answered with "others" were included in the data analysis.

Table 2. Characteristics of the end users (n=187).

Characteristics	End users, n (%)			Total, n (%)
	Patients (n=90)	Caregivers (n=77)	HCPs[a] (n=20)	
Age (years)				
<60	N/A[b]	44 (57.1)	20 (100)	64 (34.2)
61-75	46 (51.1)	25 (32.5)	N/A	71 (38)
>75	44 (48.9)	8 (10.4)	N/A	52 (27.8)
Sex				
Male	50 (55.6)	23 (29.9)	2 (10)	75 (40.1)
Female	40 (44.4)	54 (70.1)	18 (90)	112 (59.9)
Patients				
Diagnosis				
Parkinson disease	64 (71.1)	N/A	N/A	N/A
Alzheimer disease	5 (5.6)	N/A	N/A	N/A
Other dementias	21 (23.3)	N/A	N/A	N/A
Living situation				
At home	86 (95.6)	N/A	N/A	N/A
At home and temporarily at a day care center	2 (2.2)	N/A	N/A	N/A
At a residential center	1 (1.1)	N/A	N/A	N/A
Not reported	1 (1.1)	N/A	N/A	N/A
General health status (4 weeks)				
Very poor	4 (4.4)	N/A	N/A	N/A
Poor	18 (20)	N/A	N/A	N/A
Fair	32 (35.6)	N/A	N/A	N/A
Good	24 (26.7)	N/A	N/A	N/A
Very good	7 (7.8)	N/A	N/A	N/A
Not reported	5 (5.6)	N/A	N/A	N/A
Caregivers				
Type of caregiver				
Informal	N/A	68 (88.3)	N/A	N/A
Formal	N/A	7 (9.1)	N/A	N/A
Not reported	N/A	2 (2.6)	N/A	N/A
Living with the person you care for				
Yes	N/A	38 (49.3)	N/A	N/A
No	N/A	28 (36.4)	N/A	N/A
Partially	N/A	8 (10.4)	N/A	N/A
Not reported	N/A	3 (3.9)	N/A	N/A
HCPs				
Specialty				
Neurologists	N/A	N/A	1 (5)	N/A
Nurses	N/A	N/A	3 (15)	N/A
Psychologists	N/A	N/A	3 (15)	N/A
Physiotherapists	N/A	N/A	5 (25)	N/A

Characteristics	End users, n (%)			Total, n (%)
	Patients (n=90)	Caregivers (n=77)	HCPs[a] (n=20)	
Speech therapists	N/A	N/A	3 (15)	N/A
Music therapists	N/A	N/A	1 (5)	N/A
Social workers	N/A	N/A	2 (10)	N/A
Educational trainers	N/A	N/A	2 (10)	N/A

[a]HCP: health care professional.

[b]N/A: not applicable.

Emerged Themes

Overview

In this section, we present the 3 emerged themes developed based on merging and interpreting the qualitative and quantitative data, namely experiences associated with illness (theme 1), thoughts about the platform technology (theme 2), and desired properties (theme 3). For every emerged theme, different STs were identified (Figure 2). We first present the detailed qualitative findings for each ST, followed by the relevant supportive quantitative findings. In addition, Table 3 illustrates a summary of the mixed methods analysis for all 3 themes and STs.

Figure 2. Overview of the emerged themes and subthemes.

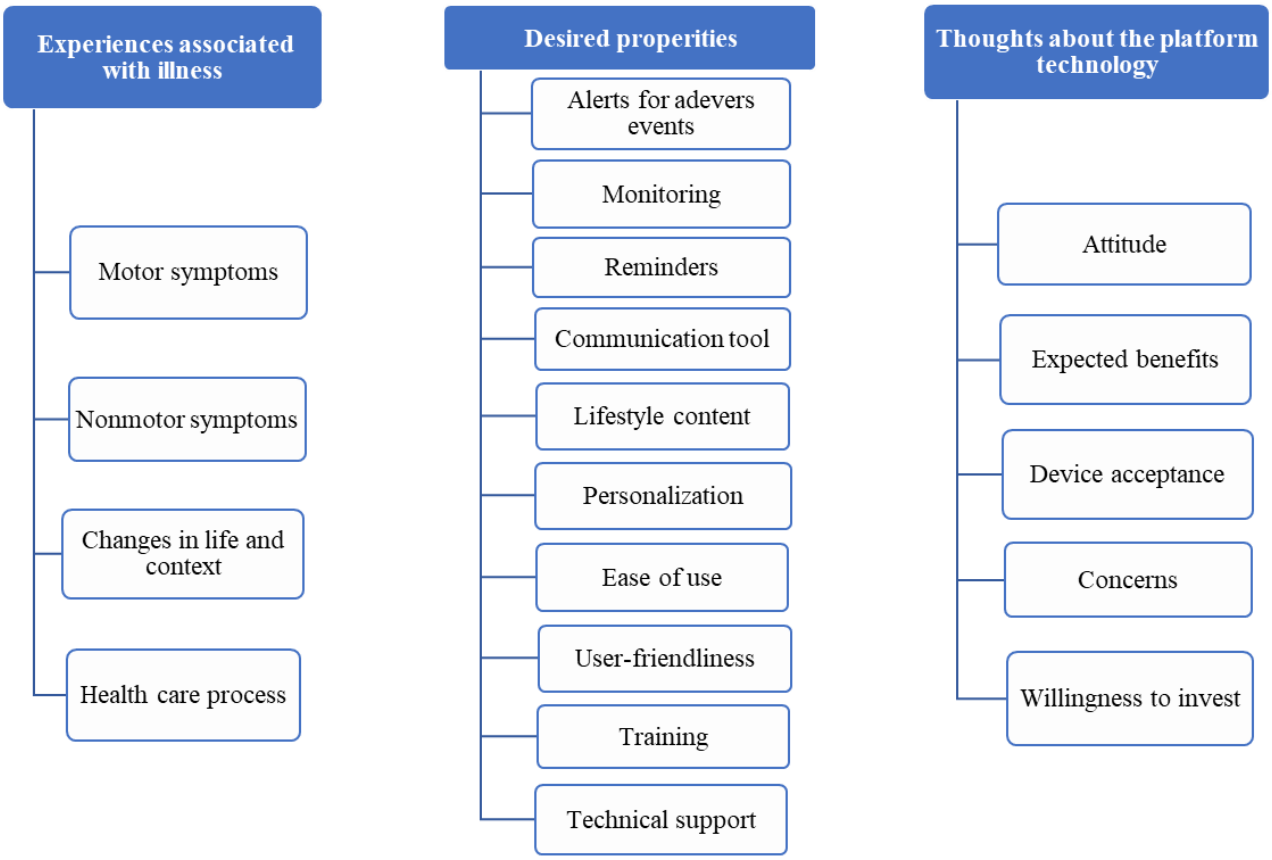


Table 3. Mixed methods analysis (some ratios are approximated).

T[a] and ST[b]	Qualitative data codes	Supportive quantitative results (survey results)
T1: experiences associated with illness		

ST1.1: motor symptoms	• Stiffness
	• Loss of balance
	• Frequent falls and injuries
	• Feeling of insecurity and disability
	• Limited mobility, and physical limitations
		• Most reported motor symptoms:
		  • Stiffness was reported by 78% (66/85) of the patients and 83% (60/72) of the caregivers
		  • Loss of balance was reported by 66% (56/85) of the caregivers and 83% (60/72) of the caregivers

ST1.2: nonmotor symptoms	• Concentration problems • Memory problems • Difficulties in communication • Risks of medication misuse • Missing meals • Disorientation	• Most reported nonmotor symptom is as follows: • Difficulties in concentration was reported by 58% (49/85) of the patients and 63% (45/72) of the caregivers

ST1.3: changes in life and context	• Difficulties in ADL[c]
	• Daily struggles
	• Patients need support in everything
	• Feelings of isolation
	• Distressed and overworked caregivers
	• Coping strategies
		• Difficulties in performing ADL and caregiver burden scale are as follows:
		  • Patients had difficulties in performing 15 of the ADL listed in the survey, as reported by >50% (38/72) of the caregivers
		  • In the 7-item abbreviated Zarit scale, 57% (41/72) of the caregivers reported to have family burden

ST1.4: health care process	• Complex health care process
	• Shortage in the number of HCPs[d]
	• Long waiting time for patients
	• Limited time offered by HCPs
	• Difficulties in accessing health care sites by the patients
		• Reported difficulties in performing the following activities:
		  • Difficulties in accessing therapy sites was reported by 43% (37/85) of the patients and 79% (57/72) of the caregivers
		  • Difficulties in accessing rehab sites was reported by 43% (37/85) of the patients and 68% (49/72) of the caregivers

T2: thoughts about the platform technology		
ST2.1: attitude	• Positive: supportive, needed, loving it, and specifically needed and accepted at pandemic situation • Negative: complicated, stressful, difficult, and not for everyone • Aid and not replacement	• N/A[e]

ST2.2: device acceptance	• Smartphones, wearables, and tablets are highly accepted Smart television is complicated
	• Cameras are invasive
		• Most accepted devices are as follows:
		  • Wearables: 57% (48/85) of the patients would like or love to use it
		  • Smartphones and tablets: 69% (50/72) of the caregivers would like or love to use it
		• Least accepted devices are as follows:
		  • Cameras: 33% (28/85) of the patients and 31% (22/72) of the caregivers would like or love to use it

T[a] and ST[b]	Qualitative data codes	Supportive quantitative results (survey results)
ST2.3: expected benefits	• Improving communication between patients, caregivers, and HCPs • Supporting the integrated care approach • Improving work efficiency • Supporting patient independency, caregiver engagement, and relationship between patients and HCPs • Improving the health care process	• The following expected benefits were agreed upon by >50% (43/85) of the patients and around 50% (34/72) caregivers: • Increase the feelings of safety or autonomy of patients in their homes • Improve the patient's mental or physical condition • Increase the patient's perception of empowerment • Improve the patient's social participation • Valid tool to respond to your needs
ST2.4: concerns	• Privacy, data protection • Costs • Handling abilities specially by the patients	• N/A
ST2.5: willingness to invest	• Financial investment needs organizational support • Individuals pay when there are benefits • End users are willing to invest time and effort to learn using the platform	• Willingness to pay: 51% (43/85) of the patients and 49% (35/72) of the caregivers were not sure if they would pay for the platform
T3: desired properties		

T[a] and ST[b]	Qualitative data codes	Supportive quantitative results (survey results)
ST3.1: alerts for adverse events	• Detecting and recording hazardous situations • Relief for caregivers • Need a supportive infrastructure	• Alerts for adverse events was reported as a desired functionality by 80% (68/85) of the patients and 90% (65/72) of the caregivers
ST3.2: monitoring	• Continuous monitoring of vital signs • Health status measures • Movement and gait changes in patients with Parkinson disease • Sleep disorders • Symptoms evolution • Medication side effects • Monitoring in real time	• Monitoring tool was reported as a desired functionality by 85% (72/85) of the patients and 82% (59/72) of the caregivers • Real-time information was reported as a desired feature by 80% (68/85) of the patients and 86% (62/72) of the caregivers
ST3.3: communication tool	• Communication with HCPs • Communication with peers • Chat tool • Get information at home • Digital interviews, interventions, and follow-up sessions • The need for in-person contact	• Communication tool was reported as a desired functionality by 79% (67/85) of the patients and 88% (63/72) of the caregivers
ST3.4: reminders	• Appointments • Medications • Drinking and mealtime reminders • Relief for families	• Reminders and a tool to organize appointments was reported as a desired functionality by 67% (57/85) of the patients and 75% (54/72) of the caregivers
ST3.5: lifestyle content	• The need for PA[f] and nutrition recommendations Cognitive games	• Social networking tool was reported as a desired functionality by 60% (51/85) of the patients and 68% (49/72) of the caregivers • Lifestyle recommendations (nutrition and PA) was reported as a desired functionality by 75% (64/85) of the patients and 79% (57/72) of the caregivers
ST3.6: ease of use	• Simple, passive platform • Low interaction • Easily retrievable information • No overwhelming emails and requests	• Easy to set up and start platform was reported as a desired feature by 75% (64/85) of the patients and 89% (52/72) of the caregivers[g] • Few steps to get the functionality you want was reported as a desired feature by 79% (67/85) of the patients and 88% (63/72) of the caregivers[g]
ST3.7: personalization	• Platform that is adapted to users' skills and cognitive abilities • Platform that considers the different target users • Platform that provides relevant information.	• Platform adapted to users' skills was reported as desired feature by 74% (63/85) of the patients and 83% (60/72) of the caregivers[g]
ST3.8: user-friendliness	• Comfortable wearables • Predefined layout • Less text • More graphs and diagrams	• Comfortable wearables was reported as a desired feature by 72% (61/85) of the patients and 81% (58/72) of the caregivers[g]
ST3.9: training	• Provide training for end users • Supportive manuals • One-on-one training sessions	• N/A
ST3.10: technical support	• Provide supportive infrastructure (eg, networks, WiFi, and 5G) • Hotline for technical support • Automatic updates and backups	• N/A

[a]T: theme.

[b]ST: subtheme.

[c]ADL: activities of daily life.

[d]HCP: health care professional.

[e]N/A: not applicable.

[f]PA: physical activity.

[g]Different features as listed in the surveys with answers of very important or important (%).

Theme 1: Experiences Associated With Illness

In this theme, we present what the participants expressed regarding the NDD symptoms, how the illness affected their everyday lives, and the difficulties encountered within the health care services.

ST1.1: Motor Symptoms

Among the different motor symptoms associated with NDDs, stiffness and loss of balance were frequently mentioned. Most patients and caregivers expressed their concerns regarding the consequences of motor symptoms, such as frequent falls and injuries:

> *It is normal, but I am concerned because consequences can be severe. Fear of consequential damage (broken bone, etc.)* [Caregiver, Germany]

Furthermore, patients' mobility becomes limited, and moving around becomes problematic and physically demanding, in particular for patients with PD:

> *At least from what I have seen with Parkinson patients, having to move and go somewhere is very physically demanding.* [Academia 1, Germany]

The consequences were not only physical; patients also explained that because of motor symptoms, they feel hindered and insecure:

> *I couldn't move myself without being looked at and wobbled around. He is drunk or something, right? That hindered me a lot. I don't want to say handicapped, but very upset...That's what bothers me the most. Being insecure, it's so bad that I feel really bad.* [Patient with OD, Germany]

ST1.2: Nonmotor Symptoms

Nonmotor symptoms represent another clinical spectrum of NDDs. For most patients and caregivers, disorientation and difficulties in concentration are worrisome, as patients can get lost:

> *The other day I got lost while I was going to the association. I went through a different street and suddenly, I did not know where I was.* [Patient 1 with PD, Spain]

Notably, memory problems are the hardest to deal with:

> *The hardest thing is to deal with memory problems.* [Caregiver, Spain]

They mentioned how memory problems could induce other challenges, such as difficulties in communication, as patients can forget the topics they are discussing or fail to identify with whom they are talking:

> *Sometimes I can't remember their names, and in the middle of talking, I just stop. Sometimes it happens that I don't even remember what I wanted to say. I don't know where it comes from but it happens. And it's not good, and I'm a little worried.* [Patient with OD, Germany]

Furthermore, owing to memory issues, patients can experience risky situations, such as an overdose or underdose of their medications, as well as missing mealtimes:

> *She [the patient] would forget to eat, saying she's full; I don't leave medication within her reach as she would either not take it or take more than actually prescribed.* [Caregiver, Romania]

ST1.3: Changes in Life and Context

After being diagnosed with NDDs, both patients and caregivers reported feeling isolated. Although patients felt apathetic and sad at home, caregivers had to rearrange their daily routines to focus all their efforts on taking care of the patients. In fact, one of the caregivers described caring for patients with NDDs as caring for a grown-up child, as they need support in almost all their activities:

> *It's complicated, emotionally and logistically. Since the diagnosis, I stopped what I used to do. I only leave the house for urgent things.* [Caregiver, Portugal]

In their attempts to live with the illness, both patients and caregivers reported some coping strategies. For example, one of the patients with PD mentioned engaging in sports and being more physically active:

> *Try and compensate for the effects of the disease, I do sport, games, computer activities, keep house accounts. I keep active in general and I practice sports in particular.* [Patient 2 with PD, Spain]

Another patient with PD reported making to-do and shopping lists, which can help organize daily tasks:

> *I make lists and lists about money, about other things.* [Patient 1 with PD, Spain]

Meanwhile, caregivers also need to cope and make symptom-specific adjustments, such as adjustments to the living place or nutritional considerations:

> *Of course, you try to make the most of it. So, for example, in the case of swallowing disorders, you look at whether you are thickening or purify the food, and I just see that it is nicely prepared...To reduce the risk of falls, for example I've taken a carpet out of the living room before to avoid a risk of falling. Such things. So, I'm just advising a lot. So, even if certain things are worrying, we still have to deal with them.* [Caregiver, Germany]

ST1.4: Health Care Process

The health care process for patients with NDDs usually involves >1 specialty and requires many visits. Considering their *motor symptoms*, going to all these visits is demanding for both patients and caregivers, particularly those who live in rural areas:

> *You have the foot care, the pedologists—rarely now, because they are all fully booked. And just this whole medical complex that works together. You still have one or the other family doctor who still makes home visits. And then it stops. Because occupational therapists, speech therapists are rather rare and difficult to get here in the countryside.* [Caregiver, Germany]

Other challenges regarding the health care process were reported, such as a poor physician-to-patient ratio. On one hand, this can be stressful for HCPs as they have to manage extra numbers:

> *Doctors have more patients than they can handle and are late in seeing patients.* [Patient 1 with PD, Spain]

On the other hand, patients have to wait longer to get their appointments:

> *Right now, in Spain for patients to be seen by a specialist health professional, they have a 6 months period wait, and to get a social worker appointment it's more.* [HCP, Spain]

Patients are usually not satisfied with the time offered to them by HCPs:

> *I'd say that nursing hours are too short. This makes the patients very disappointed that they can't even talk to them a little. And then they (patients) are very sad.* [Patient with OD, Germany]

Supportive Quantitative Results for Theme 1

In the web-based surveys, both patients and caregivers were asked to report about the symptoms experienced by patients and whether they were worried about them; the difficulties in performing activities of daily living (ADL); and family burden for caregivers.

Stiffness was the most experienced symptom as reported by both patients (66/85, 78%) and caregivers (60/72, 83%), whereas stumbles and falls were the most frequent symptoms patients expressed that they were worried about (19/33, 58%). Regarding the nonmotor symptoms, feeling sad was the most common symptom reported by patients (61/85, 72%), whereas feeling anxious or nervous was reported by 69% (50/72) of the caregivers. Difficulties in communication was reported by 55% (47/85) of the patients and 68% (49/72) of the caregivers.

As for ADL, dressing and undressing was the most difficult activity for the patients (46/85, 54%), whereas accessing therapy sites and moving outside the house was the most difficult for the patients as reported by 79% (56/72) of the caregivers.

On the basis of the results from the 7-item abbreviated Zarit scale, 57% (41/72) of the caregivers reported having family burden.

All the detailed results from the surveys regarding this theme are provided in the Multimedia Appendix 1.

Theme 2: Thoughts About the Platform Technology

In this theme, the opinions of the participants regarding the proposed platform technology were gathered, including their attitudes toward different aspects of the platform technology, device preferences, their expected benefits and main concerns, and their willingness to use it and pay for it.

ST2.1: Attitude

On being introduced to the concept of the PROCare4Life platform and its main objectives, the participants showed varied attitudes. Some were positive about the initial platform design; in fact, several patients and caregivers said they personally loved it. HCPs and key stakeholders found it to be helpful, interesting, and required in the health care process:

> *Not only interesting but also very much needed.* [Media actor, Portugal]

Notably, most of the participants referred to the COVID-19 pandemic and the subsequent lockdown in most of the European countries as a reason for the increased interest in digital integrated communication platforms in health care. Patients have become more flexible about using ICTs:

> *I would use a tablet, for example, if we have a pandemic or something like that.* [Patient with OD, Germany]

HCPs thought that PROCare4Life is needed to continue providing services to their patients in situations where access to health institutions and facilities was limited for emergencies only:

> *All this COVID 19 situation changed everybody's perspective. Not being able to be with people but still wanting to care for them. If we could have a digital system that allowed us to monitor someone at a distance and that also allowed us to be in contact and interact with them, that would be very important.* [Market actor 2, Portugal]

Furthermore, the pandemic was thought to be a catalyst in developing the market of digital integrated care platforms.

A negative attitude was also reported: one of the patients referred to old age as a challenge for accepting technological devices and benefiting from it. Some caregivers found the platform complicated, in particular for patients with advanced dementia, when the abilities to use any technical device become questionable:

> *This only works if there is no disease that does not affect it. With advanced dementia, the use of such devices no longer works at all.* [Caregiver, Germany]

In addition, a negative attitude from some HCPs was based on the opinion that such a platform could be an additional burden to their work, both time-consuming and stressful:

> *It can cause some stress to the team since there's an additional pressure and responsibilities.* [HCP, Portugal]

Finally, participants pointed out that health care technological solutions such as PROCare4Life should be only a support and not a replacement for physical contact and in-person interaction:

> *Technology can be considered a support and an aid, but not a replacement.* [HCP, Italy]

ST2.2: Device Acceptance

The initial design of the platform and different devices to be integrated were explained. Most patients and caregivers preferred using wearables and smartphones:

> *With an explanation and knowing the objective. Yes, she loves wearing devices and wanted an Android smartphone, she wears tele-assistance and likes it.* [Caregiver, Spain]

However, during the workshop in Germany, some HCPs preferred tablets over smartphones, referring to the negative experience they had regarding the smartphones' usability:

> *With smartphones it was certainly the case-so the feedback that it was difficult to use because people didn't understand it well and the volume was so low.* [HCP, Germany]

In addition, including stationary devices was seen as helpful, as it can ensure continuous monitoring of the patients in case they forgot to wear their wearables:

> *The strength of having sensors at home is that if they have a wearable system, people might forget to put it on. Fixed sensors will be better, because they will always be present.* [Media actor, Portugal]

Conversely, some devices were less accepted, such as a smart television, which was thought to be difficult for patients with dementia, and cameras, owing to data protection and privacy intrusion worries:

> *Dealing with this [smart TV] is difficult...I would prefer not to use it [cameras], even if it doesn't record images, I see a privacy problem, although it might make sense. But it's data protection difficult and I don't know if I want to be monitored by technology all day long.* [Caregiver, Germany]

ST2.3: Expected Benefits

The participants expected several benefits when using the platform. Improvements in communication among patients, caregivers, and HCPs, as well as among HCPs, was frequently mentioned:

> *Communication between specialists could probably be better, and between the specialists and the patients also.* [Caregiver, Spain]

This, in turn, enables information sharing among all the stakeholders involved in the health care process and supports the multidisciplinary approach for patients:

> *It is good for both parties. For example, if someone comes to the hospital, there would be direct information about that person. Then this saves them from repeating their medical history, especially if this does not work well anymore due to an illness. The same is the case with institutions. If I imagine from my professional field in the rehabilitation clinic to have this possibility and to be able to access the information directly, that would be very practical. After all, you want to provide the best possible care for people. By the bundled information you would have a good impression of the person and then you can give more individual advice regarding the future of the patients.* [Market actor 1, Germany]

For health care teams, the PROCare4Life platform was thought to help saving time and effort per patient, reduce the workload on the nursing staff, and subsequently improve the overall work efficiency:

> *I think, that I would benefit from PC4L in terms of working efficiency, saving time devoted for each patient and reduce the reachability time.* [HCP, Italy]

On an individual level, the platform was seen to help patients to live independently at home:

> *People can remain as long as possible in their own coziness, while they are still independent as long as possible.* [HCP, Germany]

It can ensure the engagement of caregivers in the health process and improve the relationship between patients and HCPs:

> *A positive aspect is the involvement of the caregiver, that usually is unfairly underestimated [...] enhancing the relationship of trust between patient and health professionals.* [HCP, Italy]

ST2.4: Concerns

Most of the concerns expressed by the participants regarding the proposed platform were related to privacy and data protection. Owing to the nature of such platforms that require sharing personal data, explaining everything to the user and obtaining their consent were considered a must. Meanwhile, concerns related to the security of the platform and the protection measures followed to secure the data were mentioned:

> *How secure is all this? So how secure is this server? So that's what I always think. It is also very sensitive data. And we know: data protection, hackers—a lot can happen and you have to be aware of that.* [HCP, Germany]

Costs and the price of the platform were seen as a typical barrier not only for PROCare4Life but also for any digital health care solution:

> *The typical barrier I would say it is the price that it takes to be implemented.* [Media actor, Italy]

Finally, the questionable abilities of the patients to handle the proposed platform was reported as a concern:

> *I don't think patients can handle it. They don't know the technology well.* [Caregiver, Germany]

ST2.5: Willingness to Invest

Regarding the willingness to invest in the proposed platform, participants mainly commented on 3 areas of investment (financial, time, and effort). Most patients and caregivers were

unsure about investing money to pay for this platform. Patients mostly wanted to see the benefits they would get from the platform before deciding, whereas formal caregivers thought that the platform should be financed by patients, their relatives, or health insurance authorities:

> *Depends on the perspective. I as a nurse no. This is what relatives and patients should pay for. Or actually the health insurance companies.* [Caregiver, Germany]

On the other hand, HCPs reported that health institutions would pay for such platforms if they are beneficial:

> *Yes, if it's something new and that we can all benefit from that, we think our organization would pay for that.* [HCP, Portugal]

Another HCP suggested the platform to be financed by a third party or providing it as a rent service:

> *You think about who has to pay for a platform like that, you can simply offer it as a service. That can be financed by a project, or any other financing possibility than the user. Maybe in a rent form, I don't know, 60 or even 300 Euros per month. As long as its value is favorable.* [HCP, Germany]

Regarding the time and effort to learn, the end users were more willing to invest in the proposed platform, stating that it pays off eventually:

> *It takes an initial time to get used to all the tools but it pays off in long-term.* [HCP, Portugal]

Supportive Quantitative Results for Theme 2

Questions related to device acceptance, expected benefits, and willingness to pay for the proposed platform were asked in the web-based surveys.

Patients and caregivers were asked about their acceptance for several technological devices that are thought to be included in the platform. On a Likert item questionnaire, wearable devices were highly accepted among patients; in fact, 57% (48/85) of them would like or love to use it, whereas mobile or tablet was more accepted by caregivers (50/72, 69%). Cameras were the least accepted, as only 33% (28/85) of the patients and 31% (22/72) of caregivers would like or love to use them.

Among a list of different expected benefits from the proposed platform, increasing feelings of safety or autonomy of patients in their homes was the most agreed upon by both patients (66/85, 78%) and caregivers (50/72, 69%).

When asked about their willingness to pay for the proposed platform, around half of the patients (43/85, 51%) and caregivers (35/72, 49%) answered "I don't know."

More detailed quantitative findings can be found in the Multimedia Appendix 2.

Theme 3: Desired Platform Properties and Supportive Measures

This theme reports about participants' answers regarding the needed features and supportive functionalities to be included in PROCare4Life.

ST3.1: Alerts for Adverse Events

Some symptoms related to NDDs can appear suddenly. Including a function that detects adverse events and informs a responsible party about them was thought to be useful and a relief for patients and caregivers:

> *It [adverse events alert] would be very useful. So that I know if there's something wrong with her and what it could be. It would leave me in peace.* [Caregiver, Spain]

Meanwhile, adverse event alerts support HCPs in detecting the daily irregularities or the symptom-induced hazardous situations of the patients and reporting them to the health care team:

> *A tool to measure, record and analyze on/off stages, swallowing, activities of daily living, falls, dangerous behavior, quality of sleep. With the possibility of reporting the adverse effects to the nurse.* [HCP, Spain]

However, one of the caregivers stated that for adverse event alert to be effective, supportive infrastructure is needed:

> *We often get calls that a patient in a city several hundred kilometers away. And we have nothing to do with that. It just makes us crazy and doesn't help people. The infrastructure for this has not yet been properly developed.* [Caregiver, Germany]

ST3.2: Monitoring

Having a tool that monitors the patients' health status was seen as very useful by most of the participants, as it provides continuous and objective information about the patients. It can be used to monitor the patient's vital signs and other health-related measures:

> *So these systems could be used for monitoring pulse and sugar status and vitamin status and nutritional status and exercise status.* [Market actor 2, Germany]

HCPs were interested in monitoring the symptom evolution and the side effects of certain medications. Furthermore, monitoring the movement and gait patterns to detect any changes (ie, fall detection), particularly in patients with PD, was thought to be important:

> *One could also install sensors in the room and measure the movement patterns of the residents with Parkinson's disease. This would also make it possible to recognize early on, for example, if the person has a certain movement pattern that he or she will soon fall. And warning signals can also be sent accordingly.* [HCP, Germany]

In addition, monitoring patients in real time was thought to be important:

> *The perspective of the information in a real moment it's really important and the possibility to also maintain informed people of interest related to the caregivers or people who are monitoring them.* [Academia 2, Germany]

One of the caregivers stated that if patients were monitoring themselves all the time, without knowing how to interpret the values, it would be stressful for the patients:

> *If the patients are monitoring themselves, I find it horrible...And if patients always monitor themselves, I think they're just afraid of the values and then it's the bracelet or the fear of the values that stress the patients, not the disease itself.* [Caregiver, Germany]

This could mean that monitoring needs to be controlled or adapted to the users.

ST3.3: Communication Tools

Inclusion of communication tools that could work in different ways was frequently reported. First, the participants thought that a tool for communication that enables patients and caregivers to ask questions regarding their medications or issuing health reports and sick leave without in-person appointments with HCPs would be useful:

> *For example, information about drugs contraindicated in the disease or on aspects such as a sick leave, some simple information that would avoid me having to go to my health center.* [Patient 1 with PD, Spain]

In addition, HCPs can launch digital interviews and interventions, for instance, physiotherapy sessions that could be held on the web when the patients are not able to personally visit the health care institutions:

> *A tool that would allow video calls: Digital interviews, interventions.* [HCP, Spain]

Follow-up sessions could also be held through such tools:

> *For some phases video calls can be useful, I don't know about therapy but I can imagine that you can have it where you meet physically but also have check-up during the week over a video call.* [Academia 1, Germany]

Second, one of the HCPs thought that a communication tool that facilitates interaction among HCPs from different specialties regarding their patients' health conditions and treatments would be useful:

> *It would be great to have a two-way access and communication service with other professionals.* [HCP, Spain]

Another way for communication was suggested by one of the caregivers, which is implementing a chat tool or system where different caregivers and involved persons could share knowledge and interact:

> *A chat system! So that either a caregiver or qualified staff could answer questions, I mean, it should somehow serve as an interactive chat among all users, or a solution should be sought somehow; I mean, if I need to pose a question, it should pop up, as it would with a forum, and anyone should be able to answer that question, that is, if the site developer is not available right there and then, if someone else is nevertheless available to talk to me and share their opinion, then from the opinions of two or three participants I might extract a conclusion guiding me this or that way.* [Caregiver, Romania]

However, some of the local organizations participating in this project already have their own internal communication channels using WhatsApp or Facebook, through which they can interact with workers in their organization or organize some social events with their peers:

> *I am already in the WhatsApp group of the Parkinson sports team and also in the Facebook Parkinson care group.* [Patient 1 with PD, Spain]

Another point that a communication tool within the platform should not replace in-person communication, particularly in the case of dementia where personal contact is crucial, was highlighted:

> *I think with older people and with people with dementia is still, I think the personal conversation is still more important, maybe people with dementia don't understand that it goes over a screen.* [HCP, Germany]

ST3.4: Reminders

With the memory problems associated with NDDs, integrating reminders that inform the patients of what they have done or suggest what they should do was frequently reported in the interviews for providing support to the caregivers and patients' families:

> *The relatives are also relieved by this. Because they also think a lot for the patients. And when they no longer have to take over this reminder function, they are also relieved.* [HCP, Germany]

From the patients' and caregivers' perspectives, reminders can be sent for medication and mealtimes:

> *Maybe a device to tell her, to remind her of medication or meal hours. Or one that reminds me of all this when I've got my hands full at work, so that I can then give her reminder calls myself. For no matter how many memos I make, I get buried in my work and forget about them* [Caregiver, Romania]

For HCPs, reminders can help patients remember their medical appointments and ensure that their daily water intake is consumed:

> *A memory of drinking. Or we do—today at 11:00 o'clock–10 minutes of gymnastics.* [HCP, Germany]

ST3.5: Lifestyle Content

Inclusion of some support related to the lifestyle activities of both patients and caregivers was mentioned by the participants of this study. The areas of support reported were nutrition and physical activity:

> *Better implement an online nutrition program. I would always try to add sports and movement offers. Just as well as cultural aspects.* [Academia 2, Germany]

Sociocultural events and cognitive games were also preferred:

> *Yes, I would love to use it [games].* [Patient 1 with PD, Spain]

ST3.6: Ease of Use

The proposed platform needs to be simple and practical as stated by one of the patients with PD:

> *For me it is important that the system is practical and as simple as possible.* [Patient 2 with PD, Spain]

In the workshops, one of the HCPs mentioned that it is even better if the platform is passive, considering the cognitive abilities of the patients with NDDs:

> *In the case of people with cognitive impairment, it will also be important that the technology is not only easy but even passive.* [HCP, Spain]

Different criteria for ease of use were reported by all the participants, such as low interaction with the platform, easy navigation of the platform and retrieval of information, and fewer approvals and requests for setting up the platform.

One of the key stakeholders, who is also a clinical professional, stated that for clinical professionals, the easier the platform, the more it will be accepted and used:

> *For us as clinical practitioners I think that must be a solution where we can easily access the information and the information must be as simple as we can get. So, I think we need to have something that we don't need to learn to use it, or at least if we have to, just learn a minimum of sets to use it because if it's complex system, I think we are going to let it go. The easier to use, the easier it will be.* [Academia 4, Portugal]

ST3.7: Personalization

A relevant aspect of the platform is the ability to be customized to fit the users' needs and preferences, especially those needs related to the cognitive problems of the patients:

> *Adapted to the needs of the person according to his/her cognitive state and preferences.* [HCP, Spain]

Personalization can be achieved by adapting the platform's tools according to the target group's needs. "*Reminders*" was highlighted as one of the desired tools. For example, a water intake reminder can be developed, which is important for older people with dementia who often forget to drink hydrating fluids:

> *People with dementia often have the problem that they do not drink enough. And maybe you could install a water dispenser for them.* [HCP, Germany]

It is important that the platform considers the various needs among different end user groups and that it offers flexibility of its functions so that the end users can adjust it according to their interests and needs:

> *Also, the personalization...not only by the professional but also by other users.* [HCP, Spain]

> *I think this is something individual. It's not something you get off the shelf but like with the apps you can buy when you need something. That you can expand the system depending on the degree of illness you have or the need you have.* [Market actor 2, Germany]

Furthermore, the platform can support HCPs in personalizing their treatment plans when disease-specific information about the patients is available:

> *In my case as a physiotherapist, I would like to have access to information about habitual displacements or activities in which the person presents motor difficulty, to be able to focus the treatment towards a more functional objective. In this way we would achieve more personalized treatments.* [HCP, Spain]

ST3.8: User-friendliness

For the platform to be user-friendly, patients would like to have comfortable wearables that do not irritate their skin and familiar sensors that do not cause them anxiety:

> *Very important that it [wearables] doesn't irritate.* [Caregiver, Spain]

For the HCPs, the user-friendly criteria were that the platform layout is predefined, has less text presentation, and focuses more on visualizing the information using symbols and graphs:

> *No text. I think most people don't like to read text...I would like to see this in a traffic light system. So that the green area is everything ok, everything is good, with red something has to be done.* [HCP, Germany]

ST3.9: Training

HCPs and other key stakeholders pointed out the need for the end users to be trained on how to use the platform. This includes educating them about the different devices and providing training sessions on their use. Furthermore, training should not only rely on manuals but also provide some interactive training sessions:

> *I can learn it myself, but it is difficult for me to learn new programs by reading manual. As far as such new program I need someone who sits next to me and introduces me to the program.* [Market actor 1, Germany]

ST3.10: Technical Support

With a digital platform, technical issues can always arise. Therefore, the presence of technical support was seen as a need by HCPs and other key stakeholders. This support can be in the form of automatic backups of the stored information or as supportive infrastructure and networking:

> *If the system fails, having an automatic backup to prevent loss of information or any delays on the reports.* [HCP, Portugal]

In addition, a service hotline to report urgent technical issues is required:

> *A service hotline would also make sense. If problems arise, it is important to reach someone.* [HCP, Germany]

Supportive Quantitative Results for Theme 3

The web-based surveys included questions about several features and functionalities to be included in the platform. On a 3-point Likert-items questionnaire, a tool to monitor symptoms and activities was the highest-rated functionality, as 85% (72/85)

of the patients found it very useful or useful. Most caregivers (65/72, 90%) found that a tool to detect adverse events, unusual activities, or movements was useful.

In addition, when asked about the importance of different features of the platform, 80% (68/85) of the patients reported the real-time information feature as very important, whereas approximately 89% (64/72) of the caregivers reported an easy-to-set-up platform as very important. Detailed results from the surveys regarding this theme are provided in the Multimedia Appendix 3.

Discussion

Principal Findings

This study aimed to identify the needs and requirements for an ICT-based integrated care platform in supporting its potential end users and other stakeholders involved in the health care process of NDDs by exploring the participants' opinions regarding the health care process and digital health care solutions. Findings on experiences associated with NDD symptoms, the challenges faced by all the potential end users in health care services, and the inevitable changes to life and its context have been well documented [2,3]. This is perhaps why participants were so forthcoming with wanting to share their experience in the hope of a supportive technology solution.

Although both positive and negative attitudes of the participants toward PROCare4Life were explored, the participants expected several benefits when using the platform. These varied between supporting the patients' empowerment and independence, increasing caregivers' and HCPs' work efficiency, and ultimately improving health care services. Notably, influenced by the COVID-19 pandemic, most participants referred to the platform as a need, with most of them being willing to use it. This finding is of particular importance, as older people with chronic illnesses have been identified as a vulnerable group, who require special consideration for encouraging them to use ICT in health care during pandemic times [46].

Most of the desired functionalities reported here confirm findings from previous research, such as the need for medication intake reminders, monitoring, and communication tools [32,47]. We provide more insights into the specific needs related to these functionalities, such as the patients' constant need for meals, water intake, and appointment reminders. In addition, a monitoring functionality that provides objective data in real time for both caregivers and HCPs was reported, with emphasis on monitoring motor symptoms and gait patterns and installing a fall detection system. A communication tool was valued by all the potential end users, believing that it improves the relationship among patients, caregivers, and HCPs and facilitates the overall health care process. It is known that ICTs are useful in reducing social isolation and providing opportunities for older adults to keep in touch with the outside world [28,48,49]. Therefore, implementing a communication tool could alleviate the isolation and feeling of distress reported by the patients and caregivers in this study. Furthermore, as stated by one of the caregivers, a chat tool facilities interaction between peers and knowledge exchange among all the parties involved in the care process.

Another important finding was the need for personalization, which stems from the individuality of each user, including disease severity, experiences, and preferences. As the saying "One size does not fit all" goes, as this platform targets different end user groups, a flexible design should be considered. In addition, with the progressive nature of NDDs, there is a need to offer tailored functionalities that match the stage of the disease and patients' abilities. Having too many functionalities and options might cause confusion for the end users, particularly for patients and caregivers. Similar to what has been reported by Boman et al [33], some of our participants pointed out the need to provide >1 version or package of the platform to facilitate the customization of the product.

The diversity of needs, along with a few contradictions between the end user groups, was the main challenge in our work. Although patients and caregivers were concerned about using stationary devices and cameras, HCPs expressed that including these devices is important for ensuring the real-time monitoring functionality of the platform. In addition, the results show how end user groups tend to have different perspectives when identifying a specific property. For instance, although all the participants highlighted the need for the platform to be easy to use and user-friendly, both patients and caregivers viewed this as having familiar objects and comfortable wearables, whereas HCPs cared more for the visualization (ie, having a traffic color system) and the layout of the platform. Furthermore, HCPs emphasized that a platform that is easy to navigate through and retrieve information from is more likely to be used. It is known that simplicity, ease of use, and understandable features increase the possibility of older people with impaired cognitive abilities using digital devices independently and for a long term [29,33,50-54].

The need for providing education about the platform and training on the devices to be used were stated by HCPs and other key stakeholders. Therefore, in addition to patients, offering training opportunities for HCPs and caregivers is of great importance. Staff who are familiar with the platform, have experience, and are interested in using it play an important role in encouraging their patients to use it [55,56].

It is also noteworthy that potential end users found it difficult to decide whether they would pay for the PROCare4Life platform and by how much. Although HCPs referred to this point as one of the typical barriers for ICT health care platforms, they were more willing to pay compared with patients and caregivers and believed that their institutions would be willing too. Identical to what was reported by Contreras-Somoza et al [48], who studied the acceptability of an ICT device for older adults with mild cognitive impairment, formal caregivers in this study expected PROCare4Life to be financed by the patients themselves, relatives, or health insurance companies. Furthermore, the patients in our study needed to grasp the benefits of the platform first before deciding on paying for it. HCPs suggested providing renting offers of the platform to ease the cost burden for those who would like to use it, in case no other financing possibility was available. Therefore, there is a

need for providing a better investment in digital solutions that support healthy and independent aging, which relies on collaboration between the government, organizations, and the private health sector [48,57].

When developing an ICT platform that supports all the parties involved in the care process of older people with NDDs, identifying the users' needs, interests, and abilities is crucial [58], regardless of being challenged with the diversity of needs and different priorities of the participants in this study. Combining different views from different perspectives (eg, patients, caregivers, HCPs, and other stakeholders) is thought to prevent individual concerns such as privacy issues from becoming a barrier for using technology [59,60] and ultimately increase the acceptability of technology in health care [47].

In addition, our participants valued this study and appreciated the idea of trying to include them in the development process and understanding their needs:

> *The strongest point for me is doing this interview, to really start asking what do people need and want and do they think they can use it. I think that it relates directly to usability, and the user experience which is crucial. I worked together with therapists to develop a system back there, but no matter how good the system was, if it's not being used, there is nothing you can do about it. So, I think what you are doing is very important for the acceptance of the people for this system.* [Academic Researcher 1, Germany]

For this purpose, encapsulating a UCD approach in the development stages ensures addressing the real needs and avoiding poor final acceptance. Indeed, the rationale behind UCD is that the "purpose of any design is to serve the user, not to use a specific technology or to be an elegant piece" [61].

Implications on PROCare4Life

The data collected and knowledge gained from this study were transferred to the development and research team of PROCare4Life with the purpose of designing and redesigning the different services of the platform. What ultimately emerged from this initial phase of the project supported the efforts to raise awareness about the major areas of users' needs, where technological aspects of the platform could be more valuable. The identified "desired properties" represent the main core of the final platform solution based on a realistic idea of the problem and a better vision of what to prioritize for each of the properties, providing an insight into how the scheme of operation of the system was to be shaped. In addition, the work on specifying areas of technical interest to users is an ongoing process that shall continue throughout the project pilots as new needs emerge followed by further technical developments.

Furthermore, this study identified several prerequisites for the acceptance of this platform, for instance, training and technical support. Therefore, training manuals for users containing instructions for the implementation and use of the system have been produced and will be continuously adapted and improved until the end of the project for personalized configuration along with more training opportunities, such as interactive sessions and e-learning, which are currently being assessed in the framework of the pilots. In addition, each clinical site was assigned a technical partner for support and guidance in the event of any technical difficulty.

Regarding privacy and protection in the use of cameras, the technical team worked on a code with the depth camera and the real-time software to avoid the storage of any patient's images. In addition, a simplified illustrative video was distributed among the users, explaining how the images acquired by the camera are directly processed by a software that generates an output composed of an 18-point skeleton in combination with the depth information. Thus, no images are saved, avoiding privacy issues.

Strengths and Limitations

Developing a successful product means that the needs of the target groups are included [62]. This research was based on a large sample of participants, including different perspectives presented by patients, caregivers, HCPs, and other related key stakeholders. This ensured that all views on different needs and challenges are considered during the development and future pilot phases. The inclusion of patients with different NDDs (ie, dementia and PD) and the multisite (ie, 5 different European countries) approach allow the findings to be generalized so that the PROCare4Life platform solution can be applied to other chronic diseases and facilitates the exploitability of the achieved results in the long run. Furthermore, the mixed methodology study enriched the comprehensive data that reflected the diverse needs of the participants.

One of the main challenges this study faced was the COVID-19 pandemic, which directly impacted the dynamics and schedule of the PROCare4Life project in its first months of development. Confinement and social distancing measures implemented by all member states and local authorities in the different European countries required an adaptation in approaching and recruiting the participants involved in the study. Most data were collected remotely through phone or digital interviews and on the web; however, a few meetings were held face to face. This rearrangement required extra work for the researchers involved in this study, ensuring that the prevention and control measures to avoid infections were strictly applied. Surely, this situation created some challenges and limited in-person contact with the participants. Nevertheless, this study was able to approach the intended target groups and numbers. In addition, the PROCare4Life approach to the current pandemic situation was continuously documented in clinical partners' countries to ensure standardized preventive measures. This exchange led to the publication of a COVID-19 protocol that shares the instructions and suggestions of the respective national and local health authorities, which is available on the web [63]. In addition, the awareness of the pandemic seemed to open perspectives to the usefulness of the platform (eg, ST2.3).

By the time of the study, the design of the functionalities in the mock-ups were not developed yet. Although this was expected, considering the aim of the study and the early stage of the development process, some of the participants, in particular patients, were not able to imagine the design of the platform. It has been pointed out in the literature that patients with declined cognitive ability might have difficulty in imagining the things they cannot see or articulating their perceptions of the device

intended to be used [33-47]. Considering this, HCPs and caregivers who were familiar with the patients were included. Furthermore, the early involvement of users in the process ensures the suitability of the product for its intended target group and purpose [64].

Another limitation is that this study prioritized qualitative data, which could mean losing some important quantitative results, in particular those related to patients and caregivers. However, the study design allowed the sharing of the same topics in both study methods. During the merging of the results, all the supportive quantitative data were considered. Finally, there was an overrepresentation of patients with PD in the study, which might have caused bias in the results toward motor symptoms as well as the inability of the study to draw conclusions on the specific needs of patients with other NDDs. The reason could be that patients with PD, especially in the early stages of the disease, are less likely to experience cognitive decline, and therefore, they are more easily engaged in this kind of study where they have to sustain attention and communicate their opinions clearly. In addition, in the countries involved in this study, associations of patients with PD are very active and eager to become involved. Nevertheless, samples were free from selection bias and were naturalistic under the overall umbrella of NDDs. For future studies, there might be a need for such kind of platforms to be designed for each neurodegenerative condition according to its own peculiarities.

Conclusions

In this study, the needs of all the parties involved in the health care process of NDDs regarding an ICT-based health care platform were explored. The pandemic situation highlighted opportunities for digitalization in health care. The mixed methods approach yielded mostly consistent results, which were in line with findings from the literature. The collected data were useful for the development of the PROCare4Life platform.

Although the combination and collection of features for diverse user groups are typical for integrated care platforms, this results in exponential complexity for designers, developers, and users. Contradicting opinions and several concerns in this study demonstrate that an integrated care platform should not promise too much for too many. Instead, selection, focus, and, sometimes, restriction to the essentials are necessary. Users and other stakeholders should be involved in these decisions.

Acknowledgments

The authors would like to thank the patients, caregivers, and professionals who agreed to participate in this study. They would also like to thank the researchers who supported data collection from the study sites; all the consortium members of the Personalized Integrated Care Promoting Quality of Life for Older People (PROCare4Life) project; Filipa Pona-Ferreira and Mariana Leitão from Campus Neurológico Sénior, Torres Vedras-Portugal, for their assistance with data collection; and Gina Krause from Münster University-Germany for supporting the qualitative data analysis process. PROCare4Life is funded by the European Commission via the Horizon 2020 program (grant 875221).

Authors' Contributions

This study was conceptualized by MA, MM, and MB. All the authors contributed substantially to the manuscript and approved its final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Results from the web-based surveys corresponding to theme 1 “experiences associated with illness.”.
[PDF File (Adobe PDF File), 240 KB - formative_v6i11e39199_app1.pdf]

Multimedia Appendix 2
Results from the web-based surveys corresponding to theme 2 “thoughts about the platform technology.”.
[PDF File (Adobe PDF File), 208 KB - formative_v6i11e39199_app2.pdf]

Multimedia Appendix 3
Results from the web-based surveys corresponding to theme 3 “desired properties.”.
[PDF File (Adobe PDF File), 248 KB - formative_v6i11e39199_app3.pdf]

References

1. Global Health and Aging. World Health Organization. Bethesda, MA, USA: National Institute on Aging, National Institutes of Health; 2011 Oct. URL: https://www.nia.nih.gov/sites/default/files/2017-06/global_health_aging.pdf [accessed 2020-05-23]

2. De Marchi F, Contaldi E, Magistrelli L, Cantello R, Comi C, Mazzini L. Telehealth in neurodegenerative diseases: opportunities and challenges for patients and physicians. Brain Sci 2021 Mar 13;11(2):237 [FREE Full text] [doi: 10.3390/brainsci11020237] [Medline: 33668641]
3. Olazarán J, Muñiz R, Reisberg B, Peña-Casanova J, del Ser T, Cruz-Jentoft AJ, et al. Benefits of cognitive-motor intervention in MCI and mild to moderate Alzheimer disease. Neurology 2004 Dec 28;63(12):2348-2353. [doi: 10.1212/01.wnl.0000147478.03911.28] [Medline: 15623698]
4. Scalvini S, Giordano A. Heart failure. Optimal postdischarge management of chronic HF. Nat Rev Cardiol 2013 Jan;10(1):9-10. [doi: 10.1038/nrcardio.2012.161] [Medline: 23165075]
5. Bernini S, Stasolla F, Panzarasa S, Quaglini S, Sinforiani E, Sandrini G, et al. Cognitive telerehabilitation for older adults with neurodegenerative diseases in the COVID-19 era: a perspective study. Front Neurol 2020 Jan 14;11:623933 [FREE Full text] [doi: 10.3389/fneur.2020.623933] [Medline: 33519704]
6. Jia J, Wei C, Chen S, Li F, Tang Y, Qin W, et al. The cost of Alzheimer's disease in China and re-estimation of costs worldwide. Alzheimers Dement 2018 Apr;14(4):483-491. [doi: 10.1016/j.jalz.2017.12.006] [Medline: 29433981]
7. Rodríguez-Blázquez C, Forjaz MJ, Lizán L, Paz S, Martínez-Martín P. Estimating the direct and indirect costs associated with Parkinson's disease. Expert Rev Pharmacoecon Outcomes Res 2015;15(6):889-911. [doi: 10.1586/14737167.2015.1103184] [Medline: 26511768]
8. Oh J, An JW, Oh SI, Oh KW, Kim JA, Lee JS, et al. Socioeconomic costs of amyotrophic lateral sclerosis according to staging system. Amyotroph Lateral Scler Frontotemporal Degener 2015 Jun;16(3-4):202-208. [doi: 10.3109/21678421.2014.999791] [Medline: 25646865]
9. Wake E, Atkins H, Willock A, Hawkes A, Dawber J, Weir KA. Telehealth in trauma: a scoping review. J Telemed Telecare 2022 Jul;28(6):412-422. [doi: 10.1177/1357633X20940868] [Medline: 32715866]
10. Broese van Groenou MI, De Boer A. Providing informal care in a changing society. Eur J Ageing 2016;13(3):271-279 [FREE Full text] [doi: 10.1007/s10433-016-0370-7] [Medline: 27610055]
11. Kuluski K, Ho JW, Hans PK, Nelson ML. Community care for people with complex care needs: bridging the gap between health and social care. Int J Integr Care 2017 Jul 21;17(4):2 [FREE Full text] [doi: 10.5334/ijic.2944] [Medline: 28970760]
12. Kirst M, Im J, Burns T, Baker GR, Goldhar J, O'Campo P, et al. What works in implementation of integrated care programs for older adults with complex needs? A realist review. Int J Qual Health Care 2017 Oct 01;29(5):612-624 [FREE Full text] [doi: 10.1093/intqhc/mzx095] [Medline: 28992156]
13. INclusive INtroduction of INtegrated CAre (IN3CA). European Commission. Brussels, Belgium: European Commission; 2017. URL: https://cordis.europa.eu/project/id/621006 [accessed 2022-02-28]
14. The European Innovation Partnership on Active and Healthy Ageing - Integrated Care and Chronic Disease Management. European Commission. 2012. URL: https://www.idieikon.com/inca/integrated-care.html [accessed 2022-02-27]
15. Carretero S. Technology-enabled Services for Older People Living at Home Independently: Lessons for public long-term care authorities in the EU Member States. European Commission. Luxembourg, Luxembourg: Publications Office of the European Union; 2015. URL: https://op.europa.eu/en/publication-detail/-/publication/3b19e4d2-199a-4019-9b31-2d8d539d9bea/language-en [accessed 2021-11-30]
16. Bennett B, McDonald F, Beattie E, Carney T, Freckelton I, White B, et al. Assistive technologies for people with dementia: ethical considerations. Bull World Health Organ 2017 Dec 01;95(11):749-755 [FREE Full text] [doi: 10.2471/BLT.16.187484] [Medline: 29147055]
17. Fares N, Sherratt RS, Elhajj IH. Directing and orienting ICT healthcare solutions to address the needs of the aging population. Healthcare (Basel) 2021 Mar 02;9(2):147 [FREE Full text] [doi: 10.3390/healthcare9020147] [Medline: 33540510]
18. Rajagopalan R, Litvan I, Jung TP. Fall prediction and prevention systems: recent trends, challenges, and future research directions. Sensors (Basel) 2017 Dec 01;17(11):2509 [FREE Full text] [doi: 10.3390/s17112509] [Medline: 29104256]
19. Lindberg B, Nilsson C, Zotterman D, Söderberg S, Skär L. Using information and communication technology in home care for communication between patients, family members, and healthcare professionals: a systematic review. Int J Telemed Appl 2013;2013:461829 [FREE Full text] [doi: 10.1155/2013/461829] [Medline: 23690763]
20. Guisado-Fernández E, Giunti G, Mackey LM, Blake C, Caulfield BM. Factors influencing the adoption of smart health technologies for people with dementia and their informal caregivers: scoping review and design framework. JMIR Aging 2019 May 30;2(1):e12192 [FREE Full text] [doi: 10.2196/12192] [Medline: 31518262]
21. American Psychiatric Association. Diagnostic and statistical manual of mental disorders (DSM-IV). 4th edition. Washington, DC, USA: American Psychiatric Association; 2000.
22. Diaz-Orueta U, Hopper L, Konstantinidis E. Shaping technologies for older adults with and without dementia: reflections on ethics and preferences. Health Informatics J 2020 Dec;26(4):3215-3230 [FREE Full text] [doi: 10.1177/1460458219899590] [Medline: 31969045]
23. Committee on the Learning Health Care System in America, Institute of Medicine. In: Smith M, Saunders R, Stuckhardt L, McGinnis JM, editors. Best Care at Lower Cost: The Path to Continuously Learning Health Care in America. Washington, DC, USA: National Academies Press (US); 2013.
24. Ahmed M, Marín M, Bouça-Machado R, How D, Judica E, Tropea P, et al. Investigating users' and other stakeholders' needs in the development of a personalized integrated care platform (PROCare4Life) for older people with dementia or

Parkinson disease: protocol for a mixed methods study. JMIR Res Protoc 2021 Jan 12;10(1):e22463 [FREE Full text] [doi: 10.2196/22463] [Medline: 33433394]
25. Ergonomics of Human-System Interaction -- Part 210: Human-Centred Design for Interactive Systems. International Organization for Standardization. 2019. URL: https://www.iso.org/obp/ui/#iso:std:iso:9241:-210:ed-2:v1:en [accessed 2022-01-03]
26. Steele Gray C, Barnsley J, Gagnon D, Belzile L, Kenealy T, Shaw J, et al. Using information communication technology in models of integrated community-based primary health care: learning from the iCOACH case studies. Implement Sci 2018 Jun 26;13(1):87 [FREE Full text] [doi: 10.1186/s13012-018-0780-3] [Medline: 29940992]
27. Maresova P, Krejcar O, Barakovic S, Barakovic Husic J, Lameski P, Zdravevski E, et al. Health–related ICT solutions of smart environments for elderly–systematic review. IEEE Access 2020 Mar 17;8:54574-54600. [doi: 10.1109/access.2020.2981315]
28. Meiland F, Innes A, Mountain G, Robinson L, van der Roest H, García-Casal JA, et al. Technologies to support community-dwelling persons with dementia: a position paper on issues regarding development, usability, effectiveness and cost-effectiveness, deployment, and ethics. JMIR Rehabil Assist Technol 2017 Jan 16;4(1):e1 [FREE Full text] [doi: 10.2196/rehab.6376] [Medline: 28582262]
29. Vermeer Y, Higgs P, Charlesworth G. What do we require from surveillance technology? A review of the needs of people with dementia and informal caregivers. J Rehabil Assist Technol Eng 2019 Dec 2;6:2055668319869517 [FREE Full text] [doi: 10.1177/2055668319869517] [Medline: 31832230]
30. van der Roest HG, Meiland FJ, Comijs HC, Derksen E, Jansen AP, van Hout HP, et al. What do community-dwelling people with dementia need? A survey of those who are known to care and welfare services. Int Psychogeriatr 2009 Oct;21(5):949-965. [doi: 10.1017/S1041610209990147] [Medline: 19602305]
31. Miranda-Castillo C, Woods B, Orrell M. The needs of people with dementia living at home from user, caregiver and professional perspectives: a cross-sectional survey. BMC Health Serv Res 2013 Mar 04;13:43 [FREE Full text] [doi: 10.1186/1472-6963-13-43] [Medline: 23379786]
32. Lauriks S, Reinersmann A, Van der Roest HG, Meiland FJ, Davies RJ, Moelaert F, et al. Review of ICT-based services for identified unmet needs in people with dementia. Ageing Res Rev 2007 Oct;6(3):223-246. [doi: 10.1016/j.arr.2007.07.002] [Medline: 17869590]
33. Boman IL, Persson AC, Bartfai A. First steps in designing an all-in-one ICT-based device for persons with cognitive impairment: evaluation of the first mock-up. BMC Geriatr 2016 Mar 07;16:61 [FREE Full text] [doi: 10.1186/s12877-016-0238-x] [Medline: 26951709]
34. Nieuwboer A, De Weerdt W, Dom R, Lesaffre E. A frequency and correlation analysis of motor deficits in Parkinson patients. Disabil Rehabil 1998 May;20(4):142-150. [doi: 10.3109/09638289809166074] [Medline: 9571381]
35. Espay AJ, Bonato P, Nahab FB, Maetzler W, Dean JM, Klucken J, Movement Disorders Society Task Force on Technology. Technology in Parkinson's disease: challenges and opportunities. Mov Disord 2016 Sep;31(9):1272-1282 [FREE Full text] [doi: 10.1002/mds.26642] [Medline: 27125836]
36. Setareh Forouzan A, Mahmoodi A, Jorjoran Shushtari Z, Salimi Y, Sajjadi H, Mahmoodi Z. Perceived social support among people with physical disability. Iran Red Crescent Med J 2013 Aug;15(8):663-667 [FREE Full text] [doi: 10.5812/ircmj.12500] [Medline: 24578832]
37. Lundberg S. The results from a two-year case study of an information and communication technology support system for family caregivers. Disabil Rehabil Assist Technol 2014 Jul;9(4):353-358. [doi: 10.3109/17483107.2013.814170] [Medline: 23855387]
38. Creswell JW, Plano Clark VL. Designing and Conducting Mixed Methods Research. 2nd edition. Thousand Oaks, CA, USA: Sage Publications; 2010.
39. PROCare4Life. URL: https://procare4life.eu/ [accessed 2022-04-15]
40. Spanish Society of Neurology. Movement Disorders Study Group. URL: https://getm.sen.es/ [accessed 2022-06-12]
41. Mahoney FI, Barthel DW. Functional evaluation: the Barthel index. Md State Med J 1965 Mar;14:61-65. [Medline: 14258950]
42. Gort AM, March J, Gómez X, de Miguel M, Mazarico S, Ballesté J. Short Zarit scale in palliative care. Med Clin (Barc) 2005 May 07;124(17):651-653. [doi: 10.1157/13074742] [Medline: 15882512]
43. Gort AM, Mingot M, Gomez X, Soler T, Torres G, Sacristán O, et al. Use of the Zarit scale for assessing caregiver burden and collapse in caregiving at home in dementias. Int J Geriatr Psychiatry 2007 Oct;22(10):957-962. [doi: 10.1002/gps.1770] [Medline: 17299807]
44. Boyatzis RE. Transforming Qualitative Information: Thematic Analysis and Code Development. Thousand Oaks, CA, USA: Sage Publications; 1998.
45. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
46. Bajpai N, Biberman J, Ye Y. ICTs and Public Health in the Context of COVID-19. Center for Sustainable Development, Columbia University. 2020 Apr. URL: https://csd.columbia.edu/sites/default/files/content/docs/ICT%20India/Papers/ICT_India_Working_Paper_30_0.pdf [accessed 2022-01-23]

47. Meiland FJ, Hattink BJ, Overmars-Marx T, de Boer ME, Jedlitschka A, Ebben PW, et al. Participation of end users in the design of assistive technology for people with mild to severe cognitive problems; the European Rosetta project. Int Psychogeriatr 2014 May;26(5):769-779. [doi: 10.1017/S1041610214000088] [Medline: 24507571]
48. Contreras-Somoza LM, Irazoki E, Castilla D, Botella C, Toribio-Guzmán JM, Parra-Vidales E, et al. Study on the acceptability of an ICT platform for older adults with mild cognitive impairment. J Med Syst 2020 May 25;44(7):120. [doi: 10.1007/s10916-020-01566-x] [Medline: 32451740]
49. Chen YR, Schulz PJ. The effect of information communication technology interventions on reducing social isolation in the elderly: a systematic review. J Med Internet Res 2016 Jan 28;18(1):e18 [FREE Full text] [doi: 10.2196/jmir.4596] [Medline: 26822073]
50. Cahill S, Begley E, Faulkner JP, Hagen I. "It gives me a sense of independence" – findings from Ireland on the use and usefulness of assistive technology for people with dementia. Technol Disabil 2007 Apr 13;19(2-3):133-142. [doi: 10.3233/tad-2007-192-310]
51. Evald L. Prospective memory rehabilitation using smartphones in patients with TBI: what do participants report? Neuropsychol Rehabil 2015;25(2):283-297. [doi: 10.1080/09602011.2014.970557] [Medline: 25333304]
52. Arthanat S, Simmons CD, Favreau M. Exploring occupational justice in consumer perspectives on assistive technology. Can J Occup Ther 2012 Dec;79(5):309-319. [doi: 10.2182/CJOT.2012.79.5.7] [Medline: 23539776]
53. Jamieson M, Cullen B, McGee-Lennon M, Brewster S, Evans JJ. The efficacy of cognitive prosthetic technology for people with memory impairments: a systematic review and meta-analysis. Neuropsychol Rehabil 2014;24(3-4):419-444. [doi: 10.1080/09602011.2013.825632] [Medline: 23957379]
54. Kintsch A, DePaula R. A framework for the adoption of assistive technology. University of Colorado at Boulder. Colorado, USA: Supporting Learning Through Assistive Technology; 2002. URL: http://l3d.cs.colorado.edu/clever/assets/pdf/ak-SWAAAC02.pdf [accessed 2022-01-02]
55. Nordin S, Sturge J, Ayoub M, Jones A, McKee K, Dahlberg L, et al. The role of information and communication technology (ICT) for older adults' decision-making related to health, and health and social care services in daily life-a scoping review. Int J Environ Res Public Health 2021 Dec 23;19(1):151 [FREE Full text] [doi: 10.3390/ijerph19010151] [Medline: 35010408]
56. Gil AP, Moniz AB, de São Jose J. Technologies in care for older people - EPTA Report 2019. European Parliamentary Technology Assessment. The Hague, The Netherlands: European Parliamentary Technology Assessment; 2019. URL: https://eptanetwork.org/images/documents/minutes/EPTA_report_2019.pdf [accessed 2022-02-07]
57. Cuarto examen y evaluación de la aplicación del Plan de Acción Internacional de Madrid sobre el Envejecimiento: Impulsando la agenda de envejecimiento. United Nations. 2020. URL: https://www.cepal.org/es/notas/cuarto-examen-evaluacion-la-aplicacion-plan-accion-internacional-madrid-envejecimiento [accessed 2022-03-15]
58. García-Betances RI, Cabrera-Umpiérrez MF, Ottaviano M, Pastorino M, Arredondo MT. Parametric cognitive modeling of information and computer technology usage by people with aging- and disability-derived functional impairments. Sensors (Basel) 2016 Mar 22;16(2):266 [FREE Full text] [doi: 10.3390/s16020266] [Medline: 26907296]
59. van Boekel LC, Wouters EJ, Grimberg BM, van der Meer NJ, Luijkx KG. Perspectives of stakeholders on technology use in the care of community-living older adults with dementia: a systematic literature review. Healthcare (Basel) 2019 May 28;7(2):73 [FREE Full text] [doi: 10.3390/healthcare7020073] [Medline: 31141999]
60. Olsson A, Persson AC, Bartfai A, Boman IL. Sensor technology more than a support. Scand J Occup Ther 2018 Mar;25(2):79-87. [doi: 10.1080/11038128.2017.1293155] [Medline: 28276963]
61. Mubashra A. User Centred Design: a new essential for effective designs for the physically challenged. Int J Innov Creativity Change 2021;15(7):915-923.
62. McCabe L, Innes A. Supporting safe walking for people with dementia: user participation in the development of new technology. Gerontechnology 2013;12(1):4-15. [doi: 10.4017/gt.2013.12.1.006.00]
63. COVID-19 Protocol. PROCare4Life. 2021 May. URL: https://procare4life.eu/wp-content/uploads/2021/08/IFIC3503-PROCare4Life-COVID-19-Protocol.pdf [accessed 2022-04-22]
64. Duarte J, Guerra A. User-centered healthcare design. Procedia Comput Sci 2012;14:189-197. [doi: 10.1016/j.procs.2012.10.022]

Abbreviations

ADL: activities of daily living
HCP: health care professional
ICT: information and communication technology
NDD: neurodegenerative disease
OD: other dementias
PD: Parkinson disease
PROCare4Life: Personalized Integrated Care Promoting Quality of Life for Older People
ST: subtheme
T: theme

UCD: user-centered design

Edited by A Mavragani; submitted 02.05.22; peer-reviewed by A Olsson, J Hernández-Vara; comments to author 27.06.22; revised version received 18.07.22; accepted 17.08.22; published 30.11.22.

Please cite as:
Ahmed M, Marín M, How D, Judica E, Tropea P, Bentlage E, J Ferreira J, Bouça-Machado R, Brach M
End Users' and Other Stakeholders' Needs and Requirements in the Development of a Personalized Integrated Care Platform (PROCare4Life) for Older People With Dementia or Parkinson Disease: Mixed Methods Study
JMIR Form Res 2022;6(11):e39199
URL: https://formative.jmir.org/2022/11/e39199
doi:10.2196/39199
PMID:36449340

Original Paper

The Passive Monitoring of Depression and Anxiety Among Workers Using Digital Biomarkers Based on Their Physical Activity and Working Conditions: 2-Week Longitudinal Study

Kazuhiro Watanabe[1], PhD; Akizumi Tsutsumi[1], MD

Department of Public Health, Kitasato University School of Medicine, Sagamihara, Japan

Corresponding Author:
Kazuhiro Watanabe, PhD
Department of Public Health
Kitasato University School of Medicine
1-15-1 Kitazato
Minami-ku
Sagamihara, 252-0374
Japan
Phone: 81 42 778 9352
Fax: 81 42 778 9257
Email: kzwatanabe-tky@umin.ac.jp

Abstract

Background: Digital data on physical activity are useful for self-monitoring and preventing depression and anxiety. Although previous studies have reported machine or deep learning models that use physical activity for passive monitoring of depression and anxiety, there are no models for workers. The working population has different physical activity patterns from other populations, which is based on commuting, holiday patterns, physical demands, occupations, and industries. These working conditions are useful in optimizing the model used in predicting depression and anxiety. Further, recurrent neural networks increase predictive accuracy by using previous inputs on physical activity, depression, and anxiety.

Objective: This study evaluated the performance of a deep learning model optimized for predicting depression and anxiety in workers. Psychological distress was considered a depression and anxiety indicator.

Methods: A 2-week longitudinal study was conducted with workers in urban areas in Japan. Absent workers were excluded. In a daily survey, psychological distress was measured using a self-reported questionnaire. As features, activity time by intensity was determined using the Google Fit application. Additionally, we measured age, gender, occupations, employment status, work shift types, working hours, and whether the response date was a working or nonworking day. A deep learning model, using long short-term memory, was developed and validated to predict psychological distress the next day, using features of the previous day. Further, a 5-fold cross-validation method was used to evaluate the performance of the aforementioned model. As the primary indicator of performance, classification accuracy for the severity of the psychological distress (light, subthreshold, and severe) was considered.

Results: A total of 1661 days of supervised data were obtained from 236 workers, who were aged between 20 and 69 years. The overall classification accuracy for psychological distress was 76.3% (SD 0.04%). The classification accuracy for severe-, subthreshold-, and light-level psychological distress was 51.1% (SD 0.05%), 60.6% (SD 0.05%), and 81.6% (SD 0.04%), respectively. The model predicted a light-level psychological distress the next day after the participants had been involved in 3 peaks of activity (in the morning, noon, and evening) on the previous day. Lower activity levels were predicted as subthreshold- and severe-level psychological distress. Different predictive results were observed on the basis of occupations and whether the previous day was a working or nonworking day.

Conclusions: The developed deep learning model showed a similar performance as in previous studies and, in particular, high accuracy for light-level psychological distress. Working conditions and long short-term memory were useful in maintaining the model performance for monitoring depression and anxiety, using digitally recorded physical activity in workers. The developed model can be implemented in mobile apps and may further be practically used by workers to self-monitor and maintain their mental health state.

(JMIR Form Res 2022;6(11):e40339) doi:10.2196/40339

KEYWORDS

digital biomarkers; mobile health; mental health; psychological distress; depression; anxiety; physical activity

Introduction

Physical activity is an important health-related bodily activity for treating and preventing depression and anxiety [1]. Daily physical activity protects against depressive moods experienced by individuals in their daily lives [2]. In the working population, physical inactivity is widespread [3] owing to an increase in work involving low-intensity activities [4]. Based on its effectiveness in preventing depression and anxiety, which are common among workers [5,6], promoting physical activity is the most obvious intervention to effectively prevent common mental disorders in the workplace [7]. For healthy working people, information on their physical activity is useful to self-monitor and maintain their mental health.

Recent studies have shown that physical activity measured using digital tools serve as digital biomarkers in the passive monitoring of depression and anxiety. Furthermore, novel digital technologies and machine or deep learning models have used physical activity to predict depression and anxiety [8-19]. A recent systematic review [8] reported that 19 studies measured physical activity using smartphones and wearable devices to passively monitor depression. It was discovered that activity time, level, intensity, movement speed, and step counts were indicators that are significantly correlated with depression. These indicators predominantly negatively correlated with depression. Support vector machines [9,10], random forests [11,12], regression trees [15,16], regression models [13,18], and ensemble learning were used as learning models. The predictive performance of these models showed moderate to strong correlations with depression and approximately 80% accuracy for depression severity.

However, no studies have developed models that use working conditions to passively monitor depression and anxiety in workers by measuring physical activity using digital tools. A previous study [18] sampled workers and used physical activity, sleep, and the heart rate as features. However, the study did not use information on the working conditions. Other models did not target workers but were based on college or undergraduate students or patients. Regarding the employed population, patterns of physical activity reflected the differences between other populations and that within the population based on commuting or holiday patterns, physical demands, occupations, and industries [20]. These working conditions would explain the variations in physical activities and was useful in optimizing the model to monitor depression and anxiety in workers. Additionally, deep learning models, such as recurrent neural networks, which have feedback connections with successive inputs, would increase the prediction accuracy because the information in the previous inputs regarding physical activity and the state of depression and anxiety explain subsequent inputs.

This study aims to evaluate the predictive and classification performance of a deep learning model for analyzing depression and anxiety, that is, psychological distress, which has been optimized for workers. We used workers' physical activity time (measured using a smartphone app) and the psychological distress state from the previous day as features to monitor their psychological distress on the next day. Additionally, working conditions (information on whether the previous day was a working day or not), occupation, employment status, shift type, and working hours were considered features. The long short-term memory (LSTM) model was used as a deep learning model. We hypothesized that the deep learning model developed using the abovementioned characteristics would have similar or better classification performances than models used in previous studies [8-19]. Additionally, the model is expected to have a strong correlation with the measured levels of psychological distress. This study is the first investigation to develop an optimized model using working conditions and LSTM to predict depression and anxiety in workers, and it could be useful to workers to self-monitor and be a primary prevention from depression and anxiety.

Methods

Study Design and Settings

A 2-week longitudinal study with workers was conducted from November 2021 to April 2022 to measure their daily psychological distress and obtain digital data on their physical activity and working conditions. Participating workers in Japan were recruited from private companies in the Kanto region and the social networking platform Twitter. Recruitments, collection of informed consent, and data collection, which included conducting surveys, were carried out digitally via email. The following were the eligibility criteria for the study participants: (1) working in a public- or private-sector organization, (2) living or working in urban areas, and (3) having a personal smartphone. We excluded workers who were absent at baseline.

A total of 236 workers, aged between 20 and 69 years, participated in this study. Table 1 presents the characteristics of the participants.

Initially, the participants were asked to complete a baseline survey to assess their working conditions. The first page of the baseline survey explained the terms and conditions of the study, which participants had to read and approve before proceeding to the next page. Additionally, they were asked to install the Google Fit app [21,22] (available on both Android and iOS) on their smartphones. During the observational study period, surveys were distributed from 5 to 6 PM daily to measure the psychological distress of the participants and check whether the response was provided on a working or nonworking day. After this 2-week period, the participants exported and sent their Google Fit app data to us.

Table 1. Characteristics of participants at baseline (N=236).

Characteristics	Participants, n (%)
Age (years)	
<20	0 (0)
20-29	40 (16.9)
30-39	72 (30.5)
40-49	64 (27.1)
50-59	53 (22.5)
60-69	7 (3.0)
≥70	0 (0)
Gender	
Male	132 (55.9)
Female	104 (44.1)
Others or not responded	0 (0)
Occupation	
Managers	45 (23.3)
Professions, engineers, or academics	68 (27.9)
Clerks	64 (23.3)
Sales	27 (11.4)
Services	14 (5.9)
Transportation	2 (0.8)
Construction	0 (0)
Production/Skilled	5 (2.1)
Agriculture/Forestry/Fisheries	0 (0)
Others	11 (4.7)
Employment status	
Full-time	202 (85.6)
Part-time	12 (5.1)
Dispatched	3 (1.3)
Contract	10 (4.2)
Others	9 (3.8)
Shift type	
Day shift	223 (94.5)
Rotation shift	7 (2.9)
Night shift	0 (0)
Others	6 (2.5)
Working hours (per week; hours)	
1-34	21 (8.9)
35-40	60 (25.4)
41-50	114 (48.3)
51-60	34 (14.4)
61-65	4 (1.7)
66-70	1 (0.4)
≥71	2 (0.8)

Measurement of Supervised Data

Psychological Distress

In the daily survey, psychological distress of the participants was measured using the Japanese version of the K6 scale [23]. The K6 scale comprises 6 items that assess how often people experience symptoms of depression and anxiety; these items were rated on a 5-point Likert scale (0=none of the times and 4=all the times). The reliability and validity of the Japanese K6 scale was confirmed [23], and the internal consistency in this study was quite high (Cronbach α=.92). In this study, the total score on the K6 scale was considered a psychological distress indicator. Additionally, the participants were divided into 3 groups based on the cutoff scores under the K6 scale: light level (<5), subthreshold level (≥5 and <13), and severe level (≥13) [24,25]. This grouping was used to evaluate the classification performance of the deep learning model.

Physical Activity

Digital data on physical activity were obtained using the Google Fit app. Based on a previous systematic review [8], we adopted activity time by intensity as a physical activity feature. The Google Fit app could store 24 h of physical activity every 15 minutes in accordance with intensity (light and moderate to vigorous). Details of the measurement methodology and reliability of the data are described elsewhere [26]. The app used smartphone sensors or a heart rate monitor to track the physical activities of the workers and categorize the intensity of these activities on the basis of their heart rate. The maximum heart rate was calculated using the following formula: maximum heart rate=205.8–(0.685×age). If the heart rate was ≥50% that of the maximum value, the activity was categorized as vigorous.

Working Conditions

As an additional feature, working conditions were measured under the baseline and daily surveys. In the baseline survey, we obtained data on the age (<20, 20-29, 30-39, 40-49, 50-59, 60-69, and ≥70 years), gender (male, female, others, and ones with no response), type of occupation (management, engineering or education, general office tasks, sales, services, transportation, construction, and production or skill, and agriculture, forestry, or fisheries), employment status (full-time, part-time, dispatched, and contract), work shift types (day shift, rotation, and night shift), and working hours per week (1-34, 35-40, 41-50, 51-60, 61-65, 66-70, and ≥71 hours). In the daily survey, we measured whether the response date was that of a working or nonworking day.

Analysis

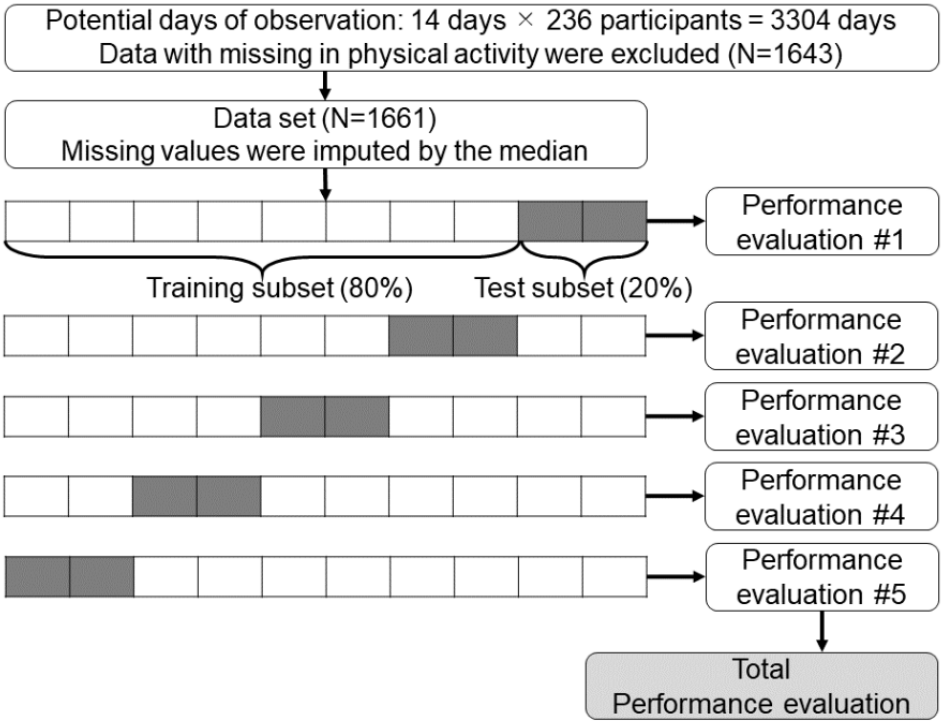
Procedure for Validation of Deep Learners

A deep learning model was developed to predict the log-transformed psychological distress in the participants on the day after their features were collected. Figure 1 shows the procedure for developing, validating, and evaluating the deep learning algorithms. A total of 1661 supervised data were used for validating and evaluating the model, which comprised data from the previous day (physical activity, working conditions, and psychological distress) and data from the next day (psychological distress). Of potential days of observation among 236 participants in the 2 weeks (3304 days), data with missing information on physical activity were excluded because the participants could not obtain their Google Fit app data. Missing values in the daily survey were imputed using medians.

A 5-fold cross-validation method was used to develop and evaluate the deep learning algorithm, using the K-Folds cross-validator in scikit-learn. The daily supervised data were randomly divided into 5 subsets, and each subset was used as the training and test data sets in rotation. Performance evaluation was conducted for each trained model using the test subsets, and the overall model evaluation was calculated as an average of the 5 performance scores.

As the deep learning model, the LSTM model was used, which is a recurrent neural network model that has feedback connections for handling consecutive inputs [27], such as the activity times in the previous periods. The length of the input was set to 96, which was equal to the length of the data on the physical activity of a worker for a single day (24 hours for every 15 minutes). Scores of K6 from the daily survey and information on the working conditions at the baseline survey were expanded to 96 lengths to align with the length of the physical activity. Therefore, the shape of the input data was a 3-rank tensor: the number of observational days (n=1661), 96, and the number of features (n=17). Occupations were one-hot encoded. The output was a log-transformed score on the K6 scale on the next day. The number of hidden layers in the LSTM model was set to 27. Node and recurrent dropout rates were set at 20%. Regarding the activation functions, the rectified linear unit, sigmoid, and linear activations were used as the hidden, recurrent, and output layers, respectively. The Adam optimization algorithm was adopted [28] with a learning rate of 0.0002. The mean squared error was used as the loss function. The batch size was set to 10, and the epoch number was set to 250. To avoid overfitting, early stopping was implemented when the loss values for the current epoch exceeded those for the 10 previous epochs. Data handling was implemented using NumPy (version 1.19.4) [29] and scikit-learn (version 0.23.1) [30]. The model development, training, and validation processes were implemented using Keras (version 2.6.0) [31].

Figure 1. Procedure for developing, validating, and evaluating the deep learning algorithms.



Model Evaluation

As the primary indicator of performance, classification accuracy for the severity of the psychological distress was considered: light level (<5), subthreshold level (≥5 and <13), and severe level (≥13). As secondary indicators, the Pearson correlation coefficient (r) and R^2 values were also calculated between the predicted and measured values for psychological distress.

Model Interpretation

To interpret how the deep learning model predicted and classified the scores and severity of psychological distress, the means of the activity time by intensity were depicted and stratified in accordance with the classification results (light, subthreshold, or severe level) in the training data.

Ethical Considerations

The study protocol was approved by the Kitasato University Medical Ethics Organization (B21-119). Informed consent was obtained before the baseline survey. On the web page, potential participants were asked to read and approve the terms and conditions of the study before proceeding to the baseline survey page. The terms and conditions stated that the research team would protect the privacy and confidentiality of the collected data and that the study data would be deidentified before analyses. The participants did not receive any compensation in this study.

Results

Performance Evaluation of the Deep Learning Model

A total of 1661 days of supervised data were obtained from 236 participants. The mean K6 score was 2.78 (SD 4.27). A total of 131 days of missing values for the K6 scores and the type of day (ie, working or nonworking) were imputed using the median. The median of K6 score was 1, and the median type of day was working. The cross-validation method divided the data into 5 sub–data sets: 4 subsets comprised 332 days of data and the others comprised 333 days of data. Of these, 96.40 (SD 6.0) days (29.02%, SD 1.8%) were nonworking days. For the learning process, 11 to 81 epochs were required to complete the process (Multimedia Appendix 1). The mean squared error values of the training data during the training and validation of the deep learning model gradually declined as the number of epochs increased. The mean score of the mean squared error of the test data was 0.386 (SD 0.06).

Table 2 shows the matrix of the predicted and measured classifications of psychological distress severity in the 5 test sub–data sets. The overall categorization accuracy of the deep learning model was 76.3% (SD 0.04%). The classification accuracies for the severe-, subthreshold-, and light-level psychological distress were 51.1% (SD 0.05%), 60.6% (SD 0.05%), and 81.6% (SD 0.04%), respectively. The correlation coefficient between the predicted and measured values for psychological distress was 0.679 (SD 0.05; R^2=0.463, SD 0.07).

Table 2. Classification performance of the 5 test sub–data sets.

Number of classified data (sub–data sets #1, #2, #3, #4, and #5)	Measured psychological distress (days), n			Total (days), n
	Severe (K6 score≥13)	Subthreshold (K6 score≥5 and <13)	Light (K6 score<5)	
Predicted psychological distress				
Severe (K6 score≥13)				
Sub–data set 1	*8* [a]	3	2	13
Sub–data set 2	*7*	5	0	12
Sub–data set 3	*11*	9	2	22
Sub–data set 4	*5*	7	0	12
Sub–data set 5	*7*	4	1	12
Subthreshold (K6 score≥5 and <13)				
Sub–data set 1	7	*39*	59	105
Sub–data set 2	3	*44*	41	88
Sub–data set 3	5	*32*	43	80
Sub–data set 4	4	*44*	36	84
Sub–data set 5	5	*33*	49	87
Light (K6 score<5)				
Sub–data set 1	3	24	*188*	215
Sub–data set 2	3	21	*208*	232
Sub–data set 3	3	16	*211*	230
Sub–data set 4	2	13	*221*	236
Sub–data set 5	1	22	*210*	233
Total				
Sub–data set 1	18	66	249	333
Sub–data set 2	13	70	249	332
Sub–data set 3	19	57	256	332
Sub–data set 4	11	64	257	332
Sub–data set 5	13	59	260	332

[a]Cells with italicized values indicate accurate classification.

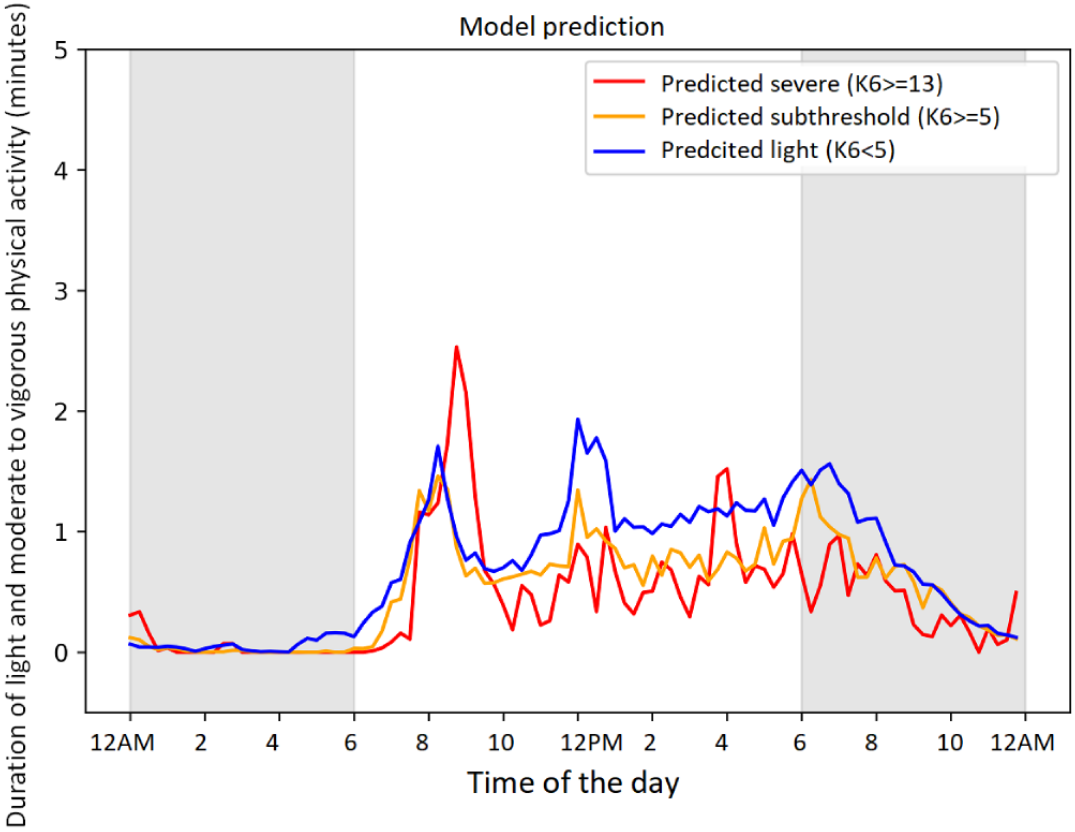
Model Interpretation

Figure 2 shows the mean of the activity time (in minutes) of light and moderate to vigorous physical activity stratified in accordance with the prediction of the psychological distress severity. The developed deep learning model predicted a light-level psychological distress on the next day when the participants had 3 peaks of activity (from 10 to 20 minutes) in a day; that is, in the morning (8 AM), at noon (12 PM), and in the evening (6 PM). The patterns that were predicted as the subthreshold and severe levels of psychological distress followed a similar trend. However, several levels of activity were relatively lower than light distress levels. The pattern predicted for the severe psychological distress level showed a longer activity time in the morning.

The mean of the activity time, which was stratified in accordance with working conditions, showed different patterns based on the prediction of the psychological distress severity. The activity time on a working day showed a similar trend to the overall results. Contrarily, those on a nonworking day exhibited very different results (Multimedia Appendix 2). On the one hand, the activity time patterns predicted for light level psychological distress had a smaller peak for the activity times in the morning and a plateau in the afternoon rather than those on a working day. On the other hand, shifted peaks in the morning (9 AM) and afternoon (4 PM) were predicted as indicators for a severe level of psychological distress the next day. Differences in activity-time patterns were also observed when stratified in accordance with occupations (Multimedia Appendix 3). The 3 predicted peaks for the activity time were similar when predicted as a light level of psychological distress. However, spikes among managers, sales, and service workers were predicted as subthreshold- and severe-level psychological distress. Moreover, numerous peaks were observed during the day for workers engaged in service occupations. Among them, levels of activity were relatively higher when patterns were predicted as the

subthreshold level of psychological distress than those of light-level distress.

Figure 2. Duration of light and moderate to vigorous physical activity (minutes) within a day stratified by the severity of psychological distress.



Discussion

Principal Findings

The deep learning model developed using LSTM, based on the physical activities and working conditions, revealed a similar performance as reported in previous studies [8-19]. The accuracy in classifying the severity of psychological distress was high, indicating a relatively high performance in classifying light-level distress. The predicted values for psychological distress correlated strongly with the measured values, and over 45% of the variance in distress was explained. The results suggest that the performance was maintained using the information on the working conditions and model, which had feedback connections to process a range of information, even when other different digital biomarkers were not considered (eg, sleep, heart rate, sociability, location, and smartphone use). Working conditions and LSTM were useful in maintaining the model's performance for monitoring depression and anxiety by using digitally recorded physical activities in workers.

The prediction and classification performance of the deep learning model was similar to that in a previous study [18] that targeted a working population. A significant advantage of the proposed model was its high accuracy for detecting light-level psychological distress, which implies that the model accurately assessed mentally healthy workers on the basis of the features obtained. In other words, workers can take care of their physical activities to prevent mental disorders before they even occurred, by using the model. This was useful for workers to maintain their physical activities and monitor whether the activity time patterns were good for their mental health. Good patterns showed 3 peaks of 10-20 minutes of physical activity in the morning, during midday, and in the evening. These patterns reflect arrival and departure from work and lunch and closely resembled those previously surveyed among workers in Singapore [18], which suggests that these patterns represent a basic and healthy rhythm among workers. Contrarily, lower levels of physical activity were associated with subthreshold- and severe-level psychological distress. These findings correspond with those of previous studies, which showed that the digital biomarkers of physical activity were predominantly negatively correlated with depression [8]. The patterns classified as severe-level psychological distress showed no peak in the noon and evening; this might reflect low-level activity owing to long working hours or high workloads. The activity patterns reflected workers' rhythm of life, which was influenced not only by their own behaviors but also factors at work.

Interestingly, there were predictive differences in working conditions. Shiftable peaks in the later activities and high levels of activity on nonworking days led to high psychological distress. Additionally, excessive workload peaks among managers, sales, and service workers were predicted to be subthreshold- and severe-level psychological distress, respectively. On the one hand, among managers, physical activity at night (after 8 PM) might be related to the severe-level distress. Service workers had several peaks of physical activity, and higher-level peaks were predicted as subthreshold-level distress. On the other hand, clerks and professionals did not

have a similar trend: lower whole levels of physical activities were associated with subthreshold- and severe-level psychological distress. These differences might depend on the activity levels of their demanding work. Managerial, sales, and service jobs are qualitatively different from clerical and professional jobs and tend to be more physically demanding. Excessive physical activity could affect the psychological distress of workers if the activities at work are physically demanding. These findings may be attributed to poor sleep and circadian rhythms [32-34], long working hours [35,36], or the physical demands of work activity [37]. A previous study showed that nighttime heart rate variations were the most important feature in their machine learning model, and discussed the circadian and rest rhythms as possible underlying mechanisms [18].

The classification performance of the developed model for the severe-level psychological distress was not high. Misclassified samples included more data from rotation and other shift workers (12%) than from the whole sample. Workers who were engaged in shift work had a different rhythm from day-shift workers. Hence, they needed different algorithms to predict the level of psychological distress. This study did not cover much data on shift workers. Consequently, further studies are needed to tune the model.

Limitations

Several limitations limited the validity and generalizability of the study. The lack of numbers and variations in the data monitored led to low generalizability. Particularly, severely distressed night-shift workers were lacking. The results could not be directly compared with those of previous studies because the latter used different measurements and digital biomarkers. Previous studies surveyed students and patients and mainly used the Patient Health Questionnaire as an indicator for depression. Information on sleep, an important digital biomarker used in the previous studies [14,17,18], was measured in this study. Daily information on the working hours and workloads were lacking in this study, which contributed considerably to predicting psychological distress. Some participants did not have their smartphones with them during work hours because of their work preferences or rules, thereby leading to a biased estimate regarding their activity time.

Conclusions

In conclusion, the developed deep learning model performed similarly to those reported in previous studies and had high accuracy in determining light-level psychological distress as a function of physical activity and working conditions. The collected information on the working conditions was useful in passively monitoring the depression and anxiety status of workers. It further contributed in determining the mental health status of workers by using digital biomarkers. The developed model can be used in mobile apps and among workers to self-monitor and maintain their mental health state.

Acknowledgments

This study was supported by the Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (JP20K19671) and the Japan Agency for Medical Research and Development (JP21de0107006). The funders neither played any role in the study design, collection, analysis, and interpretation of data, nor did they participate in the writing of the report and in the decision to submit the article for publication. We would like to thank Cactus Communications Co, Tokyo for the English language editing.

Data Availability

The data sets generated and analyzed in this study are available with the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Mean squared error as a loss function during training in the five test subsets.
[PNG File , 308 KB - formative_v6i11e40339_app1.png]

Multimedia Appendix 2
Duration of light and moderate to vigorous physical activity (minutes) within a day stratified by the severity of psychological distress for working and non-working days.
[PNG File , 768 KB - formative_v6i11e40339_app2.png]

Multimedia Appendix 3
Duration of light and moderate to vigorous physical activity (minutes) within a day stratified by the severity of psychological distress for different occupations.
[PNG File , 2360 KB - formative_v6i11e40339_app3.png]

References

1. Rebar AL, Stanton R, Geard D, Short C, Duncan MJ, Vandelanotte C. A meta-meta-analysis of the effect of physical activity on depression and anxiety in non-clinical adult populations. Health Psychol Rev 2015;9(3):366-378. [doi: 10.1080/17437199.2015.1022901] [Medline: 25739893]
2. Pemberton R, Fuller Tyszkiewicz MD. Factors contributing to depressive mood states in everyday life: a systematic review. J Affect Disord 2016 Aug;200:103-110. [doi: 10.1016/j.jad.2016.04.023] [Medline: 27131503]
3. Church TS, Thomas DM, Tudor-Locke C, Katzmarzyk PT, Earnest CP, Rodarte RQ, et al. Trends over 5 decades in U.S. occupation-related physical activity and their associations with obesity. PLoS One 2011;6(5):e19657 [FREE Full text] [doi: 10.1371/journal.pone.0019657] [Medline: 21647427]
4. Kirk MA, Rhodes RE. Occupation correlates of adults' participation in leisure-time physical activity: a systematic review. Am J Prev Med 2011 Apr;40(4):476-485. [doi: 10.1016/j.amepre.2010.12.015] [Medline: 21406284]
5. Kessler RC, Akiskal HS, Ames M, Birnbaum H, Greenberg P, Hirschfeld RMA, et al. Prevalence and effects of mood disorders on work performance in a nationally representative sample of U.S. workers. Am J Psychiatry 2006 Sep;163(9):1561-1568 [FREE Full text] [doi: 10.1176/ajp.2006.163.9.1561] [Medline: 16946181]
6. Tsuchiya M, Kawakami N, Ono Y, Nakane Y, Nakamura Y, Fukao A, et al. Impact of mental disorders on work performance in a community sample of workers in Japan: the World Mental Health Japan Survey 2002-2005. Psychiatry Res 2012 Jun 30;198(1):140-145. [doi: 10.1016/j.psychres.2011.10.014] [Medline: 22374551]
7. Joyce S, Modini M, Christensen H, Mykletun A, Bryant R, Mitchell PB, et al. Workplace interventions for common mental disorders: a systematic meta-review. Psychol Med 2016 Mar;46(4):683-697. [doi: 10.1017/S0033291715002408] [Medline: 26620157]
8. De Angel V, Lewis S, White K, Oetzmann C, Leightley D, Oprea E, et al. Digital health tools for the passive monitoring of depression: a systematic review of methods. NPJ Digit Med 2022 Jan 11;5(1):3 [FREE Full text] [doi: 10.1038/s41746-021-00548-8] [Medline: 35017634]
9. Sano A, Taylor S, McHill AW, Phillips AJ, Barger LK, Klerman E, et al. Identifying objective physiological markers and modifiable behaviors for self-reported stress and mental health status using wearable sensors and mobile phones: observational study. J Med Internet Res 2018 Jun 08;20(6):e210 [FREE Full text] [doi: 10.2196/jmir.9410] [Medline: 29884610]
10. Yue C, Ware S, Morillo R, Lu J, Shang C, Bi J, et al. Fusing location data for depression prediction. IEEE Trans Big Data 2021 Jun;7(2):355-370 [FREE Full text] [doi: 10.1109/tbdata.2018.2872569] [Medline: 35498556]
11. Wahle F, Kowatsch T, Fleisch E, Rufer M, Weidt S. Mobile sensing and support for people with depression: a pilot trial in the wild. JMIR Mhealth Uhealth 2016 Sep 21;4(3):e111 [FREE Full text] [doi: 10.2196/mhealth.5960] [Medline: 27655245]
12. Pratap A, Atkins DC, Renn BN, Tanana MJ, Mooney SD, Anguera JA, et al. The accuracy of passive phone sensors in predicting daily mood. Depress Anxiety 2019 Jan;36(1):72-81 [FREE Full text] [doi: 10.1002/da.22822] [Medline: 30129691]
13. Wang R, Wang W, daSilva A, Huckins JF, Kelley WM, Heatherton TF, et al. Tracking depression dynamics in college students using mobile phone and wearable sensing. In: Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies. New York, NY: Association for Computing Machinery; Mar 26, 2018:1-26.
14. Lu J, Shang C, Yue C, Morillo R, Ware S, Kamath J, et al. Joint modeling of heterogeneous sensing data for depression assessment via multi-task learning. In: Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies. New York, NY: Association for Computing Machinery; Mar 26, 2018:1-21.
15. Burns MN, Begale M, Duffecy J, Gergle D, Karr CJ, Giangrande E, et al. Harnessing context sensing to develop a mobile intervention for depression. J Med Internet Res 2011 Aug 12;13(3):e55 [FREE Full text] [doi: 10.2196/jmir.1838] [Medline: 21840837]
16. Jacobson NC, Weingarden H, Wilhelm S. Using digital phenotyping to accurately detect depression severity. J Nerv Ment Dis 2019 Oct;207(10):893-896. [doi: 10.1097/NMD.0000000000001042] [Medline: 31596769]
17. Ghandeharioun A, Fedor S, Sangermano L, Ionescu D, Alpert J, Dale C, et al. Objective assessment of depressive symptoms with machine learning and wearable sensors data. 2017 Presented at: 2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII); October 23-26, 2017; San Antonio, TX. [doi: 10.1109/acii.2017.8273620]
18. Rykov Y, Thach T, Bojic I, Christopoulos G, Car J. Digital biomarkers for depression screening with wearable devices: cross-sectional study with machine learning modeling. JMIR Mhealth Uhealth 2021 Oct 25;9(10):e24872 [FREE Full text] [doi: 10.2196/24872] [Medline: 34694233]
19. Meyerhoff J, Liu T, Kording KP, Ungar LH, Kaiser SM, Karr CJ, et al. Evaluation of changes in depression, anxiety, and social anxiety using smartphone sensor features: longitudinal cohort study. J Med Internet Res 2021 Sep 03;23(9):e22844 [FREE Full text] [doi: 10.2196/22844] [Medline: 34477562]
20. Gu JK, Charles LE, Ma CC, Andrew ME, Fekedulegn D, Hartley TA, et al. Prevalence and trends of leisure-time physical activity by occupation and industry in U.S. workers: the National Health Interview Survey 2004-2014. Ann Epidemiol 2016 Oct;26(10):685-692 [FREE Full text] [doi: 10.1016/j.annepidem.2016.08.004] [Medline: 27659584]
21. Google Fit: activity tracker (Android). URL: https://play.google.com/store/apps/details?id=com.google.android.apps.fitness&hl=ja&gl=US [accessed 2022-05-20]

22. Google Fit: activity tracker (iOS). URL: https://apps.apple.com/jp/app/google-fit-%E3%83%88%E3%83%A9%E3%83%83%E3%82%AB%E3%83%BC/id1433864494 [accessed 2022-05-20]
23. Furukawa TA, Kawakami N, Saitoh M, Ono Y, Nakane Y, Nakamura Y, et al. The performance of the Japanese version of the K6 and K10 in the World Mental Health Survey Japan. Int J Methods Psychiatr Res 2008;17(3):152-158 [FREE Full text] [doi: 10.1002/mpr.257] [Medline: 18763695]
24. Kessler RC, Barker PR, Colpe LJ, Epstein JF, Gfroerer JC, Hiripi E, et al. Screening for serious mental illness in the general population. Arch Gen Psychiatry 2003 Feb;60(2):184-189. [doi: 10.1001/archpsyc.60.2.184] [Medline: 12578436]
25. Sakurai K, Nishi A, Kondo K, Yanagida K, Kawakami N. Screening performance of K6/K10 and other screening instruments for mood and anxiety disorders in Japan. Psychiatry Clin Neurosci 2011 Aug;65(5):434-441 [FREE Full text] [doi: 10.1111/j.1440-1819.2011.02236.x] [Medline: 21851452]
26. Earn "Heart Points" (strong exercise) to stay healthy. Google Fit Help. URL: https://tinyurl.com/nd2k9dnc [accessed 2022-05-20]
27. Sherstinsky A. Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. Physica D: Nonlinear Phenomena 2020 Mar;404:132306 [FREE Full text] [doi: 10.1016/j.physd.2019.132306]
28. Kingma D, Ba J. Adam: a method for stochastic optimization. arXiv. Preprint posted online December 22, 2014. [FREE Full text]
29. NumPy documentation. NumPy. URL: https://docs.scipy.org/doc/numpy/index.html [accessed 2022-05-20]
30. scikit-learn: machine learning in Python. scikit learn. URL: https://scikit-learn.org/stable/ [accessed 2022-05-20]
31. Keras. URL: https://keras.io/ [accessed 2022-05-20]
32. Baglioni C, Battagliese G, Feige B, Spiegelhalder K, Nissen C, Voderholzer U, et al. Insomnia as a predictor of depression: a meta-analytic evaluation of longitudinal epidemiological studies. J Affect Disord 2011 Dec;135(1-3):10-19. [doi: 10.1016/j.jad.2011.01.011] [Medline: 21300408]
33. Germain A, Kupfer DJ. Circadian rhythm disturbances in depression. Hum Psychopharmacol 2008 Oct;23(7):571-585 [FREE Full text] [doi: 10.1002/hup.964] [Medline: 18680211]
34. Kronfeld-Schor N, Einat H. Circadian rhythms and depression: human psychopathology and animal models. Neuropharmacology 2012 Jan;62(1):101-114. [doi: 10.1016/j.neuropharm.2011.08.020] [Medline: 21871466]
35. Watanabe K, Imamura K, Kawakami N. Working hours and the onset of depressive disorder: a systematic review and meta-analysis. Occup Environ Med 2016 Dec;73(12):877-884. [doi: 10.1136/oemed-2016-103845] [Medline: 27852890]
36. Virtanen M, Jokela M, Madsen IE, Magnusson Hanson LL, Lallukka T, Nyberg ST, et al. Long working hours and depressive symptoms: systematic review and meta-analysis of published studies and unpublished individual participant data. Scand J Work Environ Health 2018 May 01;44(3):239-250 [FREE Full text] [doi: 10.5271/sjweh.3712] [Medline: 29423526]
37. Theorell T, Hammarström A, Aronsson G, Träskman Bendz L, Grape T, Hogstedt C, et al. A systematic review including meta-analysis of work environment and depressive symptoms. BMC Public Health 2015 Aug 01;15:738 [FREE Full text] [doi: 10.1186/s12889-015-1954-4] [Medline: 26232123]

Abbreviations

LSTM: long short-term memory

Edited by A Mavragani; submitted 16.06.22; peer-reviewed by I Bojic, T Yamauchi; comments to author 18.10.22; revised version received 20.11.22; accepted 22.11.22; published 30.11.22.

Please cite as:
Watanabe K, Tsutsumi A
The Passive Monitoring of Depression and Anxiety Among Workers Using Digital Biomarkers Based on Their Physical Activity and Working Conditions: 2-Week Longitudinal Study
JMIR Form Res 2022;6(11):e40339
URL: https://formative.jmir.org/2022/11/e40339
doi:10.2196/40339
PMID:36449342

Original Paper

Web-Based Health Information–Seeking Methods and Time Since Provider Engagement: Cross-sectional Study

Eden Shaveet[1], MS; Catherine Urquhart[2], BA; Marissa Gallegos[1], BS; Olaf Dammann[1,3], MD, PhD; Laura Corlin[1,4], PhD

[1]Department of Public Health and Community Medicine, Tufts University School of Medicine, Boston, MA, United States

[2]Department of Community Health, Tufts University School of Arts and Sciences, Medford, MA, United States

[3]Department of Gynecology and Obstetrics, Hannover Medical School, Hannover, Germany

[4]Department of Civil and Environmental Engineering, Tufts University School of Engineering, Medford, MA, United States

Corresponding Author:
Eden Shaveet, MS
Department of Public Health and Community Medicine, Tufts University School of Medicine
136 Harrison Ave
Boston, MA, 02111
United States
Phone: 1 617 636 7000
Email: eden.shaveet@tufts.edu

Abstract

Background: The use of web-based methods to seek health information is increasing in popularity. As web-based health information (WHI)–seeking affects health-related decision support and chronic symptom self-management, WHI-seeking from online sources may impact health care decisions and outcomes, including care-seeking decisions. Patients who are routinely connected to physicians are more likely to receive better and more consistent care. Little is known about whether WHI-seeking impacts the frequency at which patients engage with health care providers.

Objective: Our primary objective was to describe the associations between the use of web-based methods to seek information about one's own health and the time since last engaging with a health care provider about one's own health. Additionally, we aimed to assess participants' trust in health care organizations to contextualize our findings.

Methods: We analyzed data from US adults participating in the nationally representative Tufts Equity in Health, Wealth, and Civic Engagement Survey (N=1034). Bivariate associations between demographic characteristics and health information–seeking methods were assessed with Pearson chi-squared tests. Bivariate associations of Medical Mistrust Index (MMI) scores with each health information–seeking method and time since provider engagement were assessed with *F* tests and adjusted Wald tests. We fit a multivariable logistic regression model to assess the association between WHI-seeking within the 12 months prior to survey (alone or in combination with provider-based methods versus provider only) and engagement with a provider more than 1 year prior to the time of survey, adjusting for age, race and ethnicity, sex, education, insurance coverage, and MMI.

Results: Age, race and ethnicity, educational attainment, health insurance source, MMI, and time since provider engagement were each significantly associated with the health information–seeking method in bivariate analyses. Compared to using only provider-based health information seeking methods, WHI-based methods alone or in combination with provider-based methods were associated with a 51% lower likelihood (odds ratio 0.49, 95% CI 0.27-0.87) of engaging with a provider within the previous year. Participants who used WHI-seeking methods alone and those who had not engaged with a health care provider within the previous year demonstrated a higher mean MMI score; however, MMI was not a significant predictor of time since engagement with a provider in the multivariable analysis.

Conclusions: Our findings from a nationally representative survey suggest that for those who use WHI-seeking methods (alone or in combination with provider-based information-seeking methods), there is a statistically significant lower likelihood of engaging with a provider in a year compared to those who only use provider-based methods. Future research should consider the intent of a person's visit with a provider, trust in health care systems, methods of provider engagement, and specific web-based platforms for health information.

(JMIR Form Res 2022;6(11):e42126) doi:10.2196/42126

KEYWORDS

internet; social media; information-seeking behavior; consumer health information; physician-patient relations; trust

Introduction

Routine engagement with health care providers is vital for the early detection and treatment of disease, as well as for preventive care delivery [1,2]. Most often achieved through interaction with primary care services, routine provider engagement is essential for the maintenance of individual and community health [3]. Patients who are routinely connected to physicians are more likely to receive better and more consistent care [3,4].

Barriers to accessing routine or specialty health care, such as lack of health insurance coverage, are associated with poor health outcomes [5,6]. It has been suggested that such barriers may prompt online health information–seeking behaviors for self-management [7].

Health information–seeking methods are the means by which individuals acquire information about their health, health promotion, health risks, and illness [8]. Use of internet-enabled technologies to seek health information, often referred to as web-based health information (WHI)–seeking methods, has increased exponentially across US populations since the early 2000s [9] and notably during the COVID-19 pandemic, which began at the close of 2019 [10]. Use of web-based sources to obtain health information may have considerable effects on individuals' health care decisions and outcomes [9]. These effects may include encouraging use of ambulatory services, informing decisions to self-manage symptoms, and influencing attitudes toward a disease, treatment, or procedure [9,11,12].

Prior work investigating the impact of WHI-seeking on patient-physician relationships suggested that WHI-seeking may improve relationships by promoting meaningful patient participation in conversation during appointments [13]. Other work examining the impact of WHI-seeking on patient treatment compliance concluded that encouraging patients to seek health information relating to their treatment online may improve overall compliance with recommendations [14]. However, contrary to work positing that WHI-seeking enables meaningful patient engagement with health care providers [13], other work has suggested that participation in online forums may promote mistrust in health services [15].

In considering the role of WHI-seeking on patient-provider relationships and patient well-being, it can be informative to understand how this affects the frequency at which patients engage with health care providers. Prior research suggests that WHI-seeking may lead to more frequent visits with physicians and that trust in information from health care providers may moderate this association [16,17]. Nevertheless, as WHI-seeking behaviors have increased, little research has examined its role in patient engagement with health care providers [18]. This paper is an attempt to fill this gap in the literature. Specifically, using nationally representative data, we seek to explore and generate hypotheses regarding the association between WHI-seeking behavior and length of time since last engaging with a health care provider about ones' own health.

Methods

Study Sample

Participants were recruited via a random sampling of telephone numbers and residential addresses by Ipsos, a multinational market research and social science company. As described previously, data were collected via the second deployment of the Equity in Health, Wealth, and Civic Engagement Survey designed by Tufts University [19]. The survey was fielded in English and Spanish, and was deployed between April 23, 2021, and May 3, 2021, through the Ipsos KnowledgePanel, an online, nationally representative, probability-based panel used to retrieve insights from US adults. Upon accepting the initial invitation to join the KnowledgePanel, respondents were asked to complete a short demographic survey prior to becoming active KnowledgePanel members. Eligible participants for the Equity in Health, Wealth, and Civic Engagement Survey were noninstitutionalized US adults aged 18 years and older who were proficient in English or Spanish languages. Those without access to the internet were provided a laptop and internet access by Ipsos at no cost.

Ethics Approval

All study protocols were reviewed and approved by the Social, Behavioral, and Educational Research Institutional Review Board at Tufts University, Boston, MA (protocol STUDY00000428). All participants provided informed consent to participate in the Ipsos KnowledgePanel and to complete the Tufts Equity in Health, Wealth, and Civic Engagement Survey. Participants received standard incentive payments upon survey completion (eg, 1000 points, the cash-equivalent of US $1 and an entry into the KnowledgePanel sweepstakes for completing a survey that takes longer than 15 minutes; median completion time=15 minutes). Participant responses were deidentified, and precautions were taken to ensure confidentiality and privacy for participants (eg, storing data only on secure drives) [19].

Procedures

Of the 2107 KnowledgePanel participants invited to complete the survey, 1449 (68.77%) responded to the Tufts Equity in Health, Wealth, and Civic Engagement Survey. Participants included in our analysis (n=1034) were those who met both of the following inclusion criteria: (1) provided information about the time elapsed since they had last engaged with a health care provider about their own health, and (2) indicated they had used web-based or health care provider–based methods to seek health information in the prior year, as described in our Measures subsection.

Measures

Demographic Variables

Demographic and contextual items incorporated in our analyses included self-reported age (18-24, 25-34, 35-44, 45-54, 55-64, 65-75, or >75 years), race and ethnicity (Hispanic, non-Hispanic Black, non-Hispanic White, or non-Hispanic multiracial/other),

sex (female or male), educational attainment (less than high school diploma, high school diploma/General Educational Development, some college/associate's degree, bachelor's degree, or master's degree/higher), annual household income (<US $10,000, US $10,000-24,999, US $25,000-49,000, US $50,000-74,999, US $75,000-99,999, US $100,000-149,999, or ≥ US $150,000), insurance coverage (no insurance, employer-sponsored, government-sponsored [including Medicare, Medicaid, and military], health insurance marketplace, or other), and household internet access (defined as 1 or more members of a participant's household, including themselves, having access to the internet).

Health Information–Seeking Variables

Our health information–seeking variables were derived from a survey item ("Have you used any of the following sources for health information in the past 12 months?") in which participants could select all responses that applied ("Doctor," "Pharmacist," "Nurse, nurse practitioner or physician's assistant," "Relative, friend, or co-worker," "Someone you know who has a particular medical condition," "Disease-related association or society," "Patient support group or foundation," "Educational forum at a local clinic, hospital, community center or other location," "Pharmaceutical company," "Health insurance company," "Newspapers or magazines," "Television," "The internet," "Social Media [such as Facebook, Twitter]", "Healthcare app for smartphone or tablet"). Affirmative responses of "The internet," "Social media," or "Healthcare app" were combined into a composite variable indicating the use of WHI-seeking methods. Affirmative responses of "Doctor," or "Nurse, nurse practitioner or physician's assistant" were combined into a composite variable indicating provider-based health information–seeking.

Time-Since-Provider-Engagement Variables

We developed a dichotomous variable for time since health care provider engagement based on participant responses to "How long has it been since you last saw or talked to a doctor or other healthcare professional about your own health?" Response options included "6 months or less," "More than 6 months, but not more than 1 year ago," "More than 1 year, but not more than 2 years ago," "More than 2 years, but not more than 5 years ago," "More than 5 years ago," and "Never." We chose to make this variable dichotomous (≤1 year or >1 year ago) for sample size reasons, but the full distribution of responses is shown in Multimedia Appendix 1, Table S1.

Medical Mistrust Index Variables

To aid in the interpretation of our descriptive results, we examined scores on the Medical Mistrust Index (MMI) developed to assess a participant's trust in health care organizations [20]. The MMI is a 7-item index offered on a 5-point Likert scale including the following response values: "strongly disagree," "moderately disagree," "neutral," "moderately agree," and "strongly agree" [20]. We coded responses from 0 ("strongly disagree") to 4 ("strongly agree"). Total scores could be between 0 and 28, with higher values reflecting greater perceived mistrust of health care organizations. The seven index items include the following: (1) "You'd better be cautious when dealing with healthcare organizations."(2)"Patients have sometimes been deceived or misled by healthcare organizations." (3) "When healthcare organizations make mistakes, they usually cover it up." (4) "Healthcare organizations have sometimes done harmful experiments on patients without their knowledge." (5) "Healthcare organizations don't always keep your information totally private."(6) "Sometimes I wonder if healthcare organizations really know what they are doing." (7) "Mistakes are common in healthcare organizations."

Analysis

We examined demographic characteristics of participants and MMI scores overall by category of health information–seeking method (WHI only, provider-based only, both) and by time since last provider engagement. Bivariate associations between demographic characteristics and health information–seeking method were assessed with Pearson chi-squared tests. Bivariate associations of MMI with each health information–seeking method and time since provider engagement were assessed with *F* tests (for overall associations) and adjusted Wald tests (for pairwise comparisons). These tests were selected based on their ability to support survey weighting. We fit a multivariable logistic regression model to assess the association between WHI-seeking within the 12 months prior to survey (alone or in combination with provider-based methods versus provider only) and engagement with a provider more than 1 year prior to the time of survey, adjusting for age, race and ethnicity, sex, education, insurance coverage (dichotomous; insured or uninsured), and MMI. We chose a multivariable logistic regression model to accommodate our dichotomous outcome and adjust for relevant covariates. These covariates were chosen based on significant bivariate associations with the primary exposure or outcome. All analyses were conducted using Stata 17 (StataCorp) and R 4.1.2 (The R Foundation for Statistical Computing) and applied sample weights to be more representative of the US population based on the US Census Bureau's 2019 current population estimates [19]. Sample weights varied from 0.131 to 4.827 with a median of 0.826 for the full sample.

Results

Sample Characteristics

Sample characteristics are described in Table 1. Briefly, most (70.85%) participants were between 25 and 64 years old. Most were non-Hispanic White (63.19%) and approximately half (53.47%) were female. Most participants (68.41%) had more than a high school level of education, and most (56.06%) had an annual household income of at least US $75,000. Nearly all participants (95.42%) had health insurance, and nearly all (99.07%) had internet access. Nearly half of participants (47.20%) used both WHI and provider-based seeking methods, with only 16.96% using only WHI-based methods. Whereas 87.09% of participants had engaged with a health care provider in the previous year, only 83.04% of participants reported provider-based or web- and provider-based information-seeking methods.

Table 1. Sample characteristics overall, by health information–seeking method, and by time since provider engagement.

Characteristic	Overall unweighted n (weighted %)	Health information–seeking method (weighted row %)				Time since provider engagement (weighted row %)		
		Web-based only	Provider-based only	Web- and provider-based	*P* value[a]	≤1 year	>1 year	*P* value[a]
Overall	1034 (100)	16.96	35.84	47.20		87.09	12.91	
Age in years (N=1034)					.005			<.001
18-24	33 (6.15)	17.33	14.46	68.21		80.29	19.71	
25-34	122 (18.01)	17.25	35.07	47.68		78.16	21.84	
35-44	157 (16.87)	19.55	31.65	48.80		81.21	18.79	
45-54	145 (14.19)	19.46	34.31	46.23		85.99	14.01	
55-64	255 (21.78)	21.02	34.95	44.03		92.45	7.55	
65-74	207 (14.12)	10.15	40.68	49.17		95.39	4.61	
≥75	115 (8.88)	8.02	57.13	34.85		96.49	3.51	
Race and ethnicity (N=1034)					.001			.001
White, non-Hispanic	505 (63.19)	14.31	36.57	49.12		90.65	9.35	
Black, non-Hispanic	235 (11.18)	18.09	42.00	39.91		89.61	10.39	
Hispanic	231 (15.22)	28.07	37.11	34.81		78.14	21.86	
Other or 2+ races, non-Hispanic	63 (10.42)	15.57	22.94	61.48		75.88	24.12	
Sex (N=1034)					.16			.003
Female	518 (53.47)	15.35	34.19	50.46		90.99	9.01	
Male	516 (46.53)	18.80	37.74	43.46		82.61	17.39	
Education (N=1034)					.001			.16
Less than a high school diploma	90 (8.76)	29.68	33.34	36.98		81.67	18.33	
High school diploma/GED[b]	258 (22.83)	21.28	41.61	37.11		83.84	16.16	
Some college or associate's degree	282 (31.40)	15.04	38.16	46.80		90.93	9.07	
Bachelor's degree	230 (21.44)	13.01	34.84	52.14		84.65	15.35	
Master's degree or higher	174 (15.58)	12.75	25.51	61.74		90.53	9.47	
Annual household income (N=1034)					.34			.63
<US $10,000	26 (2.17)	13.05	31.72	55.22		91.56	8.44	
US $10,000-24,999	93 (8.83)	19.87	38.80	41.33		86.32	13.68	
US $25,000-$49,999	177 (14.89)	23.58	30.88	45.54		88.67	11.33	
US $50,000-74,999	173 (18.05)	17.00	33.30	49.69		84.22	15.78	
US $75,000-99,999	159 (14.69)	20.07	36.32	43.61		82.95	17.05	
US $100,000-149,999	203 (19.59)	14.01	42.68	43.31		90.07	9.93	
≥US $150,000	203 (21.78)	12.15	34.08	53.77		88.38	11.62	
Health insurance source (N=1031)					<.001			<.001
No insurance	54 (4.58)	52.74	20.74	26.51		46.50	53.50	
Employer	546 (57.39)	15.09	34.65	50.26		87.71	12.29	
Government[c]	346 (30.02)	10.67	41.91	47.42		93.71	6.29	
Health insurance marketplace	48 (4.26)	40.41	24.67	34.92		76.17	23.83	
Other source	37 (3.74)	28.02	29.05	42.92		84.75	15.25	
Internet access[d] (N=1029)					.37			.74
Yes	1017 (99.07)	16.97	35.48	47.56		87.14	12.86	

Characteristic	Overall unweighted n (weighted %)	Health information–seeking method (weighted row %)				Time since provider engagement (weighted row %)		
		Web-based only	Provider-based only	Web- and provider-based	*P* value[a]	≤1 year	>1 year	*P* value[a]
No	12 (0.93)	15.60	57.81	26.60		83.78	16.22	
Medical Mistrust Index (N = 1034)[e]	14.6 (14.1-15.1)	16.4 (15.3-17.5)	14.3 (13.5-15.2)	14.2 (13.5-14.9)	.002	14.4 (13.9-14.9)	16.1 (14.6-17.6)	.04
Time since provider engagement (N=1034)					<.001			
≤1 year	910 (87.09)	12.41	38.12	49.47				
>1 year	124 (12.91)	47.61	20.51	31.88				
Health information–seeking method (N=1034)								<.001
Web-based only	190 (16.96)	N/A[f]	N/A	N/A		63.75	36.25	
Provider-based only	388 (35.84)	N/A	N/A	N/A		92.61	7.39	
Web- and provider-based	456 (47.20)	N/A	N/A	N/A		91.28	8.72	

[a]*P* value for Pearson chi-squared test or *F* test (for Medical Mistrust Index).

[b]GED: General Education Development.

[c]Includes Medicare, Medicaid, and military insurance.

[d]Defined as 1 or more members of a participant's household having access to the internet.

[e]Values in this row are reported as mean (95% CIs). *P* values are for *F* test.

[f]N/A: not applicable. Corresponding rows and columns refer to the same groups.

Bivariate Associations

Age, race and ethnicity, educational attainment, health insurance source, MMI, and time since provider engagement were each significantly associated with the health information–seeking methods in bivariate analyses (Table 1). For example, MMI scores were significantly higher for people who used WHI-based methods alone compared to either provider-based methods alone (*P*=.003) or both WHI- and provider-based methods (*P*=.001). Similarly, age, race and ethnicity, sex, MMI, and health insurance source were each significantly associated with time since provider engagement in bivariate analyses (Table 1). For example, MMI scores were significantly higher for people who had not engaged with a provider in the previous year (*P*=.04). However, among those who had used WHI-based methods (alone or in combination with provider-based methods; n=646), there was no significant difference in mean MMI scores (*P*=.08).

Multivariable Associations

Compared to using only provider-based health information–seeking methods, using WHI-based methods alone or in combination with provider-based methods was associated with a 51% lower likelihood (odds ratio 0.49, 95% CI 0.27-0.87) of engaging with a provider within the previous year (Table 2). Being female or insured was associated with an increased likelihood of engaging with a provider within the previous year, whereas identifying as non-Hispanic and more than 2 races or a race other than White or Black was associated with a lower likelihood (Table 2).

Table 2. Association between web-based health information–seeking within the prior year and engagement with a provider more than 1 year prior (N=1031).

	Engagement with a provider within the past year (referent: >1 year) odds ratio (95% CI)
Health information–seeking (referent: only provider-based information-seeking)	
Web-based only or web-based and provider-based	0.49 (0.27, 0.87)[a]
Age group (referent: 18-24 years)	
25-34 years	0.66 (0.18-2.45)
35-44 years	0.85 (0.25-2.92)
45-54 years	1.14 (0.31-4.24)
55-64 years	2.11 (0.62-7.18)
65-74 years	2.87 (0.79-10.41)
≥75 years	3.60 (0.82-15.89)
Race and ethnicity (referent: non-Hispanic White)	
Black, non-Hispanic	1.33 (0.65-2.73)
Hispanic	0.63 (0.35-1.12)
Other or 2+ races, non-Hispanic	0.36 (0.16-0.84)[a]
Sex (referent: male)	
Female	2.28 (1.32-3.92)[a]
Education (referent: less than high school diploma)	
High school diploma/GED[b]	0.63 (0.28-1.45)
Some college or associate's degree	1.21 (0.51-2.90)
Bachelor's degree	0.84 (0.34-2.08)
Master's degree or higher	1.50 (0.54-4.19)
Insurance coverage (referent: uninsured)	
Insured	5.73 (2.71-12.15)[a]
Medical Mistrust Index	0.97 (0.93-1.01)

[a]*P*<.05.

[b]GED: General Education Development.

Discussion

Principal Results

Our nationally representative cross-sectional study sought to further the discussion about how the use of WHI-seeking methods relate to the time duration since last engaging with a health care provider about one's own health. We observed that for those who use WHI-seeking methods (alone or in combination with provider-based information-seeking methods), there is a lower likelihood of engaging with a provider in a year compared to those who only use provider-based methods. Given that more than half of US adults use the internet as their primary source of health information [18], our findings, paired with literature suggesting a decline in the frequency at which commercially insured US adults receive primary care [21], point to potentially novel shifts in patient engagement for individuals who seek WHI. These interpretations are consistent with prior work suggesting that WHI-seeking may influence a person's medical treatment decisions, including whether to visit a physician or not [22]. These interpretations are also consistent with prior work suggesting that WHI-seeking influences patients' trust in health care providers [23-25], which is in turn associated with the frequency at which patients engage with their providers [26,27]. These findings differ, however, from previous work suggesting that WHI-seeking leads to more frequent visits with physicians [16] and that this effect is larger for those who exhibit lower trust in information offered by health care providers [17].

Medical mistrust remains a complex challenge that inhibits access to care by dissuading utilization of and participation in health services [20]. Recognized as a social determinant of health driving health disparities for marginalized groups [28], medical mistrust reflects the influence of both health misinformation and longstanding struggles to restore trust following historical medical misdeeds and mistakes [29]. Prior work suggests that some WHI-seeking methods may promote patient mistrust in health services by virtue of repeat false or

misleading portrayals of health care, such as through online forums or social media posts [15,30].

In our study, statistically significant differences were observed in mean MMI scores by health information method used in the previous year and with time since engagement with a health care provider, with those who used WHI-seeking methods alone and those who had not engaged with a health care provider within the previous year demonstrating a higher mean MMI score. Whereas this may suggest a relationship between use of WHI-seeking methods alone and higher perceived mistrust of health care organizations, MMI was not a significant predictor of time since engagement with a provider in the multivariable analysis. Additional work is needed to identify the role of MMI in the relationship between WHI-seeking and health care provider engagement. Nuances in these relationships—such as differences by specific web-based platforms—may be especially salient given literature suggesting that the availability and use of WHI by patients can be beneficial to patient-provider relationships [13,14,23]. For example, several studies assert that WHI-seeking may promote health literacy among patients and improve communication with providers [13,23]. Others find that physician-encouraged WHI-seeking by patients may improve patient compliance with treatment recommendations [14]. Some works speculate that the impact of WHI-seeking on patient-provider relationships is contingent on several factors, including the quality of WHI retrieved by patients [23,31], the willingness of providers to discuss WHI brought forth by a patient [32], a provider's reactions to a patient's presentation of WHI with respect to their treatment [13,33], and whether the WHI serves to complement or challenge a provider's medical expertise [34].

Lack of health insurance coverage, a well-documented barrier to care, was also a significant predictor of time since engagement with a health care provider. Whereas the plurality of respondents in our sample used a combination of web-based and provider-based methods to seek health information in the previous year, over half of respondents without insurance used web-based methods only and over half had not engaged with a health care provider in the previous year. These findings suggest that barriers to routine care, such as lack of insurance or other financial constraints, might prompt the use of WHI-seeking methods for condition self-management [7,35].

Limitations and Strengths

Our study had several limitations. First, our cross-sectional survey was fielded during the COVID-19 pandemic, and the results may not be generalizable beyond this time frame. Disruption of routine health care service delivery and escalation of novel service delivery methods throughout the pandemic, such as the increased use of telehealth [36], impacted the frequency and means by which patients engaged with health care providers overall [37]. As respondents were not asked about how they engaged with health care providers, this study does not explicitly specify whether the 87.09% of respondents who engaged with providers used telehealth services or engaged with providers face to face (which could explain the counterintuitive statistic in our sample that 83.04% engaged with provider-based or web- and provider-based information-seeking methods). It is also possible that respondents interpreted response options of "the internet" or "Healthcare app" to include use of telehealth services or patient portals. Second, there might have been social desirability bias in terms due to the self-reported nature of when participants last engaged with health care providers. Third, our sample was limited to individuals who sought WHI or information from provider-based sources in the 12 months prior to the survey and excluded those who only sought health information from other sources during this time frame. Fourth, although participants without access to the internet were provided a laptop and internet connection by Ipsos at no cost, this alone does not guarantee that the sample is adequately representative of those without internet access. Similarly, respondents were recruited via random sampling of telephone numbers and residential addresses. Thus, those without telephone numbers or residential addresses may be underrepresented in our sample. These issues and others might have affected the ability of our survey to be truly nationally representative. Finally, our sample size was relatively small, and we therefore dichotomized certain variables, such as time since provider engagement. We might have lost nuances in trends because of this decision, and our multivariable regression model might have been underpowered.

Despite these limitations, our study presents several strengths, including its sampling methodology coupled with use of sample weights to maximize its representativeness of the US population. Collection of information about health information–seeking methods and engagement with health care providers at this scale offers novel insight and corroborates previously suggested relationships between WHI-seeking behaviors and health care service utilization. It is important to add this current evidence, as patients' relationships with the internet as a source of health information evolves as societal norms and ways of accessing information change. Our work also provides novel insight into several characteristics that had not been investigated together in relation to WHI-seeking and time since provider engagement, such as insurance coverage. In addition, we were able to assess the role of medical mistrust, a potentially critical factor in the relationship between health information–seeking methods and engagement with health care providers. These findings present key insights for an emerging area of research and offer directions for inquiry in future research.

Implications and Future Research

The association between use of WHI-seeking methods and time since provider engagement suggests questions for future research. For example, an investigation could be conducted to identify associations between use of WHI-seeking methods and time since provider engagement by service rendered (eg, routine or preventive versus emergency). Whereas delays in provider engagement for routine or preventive care are potentially detrimental to individual and community health, reducing unnecessary emergency provider visits and associated per capita costs is desirable for patients and health systems [38]. Additionally, although the use of WHI-seeking methods has demonstrable benefits [13,14,39,40], concerns arise around exposure to online health misinformation [41]. Misinformation and disinformation around health have received increased public attention in recent years due to the negative impact they have

had on individual and public health, notably during the COVID-19 pandemic [42]. Although the popularity of online information and social media platforms has contributed to the prevalence of online misinformation [43], it is unknown whether those who use the internet as a primary source of health information are disproportionately exposed to health misinformation. Future research should investigate whether exposure to health misinformation or other factors underlie the associations we observed between the method of health information–seeking and time since provider engagement.

Conclusions

Our findings from a nationally representative survey suggest that for those who use WHI-seeking methods (alone or in combination with provider-based information-seeking methods), there is a statistically significant lower likelihood of engaging with a provider in a year compared to those who only use provider-based methods. Future research should consider the intent of a person's visit with a provider, trust in health care systems, methods of provider engagement, and specific web-based platforms for health information.

Acknowledgments

The authors thank the members of the Tufts Research Group on Equity in Health, Wealth, and Civic Engagement for providing the data set used in this research. The authors also thank members of the Leveraging Environmental Epidemiology for Equity (LEEFE) Lab at Tufts University for their feedback on the initial conceptualization of this work and presentation of prepublication findings.

This research was supported by the Tufts University Office of the Vice Provost for Research and by the Tufts/TMC COVID-19 Rapid Response Seed Funding Program. LC is supported by Eunice Kennedy Shriver National Institute of Child Health & Human Development (NICHD; grant #K12HD092535).

Data Availability

A deidentified version of the study data will be made available upon reasonable request to Peter Levine. Peter Levine is associate dean of academic affairs and Lincoln Filene professor of citizenship and public affairs in Tufts University's Jonathan Tisch College of Civic Life. He is also a coprincipal investigator of the Equity in Health, Wealth and Civic Engagement Priority Research Cluster. Exploration of some of the salient measures collected in our survey, as well as bivariate comparisons and visualizations of results, are available at our public-facing website.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Table displaying time since provider engagement for study sample (N=1034; weighted proportions).
[DOCX File , 13 KB - formative_v6i11e42126_app1.docx]

References

1. Starfield B, Shi L, Macinko J. Contribution of primary care to health systems and health. Milbank Q 2005;83(3):457-502 [FREE Full text] [doi: 10.1111/j.1468-0009.2005.00409.x] [Medline: 16202000]
2. Kravet SJ, Shore AD, Miller R, Green GB, Kolodner K, Wright SM. Health care utilization and the proportion of primary care physicians. Am J Med 2008 Feb;121(2):142-148. [doi: 10.1016/j.amjmed.2007.10.021] [Medline: 18261503]
3. Shi L. The impact of primary care: a focused review. Scientifica (Cairo) 2012;2012:432892 [FREE Full text] [doi: 10.6064/2012/432892] [Medline: 24278694]
4. Atlas SJ, Grant RW, Ferris TG, Chang Y, Barry MJ. Patient-physician connectedness and quality of primary care. Ann Intern Med 2009 Mar 03;150(5):325-335 [FREE Full text] [doi: 10.7326/0003-4819-150-5-200903030-00008] [Medline: 19258560]
5. Friedberg MW, Hussey PS, Schneider EC. Primary care: a critical review of the evidence on quality and costs of health care. Health Aff (Millwood) 2010 May;29(5):766-772. [doi: 10.1377/hlthaff.2010.0025] [Medline: 20439859]
6. Lazar M, Davenport L. Barriers to health care access for low income families: a review of literature. J Community Health Nurs 2018;35(1):28-37. [doi: 10.1080/07370016.2018.1404832] [Medline: 29323941]
7. Bhandari N, Shi Y, Jung K. Seeking health information online: does limited healthcare access matter? J Am Med Inform Assoc 2014 Nov;21(6):1113-1117 [FREE Full text] [doi: 10.1136/amiajnl-2013-002350] [Medline: 24948558]
8. Lambert SD, Loiselle CG. Health information seeking behavior. Qual Health Res 2007 Oct;17(8):1006-1019. [doi: 10.1177/1049732307305199] [Medline: 17928475]
9. Sbaffi L, Rowley J. Trust and credibility in web-based health information: a review and agenda for future research. J Med Internet Res 2017 Jun 19;19(6):e218 [FREE Full text] [doi: 10.2196/jmir.7579] [Medline: 28630033]

10. Cuan-Baltazar JY, Muñoz-Perez MJ, Robledo-Vega C, Pérez-Zepeda MF, Soto-Vega E. Misinformation of COVID-19 on the Internet: infodemiology study. JMIR Public Health Surveill 2020 Apr 09;6(2):e18444 [FREE Full text] [doi: 10.2196/18444] [Medline: 32250960]
11. Chen Y, Li C, Liang J, Tsai C. Health Information Obtained From the Internet and Changes in Medical Decision Making: Questionnaire Development and Cross-Sectional Survey. J Med Internet Res 2018 Feb 12;20(2):e47 [FREE Full text] [doi: 10.2196/jmir.9370] [Medline: 29434017]
12. Aoun L, Lakkis N, Antoun J. Prevalence and Outcomes of Web-Based Health Information Seeking for Acute Symptoms: Cross-Sectional Study. J Med Internet Res 2020 Jan 10;22(1):e15148 [FREE Full text] [doi: 10.2196/15148] [Medline: 31922490]
13. Tan SS, Goonawardene N. Internet Health Information Seeking and the Patient-Physician Relationship: A Systematic Review. J Med Internet Res 2017 Jan 19;19(1):e9 [FREE Full text] [doi: 10.2196/jmir.5729] [Medline: 28104579]
14. Zhang Z, Yang H, He J, Lu X, Zhang R. The Impact of Treatment-Related Internet Health Information Seeking on Patient Compliance. Telemed J E Health 2021 May 01;27(5):513-524. [doi: 10.1089/tmj.2020.0081] [Medline: 32907505]
15. Terbeck S, Chesterman L. Parents, ADHD and the internet. Atten Defic Hyperact Disord 2012 Sep;4(3):159-166. [doi: 10.1007/s12402-012-0081-7] [Medline: 22736404]
16. Lee C. Does the internet displace health professionals? J Health Commun 2008;13(5):450-464. [doi: 10.1080/10810730802198839] [Medline: 18661387]
17. Lee C, Hornik RC. Physician trust moderates the Internet use and physician visit relationship. J Health Commun 2009;14(1):70-76 [FREE Full text] [doi: 10.1080/10810730802592262] [Medline: 19180372]
18. Wang X, Shi J, Kong H. Online Health Information Seeking: A Review and Meta-Analysis. Health Commun 2020 Apr 14:1163-1175. [doi: 10.1080/10410236.2020.1748829] [Medline: 32290679]
19. Stopka TJ, Feng W, Corlin L, King E, Mistry J, Mansfield W, et al. Assessing equity in health, wealth, and civic engagement: a nationally representative survey, United States, 2020. Int J Equity Health 2022 Jan 28;21(1):12 [FREE Full text] [doi: 10.1186/s12939-021-01609-w] [Medline: 35090465]
20. LaVeist T, Isaac L, Williams K. Mistrust of health care organizations is associated with underutilization of health services. Health Serv Res 2009 Dec;44(6):2093-2105 [FREE Full text] [doi: 10.1111/j.1475-6773.2009.01017.x] [Medline: 19732170]
21. Ganguli I, Shi Z, Orav EJ, Rao A, Ray KN, Mehrotra A. Declining Use of Primary Care Among Commercially Insured Adults in the United States, 2008–2016. Ann Intern Med 2020 Feb 04;172(4):240-247. [doi: 10.7326/m19-1834]
22. Hanif F, Read JC, Goodacre JA, Chaudhry A, Gibbs P. The role of quality tools in assessing reliability of the internet for health information. Inform Health Soc Care 2009 Dec 01;34(4):231-243. [doi: 10.3109/17538150903359030] [Medline: 19919300]
23. Luo A, Qin L, Yuan Y, Yang Z, Liu F, Huang P, et al. The Effect of Online Health Information Seeking on Physician-Patient Relationships: Systematic Review. J Med Internet Res 2022 Feb 10;24(2):e23354 [FREE Full text] [doi: 10.2196/23354] [Medline: 35142620]
24. Lu T, Chen H, Xu Y, Zhang C. Internet usage, physician performances and patient's trust in physician during diagnoses: investigating both pre-use and not-use internet groups. : Jan 5-8, 2016; 2016 Presented at: 49th Hawaii International Conference on System Sciences (HICSS); 2016; Koloa, Hawaii, USA p. 3189-3198. [doi: 10.1109/hicss.2016.400]
25. Zhang R, Lu X, Wu W, Shang X, Liu M. Mature or Emerging? The Impact of Treatment-Related Internet Health Information Seeking on Patients' Trust in Physicians. Int J Environ Res Public Health 2018 Aug 28;15(9):1-17 [FREE Full text] [doi: 10.3390/ijerph15091855] [Medline: 30154309]
26. Bonds DE, Foley KL, Dugan E, Hall MA, Extrom P. An exploration of patients' trust in physicians in training. J Health Care Poor Underserved 2004 May;15(2):294-306. [doi: 10.1353/hpu.2004.0018] [Medline: 15253380]
27. Hall MA, Zheng B, Dugan E, Camacho F, Kidd KE, Mishra A, et al. Measuring patients' trust in their primary care providers. Med Care Res Rev 2002 Sep 01;59(3):293-318. [doi: 10.1177/1077558702059003004] [Medline: 12205830]
28. Benkert R, Cuevas A, Thompson HS, Dove-Meadows E, Knuckles D. Ubiquitous Yet Unclear: A Systematic Review of Medical Mistrust. Behav Med 2019 Apr 03;45(2):86-101 [FREE Full text] [doi: 10.1080/08964289.2019.1588220] [Medline: 31343961]
29. Jaiswal J, Halkitis PN. Towards a More Inclusive and Dynamic Understanding of Medical Mistrust Informed by Science. Behav Med 2019 Apr 03;45(2):79-85. [doi: 10.1080/08964289.2019.1619511] [Medline: 31343962]
30. Joseph AM, Fernandez V, Kritzman S, Eaddy I, Cook OM, Lambros S, et al. COVID-19 Misinformation on Social Media: A Scoping Review. Cureus 2022 Apr;14(4):e24601 [FREE Full text] [doi: 10.7759/cureus.24601] [Medline: 35664409]
31. Silver MP. Patient perspectives on online health information and communication with doctors: a qualitative study of patients 50 years old and over. J Med Internet Res 2015 Jan 13;17(1):e19 [FREE Full text] [doi: 10.2196/jmir.3588] [Medline: 25586865]
32. Iverson SA, Howard KB, Penney BK. Impact of internet use on health-related behaviors and the patient-physician relationship: a survey-based study and review. J Am Osteopath Assoc 2008 Dec;108(12):699-711. [Medline: 19075034]
33. Murray E, Lo B, Pollack L, Donelan K, Catania J, White M, et al. The impact of health information on the internet on the physician-patient relationship: patient perceptions. Arch Intern Med 2003 Jul 28;163(14):1727-1734. [doi: 10.1001/archinte.163.14.1727] [Medline: 12885689]

34. Kivits J. Informed patients and the internet: a mediated context for consultations with health professionals. J Health Psychol 2006 Mar 01;11(2):269-282. [doi: 10.1177/1359105306061186] [Medline: 16464924]
35. Rod K. Finding Ways to Lift Barriers to Care for Chronic Pain Patients: Outcomes of Using Internet-Based Self-Management Activities to Reduce Pain and Improve Quality of Life. Pain Res Manag 2016;2016:8714785 [FREE Full text] [doi: 10.1155/2016/8714785] [Medline: 27445632]
36. Schofield M. Regulatory and legislative issues on telehealth. Nutr Clin Pract 2021 Aug;36(4):729-738. [doi: 10.1002/ncp.10740] [Medline: 34159667]
37. Riera R, Bagattini, Pacheco RL, Pachito DV, Roitberg F, Ilbawi A. Delays and Disruptions in Cancer Health Care Due to COVID-19 Pandemic: Systematic Review. JCO Glob Oncol 2021 Feb;7:311-323 [FREE Full text] [doi: 10.1200/GO.20.00639] [Medline: 33617304]
38. Hewner S, Sullivan SS, Yu G. Reducing Emergency Room Visits and In-Hospitalizations by Implementing Best Practice for Transitional Care Using Innovative Technology and Big Data. Worldviews Evid Based Nurs 2018 Jun;15(3):170-177 [FREE Full text] [doi: 10.1111/wvn.12286] [Medline: 29569327]
39. Cline R, Haynes K. Consumer health information seeking on the Internet: the state of the art. Health Educ Res 2001 Dec 01;16(6):671-692. [doi: 10.1093/her/16.6.671] [Medline: 11780707]
40. Jia X, Pang Y, Liu LS. Online Health Information Seeking Behavior: A Systematic Review. Healthcare (Basel) 2021 Dec 16;9(12):1740 [FREE Full text] [doi: 10.3390/healthcare9121740] [Medline: 34946466]
41. Swire-Thompson B, Lazer D. Public Health and Online Misinformation: Challenges and Recommendations. Annu Rev Public Health 2020 Apr 02;41:433-451. [doi: 10.1146/annurev-publhealth-040119-094127] [Medline: 31874069]
42. van Huijstee D, Vermeulen I, Kerkhof P, Droog E. Continued influence of misinformation in times of COVID-19. Int J Psychol 2022 Feb 26;57(1):136-145 [FREE Full text] [doi: 10.1002/ijop.12805] [Medline: 34448200]
43. Pan W, Liu D, Fang J. An Examination of Factors Contributing to the Acceptance of Online Health Misinformation. Front Psychol 2021;12:630268 [FREE Full text] [doi: 10.3389/fpsyg.2021.630268] [Medline: 33732192]

Abbreviations

MMI: Medical Mistrust Index
WHI: web-based health information

Edited by A Mavragani; submitted 23.08.22; peer-reviewed by A Joseph, PCI Pang; comments to author 11.10.22; revised version received 30.10.22; accepted 31.10.22; published 30.11.22.

Please cite as:
Shaveet E, Urquhart C, Gallegos M, Dammann O, Corlin L
Web-Based Health Information–Seeking Methods and Time Since Provider Engagement: Cross-sectional Study
JMIR Form Res 2022;6(11):e42126
URL: https://formative.jmir.org/2022/11/e42126
doi:10.2196/42126
PMID:36449328

Original Paper

Psychological Disorders of Patients With Allergic Rhinitis in Chengdu, China: Exploratory Research

Heyin Huang[1*], MD; Yichen Wang[2*], MM; Lanzhi Zhang[3], BMS; Qinxiu Zhang[4], MD; Xiaojuan Wu[1], MM; Hengsheng He[1], BMS

[1]Department of Otolaryngology, Sichuan Integrative Medicine Hospital, Chengdu, China

[2]Department of Geriatric Orthopedics, Sichuan Province Orthopaedic Hospital, Chengdu, China

[3]Department of Medical Cosmetology, Affiliated Hospital of Chengdu University of Traditional Chinese Medicine, Chengdu, China

[4]Department of Otolaryngology, Affiliated Hospital of Chengdu University of Traditional Chinese Medicine, Chengdu, China

[*]these authors contributed equally

Corresponding Author:
Qinxiu Zhang, MD
Department of Otolaryngology
Affiliated Hospital of Chengdu University of Traditional Chinese Medicine
No. 39-41, Shierqiao Road, Jinniu District
Chengdu, 610000
China
Phone: 86 18980880173
Email: zhqinxiu@163.com

Abstract

Background: The number of patients with allergic rhinitis (AR) has exceeded 500 million worldwide due to the unstable curative effect that can easily produce mental and psychological disorders. However, most of the relevant existing literature is one-on-one retrospective analyses or targeted meta-analyses of AR with psychological disorders like irritability, depression, and anxiety, while “multi-hospital + interdisciplinary” multiple regression analyses are scarce.

Objective: This study aims to precisely identify the psychological disorders of patients with AR who were diagnosed and treated in the five most renowned hospitals in Chengdu, China over the past 5 years using 10 classification methods so as to attract attention and care from otolaryngologists.

Methods: The Symptom Checklist 90 (SCL-90) was used to group and score the mental state of 827 strictly screened patients with AR according to 9 classification criteria. The scores were then compared within groups. Intergroup comparisons were made between the study group and the Chinese norm, and the positive factors for psychological disorders were extracted. Four symptoms in the study group, that is, nasal itching, sneezing, clear discharge, and nasal congestion, were scored on a visual analog scale. Partial correlation analysis was performed between the extracted positive factors for psychological disorders and the symptom scores by the multiple regression statistical method.

Results: Among 827 patients, 124 (15%) had no mental health impairments, 176 (21.3%) had mild impairments, 474 (57.3%) had mild to moderate impairments, 41 (5%) had moderate to severe impairments, and 12 (1.4%) had severe impairments. The average score of the SCL-90 for all 827 patients was 2.64 (SD 0.25), which corresponded to mild to moderate mental health impairments. The 827 patients scored significantly higher for the 4 positive factors: depression, anxiety, psychosis, and other (sleep, diet). Depression was positively correlated with sneezing and clear discharge, anxiety was positively correlated with nasal itching and congestion, psychosis was positively correlated with nasal itching and sneezing, and other (sleep, diet) was positively correlated with clear discharge and nasal congestion.

Conclusions: Patients with AR have mild to moderate mental health impairments, with women and those with abnormal BMI, aged ≥45 years, with a monthly salary <¥5110 (US $700), with a disease duration <13 years, residing in urban areas, with a high school or above education, or who are indoor laborers being at high risk and requiring more care, follow-up, and comprehensive therapy from otolaryngologists.

(JMIR Form Res 2022;6(11):e37101) doi:10.2196/37101

KEYWORDS

psychological disorders; allergic rhinitis; Chengdu; China

Introduction

According to the guidelines of the World Health Organization (WHO) and Allergic Rhinitis and its Impact on Asthma, the number of patients with allergic rhinitis (AR) has exceeded 500 million globally [1]. Due to the unstable curative effect, symptoms like nasal itching, sneezing, clear discharge, and nasal congestion recur easily, which are hardly eradicated, thus readily generating mental/psychological disorders [2]. In China from 2007 to 2016, a total of 6 ear, nose, and throat doctors were hacked to death by patients with knives, and 3 doctors were seriously injured. It is thus imperative to study the correlation between AR and psychological disorders in depth, which is in line with the "biological-psychological-social" international medical model.

However, most of the relevant existing literature are one-on-one retrospective analyses or targeted meta-analyses of AR with psychological disorders like irritability, depression, and anxiety [3], while "multi-hospital + interdisciplinary" multiple regression analyses are scarce. To this end, this study makes the following 3 innovations by inviting the 5 most renowned hospitals, both Chinese and Western, in China's Chengdu and integrating three perspectives of otolaryngology, psychology, and psychiatry. First, a comprehensive intragroup comparison is made on the study group. The 9 classification criteria for intragroup comparison are derived from the preliminary questionnaire survey, which are the factors most likely to cause psychological disorders among patients with AR that are preferentially screened by the authors. Second, an intergroup comparison is performed between the study group and the Chinese norm to derive the 4 positive factors most associated with the patients' psychological disorders. Third, the visual analog scale (VAS) is used to score the 4 AR symptoms of nasal itching, sneezing, clear discharge, and nasal congestion, which are then subjected to the one-on-one partial correlation analysis separately with the 4 positive factors of psychological disorders.

Methods

Clinical Data

A total of 1536 patients with AR who received treatment at 5 hospitals from July 2013 to January 2018 were selected. A routine questionnaire survey was conducted strictly following the inclusion and exclusion criteria for the AR group, rejecting 709 cases and enrolling 827 cases.

Ethics Approval

All selected patients signed the informed consent and questionnaire. The study has been certified by the ethics review system of the World Federation of Traditional Chinese Medicine Societies (CAP2013 (A) 0038).

Diagnostic Criteria

We will strictly follow the latest AR diagnostic criteria promulgated by China in 2018 [4]. The following three points must be presented simultaneously: (1) having two of the four symptoms of nasal itching, sneezing, clear discharge, and nasal congestion, and symptoms last for half an hour to over an hour and occur more than 4 days weekly; (2) skin prick test parameters are positive, one of which is (++) or above, or the allergen-specific IgE is positive; and (3) the morphology of nasal mucosa shows inflammatory change.

Exclusion Criteria

The following exclusion criteria will be used: failing to complete the questionnaire carefully or completely; having abnormal nasal anatomy (eg, those definitively diagnosed with nasal sinusitis or septum deviation); having autoimmune diseases; having systemic and chronic diseases; having any history of psychogenic illness; or having an incomplete family structure, such as separation, divorce, widowhood, and loss of child or children.

Psychological Scale and Scoring Method

There are 90 items on the Symptom Checklist 90 (SCL-90) [5], each of which is graded from 1 to 5. The 90 items are randomly arranged and reflect 10 psychological factors.

Total average score = Total score of the 90 items / 90

Factor score = Total score of each item constituting a factor / Number of items constituting the factor

The following will be used for the classification of the symptom [6]: a factor score and total average score <1.5 points will be judged as asymptomatic, a factor score and total average score ≥1.5 points and <2.5 points will be judged as mild, a factor score and total average score ≥2.5 points and <3.5 points will be judged as mild to moderate, a factor score and total average score ≥3.5 points and <4.5 points will be judged as moderate to severe, a factor score and total average score ≥4.5 points will be judged as severe.

Evaluation and Quantification of Symptoms

A total of 4 AR symptoms (ie, nasal itching, sneezing, clear discharge, and nasal congestion) were evaluated separately on a 10-point VAS, with 10 representing the severest, and 0 indicating the mildest.

Mean VAS score = Total scores of the five symptoms / 5

Comparison and Research Methodologies

Intragroup and Intergroup Comparison

The 10 SCL-90 factor scores and whole average scores were compared between the AR group (study group) and the Chinese norm (healthy controls group) [7], from which the positive factors of the SCL-90 were extracted.

Partial Correlation Analysis

The 4 AR symptoms were evaluated with the VAS and recorded for scores. A partial correlation analysis [8] was made between the scores of each symptom and the positive factors through the multiple regression statistical method [9] by taking measurement

data like BMI, age, monthly salary, and disease duration as the control variables.

Statistical Analysis

If the measurement data is consistent with the normal distribution, they will be expressed by $x \pm s$, and all data will accept the normality test and homogeneity test of variance in a single-sample Kolmogorov-Smirnov test. The normal distribution data was tested by independent sample *t* or paired sample *t* test, and the nonnormal distribution data was tested by Mann-Whitney *U*. A one-way and orderly contingency table will accept the rank sum test of Kruskal-Wallis *H*. Using the multivariate regression statistical method, the control variables were first established, and then a partial correlation analysis with a partial correlation coefficient was performed. SPSS 19.0 software (IBM Corp) was used to carry out the statistical analysis of the data, and the difference has statistical significance if $P<.05$.

Results

Intragroup Comparisons

The SCL-90 was used to group and score the mental state of 827 patients with AR (Table 1 and Figures 1-4). The intragroup comparison by gender (male: n=396; female: n=431) is shown in Table 1 and Figure 5.

The intragroup comparison of BMI is shown in Table 1 and Figure 6). In accordance with the WHO criteria for BMI [10], patients were divided into two groups: normal (18.5~24.9; n=531) and abnormal (<18.5 or ≥25; n=296). The intragroup comparison of age is shown in Table 1 and Figure 7. In accordance with the WHO age classification criteria [11], patients were divided into two groups: ages <44 years (n=438) and ≥45 years (n=389). The intragroup comparison by marital status (unmarried: n=448; married: n=379) is shown in Table 1 and Figure 8). The intragroup comparison by monthly salary is shown in Table 1 and Figure 9. Based on the average monthly salary (¥5110, US $700) of China's Chengdu City in 2016 [12], patients were divided into two groups: <¥5110 (US $700; n=391) and ≥¥5110 (US $700; n=436). The intragroup comparison by disease duration is shown in Table 1 and Figure 10). Patients were divided into two age groups: <13 years (n=451) and ≥13 years (n=376).

The intragroup comparison by living environment (urban or rural) is shown in Table 1 and Figure 11. The classification was made according to the Chinese administrative units [13], with those living in provinces, cities, and districts called "Urban" residents (n=471), and those living in townships, towns, and villages called "Rural" residents (n=356). The intragroup comparison by education level is shown in Table 1 and Figure 12. Patients were divided into two groups: "Junior high school and below" (n=379) and "Senior high school and above" (n=448). The intragroup comparison by working environment (indoors or outdoors) is shown in Table 1 and Figure 13). The Labor Law of the People's Republic of China stipulates [14] that working hours should not exceed 8 hours a day. Based on this, the patients were divided into two groups: "Indoor" (average daily indoor working hours <4; n=475) and "Outdoor" (≥4 hours; n=352).

Table 1. The Symptom Checklist 90 was used to group and score the mental state of 827 patients with allergic rhinitis according to 9 classification criteria, namely, gender, BMI, age, marital status, monthly salary, disease duration, living environment (urban or rural), education level, and working environment (indoors or outdoors), and then all the scores were compared within groups.

Group	Somatization	Obsessive-compulsive	Interpersonal sensitivity	Depression	Anxiety	Hostility	Phobic anxiety	Paranoia	Psychoticism	Other (sleep, diet)	Total average score
Gender											
Men, mean (SD)	2.00 (0.45)	2.52 (0.61)	2.63 (0.51)	2.85 (0.52)	3.19 (0.38)	3.02 (0.42)	1.90 (0.27)	2.77 (0.45)	3.03 (0.41)	2.76 (0.45)	2.67 (0.33)
Women, mean (SD)	2.66 (0.51)	2.80 (0.57)	3.37 (0.47)	3.44 (0.57)	3.36 (0.67)	2.66 (0.35)	2.11 (0.60)	3.13 (0.48)	3.20 (0.32)	3.47 (0.36)	3.02 (0.45)
Z score	–5.10	–1.22	–5.51	–4.86	–0.69	–1.70	–0.99	–1.75	–0.74	–5.28	–1.65[a]
P value	*.03*[b]	>.99	*.002*	*.04*	>.99	>.99	>.99	>.99	>.99	*.01*	>.99
BMI											
Normal BMI, mean (SD)	1.91 (0.21)	2.26 (0.74)	2.09 (0.45)	2.11 (0.32)	2.17 (0.47)	1.77 (0.82)	1.80 (0.53)	1.95 (0.33)	1.75 (0.63)	2.24 (0.72)	1.99 (0.15)
Abnormal BMI, mean (SD)	2.02 (0.33)	2.71 (0.63)	2.30 (0.41)	2.41 (0.88)	2.26 (0.53)	1.90 (0.30)	1.93 (0.96)	2.61 (0.71)	2.01 (0.48)	2.41 (0.40)	2.26 (0.38)
T test (*df*)	0.61 (412)	4.10 (412)	2.26 (412)	3.24 (412)	0.36 (412)	1.38 (412)	1.25 (412)	5.29 (412)	2.95 (412)	1.77 (412)	3.13 (412)
P value	.72	*.03*	.43	.20	.84	.64	.69	*.01*	.32	.51	.26
Age (years)											
<44, mean (SD)	2.53 (0.33)	3.11 (0.42)	3.43 (0.38)	2.69 (0.28)	2.77 (0.50)	3.03 (0.35)	2.50 (0.44)	2.45 (0.67)	2.61 (0.66)	2.83 (0.52)	2.80 (0.41)
≥44, mean (SD)	3.23 (0.52)	2.81 (0.45)	2.82 (0.53)	2.24 (0.43)	2.51 (0.41)	2.52 (0.47)	2.19 (0.51)	1.93 (0.23)	2.35 (0.52)	3.47 (0.38)	2.61 (0.33)
T test (*df*)	5.22 (412)	1.70 (412)	5.06 (412)	3.36 (412)	1.08 (412)	4.09 (412)	1.70 (412)	4.22 (412)	1.10 (412)	5.10 (412)	–0.09 (412)
P value	*.02*	>.99	*.03*	.16	>.99	.08	>.99	.05	>.99	*.03*	>.99
Marital status											
Unmarried, mean (SD)	2.81 (0.26)	1.84 (0.56)	2.36 (0.44)	2.33 (0.70)	2.53 (0.80)	2.03 (0.66)	2.28 (0.34)	2.01 (0.36)	2.50 (0.31)	2.48 (0.12)	2.32 (3.70)
Married, mean (SD)	2.94 (0.11)	2.11 (0.31)	2.45 (0.34)	2.49 (0.15)	2.97 (0.69)	2.40 (0.14)	2.41 (0.50)	1.95 (0.55)	1.93 (0.41)	3.14 (0.55)	2.48 (4.05)
Z score	2.19	3.07	1.76	2.33	4.48	3.51	2.02	0.31	4.66	5.17	2.47
P value	.48	.24	.77	.39	*.04*	.18	.59	.89	*.02*	*.001*	.41
Monthly pay (¥)											
<5110[c], mean (SD)	3.05 (0.22)	3.41 (0.65)	3.13 (0.27)	3.33 (0.45)	2.69 (0.44)	1.99 (0.82)	1.79 (0.82)	2.84 (0.55)	2.53 (0.90)	3.33 (0.51)	2.81 (0.94)
≥5110, mean (SD)	3.41 (0.81)	3.24 (0.28)	2.51 (0.62)	3.09 (0.24)	2.58 (0.71)	2.16 (0.60)	2.00 (0.30)	3.24 (0.14)	2.84 (0.75)	3.44 (0.75)	2.85 (0.15)
Z score	–3.88	–1.93	–4.93	–2.91	–1.48	–2.08	–2.65	–4.02	–3.46	–1.52	–1.77
P value	.18	.55	*.02*	.33	.64	.49	.41	*.04*	.27	.60	>.99
Disease course											
<13 years, mean (SD)	2.61 (0.76)	2.77 (0.50)	3.11 (0.33)	2.85 (0.38)	2.48 (0.27)	2.70 (0.22)	2.22 (0.50)	2.79 (0.36)	2.19 (0.29)	3.34 (0.55)	2.71 (0.60)
≥13 years, mean (SD)	3.05 (0.41)	2.72 (0.38)	3.47 (0.70)	3.46 (0.59)	2.97 (0.41)	2.18 (0.34)	2.69 (0.39)	2.90 (0.88)	2.58 (0.58)	2.51 (0.24)	2.85 (0.39)
Z score	1.38	0.22	0.90	4.98	1.49	2.37	1.42	0.57	0.96	5.44	0.64
P value	.93	>.99	>.99	*.02*	.71	.65	.84	>.99	>.99	*.002*	>.99

Group	Somatization	Obsessive-compulsive	Interpersonal sensitivity	Depression	Anxiety	Hostility	Phobic anxiety	Paranoia	Psychoticism	Other (sleep, diet)	Total average score
Living environment											
City, mean (SD)	3.29 (0.61)	2.81 (0.77)	3.10 (0.19)	3.40 (0.84)	3.04 (0.64)	2.39 (0.48)	2.40 (0.67)	2.19 (0.21)	3.02 (0.55)	3.40 (0.41)	2.91 (0.11)
Township, mean (SD)	2.65 (0.50)	2.05 (0.38)	2.49 (0.67)	2.49 (0.70)	2.33 (0.51)	2.07 (0.46)	2.08 (0.40)	1.78 (0.18)	2.40 (0.77)	2.52 (0.77)	2.29 (0.40)
T test (*df*)	4.88 (412)	5.25 (412)	4.57 (412)	5.91 (412)	5.09 (412)	1.83 (412)	1.79 (412)	2.55 (412)	4.63 (412)	5.55 (412)	–4.55 (412)
P value	*.04*	*.01*	*.04*	*.003*	*.02*	.71	.79	.56	*.04*	*.001*	*.04*
Educational level											
Junior or below, mean (SD)	2.22 (0.19)	2.86 (0.78)	2.52 (0.11)	2.33 (0.94)	2.82 (0.32)	1.96 (0.60)	2.04 (0.73)	2.66 (0.46)	2.60 (0.60)	2.46 (0.84)	2.45 (0.51)
High or above, mean (SD)	2.84 (0.37)	2.41 (0.31)	3.19 (0.22)	2.50 (0.55)	2.32 (0.28)	2.25 (0.44)	2.50 (0.34)	2.47 (0.50)	2.29 (0.45)	3.03 (0.30)	2.59 (0.65)
Z score	–4.82	–2.71	–4.90	–0.33	–3.38	–1.31	–2.94	–0.38	–1.46	–4.64	–0.27[a]
P value	*.03*	.51	*.03*	>.99	.29	>.99	.46	>.99	.16	*.04*	>.99
Working environment											
Indoors, mean (SD)	3.49 (0.64)	2.83 (0.35)	3.33 (0.30)	3.21 (0.29)	2.57 (0.93)	2.69 (0.56)	2.25 (0.25)	2.71 (0.10)	3.28 (0.60)	3.50 (0.85)	2.99 (0.48)
Outdoors, mean (SD)	2.78 (0.46)	3.44 (0.40)	2.56 (0.68)	2.88 (0.14)	3.40 (0.50)	3.34 (0.88)	2.51 (0.15)	2.41 (0.32)	2.67 (0.31)	2.85 (0.37)	2.88 (0.72)
Z score	5.31	4.96	5.79	1.51	6.22	5.08	1.02	1.11	4.89	5.00	0.04
P value	*.01*	*.04*	*.002*	>.99	*.004*	*.03*	>.99	>.99	*.04*	*.03*	>.99

[a]These are *t* test values (*df* 412).

[b]Italics indicate statistical significance at the <.05 level.

[c]US $700.

Figure 1. Comparison of somatization, obsessive-compulsive symptoms, and interpersonal sensitivity on 18 items in 9 groups. SCL-90: Symptom Checklist 90.

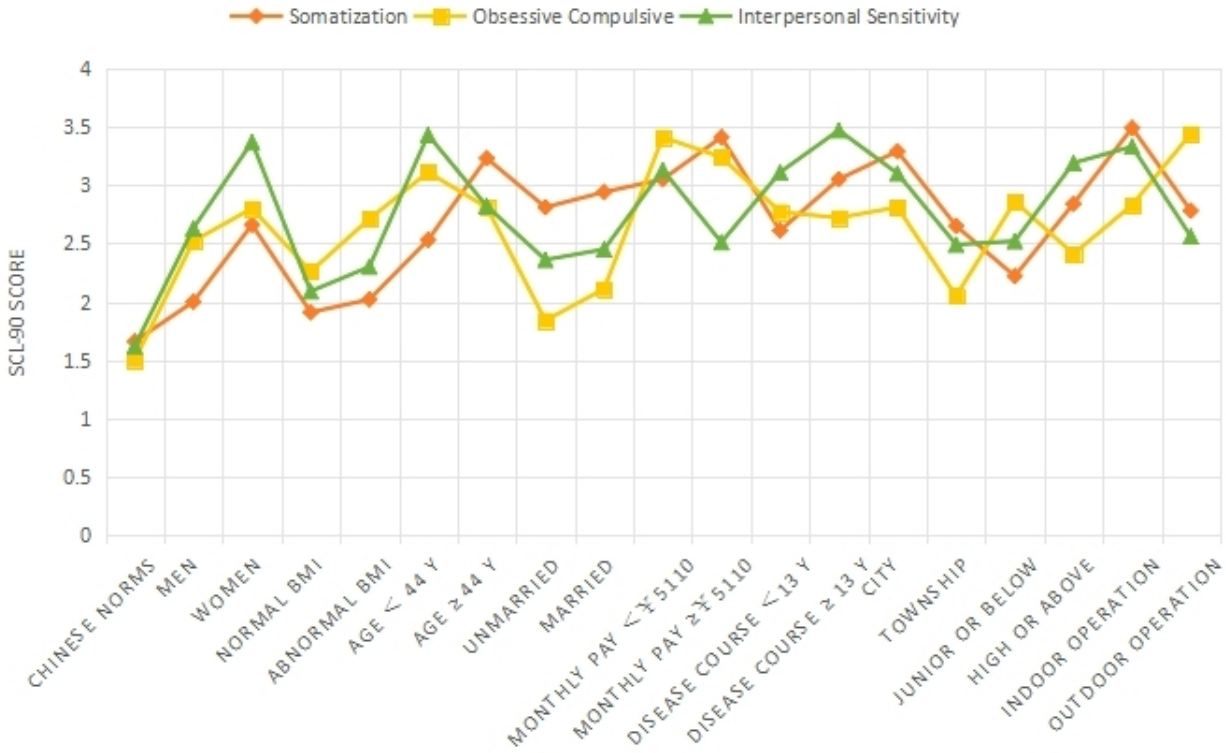


Figure 2. Comparison of depression, anxiety, and hostility on 18 items in 9 groups. SCL-90: Symptom Checklist 90.

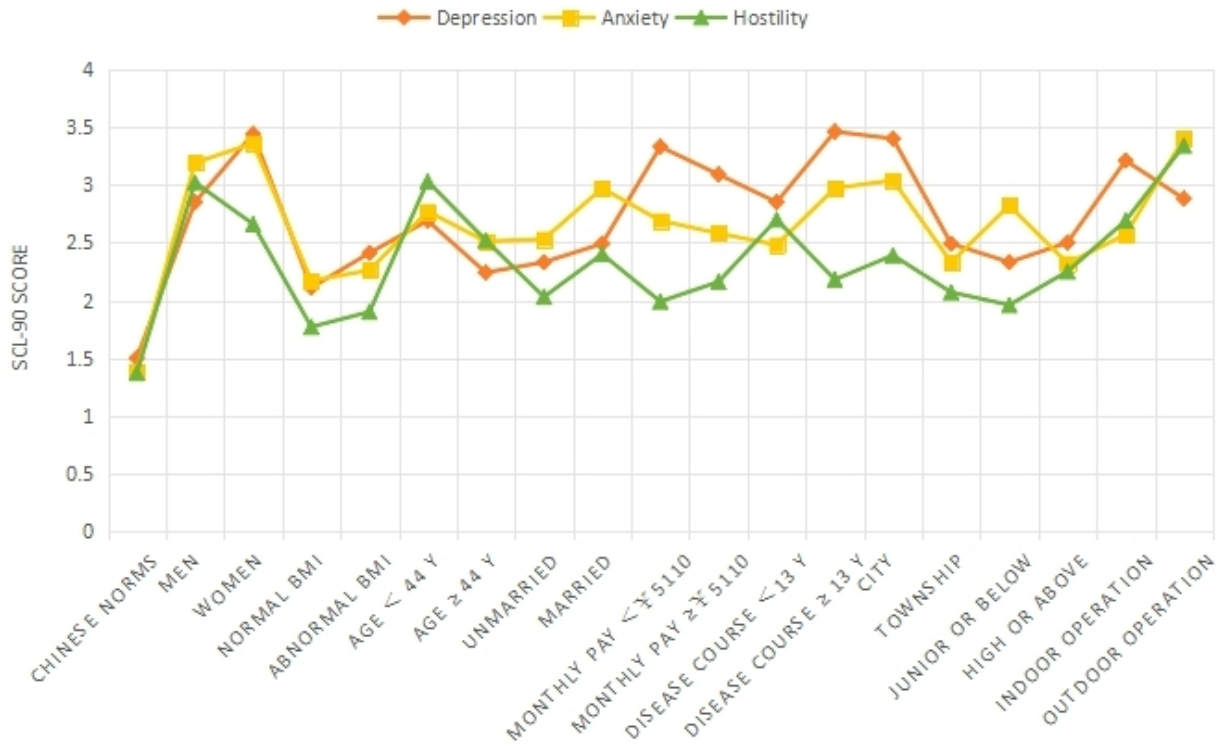


Figure 3. Comparison of phobic anxiety, paranoia, psychosis, and other (eg, sleep and diet) on 18 items in 9 groups. SCL-90: Symptom Checklist 90.

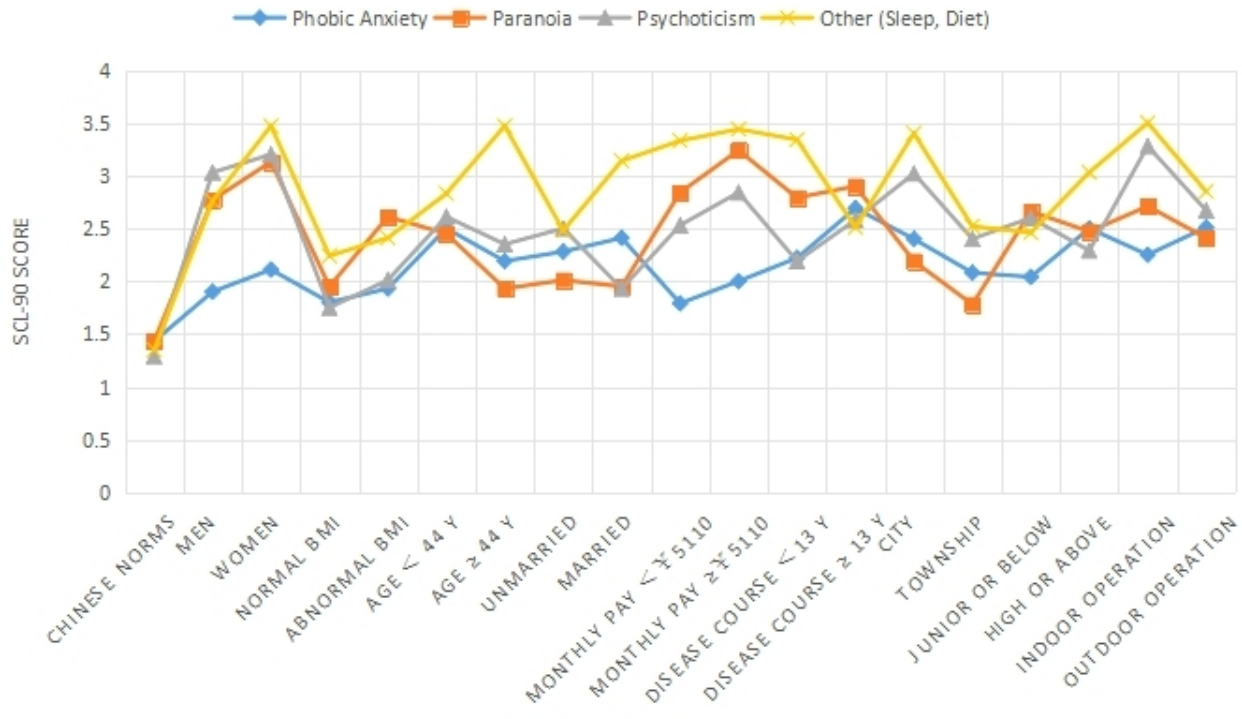


Figure 4. Comparison of 18 items in 9 groups and Chinese norms on total average score. SCL-90: Symptom Checklist 90.

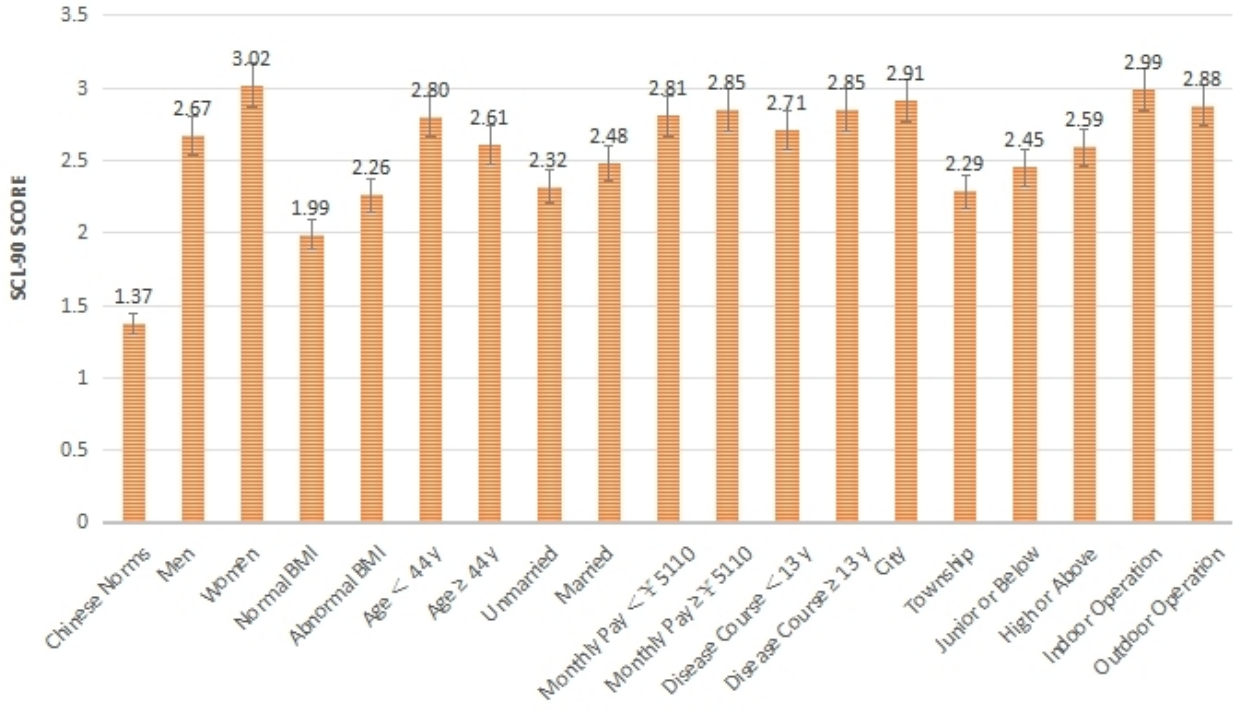


Figure 5. Intragroup comparison of gender. *P<.05.

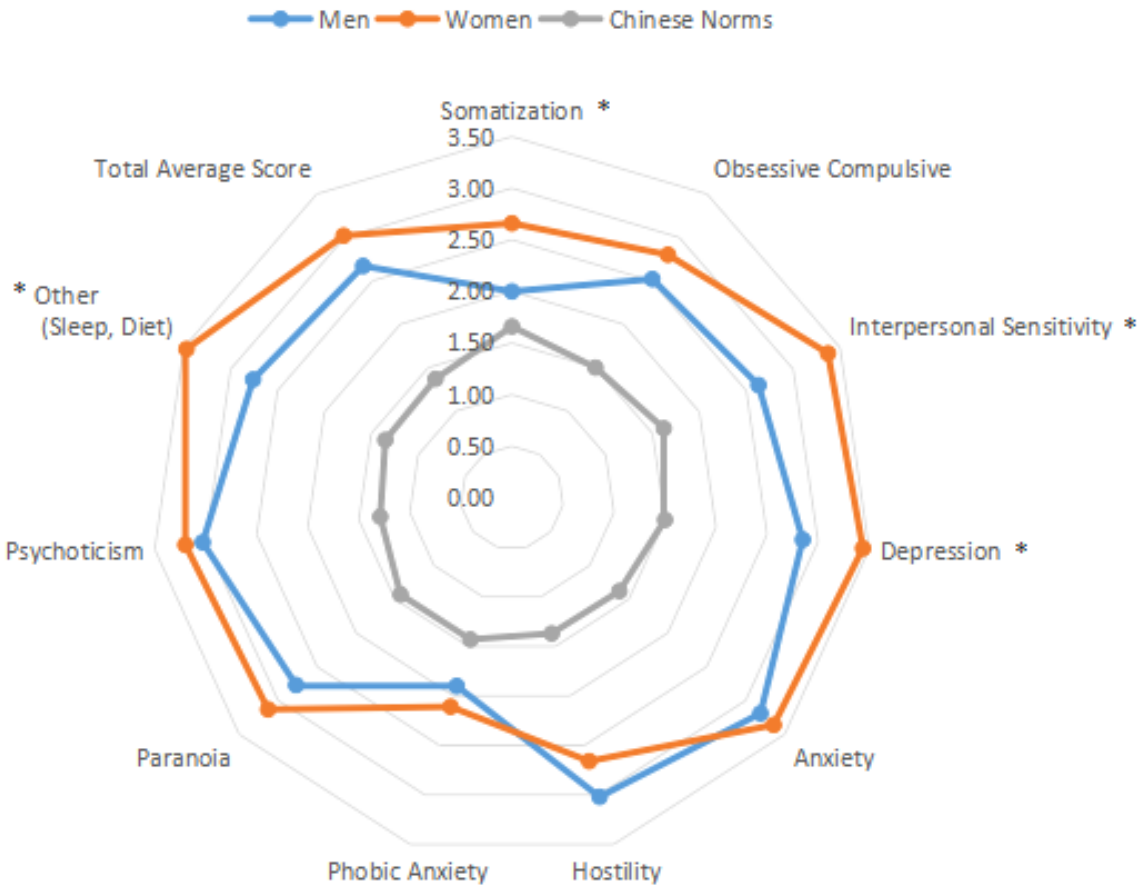


Figure 6. Intragroup comparison of BMI. *P<.05.

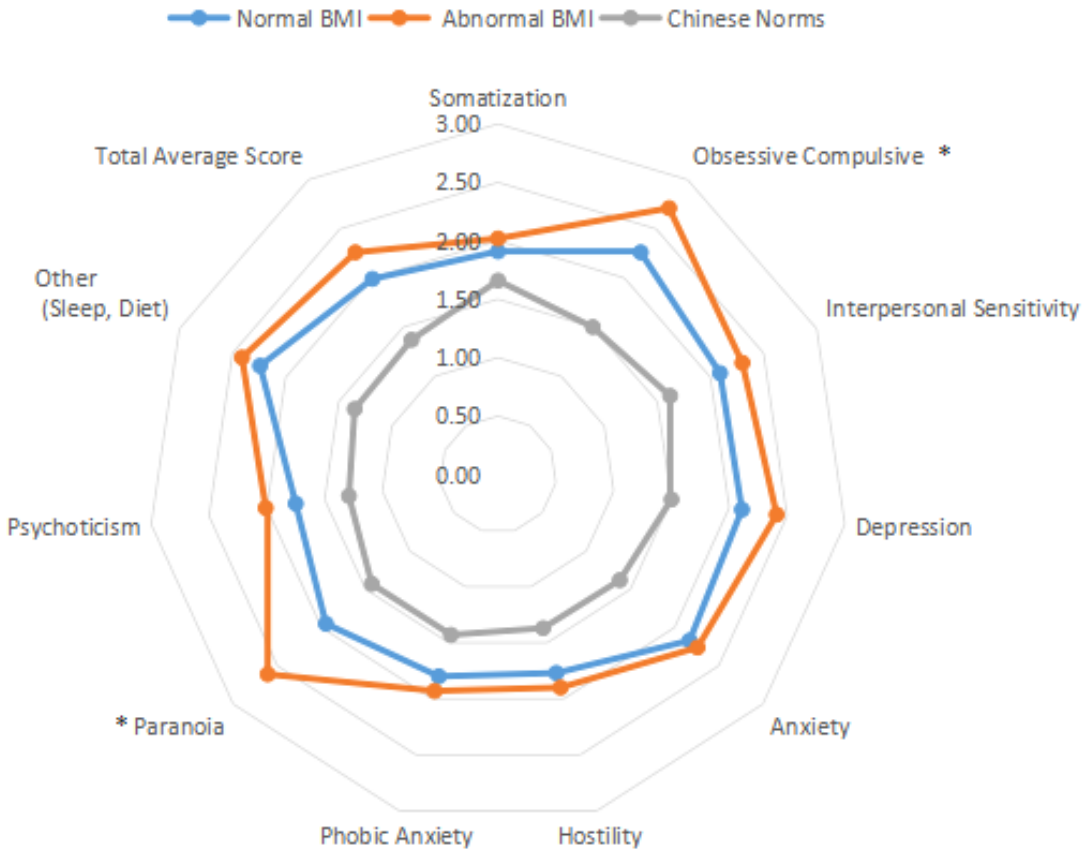


Figure 7. Intragroup comparison of age. *P<.05.

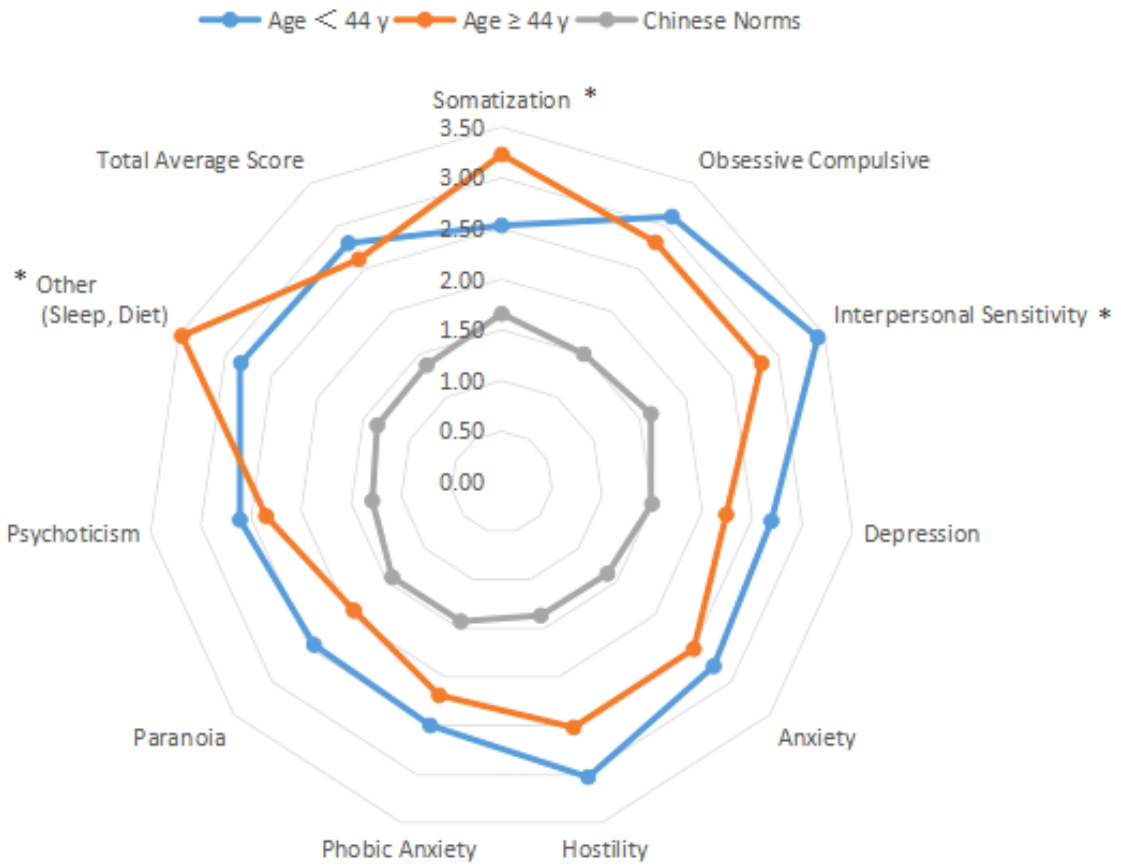


Figure 8. Intragroup comparison of marital status. *P<.05.

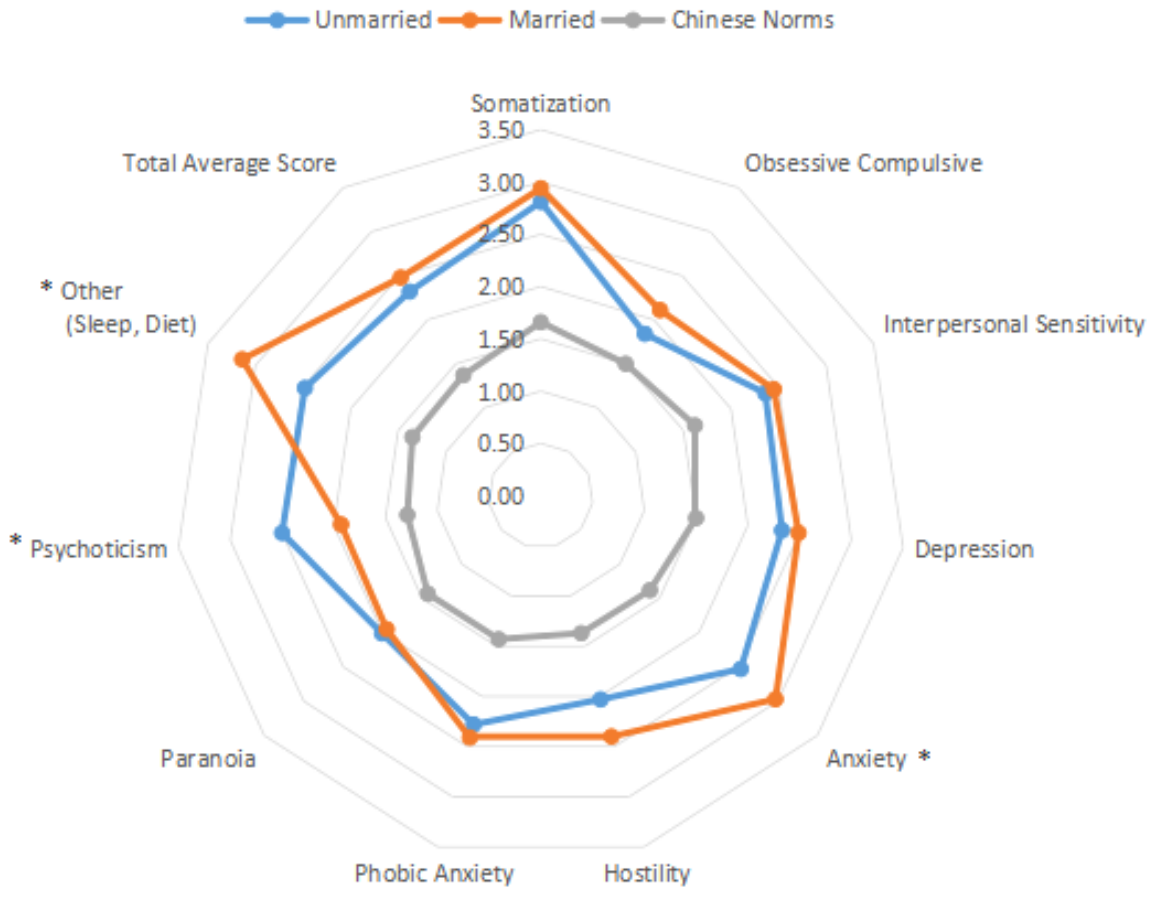


Figure 9. Intragroup comparison of monthly salary. *P<.05.

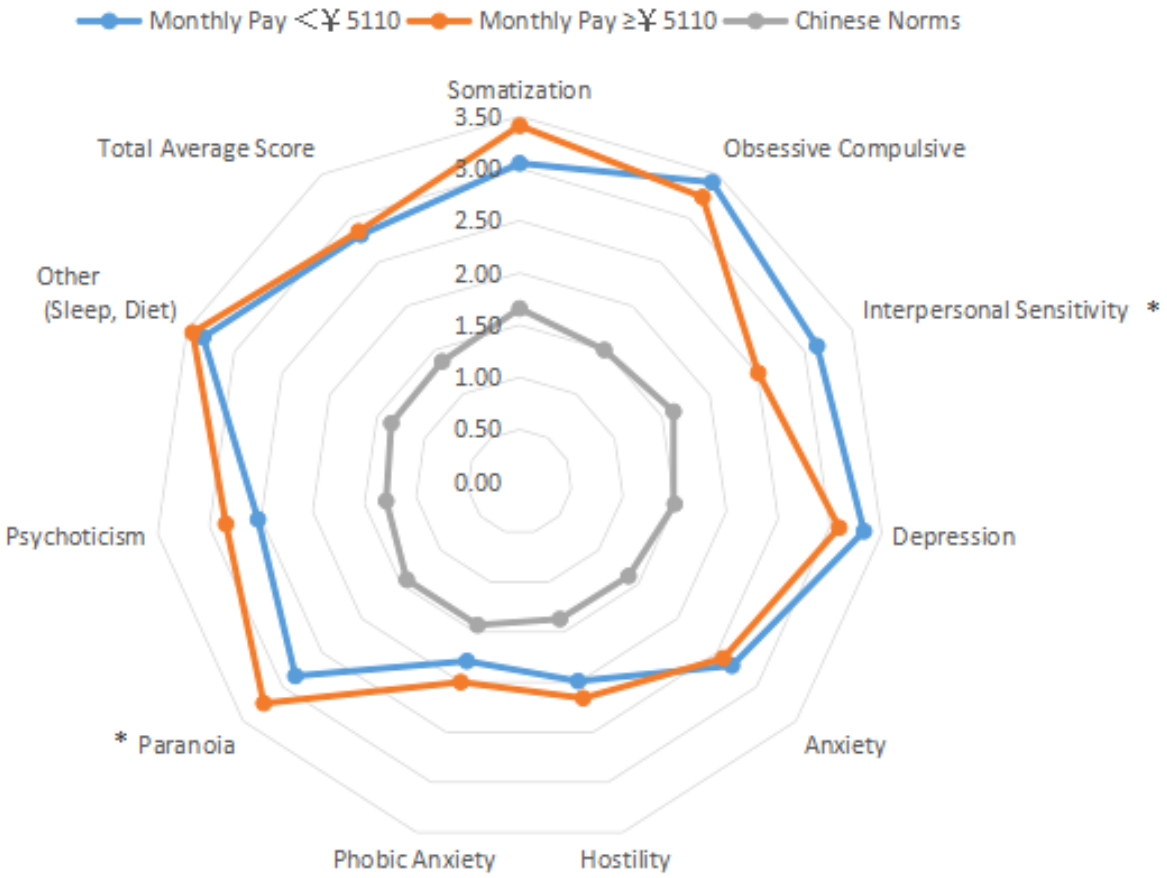


Figure 10. Intragroup comparison of disease duration. *P<.05.

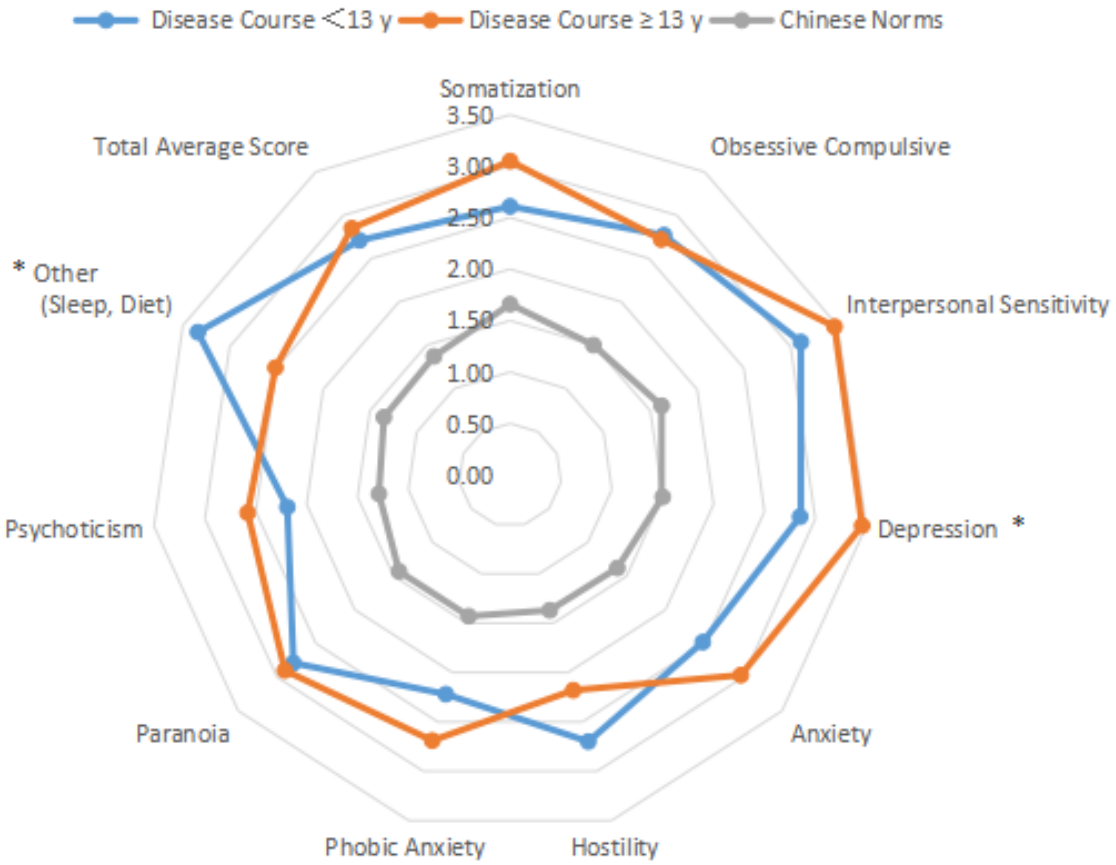


Figure 11. Intragroup comparison of living environment. *P<.05.

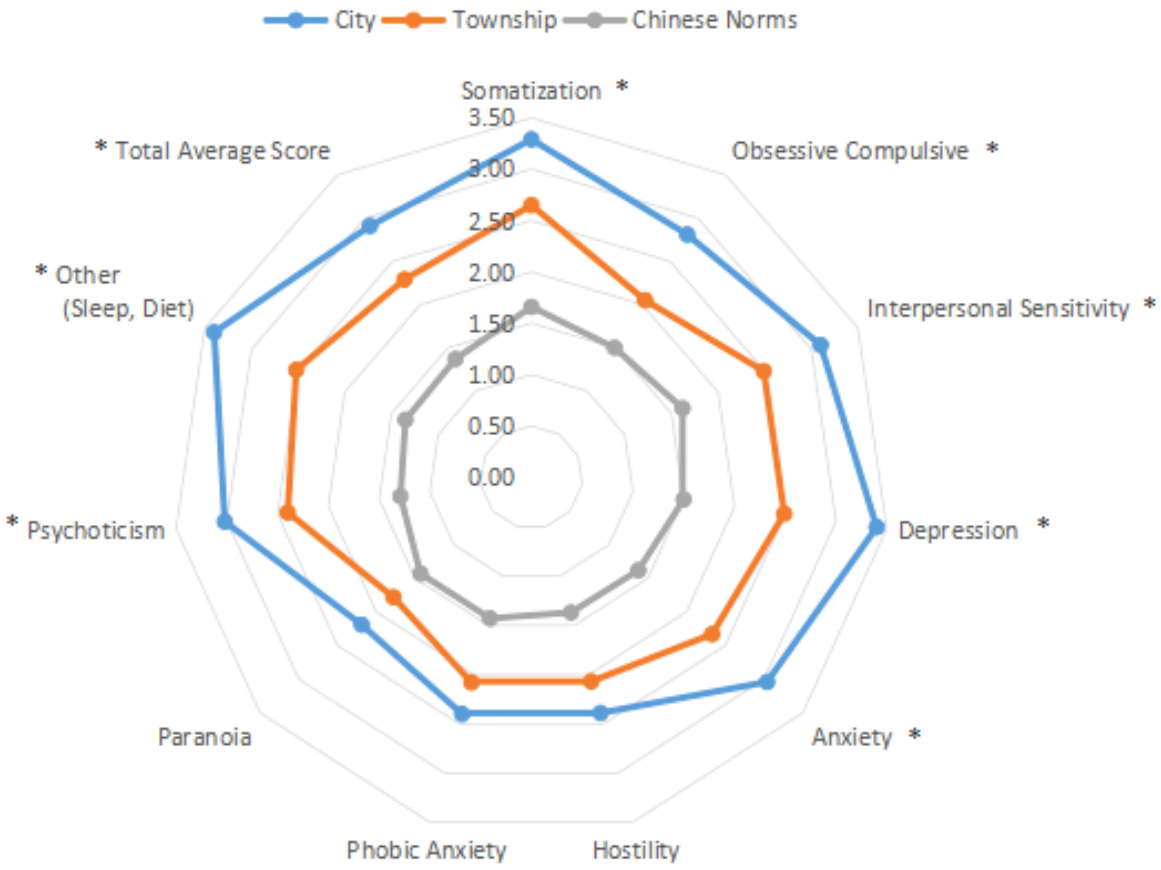


Figure 12. Intragroup comparison of education level. *P<.05.

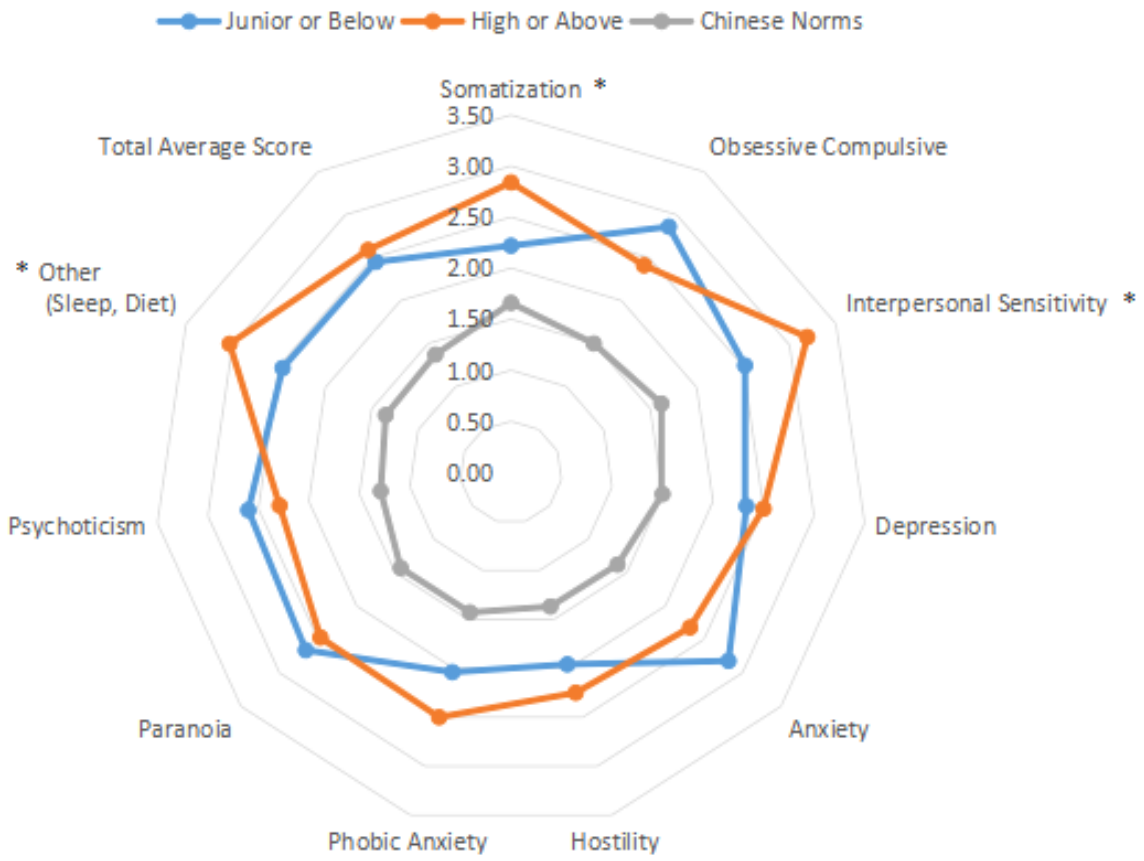
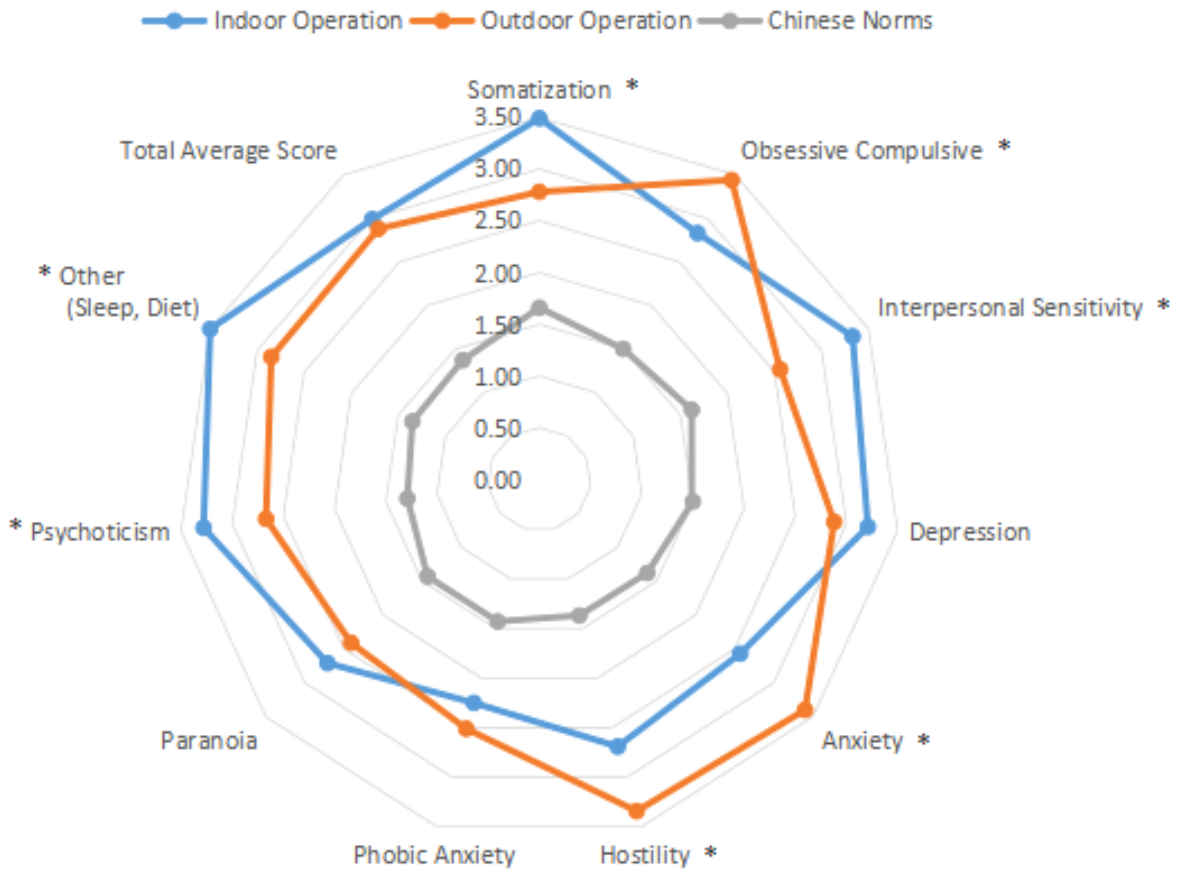


Figure 13. Intragroup comparison of working environment. *P<.05.



Intergroup Comparison

A comparison was made between the AR group and the Chinese norm, and the positive factors of the SCL-90 were extracted (Table 2). Among 827 patients, 124 (15%) had no mental health impairments, 176 (21.3%) had mild impairments, 474 (57.3%) had mild to moderate impairments, 41 (5%) had moderate to severe impairments, and 12 (1.4%) had severe impairments. The average score of the SCL-90 for all 827 patients was 2.64 (SD 0.25), which corresponded to the mild to moderate mental health impairments.

Table 2. Comparison between the AR group (study group) and the Chinese norm (healthy control group).

	AR[a], mean (SD)	Chinese norm, mean (SD)	*Z* score	*P* value
Somatization	2.75 (0.43)	1.66 (0.28)	–3.15	>.99
Obsessive-compulsive	2.71 (0.22)	1.50 (0.89)	–3.87	.13
Interpersonal sensitivity	2.83 (0.37)	1.62 (0.61)	–3.99	.10
Depression	2.80 (0.51)	1.50 (0.59)	–5.51	*.01* [b]
Anxiety	2.72 (0.40)	1.38 (0.43)	–6.03	*.002*
Hostility	2.39 (0.36)	1.37 (0.48)	–2.11	>.99
Phobic anxiety	2.20 (0.31)	1.43 (0.58)	–0.98	>.99
Paranoia	2.49 (0.62)	1.43 (0.57)	–2.74	>.99
Psychoticism	2.54 (0.56)	1.29 (0.42)	–4.61	*.02*
Other (sleep, diet)	2.95 (0.47)	1.35 (0.39)	–8.17	*<.001*
Total average score	2.64 (0.25)	1.37 (0.48)	–5.09	*.02*

[a]AR: allergic rhinitis.

[b]Italics indicate statistical significance at the <.05 level.

Partial Correlation Analysis

Partial correlation analysis was performed on the 4 positive factors of the SCL-90 and the 4 symptoms of AR (Table 3). First, the VAS was used to evaluate the 4 AR symptoms separately. Nasal itching had 7 points, sneezing 8 points, clear discharge 8 points, and nasal congestion 6 points. Second, partial correlation analysis was made between the scores of the 4 symptoms and the 4 positive factors in Table 2 through the multiple regression statistical method by taking measurement data like BMI, age, monthly salary, and disease duration as the control variables.

Table 3. Partial correlation analysis between 4 positive factors of the SCL-90 and 4 symptoms of allergic rhinitis.

Positive factors of SCL-90[a]	Nasal itching		Sneezing		Clear discharge		Nasal congestion	
	r	*P* value	*r*	*P* value	*r*	*P* value	*r*	*P* value
Depression	0.13	.27	0.37	*.02* [b]	0.47	*.001*	0.05	.70
Anxiety	0.42	*.01*	0.11	.33	0.20	.08	0.43	*.01*
Psychoticism	0.33	*.03*	0.40	*.01*	0.17	.14	0.18	.20
Other (sleep, diet)	0.15	.25	0.13	.26	0.30	*.04*	0.32	*.03*

[a]SCL-90: Symptom Checklist 90.

[b]Italics indicate statistical significance at the <.05 level.

Discussion

Principal Findings

Zheng et al [15] pointed out that 69.45% of patients with AR have depressive tendencies. While treating rhinitis, doctors should give "regular and habitual" comfort and follow-ups, and if necessary, should refer patients to psychologists. Muñoz-Cano et al [16] found by using the self-rating anxiety scale that 58.77% of AR patients' anxiety could not be ignored, and their life happiness index decreased with a prolonged course of the disease. If patients did not receive active "packaged treatment" (understanding, care, counseling, therapy, follow-up) [17], they would turn from "unhealthily free" to "unfreely anxious" [18] and, in extreme cases, might even induce violence [19]. From 2011 to 2014, a total of 4 chief otolaryngologists were killed with a knife by patients in China. Therefore, it is imperative to strictly follow the "biological-psychological-social" international medical model and to have a comprehensive understanding of the association between AR and psychological disorders.

Comparison to Prior Work

In the existing literature, despite many in-depth studies, there are few comprehensive studies in this regard due to the following reasons. First, part of the literature only made one-on-one retrospective analysis or targeted meta-analysis of AR with depression, anxiety, hostility, and other psychological disorders [20], while few used the SCL-90 scale for a comprehensive 10-factor study. Second, although some literature used the SCL-90 scale, they only investigated the patients within respective hospitals [21], resulting in small sample sizes that could only reflect the status of a district rather than a city. Third, some literature surveyed multiple hospitals with the SCL-90 scale, but the research teams were limited to the otolaryngology department only, where the cases were often collected by nurses and the doctors were responsible for the diagnosis, treatment, and final data analysis [22]. Such results would appear to be less professional and authoritative because of the lack of cooperation from psychologists and psychiatrists.

Therefore, this study addresses the above 3 shortcomings and makes the following 3 optimizations on the research design. First, the SCL-90 scale was introduced systematically, which has a good discriminating function as it includes 90 items such as consciousness, emotion, diet/sleep, sensation, and lifestyle, and comprehensively reflects the distribution characteristics of 10 psychological symptoms with 10 factors [23]. Second, this study specifically covers the 5 most renowned hospitals in Chengdu, China, which not only enlarges the sample size but also takes into account the geographical locations in all directions so that the data are more detailed and can more accurately reflect the real-world research of a city. Third, this study is fortunately sponsored by two National Natural Science Foundations, so we were able to invite psychologists and psychiatrists to carry out the experimental design and data analysis together, thus making the team configuration better and the results more authoritative.

Limitations

To our knowledge, this is the first multicenter and interdisciplinary study of patients with AR. However, there are some limitations in this study. First, a few patients inevitably pretend to cooperate when filling out the questionnaire, which will affect the accuracy of the data. Second, psychological factors are greatly influenced by subjective factors, and the psychological state of patients at the time of filling in the form cannot reflect their later psychological state. Third, this study took place before the COVID-19 pandemic, but unfortunately, epidemic factors were not included in the study as a single factor. Since the outbreak of COVID-19 in Chengdu, China, the number of patients with mental diseases has increased greatly.

Conclusions

The intergroup comparisons in this study were performed on 827 patients with AR in China's Chengdu City versus the Chinese norm, which reveal mild to moderate psychological disorders. The scores for 4 factors, namely, depression, anxiety, psychosis, and other (sleep, diet), are statistically higher than the Chinese norm, indicating that the psychological disorders

develop mainly into these 4 types among the patients with AR. Therefore, otolaryngologists should pay close attention to the above four psychological states of patients when assessing the curative efficacy of somatopathies. Specifically, this study comprises 9 intragroup comparisons and 2 intergroup comparisons. Among the 9 classification criteria for intragroup comparisons, the living environment (urban or rural) yielded 8 positive results; working environment (indoors or outdoors) yielded 7 positive results; gender yielded 4 positive results; age, marital status, and education level yielded 3 positive results each; and BMI, monthly salary, and disease duration yielded 2 positive results each. This demonstrates the influencing factors of AR patients' physical and mental health and their proportions, as well as the importance of the "biological-psychological-social" medical model. Meanwhile, these results also remind the otolaryngologists in Chengdu to offer more care and follow-ups to the following target positive populations: women and those with abnormal BMI, aged ≥45 years, with a monthly salary <¥5110 (US $700), with a disease duration <13 years, residing in urban areas, with a high school or above education, and who are indoor laborers.

Acknowledgments

This research has been funded by two National Natural Science Foundations of China, both managed by Chengdu University of Traditional Chinese Medicine, Chengdu, Sichuan, China.

Conflicts of Interest

None declared.

References

1. Head K, Snidvongs K, Glew S, Scadding G, Schilder AG, Philpott C, et al. Saline irrigation for allergic rhinitis. Cochrane Database Syst Rev 2018 Jun 22;6:CD012597 [FREE Full text] [doi: 10.1002/14651858.CD012597.pub2] [Medline: 29932206]
2. Baena-Cagnani CE, Canonica GW, Zaky Helal M, Gómez RM, Compalati E, Zernotti ME, ISMAR Study Group. The international survey on the management of allergic rhinitis by physicians and patients (ISMAR). World Allergy Organ J 2015;8(1):10 [FREE Full text] [doi: 10.1186/s40413-015-0057-0] [Medline: 25977744]
3. Lee M, Kim Y, Lee JA. Okbyungpoongsan (Yupingfeng) for treating allergic rhinitis: a protocol for the systematic review of controlled trials. Medicine (Baltimore) 2018 Nov;97(45):e13227. [doi: 10.1097/MD.0000000000013227] [Medline: 30407365]
4. Cheng L, Chen J, Fu Q, He S, Li H, Liu Z, et al. Chinese society of allergy guidelines for diagnosis and treatment of allergic rhinitis. Allergy Asthma Immunol Res 2018 Jul;10(4):300-353 [FREE Full text] [doi: 10.4168/aair.2018.10.4.300] [Medline: 29949830]
5. Tian F, Li H, Tian S, Yang J, Shao J, Tian C. Psychological symptoms of ordinary Chinese citizens based on SCL-90 during the level I emergency response to COVID-19. Psychiatry Res 2020 Jun;288:112992 [FREE Full text] [doi: 10.1016/j.psychres.2020.112992] [Medline: 32302816]
6. Wei Y, Li H, Wang H, Zhang S, Sun Y. Psychological status of volunteers in a phase I clinical trial assessed by Symptom Checklist 90 (SCL-90) and Eysenck Personality Questionnaire (EPQ). Med Sci Monit 2018 Jul 17;24:4968-4973 [FREE Full text] [doi: 10.12659/MSM.909524] [Medline: 30015333]
7. Li X, Song W, Zhang JY, Lu C, Wang YX, Zheng YX, et al. Factors associated with mental health of graduate nursing students in China. Medicine (Baltimore) 2021 Jan 22;100(3):e24247. [doi: 10.1097/MD.0000000000024247] [Medline: 33546046]
8. Mendola J, Lam J, Rosenstein M, Lewis L, Shmuel A. Partial correlation analysis reveals abnormal retinotopically organized functional connectivity of visual areas in amblyopia. Neuroimage Clin 2018;18:192-201 [FREE Full text] [doi: 10.1016/j.nicl.2018.01.022] [Medline: 29868445]
9. Rasero J, Amoroso N, La Rocca M, Tangaro S, Bellotti R, Stramaglia S, Alzheimer's Disease Neuroimaging Initiative. Multivariate regression analysis of structural MRI connectivity matrices in Alzheimer's disease. PLoS One 2017;12(11):e0187281 [FREE Full text] [doi: 10.1371/journal.pone.0187281] [Medline: 29135998]
10. Beyene HB, Olshansky G, T Smith AA, Giles C, Huynh K, Cinel M, et al. High-coverage plasma lipidomics reveals novel sex-specific lipidomic fingerprints of age and BMI: evidence from two large population cohort studies. PLoS Biol 2020 Sep;18(9):e3000870 [FREE Full text] [doi: 10.1371/journal.pbio.3000870] [Medline: 32986697]
11. Martin-Kerry JM, Knapp P, Atkin K, Bower P, Watt I, Stones C, et al. Supporting children and young people when making decisions about joining clinical trials: qualitative study to inform multimedia website development. BMJ Open 2019 Jan 09;9(1):e023984 [FREE Full text] [doi: 10.1136/bmjopen-2018-023984] [Medline: 30782720]
12. Fan M, Pena AA. Do minimum wage laws affect those who are not covered? Evidence from agricultural and non-agricultural workers. PLoS One 2019;14(10):e0221935 [FREE Full text] [doi: 10.1371/journal.pone.0221935] [Medline: 31577803]

13. Zhang J, Yuan X, Tan X, Zhang X. Delineation of the urban-rural boundary through data fusion: applications to improve urban and rural environments and promote intensive and healthy urban development. Int J Environ Res Public Health 2021 Jul 05;18(13):7180 [FREE Full text] [doi: 10.3390/ijerph18137180] [Medline: 34281117]
14. Chen T, Tian D, Deng P, Zhou E, Huang J. Study on instant delivery service riders' safety and health by the effects of labour intensity in China: a mediation analysis. Front Public Health 2022;10:907474. [doi: 10.3389/fpubh.2022.907474] [Medline: 35812478]
15. Zheng Z, Sun Z, Zhou X, Zhou Z. Efficacy of Chinese herbal medicine in treatment of allergic rhinitis in children: a meta-analysis of 19 randomized controlled trials. J Int Med Res 2018 Oct;46(10):4006-4018 [FREE Full text] [doi: 10.1177/0300060518786905] [Medline: 30084286]
16. Muñoz-Cano R, Ribó P, Araujo G, Giralt E, Sanchez-Lopez J, Valero A. Severity of allergic rhinitis impacts sleep and anxiety: results from a large Spanish cohort. Clin Transl Allergy 2018;8:23 [FREE Full text] [doi: 10.1186/s13601-018-0212-0] [Medline: 30002811]
17. Melamed I, Heffron M. Attention deficit disorder and allergic rhinitis: are they related? J Immunol Res 2016;2016:1596828. [doi: 10.1155/2016/1596828] [Medline: 27872863]
18. Canonica GW, Triggiani M, Senna G. 360 degree perspective on allergic rhinitis management in Italy: a survey of GPs, pharmacists and patients. Clin Mol Allergy 2015;13:25 [FREE Full text] [doi: 10.1186/s12948-015-0029-5] [Medline: 26528081]
19. Amritwar AU, Lowry CA, Brenner LA, Hoisington AJ, Hamilton R, Stiller JW, et al. Mental health in allergic rhinitis: depression and suicidal behavior. Curr Treat Options Allergy 2017 Mar;4(1):71-97 [FREE Full text] [doi: 10.1007/s40521-017-0110-z] [Medline: 28966902]
20. Kim DH, Han K, Kim SW. Relationship between allergic rhinitis and mental health in the general Korean adult population. Allergy Asthma Immunol Res 2016 Jan;8(1):49-54 [FREE Full text] [doi: 10.4168/aair.2016.8.1.49] [Medline: 26540501]
21. Mir E, Panjabi C, Shah A. Impact of allergic rhinitis in school going children. Asia Pac Allergy 2012 Apr;2(2):93-100 [FREE Full text] [doi: 10.5415/apallergy.2012.2.2.93] [Medline: 22701858]
22. Lee YS, Kim SH, You JH, Baek HT, Na C, Kim BN, et al. Attention deficit hyperactivity disorder like behavioral problems and parenting stress in pediatric allergic rhinitis. Psychiatry Investig 2014 Jul;11(3):266-271 [FREE Full text] [doi: 10.4306/pi.2014.11.3.266] [Medline: 25110499]
23. Li P, Wang F, Ji G, Miao L, You S, Chen X. The psychological results of 438 patients with persisting GERD symptoms by Symptom Checklist 90-Revised (SCL-90-R) questionnaire. Medicine (Baltimore) 2018 Feb;97(5):e9783. [doi: 10.1097/MD.0000000000009783] [Medline: 29384872]

Abbreviations

AR: allergic rhinitis
SCL-90: Symptom Checklist 90
VAS: visual analog scale
WHO: World Health Organization

Edited by A Mavragani; submitted 07.02.22; peer-reviewed by K Uludag, Z Jin, S Mallick; comments to author 20.07.22; revised version received 13.09.22; accepted 27.09.22; published 10.11.22.

Please cite as:
Huang H, Wang Y, Zhang L, Zhang Q, Wu X, He H
Psychological Disorders of Patients With Allergic Rhinitis in Chengdu, China: Exploratory Research
JMIR Form Res 2022;6(11):e37101
URL: https://formative.jmir.org/2022/11/e37101
doi:10.2196/37101
PMID:36355407

Original Paper

Matched Endoscopic Sleeve Gastroplasty and Laparoscopic Sleeve Gastrectomy Cases: Formative Cohort Study

Skye Marshall[1,2], PhD; Graeme G Rich[3], MBBS; Felicity Cohen[4], BA; Asha Soni[4], BNurs; Elizabeth Isenring[1], PhD

[1]Bond University Nutrition & Dietetics Research Group, Faculty of Health Sciences & Medicine, Bond University, Gold Coast, Australia

[2]Research Institute for Future Health, Gold Coast, Australia

[3]Australian National University, Sydney, Australia

[4]Weightloss Solutions Australia, Varsity Lakes, Australia

Corresponding Author:
Skye Marshall, PhD
Bond University Nutrition & Dietetics Research Group
Faculty of Health Sciences & Medicine
Bond University
University Drive
Gold Coast, 4116
Australia
Phone: 61 0411166230
Email: skye_marshall@bond.edu.au

Abstract

Background: Bariatric weight-loss surgery rates are increasing internationally. Endoscopic sleeve gastroplasty (ESG) is a novel, minimally invasive endoscopic procedure thought to mimic some of the effects of a more common surgery, laparoscopic sleeve gastrectomy (LSG). Patient factors affecting procedural choice are unexplored.

Objective: This formative study aimed to determine the preoperative and early postoperative characteristics of adults matched for age, sex, and BMI who chose ESG versus LSG.

Methods: This prospective cohort study recruited ESG and matched LSG adults in Australia. Preoperative outcomes were medical history, glycemic biomarkers, blood lipids, liver function enzymes, albumin, blood pressure, hepatic steatosis index, the Gastrointestinal Symptom Rating Scale, the Impact of Weight on Quality of Life–Lite questionnaire, and body composition via dual-energy x-ray absorptiometry. Adverse events were recorded preoperatively and up to 2 weeks postoperatively. SPSS was used to test if there were differences between cohorts by comparing means or mean ranks, and binary regression was used to understand how characteristics might predict procedure choice.

Results: A total of 50 (including 25 ESG and 25 LSG) patients were recruited, who were primarily White (45/50, 90%) and female (41/50, 82%) with a mean age of 41.7 (SD 9.4) years. Participants had a mean of 4.0 (SD 2.2) active comorbid conditions, with the most common being nonalcoholic fatty liver disease (38/50, 76%), back pain (32/50, 64%), anxiety or depression (24/50, 48%), and joint pain (23/50, 46%). The LSG cohort had higher hemoglobin A_{1c} (5.3%, SD 0.2%) than the ESG cohort (5%, SD 0.2%; *P*=.008). There was a 2.4 kg/m^2 difference in median BMI (*P*=.03) between the groups, but fat and fat-free mass had no meaningful differences. Comparing the LSG and ESG groups showed that the LSG group had lower total quality of life (49.5%, SD 10.6% vs 56.6%, SD 12.7%; *P*=.045), lower weight-related self-esteem (10.7%, IQR 3.6%-25% vs 25%, IQR 17.9%-39.3%; *P*=.02), and worse abdominal pain (38.9%, IQR 33.3%-50% vs 53.9%, SD 14.2%, *P*=.01). For every percent improvement in weight-related self-esteem, the odds for selecting ESG increased by 4.4% (95% CI 1.004-1.085; *P*=.03). For every percent worsening in hunger pain, the odds for selecting ESG decreased by 3.3% (95% CI 0.944-0.990; *P*=.004).

Conclusions: There was very little evidence that Australian adults who chose an endoscopic versus surgical sleeve had different rates of comorbidities, body fat percentage, or weight-related quality of life. There was evidence against the test hypothesis, that is, there was evidence suggesting that lower self-esteem predicted choosing a more invasive sleeve (ie, LSG rather than ESG)

Trial Registration: Australia New Zealand Clinical Trials Registry ACTRN12618000337279; https://anzctr.org.au/Trial/Registration/TrialReview.aspx?id=374595

(JMIR Form Res 2022;6(11):e29713) doi:10.2196/29713

KEYWORDS

endoscopic sleeve gastroplasty; laparoscopic sleeve gastrectomy; obesity; bariatric surgery; interdisciplinary research; cohort study; metabolic surgery; weight loss; comorbidity; body composition; surgery; gastroplasty; endoscopic surgery; gastroenterology; preoperative; postoperative; outcome; body mass index; sex; age; gastrointestinal; prospective

Introduction

Despite preventative public health policies enacted by governments around the world, the global prevalence of obesity in adults (defined as BMI ≥30 kg/m^2) has tripled since 1975, leading to an estimated 650 million individuals with obesity in 2016 [1]. Obesity is a principal metabolic risk factor for chronic comorbidities, including type 2 diabetes mellitus, cardiovascular disease, and obstructive sleep apnea [2]. Diet, exercise, and pharmacotherapy have limited long-term efficacy for weight loss; thus, the number of individuals who elect to undergo bariatric surgery, that is, surgical treatment for obesity, is increasing [3,4]. Bariatric surgery has been demonstrated to result in clinically meaningful weight loss over 1 to 5 years, with studies showing the maintenance of health improvements for up to 10 years [5,6]. Studies have also demonstrated that patients with diabetes, hyperlipidemia, hypertension, and obstructive sleep apnea experience improvement and, in some cases, complete resolution of the comorbidity [7,8].

The United States has the highest number of bariatric procedures performed each year; in 2016, there were 216,000 procedures performed [9]. Other countries with high numbers of bariatric procedures are Brazil (97,480 in 2014), France (46,960 in 2014), Argentina (36,668 in 2014), and Australia (21,043 in 2018-2019) [10,11]. Worldwide, 96% of surgeries are performed laparoscopically; the most common form of the surgery in Australia is laparoscopic sleeve gastrectomy (LSG), which comprises 71.5% of all procedures [11,12].

Despite demonstrating good weight loss efficacy, bariatric surgeries such as LSG have relatively high rates of adverse events (10%-17%), postoperative mortality (0.3%), failure (10%-20%), and weight regain (20%-30%) [13-17]. Additionally, many patients are unwilling to undergo an invasive, incisional, resective surgery, precluding them from an otherwise effective treatment. Endoscopic sleeve gastroplasty (ESG), which is minimally invasive and incisionless, is designed to be analogous to traditional LSG and is being increasingly utilized worldwide [18].

ESG is performed utilizing an endoscopic suturing system (Overstitch, Apollo Endosurgery), coupled with a double-channel flexible gastroscope (GIF2T-180 series, Olympus Optical). These instruments are passed orally into the stomach to allow for full-thickness sutures to permanently plicate the gastric lumen into a narrow “sleeve-like” tubular configuration from the incisura angularis to the fundus [18].

Studies investigating the mechanisms of action of ESG suggest that reduction in gastric volume and improved satiety through delayed gastric emptying contribute to decreased caloric intake and weight loss [5,19]. Available studies report a markedly lower postprocedure complication rate compared to traditional surgical procedures [5]. A chart audit found that among patients with similar preoperative BMI, sex, and age, the ESG cohort had lower weight loss at 12 months but also had fewer complications and a shorter hospital stay [20].

Procedure selection is a complex process involving both medical recommendations and patient preferences. It is unknown if patient demographic or medical characteristics beyond weight loss targets and the risk of complications are associated with procedure choice (ie, surgical vs endoscopic sleeve), which may limit the patient-centeredness of care and interpretation of outcomes. A recent Australian study of unmatched ESG and LSG cohorts suggested factors including preprocedural BMI, quality of life, and gastrointestinal symptoms may play a role [21]. The present formative Australian study aimed to determine the preoperative and early postoperative characteristics of adults matched for age, sex, and BMI who chose ESG versus LSG.

Methods

This is a substudy of a larger prospective cohort study (Universal Trial Number U1111-1216-8678) undertaken and reported according to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) checklist [22] and registered prospectively on March 6, 2018, at the Australia New Zealand Clinical Trials Registry (ACTRN12618000337279). The results of the larger study have been published elsewhere [21].

Ethical Considerations

This study received ethical approval from the Bond University Human Research Ethics Committee (SM02936); written informed consent was obtained from the included patients by the research team.

Eligibility Criteria, Matched Cases, and Recruitment

Patients were consecutively recruited from a privately funded outpatient medical clinic in Queensland, Australia, that offered both ESG and LSG procedures. Approximately 500 patients presented to the outpatient clinic during the recruitment phase. Adults aged ≥18 years undergoing either an ESG or LSG procedure from June 2018 to May 2019 were eligible to enroll. Patients unwilling to undergo a dual-energy x-ray absorptiometry (DXA) scan at the local study site were excluded.

Procedural selection for ESG or LSG occurred after a medical consultation with the proceduralists (surgeons or gastroenterologists) that considered the merits of the available procedures, the risks and benefits specific to the patient, and the preference of the patient. Eligibility considerations for the procedures can be found in Multimedia Appendix 1, Table S1.

Due to higher case numbers for LSG than ESG, all eligible ESG and LSG patients were recruited, and each ESG patient was then matched against the available pool of LSG patients for age, sex, and BMI. LSG patients who were not matched against an

ESG patient were excluded from inferential data analysis. Matching of patients was blinded to their characteristics and outcomes (other than age, sex, and BMI). Usual care conditions are reported in Multimedia Appendix 1.

Outcomes

The outcomes of this study were demographic characteristics, body composition, blood pressure, glycemic measures, blood lipids, liver function enzymes, albumin, hepatic steatosis, gastrointestinal symptoms, weight-related quality of life, and adverse events. All outcomes were measured preoperatively, except for perioperative characteristics, which were measured on the day of the procedure, and adverse events, which were measured perioperatively and up to 14 days postoperatively.

Data Collection Tools

Survey-style data collection tools were self-completed by participants, with all other data being recorded from the study site medical progress notes, pathology reports, and consultation letters or letters of referral from the patient's general practitioner.

Participant Demographic and Preoperative Medical Characteristics

Participant characteristics were recorded from medical records and baseline interviews; they included age, sex, diagnoses of obesity-related comorbidities (including, but not limited to, type 2 diabetes mellitus or prediabetes, hypertension, dyslipidemia, obstructive sleep apnea or use of a continuous positive airway pressure device, osteoarthritis, nonalcoholic fatty liver disease, weight-related joint pain, depression or anxiety, gestational diabetes, polycystic ovary syndrome, and gastroesophageal reflux disease), ethnicity (according to the Australian Bureau of Statistics categorizations), and area of residence (rural or metropolitan). Blood pressure was measured by a registered nurse at the patient's first preoperative appointment at the study site.

Preoperative Body Composition

Measurements of weight, height, and body composition (with DXA) were obtained by trained research assistants at the Bond University Institute of Health and Sport (Robina, Queensland, Australia) using calibrated scales (Wedderburn WM204), a wall-mounted stadiometer with high speed counter (Harpenden Model 602VR; Holtain Limited), and the Lunar Prodigy DXA (Encore Version 14.10.022; GE Medical Systems Lunar), respectively. The assessments were made in a rested, fasting state; participants wore minimal, well-fitting clothing and no jewelry, had their hair down, and had a voided bladder. Height was measured with participants standing with their back to the wall, hands by their sides, feet together, and head in the frontal plane. The height of the board was placed after a breath inhalation and recorded. Fat mass, fat-free mass, and bone mineral content were recorded for the total body, trunk, limbs, and android and gynoid regions. Due to larger body sizes, the DXA scan was implemented by scanning the left and right sides of the body in 2 separate scans to simulate a whole-body scan, rather than using estimated limb mass. For those who failed to attend the DXA scanning session, height and weight were recorded from the patients' medical records, obtained at their first appointment at the study site, and thus comprised data that were measured, but not calibrated or standardized.

Preoperative Comorbid Measures

Systolic and diastolic blood pressure (mm Hg); fasting lipid profile, including total cholesterol, low-density lipoprotein (LDL) cholesterol, high-density lipoprotein (HDL) cholesterol, and triglycerides (mmol/L); fasting serum glucose (mmol/L); hemoglobin A_{1c} (HbA_{1c}; presented as a percentage); albumin (g/L); aspartate aminotransferase (units/L); and alanine aminotransferase (units/L) were measured routinely prior to the procedure and were obtained from pathology reports. The hepatic steatosis index (HSI) was calculated using pathology, sex, and BMI and coded as no nonalcoholic fatty liver disease (NAFLD) present when the HSI was <30.0 or NAFLD present when the HSI was >36.0 [23].

Preoperative Gastrointestinal Symptoms

Gastrointestinal symptoms, including abdominal pain, reflux, diarrhea, indigestion, and constipation, were evaluated with the Gastrointestinal Symptom Rating Scale (GSRS) [24], a non–disease-specific tool commonly used to evaluate gastrointestinal symptoms following bariatric surgery [25-27]. Completed at the time of recruitment, the GSRS is a 15-item questionnaire that asks participants to rate the symptoms they have experienced in the past week on a Likert scale ranging from 1 (no symptoms) to 7 (severe/frequent symptoms). Subcategories of abdominal pain, reflux, indigestion, constipation, and diarrhea were also calculated. Total and subcategory scores were normalized and reported on a scale of 0 (worst symptoms) to 100 (no symptoms).

Preoperative Weight-Related Quality of Life

Weight-related quality of life was measured via the Impact of Weight on Quality of Life–Lite (IWQOL-Lite) tool [28] at the time of recruitment. The IWQOL-Lite is a 31-item self-reported measure of 5 domains that affect obese individuals: physical function, self-esteem, sexual life, public distress, and work. Each item reflects experiences in the past week, ranging from "never true" to "always true." The tool provides a score for each of these domains, as well as a total score, each normalized to a scale of 0 (worst quality of life) to 100 (highest quality of life). The tool is frequently used in studies evaluating bariatric surgery [29,30] and has shown strong retest reliability, internal consistency, and correlation with general quality of life measures, including the 36-Item Short Form Health Survey [28].

Peri- and Postoperative Adverse Events

All adverse events recorded in medical records were noted and categorized as (1) directly related to the procedure, (2) possibly related to the procedure, or (3) not related to the procedure. Events were further categorized as (1) minor, (2) moderate, or (3) severe, according to the National Institutes of Health guidelines [31]. Severe adverse events were those considered life-threatening, that is, those resulting in death, inpatient hospitalization or prolongation of existing hospitalization, persistent or significant disability or reduced capacity, development of short-bowel syndrome, surgical intervention,

or otherwise medically significant events. Moderate adverse events were those that were not considered significant but still required intervention by the medical team, such as prescription of antibiotics, analgesics, or antiemetics beyond the standard postoperative inpatient protocol. Minor adverse events were those that required no intervention beyond reassurance or modified lifestyle.

Statistical Analysis

All variables were assessed based on the null hypothesis, that is, "there is no difference between 2 independent groups"; therefore, the comparative test for each variable was selected based on how the data met test assumptions. Differences between the ESG and matched and unmatched LSG cases were tested using a chi-square analysis for categorical variables or the Fisher exact test if >20% of cells had an expected count <5. To compare continuous variables between groups, an independent *t* test was used. However, if the variable failed the Levene test for homogeneity of variances, the Welch *t* test was selected. If the variable was highly skewed with the same distribution shapes in both groups, the Mann-Whitney *U* test was used to compare medians. Otherwise, if the variable was highly skewed, with different distribution shapes between groups, the Mann-Whitney *U* test was used to compare rank means. Normality was evaluated using the Shapiro-Wilk test and inspection of histograms and normal Q-Q plots.

The 2 Likert scale–based surveys (ie, the GSRS and IWQOL-Lite) were tested for internal consistency with the Cronbach α, and 95% CIs were generated via the intraclass correlation coefficient [32,33]. Outcome variables that were found to be meaningfully different between procedure types were explored using binomial logistic regression to understand their ability to predict procedure choice. Analysis was performed using SPSS (version 27; IBM Corp).

Results

Medical, Nutritional, and Procedural Characteristics of Participants

This study recruited 25 ESG and 56 LSG patients. The 25 ESG patients were matched against the available LSG patients, leading to a final sample of 50 patients, including 25 with ESG and 25 with LSG. Of the matched participants, 9 with ESG and 4 with LSG cancelled their procedure. The cancellation rate was similar between the procedures (*P*=.20). There were no meaningful differences in the baseline characteristics of participants who underwent the procedure and those who did not (Multimedia Appendix 1, Table S3). The reasons for procedure cancellation were financial (1 LSG and 2 ESG patients), changing to a noneligible procedure (1 ESG and 1 LSG patient), personal commitments (1 ESG patient), choosing to delay the procedure (1 ESG patient), no longer wanting any procedure (2 LSG patients), or unexplained (4 ESG patients).

The 56 recruited LSG participants differed from the 25 ESG participants in preoperative BMI (ESG: median 33.4, IQR 30.9-36.8 kg/m^2; LSG: median 39.6, IQR 35.8-44.4 kg/m^2; *P*<.001), osteoarthritis, HbA_{1c}, and HDL cholesterol (Tables 1 and 2). After matching, participants were primarily White (45/50, 90%) and female (41/50, 82%), with a mean age of 41.7 (SD 9.4) years. Participants had a mean of 4.0 (SD 2.2) active comorbid conditions, with the most common being HIS-identified NAFLD (38/50, 76%), back pain (32/50, 64%), anxiety or depression (24/50, 48%), and osteoarthritis or joint pain (23/50, 46%) (Table 1). Participants did not have high rates of cardiovascular or type 2 diabetes risk factors and had normal blood pressure (mean 126/84.8, SD 12.5/9.0 mm Hg), fasting blood glucose (mean 5.2, SD 1.1 mmol/L), LDL cholesterol (mean 3.2, SD 0.7 mmol/L), and triglycerides (mean 1.4, SD 0.8 mmol/L) (Table 2).

Table 1. Preoperative characteristics of endoscopic sleeve gastroplasty, matched laparoscopic sleeve gastrectomy, and unmatched laparoscopic sleeve gastrectomy patients.

Characteristics	Unmatched LSG[a] (N=56)	ESG[b] (N=25)	Matched LSG (N=25)	Matched LSG vs ESG, *P* value	Unmatched LSG vs ESG, *P* value
Age (years), mean (SD)	40.4 (8.8)	42.8 (9.9)	40.7 (8.9)	.30	.27
Female, n (%)	48 (86)	20 (80)	21 (84)	.71	.52
Ethnicity, n (%)				.39	.28
White	50 (89)	22 (88)	23 (92)		
Asian	0 (0)	2 (8)	0 (0)		
Black	1 (2)	0 (0)	1 (4)		
Indigenous Australian	2 (4)	0 (0)	0 (0)		
Pacific Islander	1 (2)	0 (0)	0 (0)		
Not disclosed	2 (4)	1 (4)	1 (4)		
Rural dwelling, n (%)	3 (5)	3 (12)	0 (0)	.24	.29
Systolic blood pressure (mm Hg), mean (SD) or median (IQR)	120.0 (110.0-131.0)[c]	126.0 (11.4)	121.3 (13.5)	.19	.25
Diastolic blood pressure (mm Hg), mean (SD) or median (IQR)	82.0 (78.0-90.0)[c]	84.3 (10)	82.7 (10.3)	.57	.70
Number of comorbidities[d], mean (SD) or median (IQR)	4.0 (3.0-5.0)[c]	3.7 (2.1)	4.3 (2.3)	.33	.40
Type 2 diabetes mellitus, n (%)	2 (4)	0 (0)	1 (4)	.50	.48
Hypertension, n (%)	16 (29)	9 (36)	6 (24)	.36	.50
Dyslipidemia, n (%)	9 (16)	7 (28)	4 (16)	.25	.21
Obstructive sleep apnea, n (%)	7 (13)	0 (0)	3 (12)	.24	.07
Osteoarthritis/joint pain, n (%)	41 (73)	11 (44)	12 (48)	.78	.01
Nonalcoholic fatty liver disease, n (%)	4 (7)	2 (8)	2 (8)	.70	.61
Polycystic ovary syndrome[e], n (%)	9 (18)	2 (10)	3 (15)	.86	.32
Gastroesophageal reflux disease, n (%)	22 (39)	9 (36)	11 (44)	.56	.78
Depression or anxiety, n (%)	23 (41)	12 (48)	12 (48)	>.99	.56
Gestational diabetes mellitus[e], n (%)	9 (18)	4 (20)	3 (15)	.83	.56
Impaired fasting glucose, n (%)	1 (2)	0 (0)	0 (0)	N/A[f]	.69
Back pain, n (%)	41 (73)	15 (60)	17 (68)	.56	.19
Asthma, n (%)	19 (34)	6 (24)	10 (40)	.23	.37

[a]LSG: laparoscopic sleeve gastrectomy.

[b]ESG: endoscopic sleeve gastroplasty.

[c]Data presented as median (IQR); other values in this row are mean (SD).

[d]Comorbidities were defined as any currently active chronic disease or syndrome, excluding sporadic conditions (eg, migraines) and allergic symptoms (eg, rhinitis).

[e]Males were excluded from these analyses.

[f]N/A: not applicable.

Table 2. Preoperative biochemistry of endoscopic sleeve gastroplasty, matched laparoscopic sleeve gastrectomy, and unmatched laparoscopic sleeve gastrectomy patients[a].

Preoperative biochemistry	Unmatched LSG[b] (N=49)	ESG[c] (N=19)	Matched LSG (N=21)	Matched LSG vs ESG, *P* value	Unmatched LSG vs ESG, *P* value
Fasting blood glucose (mmol/L), mean (SD) or median (IQR)	5.2 (4.6-5.7)[d]	5.1 (0.4)	4.9 (4.6-5.5)[d]	.30	.77
HbA_{1c} (%), mean (SD) or median (IQR)	5.2 (5.0-5.3)[d]	5.0 (0.2)	5.3 (0.2)	.008	.004
Total cholesterol (mmol/L), mean (SD)	5.2 (0.7)	5.4 (1.1)	5.5 (0.4)	.10	.19
Low-density lipoprotein cholesterol (mmol/L), mean (SD)	3.3 (0.7)	3.2 (0.8)	3.4 (0.5)	.35	.96
High-density lipoprotein cholesterol (mmol/L), mean (SD) or median (IQR)	1.1 (1.0-1.5)[d]	1.5 (0.5)	1.3 (0.3)	.06	.02
Triglycerides (mmol/L), median (IQR)	1.3 (0.9-1.9)	1.1 (0.6-2.2)	1.0 (0.77-1.9)	.76	.62
Alanine aminotransferase (units/L), median (IQR)	28.0 (21.0-46.5)	30.0 (20.0-40.0)	27 (19.5-40.0)	.70	.70
Aspartate aminotransferase (units/L), mean (SD) or median (IQR)	23.0 (18.3-30.8)[d]	28.3 (11.5)	22.5 (16.8-33.8)[d]	.22	.22
Albumin (g/L), mean (SD) or median (IQR)	40.5 (37.3-42.0)[d]	40.0 (20.0-42.0)[d]	40.7 (4.1)	.36	.92
Hepatic steatosis index, mean (SD)	52.6 (6.4)	45.1 (7.4)	47.8 (4.9)	.18	.002
Hepatic steatosis index–derived nonalcoholic fatty liver disease, n (%)	48 (100)	17 (89)	21 (100)	.22	.08

[a]N values differ from Table 1 because preoperative pathology measurements were not available for all participants.

[b]LSG: laparoscopic sleeve gastrectomy.

[c]ESG: endoscopic sleeve gastroplasty.

[d]Data presented as median (IQR); other values in this row are mean (SD).

Body Composition

Although all recruited patients were booked for a DXA scan, only 29 of 50 (60%) attended (there was no meaningful difference in attendance rate by procedure type; *P*=.15). A total of 3 participants, including 1 with LSG and 2 with ESG, attended the DXA scan 2 to 12 days after commencing a very low-calorie diet (VLCD). DXA scans were completed 3 to 79 days prior to surgery. Despite being matched for age, sex, and BMI, the BMI between groups was different (the difference in medians was 2.4 kg/m^2; *U*=196.0; *P*=.02; Table 3). Body composition, measured via DXA, showed that LSG participants had higher means or medians for both fat mass and fat-free mass, a clinically meaningful difference. There was a 6.7-kg difference in mean total-body fat mass (*P*=.14) and a 7.7-kg difference in mean total-body fat-free mass (*P*=.11).

Table 3. Preoperative body composition of matched endoscopic sleeve gastroplasty and laparoscopic sleeve gastrectomy patients.

Characteristics	Endoscopic sleeve gastroplasty (N=12)	Laparoscopic sleeve gastrectomy (N=17)	*P* value
BMI (kg/m^2), median (IQR)	33.4 (30.9-36.8)	35.8 (34.9-38.2)	.02
Total-body fat mass (kg), mean (SD)	46.9 (8.7)	53.6 (13.3)	.14
Total-body fat mass[a] (%), mean (SD)	47.9 (3.5)	50.0 (7.0)	.64
Total-body fat-free mass (kg), median (IQR) or mean (SD)	44.7 (42.9-50.3)[b]	52.4 (9.9)	.11
Total-body fat-free mass[a] (%), median (IQR) or mean (SD)	49.3 (3.4)	46.5 (43.7-51.0)[b]	.16
Total-body bone mineral content (g), median (IQR) or mean (SD)	2635.5 (2460.5-2787.3)[b]	2856.1 (471.0)	.23
Total-body bone mineral content (%), mean (SD)	2.8 (0.2)	2.6 (0.4)	.20
Android fat mass (kg), mean (SD)	4.5 (1.4)	5.2 (1.4)	.18
Android fat mass[c] (%), mean (SD)	55.2 (5.0)	56.7 (6.1)	.49
Gynoid fat mass (kg), mean (SD)	8.2 (1.7)	9.1 (2.7)	.33
Gynoid fat mass[c] (%), median (IQR) or mean (SD)	50.7 (4.5)	52.3 (46.6-55.9)[b]	.31
Android to gynoid fat mass ratio, mean (SD)	0.54 (0.11)	0.60 (0.18)	.36
Android to gynoid fat[c] ratio, median (IQR) or mean (SD)	1.10 (0.13)	1.10 (1.04-1.27)[b]	.57
Trunk fat mass (kg), mean (SD)	25.1 (5.7)	28.4 (6.7)	.18
Trunk fat mass[c] (%), mean (SD)	50.4 (3.5)	52.0 (5.6)	.40
Upper limb fat mass (kg), mean (SD)	4.5 (0.6)	5.5 (1.5)	.04
Upper limb fat mass[c] (%), median (IQR) or mean (SD)	47.8 (5.6)	51.7 (45.8-54.7)[b]	.51
Lower limb fat mass (kg), mean (SD)	16.3 (3.8)	18.8 (6.2)	.20
Lower limb fat mass[c] (%), median (IQR) or mean (SD)	47.8 (4.2)	49.9 (43.9-55.5)[b]	.23

[a]Percentage of total body mass.

[b]Data presented as median (IQR); other values in this row are mean (SD).

[c]Percentage of total region mass.

Gastrointestinal Symptoms and Quality of Life

The GSRS (15 items) and IWQOL-Lite (31 items) were completed 2 to 135 days prior to the surgery date, before the VLCD was commenced (except for 1 participant). The Cronbach α for both GSRS and IWQOL-Lite scores showed that they had good and very good internal consistency, respectively (GSRS: α=.848, 95% CI 0.778-0.904; IWQOL-Lite: α=.909, 95% CI 0.862-0.945). The LSG participants reported higher levels of perceived abdominal pain in the previous week than the ESG participants. The abdominal pain subcategory score encompasses being “bothered by stomach ache or pain,” “bothered by hunger pains,” and “bothered by nausea” (Table 4). Both groups reported a median score of 100 for reflux symptoms (indicating no symptoms) and minimal constipation and diarrhea. Total gastrointestinal symptoms did not meaningfully differ between groups. In the previous week, LSG participants reported worse weight-related quality of life (*P*=.045), which appeared to be primarily driven by worse weight-related self-esteem (*P*=.02; Table 4). The groups had similar results in subcategories of weight-related physical function, public distress, and work.

Table 4. Preoperative gastrointestinal function and quality of life of matched endoscopic sleeve gastroplasty and laparoscopic sleeve gastrectomy patients.

Scores	Endoscopic sleeve gastroplasty (N=25)	Laparoscopic sleeve gastrectomy (N=25)	*P* value
Gastrointestinal function scores[a]			
Abdominal pain (%), median (IQR) or mean (SD)	53.9 (14.2)	38.9 (33.3-50)[b]	.01
Reflux (%), median (IQR)	100 (83.3-100)	100 (83.3-100)	.92
Indigestion (%), mean (SD)	62.9 (20.6)	67.8 (19.4)	.39
Constipation, (%), median (IQR)	88.9 (55.6-100)	88.9 (77.8-94.4)	.89
Diarrhea (%), median (IQR)	94.4 (77.8-94.4)	94.4 (72.2-97.2)	.19
Total gastrointestinal symptom (%), mean (SD)	71.8 (15.4)	71.2 (13.7)	.88
Quality of life scores[c]			
Weight-related physical function (%), mean (SD)	57.5 (17.7)	53.3 (18.1)	.41
Weight-related self-esteem (%), median (IQR)	25 (17.9-39.3)	10.7 (3.6-25)	.02
Weight-related sexual life (%), mean (SD)	49.7 (28.3)	42 (20.5)	.28
Weight-related public distress (%), median (IQR) or mean (SD)	70 (65-95)[b]	67.2 (21.7)	.18
Weight-related work (%), mean (SD)	68.4 (19.5)	68.5 (16.3)	.99
Total weight-related quality of life (%), mean (SD)	56.6 (12.7)	49.5 (10.6)	.045

[a]Scores normalized to 0% for worst symptoms and 100% for no symptoms.

[b]Data presented as the median (IQR); other values in this row are mean (SD).

[c]Scores normalized to 0% for worst quality of life and 100% for best quality of life.

Adverse Events

Two perioperative events were noted. One ESG participant's procedure was abandoned perioperatively due to the identification of 3 large gastric ulcers. This participant was treated with pantoprazole for 8 weeks and the procedure was then rescheduled and performed. A second participant was scheduled for an ESG; however, the procedure was abandoned perioperatively due to the identification of a large hiatus hernia. This participant was rescheduled and received an LSG from the same proceduralist. There were no procedure-related peri- or postoperative adverse events.

Procedure Choice

Simple binary logistic regression was performed for factors that had relevant differences in effect size between procedures. For every percent improvement (ie, score increase) in weight-related self-esteem, the odds for selecting ESG increased by 4.4% (odds ratio [OR] 1.044, 95% CI 1.004-1.085; R^2=0.15; *P*=.03). For every percent improvement (ie, score increase) in abdominal pain, the odds for selecting ESG decreased by 7.2% (OR 0.928, 95% CI 0.873-0.987; R^2=0.183; *P*=.02).

Ad-hoc testing was performed to understand how abdominal pain predicted procedure choice. The 3 items contributing to the abdominal pain subcategory (including stomachache, hunger pain, and nausea) were further explored. Stomachache (OR 1.007, *P*=.50) and nausea (OR 1.014, *P*=.31) were found not to be associated with procedure choice, whereas for every percent worsening in hunger pain, the odds of selecting ESG decreased by 3.3% (OR 0.967, 95% CI 0.944-0.990; R^2=0.25; *P*=.004).

Discussion

Principal Findings

This is the first study to explore differences between matched participants who elected to have an LSG or the nonsurgical alternative, ESG. For the participants, who were similar to typical bariatric patients in Australia [11], only weight-related self-esteem and perceived abdominal pain were predictors of procedure choice (ie, surgical vs endoscopic sleeve). However, ad-hoc testing of how abdominal pain predicted procedure choice suggested the relationship was not clinically relevant. In the abdominal pain subgroup, increased pain predicted the choice of LSG, but item analysis found only hunger pain, not stomachache or nausea, predicted the choice of ESG.

Despite being matched for BMI, the rank sums were different between groups. This can be explained by the recruited ESG participants having a lower median BMI than the 56 recruited LSG participants. In this study, the LSG participants that we selected for matching to the ESG patients predominantly had BMIs lower than the median for the whole sample of unmatched LSG participants. This led to an unusual distribution of LSG BMIs, and a higher median and rank sum. This is confirmed by the LSG cohort having a mean fat mass and fat-free mass 6 to 8 kg higher than the ESG cohort. Although problematic in terms of BMI, the body composition of both groups was clinically similar when expressed as a percentage. Both groups presented

with high total body fat (48% for ESG and 50% for LSG participants) that was centrally located (the android to gynoid fat percent ratio was 1.10 for both ESG and LSG participants). Importantly, the difference in BMI was not a predictor of procedure choice.

Considering the strong association with obesity of chronic diseases, such as type 2 diabetes and cardiovascular disease, it was surprising that biomarker risk factors and blood pressure measures were normal and there was a low prevalence of impaired fasting glucose, hypertension, type 2 diabetes, and cardiovascular disease. This may be a reflection of the cohort having lower BMIs than typical bariatric candidates (the Australian mean is 41.8 kg/m^2 [11], compared to the median 33 to 36 kg/m^2 in this study) and the exclusion of procedures such as Roux-en-Y gastric bypass [34]. Although HbA_{1c} was higher in the LSG participants, this was not clinically relevant, as the difference was small and values were within normal ranges.

Comparison With Prior Work

The number of cumulative active comorbidities was high (a mean of 4.0 per patient), particularly HSI-identified NAFLD, back pain, and osteoarthritis, which were all present in over 50% of participants. Although the HSI diagnostic criteria for NAFLD identified extremely high rates (90%-100%), only 4 of 50 (8%) patients had a diagnosis in their medical record, aligning with previous research exposing NAFLD underdiagnosis in primary care [35]. The high prevalence of anxiety and depression is also of relevance, highlighting the importance of the psychologist as a core part of the multidisciplinary team. Recent research has found that intensive pre- and postoperative psychological intervention in bariatric surgery patients decreases postoperative symptoms of anxiety and depression compared with standard multidisciplinary care models [36]. The importance of the psychologist as part of the multidisciplinary team is also highlighted by the participants' extremely low weight-related quality of life in all subcategories compared to nonbariatric surgery population norms for the same BMI category. The most extreme difference was self-esteem, for which population norms for adults with a BMI of 30 to 39.9 kg/m^2 are 68% to 77%, compared to 11% to 25% in the current sample [37,38]. This is further contrasted against population norms of samples with a BMI of 18.5 to 24.9 kg/m^2, who have a self-esteem norm of 88% and a total weight-related quality of life norm of 95% [37,38].

Limitations

No conclusions can be drawn from this study about predictors of procedure choice in participants who were not matched for age, sex, and BMI. Although the preoperative characteristics may be representative of Australian adults who choose an ESG procedure, they are not representative of those who choose an LSG procedure, due to the exclusion of participants who were not matched to ESG candidates. Additionally, there are likely clinically relevant predictors of procedure choice and preoperative differences between ESG and LSG candidates that were not measured in this study, such as fertility, financial situation, and the location of the study site. Under the advice of a statistician, imputation was not used to account for missing data. This decision was made as the authors hypothesized that there would be no differences in outcomes due to being matched, and there were no trends suggesting increased power was required. However, the sample size was small, and although the *P* values when testing outcomes can be considered sufficient evidence against the null hypothesis, they cannot be considered strong evidence. Further research is required to strengthen and confirm our findings. This study was limited by the selection of outcomes, as some outcomes relevant to the sample were not measured, such as fertility, nutrition biochemistry, and bone density. Nevertheless, there were still multiple outcomes measured, and the *P* values were not adjusted (due to lack of power), leading to a chance of type I errors.

Conclusions

Australian adults who chose an endoscopic or surgical sleeve had high rates of comorbidities, especially NAFLD, back pain, and osteoarthritis, and had high body fat percentage, predominantly centrally located. Most preoperative gastrointestinal symptom scores were low, but abdominal pain was prevalent. Weight-related quality of life was very low compared to weight-adjusted population norms. There was evidence against the test hypothesis, that is, there was evidence suggesting that lower self-esteem predicted choosing a more invasive sleeve (ie, LSG rather than ESG). These preoperative characteristics can be used to improve the patient-centeredness of preoperative and postoperative care and assist in the interpretation of postoperative outcomes between the 2 procedures.

Acknowledgments

The authors acknowledge Evelyne Rathbone, a statistician at Bond University, for assisting with the sample size calculation; all study site staff, especially Isabella Maimone, for providing information to help the reporting of usual care and for assistance with data collection; and Dr Vinicius Simas, a research assistant at Bond University, for performing the dual-energy x-ray absorptiometry scans and assisting in the reporting of these procedures. This work was funded by a Commonwealth of Australia Innovation Connection grant. The Commonwealth of Australia had no role in the design of the study, the data collection, the analysis or interpretation of the data, or the writing of the manuscript.

Authors' Contributions

AS and SM collected the data. SM led planning of the study design, performed data analysis, and drafted the manuscript. SM, GGR, FC, AS, and EI contributed to the study concept and revision of the manuscript. All authors have read and approved the manuscript.

Conflicts of Interest

SM (an independent research fellow who was the data analyst and data manager appointed to the study), EI (the chief investigator), and AS (a research assistant) declare no potential or existing financial or other conflicts of interest. SM and AS have been paid a salary for work performed related to the study. GGR (the endoscopic sleeve gastroplasty [ESG] proceduralist) is an ad hoc consultant for Apollo Endosurgery and is financially compensated for training specialists to perform the endoscopic sleeve gastroplasty procedure. There was no financial compensation provided by Apollo Endosurgery for this trial. As such, there is no conflict of interest. FC is the chief executive officer of the study site. FC received no salary or direct financial benefit for contributing to this study and was not involved with the implementation of the study beyond providing access for the researchers to the study site and medical records, does not have access to the data set, and was not involved in data analysis.

Multimedia Appendix 1
Supplementary materials.
[DOCX File , 35 KB - formative_v6i11e29713_app1.docx]

References

1. Fact Sheet: Obesity and Overweight. World Health Organization. URL: http://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight [accessed 2022-10-22]
2. Fact sheet: Noncommunicable diseases. World Health Organisation. 2018. URL: https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases [accessed 2022-04-01]
3. National Health Survey: First Results, 2014-15. Australian Bureau of Statistics. 2015. URL: https://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/4364.0.55.0012014-15 [accessed 2022-04-01]
4. Cleo G, Isenring E, Thomas R, Glasziou P. Could habits hold the key to weight loss maintenance? A narrative review. J Hum Nutr Diet 2017 Oct;30(5):655-664. [doi: 10.1111/jhn.12456] [Medline: 28150402]
5. Goyal D, Watson RR. Endoscopic Bariatric Therapies. Curr Gastroenterol Rep 2016 Jun;18(6):26. [doi: 10.1007/s11894-016-0501-5] [Medline: 27098813]
6. Clinical Practice Guidelines for the Management of Overweight and Obesity in Adults, Adolescents and Children in Australia Melbourne. Melbourne, Australia: National Health and Medical Research Council; 2013.
7. Buchwald H, Estok R, Fahrbach K, Banel D, Jensen MD, Pories WJ, et al. Weight and type 2 diabetes after bariatric surgery: systematic review and meta-analysis. Am J Med 2009 Mar;122(3):248-256.e5. [doi: 10.1016/j.amjmed.2008.09.041] [Medline: 19272486]
8. Buchwald H, Avidor Y, Braunwald E, Jensen MD, Pories W, Fahrbach K, et al. Bariatric surgery: a systematic review and meta-analysis. JAMA 2004 Oct 13;292(14):1724-1737. [doi: 10.1001/jama.292.14.1724] [Medline: 15479938]
9. Estimate of Bariatric Surgery Numbers, 2011-2016. American Society for Metabolic and Bariatric Surgery. 2016. URL: https://asmbs.org/resources/estimate-of-bariatric-surgery-numbers [accessed 2022-10-22]
10. Angrisani L, Santonicola A, Iovino P, Vitiello A, Zundel N, Buchwald H, et al. Bariatric surgery and endoluminal procedures: IFSO worldwide survey 2014. Obes Surg 2017 Sep;27(9):2279-2289 [FREE Full text] [doi: 10.1007/s11695-017-2666-x] [Medline: 28405878]
11. Bariatric surgery registry seventh annual report: 2018/19. Melbourne, Australia: Monash University; 2019.
12. Angrisani L, Santonicola A, Iovino P, Formisano G, Buchwald H, Scopinaro N. Bariatric surgery worldwide 2013. Obes Surg 2015 Oct;25(10):1822-1832. [doi: 10.1007/s11695-015-1657-z] [Medline: 25835983]
13. Adams TD, Gress RE, Smith SC, Halverson RC, Simper SC, Rosamond WD, et al. Long-term mortality after gastric bypass surgery. N Engl J Med 2007 Aug 23;357(8):753-761. [doi: 10.1056/NEJMoa066603] [Medline: 17715409]
14. Lauti M, Kularatna M, Hill AG, MacCormick AD. Weight regain following sleeve gastrectomy-a systematic review. Obes Surg 2016 Jun;26(6):1326-1334. [doi: 10.1007/s11695-016-2152-x] [Medline: 27048439]
15. Gadiot RPM, Biter LU, van Mil S, Zengerink HF, Apers J, Mannaerts GHH. Long-term results of laparoscopic sleeve gastrectomy for morbid obesity: 5 to 8-year results. Obes Surg 2017 Jan;27(1):59-63. [doi: 10.1007/s11695-016-2235-8] [Medline: 27178407]
16. Catheline J, Fysekidis M, Bachner I, Bihan H, Kassem A, Dbouk R, et al. Five-year results of sleeve gastrectomy. J Visc Surg 2013 Nov;150(5):307-312 [FREE Full text] [doi: 10.1016/j.jviscsurg.2013.08.008] [Medline: 24060743]
17. Chang S, Stoll CRT, Song J, Varela JE, Eagon CJ, Colditz GA. The effectiveness and risks of bariatric surgery: an updated systematic review and meta-analysis, 2003-2012. JAMA Surg 2014 Mar;149(3):275-287 [FREE Full text] [doi: 10.1001/jamasurg.2013.3654] [Medline: 24352617]

18. Pizzicannella M, Lapergola A, Fiorillo C, Spota A, Mascagni P, Vix M, et al. Does endoscopic sleeve gastroplasty stand the test of time? Objective assessment of endoscopic ESG appearance and its relation to weight loss in a large group of consecutive patients. Surg Endosc 2020 Aug;34(8):3696-3705. [doi: 10.1007/s00464-019-07329-1] [Medline: 31932925]
19. ASGE/ASMBS Task Force on Endoscopic Bariatric Therapy. A pathway to endoscopic bariatric therapies. Surg Obes Relat Dis 2011;7(6):672-682. [doi: 10.1016/j.soard.2011.09.008] [Medline: 22082971]
20. Lopez-Nava G, Asokkumar R, Bautista-Castaño I, Laster J, Negi A, Fook-Chong S, et al. Endoscopic sleeve gastroplasty, laparoscopic sleeve gastrectomy, and laparoscopic greater curve plication: do they differ at 2 years? Endoscopy 2021 Mar;53(3):235-243. [doi: 10.1055/a-1224-7231] [Medline: 32698234]
21. Carr P, Keighley T, Petocz P, Blumfield M, Rich GG, Cohen F, et al. Efficacy and safety of endoscopic sleeve gastroplasty and laparoscopic sleeve gastrectomy with 12+ months of adjuvant multidisciplinary support. BMC Prim Care 2022 Feb 05;23(1):26 [FREE Full text] [doi: 10.1186/s12875-022-01629-7] [Medline: 35123409]
22. von Elm E, Altman D, Egger M, Pocock S, Gøtzsche PC, Vandenbroucke J, STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann Intern Med 2007 Oct 16;147(8):573-577 [FREE Full text] [doi: 10.7326/0003-4819-147-8-200710160-00010] [Medline: 17938396]
23. Lee J, Kim D, Kim HJ, Lee C, Yang JI, Kim W, et al. Hepatic steatosis index: a simple screening tool reflecting nonalcoholic fatty liver disease. Dig Liver Dis 2010 Jul;42(7):503-508. [doi: 10.1016/j.dld.2009.08.002] [Medline: 19766548]
24. Revicki DA, Wood M, Wiklund I, Crawley J. Reliability and validity of the Gastrointestinal Symptom Rating Scale in patients with gastroesophageal reflux disease. Qual Life Res 1998 Jan;7(1):75-83. [doi: 10.1023/a:1008841022998] [Medline: 9481153]
25. Poylin V, Serrot FJ, Madoff RD, Ikramuddin S, Ikrumuddin S, Mellgren A, et al. Obesity and bariatric surgery: a systematic review of associations with defecatory dysfunction. Colorectal Dis 2011 Jun;13(6):e92-103. [doi: 10.1111/j.1463-1318.2011.02584.x] [Medline: 21564470]
26. Boerlage T, van de Laar AWJM, Westerlaken S, Gerdes V, Brandjes D. Gastrointestinal symptoms and food intolerance 2 years after laparoscopic Roux-en-Y gastric bypass for morbid obesity. Br J Surg 2017 Mar;104(4):393-400 [FREE Full text] [doi: 10.1002/bjs.10419] [Medline: 27990637]
27. Gys B, Gys T, Lafullarde T. The efficacy of laparoscopic Roux-En-Y gastric bypass after previous anti-reflux surgery: a single surgeon experience. Acta Chir Belg 2015;115(4):268-272. [doi: 10.1080/00015458.2015.11681110] [Medline: 26324027]
28. Kolotkin RL, Crosby RD. Psychometric evaluation of the impact of weight on quality of life-lite questionnaire (IWQOL-lite) in a community sample. Qual Life Res 2002 Mar;11(2):157-171. [doi: 10.1023/a:1015081805439] [Medline: 12018739]
29. Wee CC, Davis RB, Jones DB, Apovian CA, Chiodi S, Huskey KW, et al. Sex, race, and the quality of life factors most important to patients' well-being among those seeking bariatric surgery. Obes Surg 2016 Jun 2;26(6):1308-1316 [FREE Full text] [doi: 10.1007/s11695-015-1956-4] [Medline: 26630951]
30. Mitchell JE, King WC, Courcoulas A, Dakin G, Elder K, Engel S, et al. Eating behavior and eating disorders in adults before bariatric surgery. Int J Eat Disord 2015 Mar;48(2):215-222 [FREE Full text] [doi: 10.1002/eat.22275] [Medline: 24719222]
31. NIA Adverse Event and Serious Adverse Event Guidelines. National Institute on Aging. 2018. URL: https://www.nia.nih.gov/sites/default/files/2018-09/nia-ae-and-sae-guidelines-2018.pdf [accessed 2022-04-01]
32. Baumgartner T, Chung H. Confidence Limits for Intraclass Reliability Coefficients. Meas Phys Educ Exerc Sci 2001 Sep;5(3):179-188 [FREE Full text] [doi: 10.1207/s15327841mpee0503_4]
33. Bravo G, Potvin L. Estimating the reliability of continuous measures with Cronbach's alpha or the intraclass correlation coefficient: toward the integration of two traditions. J Clin Epidemiol 1991;44(4-5):381-390. [doi: 10.1016/0895-4356(91)90076-l] [Medline: 2010781]
34. Rubino F, Panagiotopoulos S. Surgery: metabolic surgery: the cutting edge of diabetes care. Nat Rev Gastroenterol Hepatol 2017 Jul;14(7):389-390. [doi: 10.1038/nrgastro.2017.45] [Medline: 28428637]
35. Nielsen EM, Anderson KP, Marsden J, Zhang J, Schreiner AD. nonalcoholic fatty liver disease underdiagnosis in primary care: what are we missing? J Gen Intern Med 2022 Aug;37(10):2587-2590. [doi: 10.1007/s11606-021-07197-3] [Medline: 34816326]
36. Marshall S, Mackay H, Matthews C, Maimone IR, Isenring E. Does intensive multidisciplinary intervention for adults who elect bariatric surgery improve post-operative weight loss, co-morbidities, and quality of life? A systematic review and meta-analysis. Obes Rev 2020 Jul;21(7):e13012. [doi: 10.1111/obr.13012] [Medline: 32196906]
37. Kolotkin RL, Crosby RD. Manual for the Impact of Weight on Quality of Life (IWQOL and IWQOL-LITE) Measure. Durham, NC: Quality of Life Consulting; 2016:1-51.
38. Crosby RD, Kolotkin RL, Williams GR. An integrated method to determine meaningful changes in health-related quality of life. J Clin Epidemiol 2004 Nov;57(11):1153-1160. [doi: 10.1016/j.jclinepi.2004.04.004] [Medline: 15567631]

Abbreviations

DXA: dual-energy x-ray absorptiometry
ESG: endoscopic sleeve gastroplasty
GSRS: Gastrointestinal Symptom Rating Scale
HDL: high-density lipoprotein
HbA_{1c}: Hemoglobin A_{1c}
HSI: hepatic steatosis index
IWQOL: Impact of Weight on Quality of Life
LDL: low-density lipoprotein
LSG: laparoscopic sleeve gastrectomy
NAFLD: nonalcoholic fatty liver disease
OR: odds ratio
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
VLCD: very low-calorie diet

Edited by A Mavragani; submitted 18.04.21; peer-reviewed by A Gorny, A Rovetta; comments to author 26.04.22; revised version received 30.05.22; accepted 17.08.22; published 24.11.22.

Please cite as:
Marshall S, G Rich G, Cohen F, Soni A, Isenring E
Matched Endoscopic Sleeve Gastroplasty and Laparoscopic Sleeve Gastrectomy Cases: Formative Cohort Study
JMIR Form Res 2022;6(11):e29713
URL: https://formative.jmir.org/2022/11/e29713
doi:10.2196/29713
PMID:36422944

Original Paper

Sleep Disorders and Quality of Life in Patients With Cancer: Prospective Observational Study of the Rafael Institute

Nathaniel Scher[1,2], MD; Liath Guetta[1], MD; Clément Draghi[1], PhD; Safia Yahiaoui[1,3], MD; Mathilde Terzioglu[1]; Emilie Butaye[1]; Kathy Henriques[1]; Marie Alavoine[1]; Ayala Elharar[1]; Andre Guetta[1], MD; Alain Toledano[1,4], MD

[1]Integrative Medicine, Rafael Institute, Levallois-Perret, France

[2]Hartmann Oncology Radiotherapy Group, Hartmann Radiotherapy Institute, Levallois Perret, France

[3]Department of Radiotherapy, Institut Salah Azaiz, Tunis, Tunisia

[4]Department of Integrative Medicine, National Conservatory of Arts and Crafts, Paris, France

Corresponding Author:
Nathaniel Scher, MD
Integrative Medicine
Rafael Institute
3 boulevard Bineau
Levallois-Perret, 92300
France
Phone: 33 0184007007
Email: nathaniel.scher@gmail.com

Abstract

Background: Sleep disorders are a common occurrence in the general population. Yet today, it is clearly agreed that sleep disorders represent both a cancer risk factor and a biological consequence of the of the activation of the immuno-inflammatory system induced by cancer itself.

Objective: The aim of this study was to assess the impact of sleep disorders on quality of life and identify the type of disorder and its causes in order to offer an adapted and personalized care plan.

Methods: In a survey completed during the COVID-19 lockdown, 2000 hours of interviews were collected by remote consultations. During these calls, we administered a sleep questionnaire. This questionnaire was inspired by the STOP-BANG questionnaire and enquired about 6 items. The demographic details of each patient (eg, age and sex), the nature of the pathology, their past treatments, the ongoing cancer treatment, the mood, whether or not the patient is anxious or depressed, and the use of sleeping drug pills were analyzed. A univariate analysis was performed according to the presence or absence of fatigue. Chi-square test was applied to assess possible differences of variables' link to sleep disturbance between patients complaining of fatigue and those without fatigue. The same test was then used to analyze patients on hormone therapy and those with no hormone therapy for 2 types of cancer—breast cancer and prostate cancer.

Results: A total of 905 patients were prospectively included in this study. The average age was 66.7 (5 SD) years, and 606 (67%) patients were women; 142 patients declared being overweight. Breast cancer was the most frequently reported cancer. Nocturnal awakening was reported by 70% (n=633), fatigue by 50% (n=452), difficulty falling asleep by 38% (n=343), snoring reported by an independent observer in 38% (n=343), and apnea reported by an independent observer in 9% (n=81) of the patients. The univariate analysis showed that the feeling of tiredness was significantly greater in patients reporting difficulty falling asleep (P≥.99), pain (P<.001), and frequent awakening (P<.001), as well as in patients who were not receiving cancer treatment (P<.001). The univariate analysis showed that patients who were receiving breast cancer treatment and were under hormone therapy reported difficulty falling asleep (P=.04) and pain (P=.05). In a univariate analysis of patients treated for prostate cancer, being overweight was the only factor reported that had a statistically significant value.

Conclusions: Our preliminary data support and are consistent with data in the literature regarding the importance of sleep disorders in oncology. This justifies the usefulness of a diagnosis and early treatment of sleep disorders in patients with cancer. The Rafael Institute sleep observatory will enable patients to be identified and treated.

(JMIR Form Res 2022;6(11):e37371) doi:10.2196/37371

KEYWORDS

cancer; sleep disorder; sleep; fatigue; nocturnal; oncology; cancer care; patient-centred approach; patient-centered; personalized; personalization; customized; customization; care plan; quality of life; mood; pain; cancer treatment; overweight; obese; hormone therapy; breast; prostate

Introduction

Sleep disorders are a common occurrence in the general population. The most common sleep disorders are the following: insomnia (with a prevalence of 6% to 20%) [1]; rapid eye movement sleep behavior disorder (with a prevalence of 3% to 10%); restless legs syndrome (5% to 11%); periodic limb movements during sleep (up to 30%); obstructive sleep apnea (9% to 38%); and circadian rhythm sleep-wake disorders (3% to 10%) [2-4]. The prevalence of various types of sleep disorders in patients with cancer has not been widely investigated, as sleep disorders are often not assessed in accordance with the International Classification of Sleep Disorders Third Edition in this group of patients [5]; yet today, it is clearly agreed that sleep disorders represent both a cancer risk factor and a biological consequence of the activation of the immuno-inflammatory system induced by cancer [6]. The importance of recognizing and treating sleep disorders in an appropriate manner is crucial if we consider that they can persist for several years after the cessation of antitumor treatment (eg, chemotherapy, radiotherapy, and hormonal therapy) and be a potential cause of comorbidity and impaired quality of life [7,8]. During the last decade, several authors have been particularly interested in sleep disorders and their harmful effects on each stage of cancer treatment for the patient [9,10]. It is therefore essential to acquire a better understanding of these disorders for which treatments offer promising therapeutic prospects in oncology, the aim being to enhance patients' adhesion to stressful treatments and improve their quality of life.

The Rafael Institute, a postcancer institute, is a center for integrative medicine. Nearly 80 medical and paramedical care staff accompany patients during and after cancer treatments to improve their resilience. During the past 16 months, 1350 patients have received over 11,400 care treatments. In the context of personalized care plans oriented toward nutrition, emotional well-being, physical activity, and general well-being, a sleep center was inaugurated in the Rafael Institute in March 2020. The aim of our study was to assess the impact of sleep disorders on quality of life and identify the type of disorder and its causes in order to offer an adapted and personalized care plan.

Methods

Procedure

In a survey completed during the COVID-19 lockdown, 2000 hours of interviews were collected by remote consultations. During the lockdown, Rafael Institute caregivers made calls to maintain a link with confined patients. During these calls, we administered a sleep questionnaire. All of our patients were first contacted by email. The "RAFAËL SLEEP" questionnaire was created for this purpose. This questionnaire, inspired by the STOP-BANG questionnaire enquired about the following 6 items concerning sleep disorders: BMI, diurnal fatigue, difficulty falling asleep, nocturnal awakening, as well as snoring and apnea observed by an independent observer. Patients answering at least 2 questions in the affirmative were included in the study. The demographic details of each patient (eg, age and sex), the nature of the pathology, the treatment already received, the ongoing cancer treatment, the mood, whether or not the patient is anxious or depressed, and the use of sleeping drug pills were analyzed. A univariate analysis was performed according to the presence or absence of fatigue.

Chi-square test was applied to assess possible differences of variables link to sleep disturbance between patients complaining of fatigue and those without fatigue. The same test was then used to analyze patients on hormone therapy and those with no hormone therapy for 2 types of cancer—breast cancer and prostate cancer. R (version 4.1.2; R Foundation for Statistical Computing) was used to conduct all the statistical analyses in this study, and a $P<.05$ (two-tailed P value test) was considered as statistically significant.

Ethics Approval

This observational prospective study was approved by the institutional review board of Rafael Institute and the ethical committee of Hartmann Oncology Radiotherapy Group. Informed consent was obtained from each participant. All necessary measures to safeguard participants' anonymity and confidentiality of information were respected.

Results

A total of 1084 phone numbers were initially selected for the telephone interviews on sleep disorders; 135 were excluded (29 refusals, 102 telephone errors, 15 deceased patients, and 5 for communication difficulty). We were able to include the responses of 905 patients in our study. The participation rate was 95% (Figure 1). The average age was 66.7 (SD 5) years, and 606 (67%) were women; 142 patients declared being overweight. Breast cancer was the most frequently reported type of cancer in our population, with a prevalence of 60% (Table 1).

The sleep-affecting criteria that were assessed were fatigue, difficulty falling asleep, nocturnal awakening, as well as snoring and apnea reported by an independent observer. With respect to these criteria, nocturnal awakening was reported by 70% (n=635), fatigue by 50% (n=452), difficulty falling asleep by 38% (n=343), snoring reported by an independent observer in 38% (n=343), and apnea reported by an independent observer in 9% (n=81) of the patients. The results of the assessment showed that 18% (n=163) of patients responding to the questionnaire satisfied 2 sleep disorder criteria; 24% (n=217) fulfilled 3 sleep criteria; 13% (n=118) fulfilled 4 sleep criteria; and 5% (n=45) related to 5 criteria. A total of 50% (n=452) of patients in the observation group declared that they used sleeping

pills. No significant difference was noted in the declaration of sleep disorders, whether or not the patients declared taking insomnia treatment.

The univariate analysis showed that the feeling of tiredness was significantly greater in patients reporting difficulty falling asleep (P≥.99), pain (P<.001), and frequent awakening (P<.001), as well as in patients who were not receiving cancer treatment (P<.001; Table 2). The univariate analysis showed that patients who were receiving breast cancer treatment and were under hormone therapy essentially reported difficulty falling asleep (P=.04) and pain (P=.05; Table 3). In a univariate analysis of patients treated for prostate cancer, being overweight was the only factor reported that had a statistically significant value (Table 4).

Figure 1. Flowchart of patients included in the study.

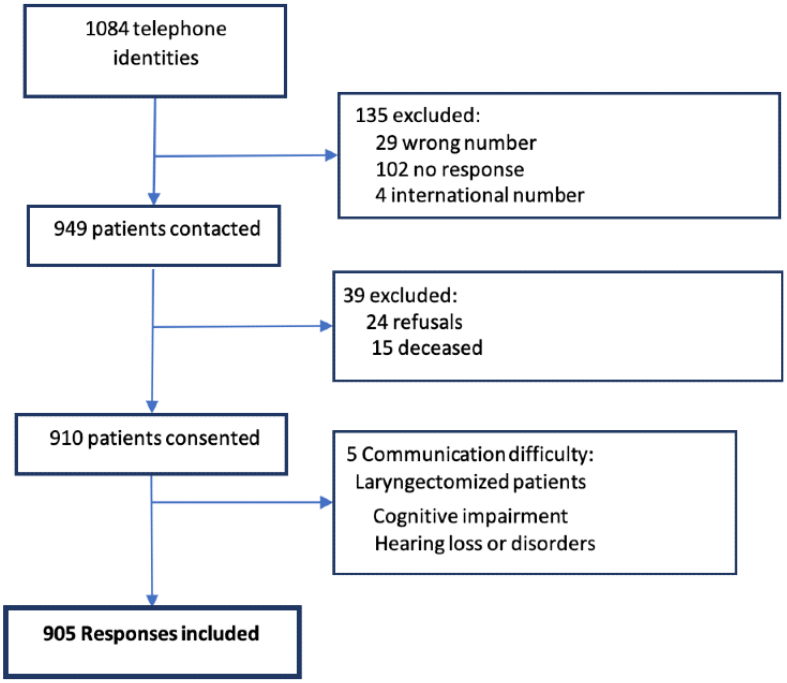


Table 1. Patient characteristics (N=905).

Characteristics	Values, n (%)
Sex	
Women	687 (76)
Men	218 (24)
Localization	
Breast	539 (60)
Prostate	107 (12)
Other	67 (7.4)
Head and neck	63 (7)
Lung	43 (4.8)
Colorectal	39 (4.3)
Uterus	17 (1.9)
Pancreas	16 (1.8)
Ovary	10 (1.1)
Day fatigue	
No	454 (50)
Yes	447 (50)
Apnea	
No	813 (90)
Yes	88 (9.8)
Appetite	
Fairly	463 (51)
A lot	261 (29)
A little	129 (14)
Not at all	48 (5.3)
Difficulties falling asleep	
No	552 (61)
Yes	349 (39)
Pain	
No pain at all	349 (39)
A little	291 (32)
Fairly	165 (18)
A lot	96 (11)
Hormonal therapy	
No	604 (67)
Yes	297 (33)
Frequent awakening	
No	266 (30)
Yes	635 (70)
Snoring	
No	548 (61)
Yes	353 (39)
Overweight	

Characteristics	Values, n (%)
No	659 (73)
Yes	242 (27)
Antitumor treatment completed	
No	418 (46)
Yes	483 (54)

Table 2. Univariate analysis of fatigue. Italicized *P* values are statistically significant.

Characteristics	Absence of diurnal fatigue (n=454)	Presence of diurnal fatigue (n=447)	Total participants, n	*P* value
Anxiety, mean (SD)	4.94 (2.56)	5.57 (2.50)	901	*<.001*
Physical form, mean (SD)	7.44 (1.55)	6.21 (1.77)	901	*<.001*
Moral, mean (SD)	7.62 (1.70)	6.63 (1.99)	901	*<.001*
Quality of life, mean (SD)	7.72 (1.46)	6.90 (1.86)	901	*<.001*
Apnea, n (%)				*≥.99*
Absent	424 (93)	389 (87)	813	
Present	30 (6.6)	58 (13)	88	
Appetite, n (%)				*<.001*
Fairly	245 (54)	218 (49)	463	
A lot	159 (35)	102 (23)	261	
A little	43 (9.5)	86 (19)	129	
Not at all	7 (1.5)	41 (9.2)	48	
Difficulty getting to sleep, n (%)				*≥.99*
No	302 (67)	250 (56)	552	
Yes	152 (33)	197 (44)	349	
Pain, n (%)				*<.001*
Not at all	226 (50)	123 (28)	349	
A little	139 (31)	152 (34)	291	
Fairly	63 (14)	102 (23)	165	
A lot	26 (5.7)	70 (16)	96	
Frequent awakening, n (%)				*<.001*
No	178 (39)	88 (20)	266	
Yes	276 (61)	359 (80)	635	
Snoring, n (%)				.6
No	280 (62)	268 (60)	548	
Yes	174 (38)	179 (40)	353	
Sex, n (%)				.45
Woman	351 (77)	336 (75)	687	
Man	103 (23)	111 (25)	214	
Overweight, n (%)				.07
No	344 (76)	315 (70)	659	
Yes	110 (24)	132 (30)	242	
Treatment finished				*≥.99*
No	187 (41)	231 (52)	418	
Yes	267 (59)	216 (48)	483	

Table 3. Univariate analysis of hormone therapy in patients treated for breast cancer.

Characteristics	No hormone therapy (n=278)	Hormone therapy (n=261)	Total number, n	*P* value
Anxiety, mean (SD)	5.28 (2.58)	5.70 (2.50)	539	.06
Sleep disorders, n (%)				.04
Absent	178 (64)	145 (56)	323	
Present	100 (36)	116 (44)	216	
Pain, n (%)				.05
A little	109 (39)	82 (31)	191	
Not at all	97 (35)	83 (32)	180	
Fairly	46 (17)	61 (23)	107	
A lot	26 (9.4)	35 (13)	61	
Frequent awakening, n (%)				.22
Absent	99 (36)	80 (31)	179	
Present	179 (64)	181 (69)	360	
Snoring, n (%)				
Absent	180 (65)	164 (63)	344	.64
Present	98 (35)	97 (37)	195	
Overweight, n (%)				.13
Absent	188 (68)	192 (74)	380	
Present	90 (32)	69 (26)	159	

Table 4. Univariate analysis of hormone therapy in patients treated for prostate cancer. Italicized *P* values are statistically significant.

Characteristics	No hormone therapy (n=78), n (%)	Hormone therapy (n=29), n (%)	Total participants, n	*P* value
Apnea				.19
No	63 (81)	20 (69)	83	
Yes	15 (19)	9 (31)	24	
Difficulty getting to sleep				.71
No	54 (69)	19 (66)	73	
Yes	24 (31%)	10 (34)	34	
Pain				>.99
Not at all	45 (58%)	17 (59)	62	
A little	20 (26)	6 (21)	26	
Fairly	7 (9)	5 (17)	12	
A lot	6 (7.7)	1 (3.4)	7	
Frequent awakening				.43
No	19 (24)	5 (17)	24	
Yes	59 (7)	24 (83)	83	
Snoring				.45
No	36 (46)	11 (38)	47	
Yes	42 (54)	18 (62)	60	
Overweight				*.04*
No	59 (76)	16 (55)	75	
Yes	19 (24)	13 (45)	32	
Treatment finished				*<.001*
No	20 (26)	21 (72)	41	
Yes	58 (74)	8 (28)	66	

Discussion

Nearly one third of French adults have sleeping problems, such as sleep deficit, trouble getting to sleep or staying asleep, disturbance of the circadian rhythm, or respiratory problems such as obstructive apnea syndrome. The effects of sleep on immune, hormonal, cardiovascular, and neurocognitive functions are well established. An association between sleep disorders and the onset or progression of many cancers has now been suggested [11]. Recent epidemiological studies report a close link between cancer and sleep apnea syndrome [12-14] or circadian rhythm sleep disorders [15-18]. Working night shifts significantly increases the risk of cancer, particularly of prostate, colon, and breast cancer [11]. The disruption of sleep-wake circadian rhythms is an independent prognostic survival factor for patients with specific types of metastatic cancers [19-21]. The dysregulation of several dozen molecular signaling pathways that control the circadian rhythm may be involved in the carcinogenesis process [22,23].

The identification and correction of circadian rhythm sleep disturbances may prevent cancers or slow their progression [24]. Sleep apnea syndrome has been particularly well studied and may be strongly involved in mortality not only of cardiovascular origin but also mortality from all other causes, including cancer [25]. This high cancer mortality rate may be more specifically related to the apnea-hypopnea index [26,27]. An index of 30, for instance, increases the relative cancer mortality risk by a factor of 4.8. This is predominantly underpinned by intermittent hypoxia-induced tumor growth as it has been demonstrated in animal models [28,29]. Sleep disorders occur in 30%-75% of newly diagnosed patients or patients who have recently undergone cancer treatment. This rate is twice of that in the general population. Patients complain of difficulties getting to sleep and staying asleep, with frequent and prolonged nocturnal awakening [30,31]. Patients report difficulties both before and after cancer treatments. Fatigue is also a major factor in this particular population, and recent data suggest that daytime fatigue is related to the sleep-wake cycle and to the quality and quantity of sleep obtained throughout the night. The high prevalence of insomnia and sleep disorders experienced by patients with cancer has been attributed to cancer treatments, in particular chemotherapy, their side effects, psychosocial factors (eg, anxiety, stress, and depression), and circadian rhythm disruptions. The initial descriptive results from our cohort of 905 patients with cancer provide preliminary data for future studies that will more specifically focus on characterizing patient profiles and proposing an overall specific care plan for sleep disorders in this patient population.

In this preliminary study, 60% (n=543) of patients satisfy at least two of the criteria surveyed in the questionnaire; this is consistent with the percentage reported in numerous other studies on sleep disorders and cancer. As suggested in several studies, the predominant symptom in patients with cancer appears to be nocturnal awakening, which is also consistent with our observation, with some 70% (n=633) of interviewed patients complaining of this symptom. Nocturnal awakening results in sleep fragmentation, which may be the main risk factor in increasing the prevalence of cancers and potentially the resistance to chemotherapy. This nocturnal awakening may be due to disruptions of sleep-wake circadian rhythms, but also due to obstructive apnea syndrome. In our study, family members of 38% (n=344) of patients reported snoring and 9% (n=81) reported apnea; 38% (n=344) of patients complained of difficulty falling asleep; this is insomnia during the first part of the night, and this percentage is the same as reported in other studies on cancer and insomnia. The COVID-19 pandemic had a significant impact on all aspects of daily life. Sleep patterns, sleep quality, as well as the diagnosis and management of sleep disorders were all profoundly affected. In this study, the patient interview was conducted during the COVID-19 lockdown, and the survey analyzed the patient's mood. Fear and anxiety of potential infection in a patient with cancer, mandatory confinements, and quarantine procedures likely combined to increase sleep dysfunction, as shown in other published studies [32]. This study does not specify the cause of the sleep disorders and does not analyze the specific impact of COVID-19 on the patients included. Nevertheless, our sleep laboratory has moved toward a new practice model, minimizing physical contact with patients and developing remote consultations to maintain contact with patients. This new effective assessment method has made it possible to structure the assessment of sleep disorders.

Conclusions from this study should be tempered in light of some limitations. First, this is an observational cohort and data were collected by remote consultation, which may have missed some important information. Second, although every attempt was made to identify all sleep disorders, the sleep-affecting criteria were reported by an independent observer and from a sleep questionnaire; therefore, it is possible that some data about sleep disorders were missed, and this might as well have affected our conclusions. The average age of our cohort was 66.7 (SD 5) years, which also corresponds to the average age of patients with cancer in France. There are markedly more women than men (n=678, 76% vs n=218, 24%). This is explained by the fact that more women adhere to the paramedical care offered at the Rafael Institute. We will therefore extend the sleep observatory to all patients with cancer who have sleep disorders, irrespective of whether or not they are being treated at the Rafael Institute.

These results have supported us in our strategy of global patient care. This will involve all new patients who come for consultations to the Rafael Institute and complete the questionnaire. If the questionnaire is positive, the patient will be offered simple respiratory polygraphy to eliminate any potential sleep apnea syndrome. Polygraphy has been chosen over ambulatory polysomnography, which is the reference examination but is more difficult to set up in the context of daily life and costs more. In the case of positive polygraphy—namely more than 30 apnea episodes per hour—patients are given a standard treatment, that is, automatically controlled continuous positive airway pressure. Otherwise, a specialized sleep consultation is proposed, with the programming of continuous positive airway pressure, if necessary. The aim of this consultation is to assess the specific sleep disorder and propose a suitable treatment. This treatment consists essentially of cognitive and behavioral therapies, the use of a therapy such as sophrology, verbalization psychotherapy, meditation, or neurofeedback. The aim is to treat sleep disorders without recourse to benzodiazepines. In the study, we found that 50% (n=452) of patients used sleeping pills. This very high percentage together with awareness of the deleterious effects of benzodiazepines encourages us to favor nonmedical treatments. Our preliminary data support and are consistent with data in the literature regarding the importance of sleep disorders in oncology. This justifies the usefulness of a diagnosis and early treatment of sleep disorders in patients with cancer. The Rafael Institute sleep observatory will enable patients to be identified and treated. The RAFAËL SLEEP questionnaire used in this study enables patients with sleep disorders to be identified. It will be complemented by sleep, anxiety, and quality of life questionnaires validated by learned societies to propose correlation studies. This will allow us to assess care objectively.

Conflicts of Interest

None declared.

References

1. Riemann D, Baglioni C, Bassetti C, Bjorvatn B, Dolenc Groselj L, Ellis JG, et al. European guideline for the diagnosis and treatment of insomnia. J Sleep Res 2017 Dec 05;26(6):675-700 [FREE Full text] [doi: 10.1111/jsr.12594] [Medline: 28875581]
2. Dauvilliers Y, Schenck CH, Postuma RB, Iranzo A, Luppi P, Plazzi G, et al. REM sleep behaviour disorder. Nat Rev Dis Primers 2018 Aug 30;4(1):19. [doi: 10.1038/s41572-018-0016-5] [Medline: 30166532]
3. McBeth J, Dixon WG, Moore SM, Hellman B, James B, Kyle SD, et al. Sleep disturbance and quality of life in rheumatoid arthritis: prospective mHealth study. J Med Internet Res 2022 Apr 22;24(4):e32825-e32810 [FREE Full text] [doi: 10.2196/32825] [Medline: 35451978]
4. Kim JH, Duffy JF. Circadian rhythm sleep-wake disorders in older adults. Sleep Med Clin 2018 Mar;13(1):39-50. [doi: 10.1016/j.jsmc.2017.09.004] [Medline: 29412982]

5. International classification of sleep disorders, third edition. American Academy of Sleep Medicine. 2014. URL: https://j2vjt3dnbra3ps7ll1clb4q2-wpengine.netdna-ssl.com/wp-content/uploads/2019/05/ICSD3-TOC.pdf [accessed 2022-10-21]
6. Lowery-Allison AE, Passik SD, Cribbet MR, Reinsel RA, O'Sullivan B, Norton L, et al. Sleep problems in breast cancer survivors 1–10 years posttreatment. Pall Supp Care 2017 May 16;16(3):325-334. [doi: 10.1017/s1478951517000311]
7. JMIR Public Health and Surveillance. JMIR publications. URL: https://www.jmir.org/search?term=Paul%20Stang&type=author&precise=true [accessed 2020-05-26]
8. Walker WH, Borniger JC. Molecular mechanisms of cancer-induced sleep disruption. Int J Mol Sci 2019 Jun 06;20(11):2780 [FREE Full text] [doi: 10.3390/ijms20112780] [Medline: 31174326]
9. JMIR Cancer. JMIR Publications. URL: https://www.jmir.org/search?term=Thomas%20Maximilian%20Deutsch&type=author&precise=true [accessed 2022-10-05]
10. Mogavero MP, DelRosso LM, Fanfulla F, Bruni O, Ferri R. Sleep disorders and cancer: state of the art and future perspectives. Sleep Med Rev 2021 Apr;56:101409. [doi: 10.1016/j.smrv.2020.101409] [Medline: 33333427]
11. Fang H, Miao N, Chen C, Sithole T, Chung M. Risk of cancer in patients with insomnia, parasomnia, and obstructive sleep apnea: a nationwide nested case-control study. J Cancer 2015;6(11):1140-1147 [FREE Full text] [doi: 10.7150/jca.12490] [Medline: 26516362]
12. Kendzerska T, Leung RS, Hawker G, Tomlinson G, Gershon AS. Obstructive sleep apnea and the prevalence and incidence of cancer. CMAJ 2014 Sep 16;186(13):985-992 [FREE Full text] [doi: 10.1503/cmaj.140238] [Medline: 25096668]
13. Martínez-García MA, Campos-Rodriguez F, Durán-Cantolla J, de la Peña M, Masdeu MJ, González M, Spanish Sleep Network. Obstructive sleep apnea is associated with cancer mortality in younger patients. Sleep Med 2014 Jul;15(7):742-748. [doi: 10.1016/j.sleep.2014.01.020] [Medline: 24907033]
14. Gozal D, Farré R, Nieto FJ. Putative links between sleep apnea and cancer: from hypotheses to evolving evidence. Chest 2015 Nov;148(5):1140-1147 [FREE Full text] [doi: 10.1378/chest.15-0634] [Medline: 26020135]
15. World Health Organization: International Agency for Research on Cancer. IARC monographs on the evaluation of carcinogenic risks to humans. 2010. URL: https://monographs.iarc.who.int/wp-content/uploads/2018/06/mono98.pdf [accessed 2022-10-21]
16. Jia Y, Lu Y, Wu K, Lin Q, Shen W, Zhu M, et al. Does night work increase the risk of breast cancer? A systematic review and meta-analysis of epidemiological studies. Cancer Epidemiol 2013 Jun;37(3):197-206. [doi: 10.1016/j.canep.2013.01.005] [Medline: 23403128]
17. Wang F, Yeung KL, Chan WC, Kwok CCH, Leung SL, Wu C, et al. A meta-analysis on dose-response relationship between night shift work and the risk of breast cancer. Ann Oncol 2013 Nov;24(11):2724-2732 [FREE Full text] [doi: 10.1093/annonc/mdt283] [Medline: 23975662]
18. Lévi F, Okyar A, Dulong S, Innominato PF, Clairambault J. Circadian timing in cancer treatments. Annu Rev Pharmacol Toxicol 2010 Feb 01;50(1):377-421. [doi: 10.1146/annurev.pharmtox.48.113006.094626] [Medline: 20055686]
19. Innominato P, Focan C, Gorlia T, Moreau T, Garufi C, Waterhouse J, Chronotherapy Group of the European Organization for ResearchTreament of Cancer. Circadian rhythm in rest and activity: a biological correlate of quality of life and a predictor of survival in patients with metastatic colorectal cancer. Cancer Res 2009 Jun 01;69(11):4700-4707 [FREE Full text] [doi: 10.1158/0008-5472.CAN-08-4747] [Medline: 19470769]
20. Filipski E, King V, Li X, Granda T, Mormont M, Liu X, et al. Host circadian clock as a control point in tumor progression. J Natl Cancer Inst 2002 May 01;94(9):690-697. [doi: 10.1093/jnci/94.9.690] [Medline: 11983758]
21. Mormont M, Waterhouse J, Bleuzen P, Giacchetti S, Jami A, Bogdan A, et al. Marked 24-h rest/activity rhythms are associated with better quality of life, better response, and longer survival in patients with metastatic colorectal cancer and good performance status. Clin Cancer Res 2000 Aug;6(8):3038-3045. [Medline: 10955782]
22. Lévi F. Circadian chronotherapy for human cancers. The Lancet Oncology 2001 May;2(5):307-315. [doi: 10.1016/s1470-2045(00)00326-0]
23. Lévi F, Focan C, Karaboué A, de la Valette V, Focan-Henrard D, Baron B, et al. Implications of circadian clocks for the rhythmic delivery of cancer therapeutics. Adv Drug Deliv Rev 2007 Aug 31;59(9-10):1015-1035. [doi: 10.1016/j.addr.2006.11.001] [Medline: 17692427]
24. Nieto FJ, Peppard PE, Young T, Finn L, Hla KM, Farré R. Sleep-disordered breathing and cancer mortality. Am J Respir Crit Care Med 2012 Jul 15;186(2):190-194. [doi: 10.1164/rccm.201201-0130oc]
25. Marshall NS, Wong KK, Cullen SR, Knuiman MW, Grunstein RR. Sleep apnea and 20-year follow-up for all-cause mortality, stroke, and cancer incidence and mortality in the Busselton Health Study cohort. J Clin Sleep Med 2014 Apr 15;10(4):355-362 [FREE Full text] [doi: 10.5664/jcsm.3600] [Medline: 24733978]
26. Gozal D, Ham SA, Mokhlesi B. Sleep apnea and cancer: analysis of a nationwide population sample. Sleep 2016 Aug 01;39(8):1493-1500 [FREE Full text] [doi: 10.5665/sleep.6004] [Medline: 27166241]
27. Gozal D, Farré R, Nieto FJ. Obstructive sleep apnea and cancer: epidemiologic links and theoretical biological constructs. Sleep Med Rev 2016 Jun;27:43-55 [FREE Full text] [doi: 10.1016/j.smrv.2015.05.006] [Medline: 26447849]
28. Almendros I, Wang Y, Becker L, Lennon FE, Zheng J, Coats BR, et al. Intermittent hypoxia-induced changes in tumor-associated macrophages and tumor malignancy in a mouse model of sleep apnea. Am J Respir Crit Care Med 2014 Mar 01;189(5):593-601. [doi: 10.1164/rccm.201310-1830oc]

29. Gozal D, Almendros I, Hakim F. Sleep apnea awakens cancer: a unifying immunological hypothesis. Oncoimmunology 2014 Mar 27;3(4):e28326 [FREE Full text] [doi: 10.4161/onci.28326] [Medline: 25050203]
30. Panjwani AA, Millar BM, Revenson TA. Tolerating uncertainty in the dark: insomnia symptoms, distress, and well-being among parents of adolescents and young adults with cancer. Int J Behav Med 2021 Feb;28(1):14-20. [doi: 10.1007/s12529-020-09869-6] [Medline: 32219773]
31. Luo J, Sands M, Wactawski-Wende J, Song Y, Margolis KL. Sleep disturbance and incidence of thyroid cancer in postmenopausal women the Women's Health Initiative. Am J Epidemiol 2013 Jan 01;177(1):42-49 [FREE Full text] [doi: 10.1093/aje/kws193] [Medline: 23221728]
32. Morin CM, Carrier J, Bastien C, Godbout R, Canadian SleepCircadian Network. Sleep and circadian rhythm in response to the COVID-19 pandemic. Can J Public Health 2020 Oct 22;111(5):654-657 [FREE Full text] [doi: 10.17269/s41997-020-00382-7] [Medline: 32700231]

Edited by A Mavragani; submitted 17.02.22; peer-reviewed by N Chalghaf, F Denis; comments to author 28.04.22; revised version received 22.08.22; accepted 29.09.22; published 24.11.22.

Please cite as:
Scher N, Guetta L, Draghi C, Yahiaoui S, Terzioglu M, Butaye E, Henriques K, Alavoine M, Elharar A, Guetta A, Toledano A
Sleep Disorders and Quality of Life in Patients With Cancer: Prospective Observational Study of the Rafael Institute
JMIR Form Res 2022;6(11):e37371
URL: https://formative.jmir.org/2022/11/e37371
doi:10.2196/37371
PMID:36422866

Original Paper

Validity of the Parsley Symptom Index—an Electronic Patient-Reported Outcomes Measure Designed for Telehealth: Prospective Cohort Study

Hants Williams[1,2], RN, PhD; Sarah Steinberg[2], MD, PhD; Kendall Leon[2], PhD; Catherine O'Shea[2], MA; Robin Berzin[2], MD; Heather Hagg[2], PhD

[1]Department of Applied Health Informatics, School of Health Professions, Stony Brook University, Stony Brook, NY, United States
[2]Parsley Health, New York, NY, United States

Corresponding Author:
Hants Williams, RN, PhD
Department of Applied Health Informatics
School of Health Professions
Stony Brook University
101 Nicolls Rd
Stony Brook, NY, 11794-8200
United States
Phone: 1 631 444 2252
Email: hantsawilliams@gmail.com

Abstract

Background: Electronic patient-reported outcomes measures (e-PROMs) are a valuable tool for the monitoring and management of chronic conditions over time. However, there are few validated tools available that capture symptoms across body systems in telehealth settings. The Parsley Symptom Index (PSI) is a recently developed symptom assessment for adults with chronic disease in telehealth settings. A previous study demonstrated the feasibility and acceptability of the PSI in a clinical telehealth setting.

Objective: The purpose of this study was to assess convergent validity between the PSI and the self-rated health (SRH) item.

Methods: This prospective cohort study took place from January 15, 2021, to December 15, 2021, among a sample of 10,519 adult patients at Parsley Health, a subscription-based holistic medical practice. The PSI and the SRH were completed by patients via an online portal. The association between the PSI and SRH was assessed via polyserial and polychoric correlations, while weighted κ scores provided information related to agreement between the PSI and SRH.

Results: From 22,748 responses, there were moderate levels of association (polyserial r=0.51; polychoric r=0.52) and agreement (weighted κ=0.46) between the PSI and SRH. In total, 74.13% (n=16,865) of responses between the PSI and SRH were relatively congruent while 36.17% (n=8229) were literally congruent.

Conclusions: The PSI demonstrates convergent validity with the SRH for adults with chronic disease in a telehealth setting. This finding further supports the validation of the PSI in a real-world clinical setting. Although it is conceptually similar to the 1-question SRH, the PSI is a 45-item PROM designed to capture quality of life and specific symptoms by body system. Future studies will compare the PSI to multi-item PROMs.

(JMIR Form Res 2022;6(11):e40063) doi:10.2196/40063

KEYWORDS

telemedicine; eHealth; mHealth; web-based

Introduction

Providing telehealth options has become indispensable to health care delivery in the United States. Even before the COVID-19 pandemic fundamentally altered the health care landscape, claim lines in the United States for nonhospital-based clinicians to patient telehealth grew 1393% [1] from 2014 to 2018. Health crisis triaging during the COVID-19 pandemic further increased demand for telehealth care [2,3], accelerating the transition from brick-and-mortar practice to the virtual interface. The pandemic spawned an entirely new telehealth industry, reducing access

and cost barriers for patients, from the rural farmer to the busy urban professional [4,5].

Having access to affordable care is especially important for the 60% of Americans that live with at least 1 chronic disease, and this group spends 2 to 4 times more on health care than do those without any chronic conditions [6]. Telehealth helps clinicians effectively manage chronic disease with increased opportunity to monitor treatments and quickly respond to patient concerns [7], which reduces costs [8] and hospitalizations [9]. Electronic patient-reported outcome measurements (e-PROMs) are tools that serve as the first patient touchpoint in a telehealth consultation generally and in particular can play a pivotal role in the clinical care of patients with chronic conditions. Completing e-PROMs allows patients to reflect on their own health, boosts patient-clinician communication, and empowers patients to steer their own health care journey [10].

Despite the fact that many Americans with chronic diseases are currently being treated via telehealth, there are limited e-PROM tools available to telehealth providers and clinics for assessing and tracking a patient's health status over time. Tools like the Patient-Reported Outcomes Measurement Information System (PROMIS) [11,12], the 36-Item Short Form Health Survey (SF-36) [13,14], and the Medical Symptom Toxicity Questionnaire (MSQ) [15] are powerful e-PROM tools for tracking a patient's health status over time, but none of them offer a single, short-form assessment that could be easily integrated into the clinician workflow or electronic medical record or that can capture symptoms across body systems like a review of systems (ROS).

As part of a larger effort to leverage new tools like e-PROMs to make the telehealth experience engaging and effective for patients with chronic diseases, a research team at Parsley Health (a subscription-based holistic medicine practice) built the Parsley Symptom Index (PSI). The PSI is a 45-item e-PROM designed specifically for use in telehealth settings to function as an ROS. When used strategically, a patient-reported outcome–driven approach can shift an ROS to a cooperative dialogue between patients and clinicians [16]. Like an ROS, the PSI focuses on bodily domains and the most commonly reported symptoms associated with chronic conditions for each domain. As a digital-first e-PROM, we built the PSI to provide immediate feedback to patients, producing data that are seamlessly adopted into the standard clinical workflow and providing the scaffold for an effective patient-clinician conversation [17]. To our knowledge, the PSI is the only existing short form e-PROM developed with preliminary validation for use within a telehealth setting for patients with chronic disease [18].

In an initial feasibility and acceptability study that assessed construct and face validity, the PSI was deployed, completed, and found helpful to both patients and clinicians [18]. Having previously described the item generation, accessibility, and interpretability in a population receiving longitudinal care, we conducted this study is to continue validation of the PSI by comparing it against the self-rated health (SRH) score, a single-item question that has been successfully used in prior research to test construct validity of patient-perceived health [19-21].

Methods

Study Design

This prospective cohort study took place at Parsley Health from January 15, 2021, to December 15, 2021, among a sample of 10,519 adult patients. Patients completed the PSI and the SRH via an online portal. The average monthly PSI completion rate was 77.21% (range 69.23%-83.44%) over the study period.

Ethics Approval

This study used patient-reported survey data that were recorded in such a manner that participants were unidentifiable to the researchers. The institutional review board at Stony Brook University considered this study exempt (IRB2020-00429) from Code of Federal Regulations Title 45 requirements [22].

Study Setting and Population

Parsley Health is a subscription-based membership model for delivering primary care and proactive chronic disease management through a holistic-medicine lens. Patients receive care from Parsley Health clinicians and health coaches in-person and remotely, with additional access to their care team via email and an online portal. Prior to the COVID-19 pandemic, Parsley Health further increased their telehealth availability to over 45 states. Inclusion criteria for this study were Parsley Health patients that had an active subscription membership plan between January 15, 2021, and December 15, 2021, and a minimum of 1 clinical encounter within their membership period. Exclusion criteria were severe psychiatric disorders (particularly psychosis and depression requiring a change in treatment in the last 30 days), age under 18 years, and being unable to speak or read English.

Parsley Symptom Index

The PSI is a 45-item, ROS-style PROM tool designed to capture chronic disease symptoms [18]. The PSI development followed the framework outlined by the Federal Drug Agency (FDA) guide for PROM development [23]. Items are grouped into 9 systems, with each containing 4 to 7 items per group that are ranked on a scale from 0 (asymptomatic) to 10 (extremely symptomatic). A total score is calculated with the following 4 cutoff ranges: 0-24, 25-43, 44-71, and greater than 71. The respective terminology for these ranges are "well" (0-24), "symptomatic" (25-43), "very symptomatic" (44-71), and "sick" (71+). Upon completing the PSI, patients can immediately view their PSI score. When they meet with their clinician, they can view it in graphical format and compare it to past responses, stratified by body systems.

Self-rated Health Item

The SRH item was administered alongside the PSI. The SRH was a mandatory item at the end of the PSI, and only complete questionnaires were included in this study. The SRH is a single question, with a 5-item Likert scale answer that reads as follows: "In general, would you say that your health is excellent, very good, good, fair, or poor?" The SRH is validated and is commonly used to demonstrate construct validity of PROMs [19-21] and allows the clinician to perform a quick global assessment of patient-perceived well-being.

Procedure

After patients scheduled a visit, they were instructed to log into an online patient portal and complete the PSI 24 to 48 hours before each clinical visit. Initial visits were rescheduled if all forms were not completed, but follow-up visits were not postponed for an incomplete PSI. For follow-up visits, patients who had not completed the PSI received an automated reminder 48 hours before the clinical visit. If the PSI was not completed after the automated prompt, another prompt was sent from the clinician or clinical operations coordinator.

When clinicians prepared for an online visit, they used a standardized note template within the electronic health record to pull the most recent PSI score into the visit note. The PSI design allowed for the results to be immediately usable: once a PSI was completed, patients received instant feedback, and clinicians could quickly import the data into the note to prepare for the patient visit. With the PSI template integrated into the beginning of the encounter note, clinicians were subtly prompted to use the PSI to discuss patient-reported symptoms and provide positive feedback to the patient for completing the PSI.

During the telehealth patient visit, the PSI score was used as a touchpoint for the patient-clinician discussion. As the PSI was previously completed, clinicians were able to ask targeted questions about symptoms and had more time to focus on burden and distribution of illness. The longitudinal PSI graph further deepened the provider's ability to identify triggers and mediators that influenced disease trajectory over time.

Association Analysis

To test the hypothesis that the SRH item would correlate with the PSI, 2 measures of association were calculated. First, a polyserial correlation was performed on the raw continuous score of the PSI (range 0-500) with the ordinal SRH categories (excellent, very good, good, fair, poor). Next, the PSI's responses were scored and translated into ordinal categories (1=great, 2=good, 3=average, 4=fair, 5=poor) to compare directly with the SRH categories and generate polychoric correlation coefficients [24]. This second analysis provided an alternative view for when the PSI is interpreted as ordinal instead of continuous.

Agreement Analysis

To determine agreement, weighted κ (quadratic) scores incorporated information about the distance between the transformed ordinal PSI and SRH ratings: ratings that were 1 category apart counted as "less disagreement" than did a pair of ratings 2 categories apart. The weighted κ method partially contributes to responses that are "near" the rating category; for example, "Very good" and "Excellent" are categorically closer than are "excellent" and "poor." To interpret the κ score, the following guidelines are used to suggest agreement [25-28]: 0=agreement equivalent to chance, 0.10-0.20=light agreement, 0.21-0.40=fair agreement, 0.41-0.60=moderate agreement, 0.61-0.80=substantial agreement, 0.81-0.99=near-perfect agreement, and 1.00=perfect agreement.

In addition, a binary interpretation of agreement results as "literally congruent" or "relatively congruent" was calculated. If the PSI and the SRH were an exact match (eg, both scored as "Very good" or both "Poor") the congruence type was scored as literal, while if an individual's responses to the PSI and SRH were not an exact match but consistent in terms of their position as either good ("Excellent," "Very good," "Good") or bad health ("Fair," "Poor"), the congruence type was scored as relative [29].

Data Analysis Software

All analyses were carried out in SAS version 9.4 (SAS Institute) [30].

Results

There were a total of 22,732 observations from 10,519 unique patients from January 15, 2021, to December 15, 2021. Only completed sociodemographic data for patients are represented in Table 1. Race and gender identity data are not complete for the entire sample and were added in late January 2020 for new members. Missing data for race and gender for members registered prior to January 2021 are still being retroactively collected by staff. Data describing race or ethnicity and gender identity refer to the segment of the population for which that data are complete (n=8042).

The distribution of responses for each scale item was skewed toward the positive (Table 2). Of the 22,748 respondents, 12.45% (n=2834) and 3.58% (n=817) reported their health as "Excellent" for the PSI and SRH, respectively; 22.85% (n=5207) and 38.65% (n=8794) rated their health as "Very good" or "Good" for the PSI, respectively, and 25.31% (n=5759) and 42.79% (n=9734) as "Very good" or "Good" for the SRH, respectively. Fewer than 25.92% (n=5897) rated their health to be "Fair" or "Poor" on the PSI and 28.23% (n=6422) did so for the SRH.

The polyserial correlation between raw PSI scores and the SRH was r=0.51, suggesting moderate association. When the PSI scores were treated as ordinal (transformed to SRH scale), the polychoric correlation coefficient was nearly identical at r=0.52, also suggesting moderate association. The weighted κ coefficient between the transformed PSI and SRH was 0.46, suggesting moderate agreement (Table 3). The agreement analysis shows approximately 74.13% (16,865/22,748) relative congruence and 36.17% (8229/22,748) literal congruence across all observations.

Although sample size diminishes with increasing visits, concordance between the PSI and SRH remains stable, even in the cells with smaller sample size. For graphic representation (Figure 1 and Figure 2), we limited our data to 1 to 3 visits for visual clarity. Lastly, in keeping with good reporting practices the Checklist for Reporting Results of Internet E-Surveys (CHERRIES) [31] is provided in Multimedia Appendix 1.

Table 1. Patient descriptives (N=10,531).

Characteristic	Value
Biological Sex, n (%)	
Female	9092 (86.33)
Male	1351 (12.82)
Other	88 (0.83)
Gender identity, n (%)[a]	
Woman	6942 (86.32)
Man	1011 (12.57)
Nonbinary	33 (0.41)
Female to male	13 (0.16)
Male to female	11 (0.13)
Nonbinary other	13 (0.16)
Transgender	6 (0.07)
Gender queer	13 (0.16)
Race, n (%)[a]	
White	6104 (75.90)
American Indian or Alaskan Native	27 (0.33)
Asian	508 (6.31)
Black or African American	560 (6.96)
Native Hawaiian or other Pacific Islander	23 (0.28)
Other	690 (8.57)
Prefer not to say	130 (1.61)
Age group, n (%)	
18-24 years	459 (4.35)
25-34 years	3931 (37.32)
35-44 years	3346 (31.77)
45-54 years	1604 (15.23)
55-64 years	783 (7.43)
65-74 years	326 (3.09)
75-84 years	74 (0.70)
85+ years	8 (0.07)
Number medical visits, mean (SD)	3.07 (3.11)
Number health coach visits, mean (SD)	2.30 (2.81)
Total membership duration, n (%)	
0-1 year	7361 (69.89)
1-2 years	1755 (16.66)
3 or more years	1415 (13.43)
Most frequent ICD[b] **codes, n (%)**	
Abdominal distension (gaseous)	3300 (31.33)
Other fatigue	3244 (30.80)
Anxiety disorder, unspecified	2554 (24.25)
Irritable bowel syndrome with diarrhea	1826 (17.33)

Characteristic	Value
Constipation, unspecified	1626 (15.44)
Insomnia, unspecified	1105 (10.49)
Hypothalamic dysfunction, not elsewhere classified	1079 (10.24)

[a]Due to missing data, N=8042 for this category.

[b]ICD: International Classification of Diseases and Related Health Problems.

Table 2. PSI[a] and SRH[b] descriptives (N=22,748).

Characteristic	Value, n (%)
Total responses	
1 response	4333 (19.04)
2 responses	5324 (23.40)
3 responses	5823 (25.59)
4 responses	3796 (16.68)
5 responses	2345 (10.30)
6 responses	786 (3.45)
7 responses	224 (0.98)
8 responses	80 (0.35)
9 responses	27 (0.11)
10 responses	10 (0.04)
Concordance: relative	
Congruent	16865 (74.19)
Incongruent	5883 (25.86)
Concordance: literal	
Congruent	8229 (36.18)
Incongruent	14519 (63.8)
Time submitted	
Daytime	5655 (24.85)
Evening	12146 (53.39)
Morning	1172 (5.15)
Night	3775 (16.59)
PSI mapped to SRH categories	
Excellent	2835 (12.46)
Very good	5209 (22.89)
Good	8801 (38.68)
Fair	2473 (10.85)
Poor	3430 (15.05)
SRH categories	
Excellent	817 (3.59)
Very good	5761 (25.28)
Good	9744 (42.76)
Fair	5105 (22.40)
Poor	1321 (5.79)

[a]PSI: Parsley Symptom Index.

[b]SRH: self-rated health.

Table 3. Association and agreement.

	Response count, n (%)	Polyserial correlation, *r*	Polychoric correlation, *r*	Relative concordance, n (%)	Literal concordance, n (%)	Weighted κ	Maximum κ
Total	22,732 (100)	0.517	0.522	16865 (74.19)	8229 (36.18)	0.460	0.754
Response 1	10,520 (46.28)	0.506	0.517	7623 (72.46)	3763 (35.77)	0.453	0.747
Response 2	6195 (27.25)	0.539	0.552	4704 (75.93)	2316 (37.38)	0.478	0.753
Response 3	3535 (15.55)	0.563	0.568	2674 (75.64)	1304 (36.89)	0.476	0.747
Response 4	1595 (7.02)	0.525	0.515	1190 (74.61)	535 (33.54)	0.421	0.715
Response 5	646 (2.84)	0.539	0.543	484 (74.92)	216 (33.44)	0.434	0.675
Response 6	177 (0.78)	0.665	0.660	138 (77.97)	65 (36.72)	0.545	0.664
Response 7	46 (0.20)	0.707	0.654	39 (84.78)	17 (36.96)	0.570	0.579
Response 8	14 (0.06)	0.687	0.569	10 (71.43)	6 (42.86)	0.446	0.897

Figure 1. PSI by SRH responses across time: two categories. PSI: Parsley Symptom Index; SRH: self-rated health.

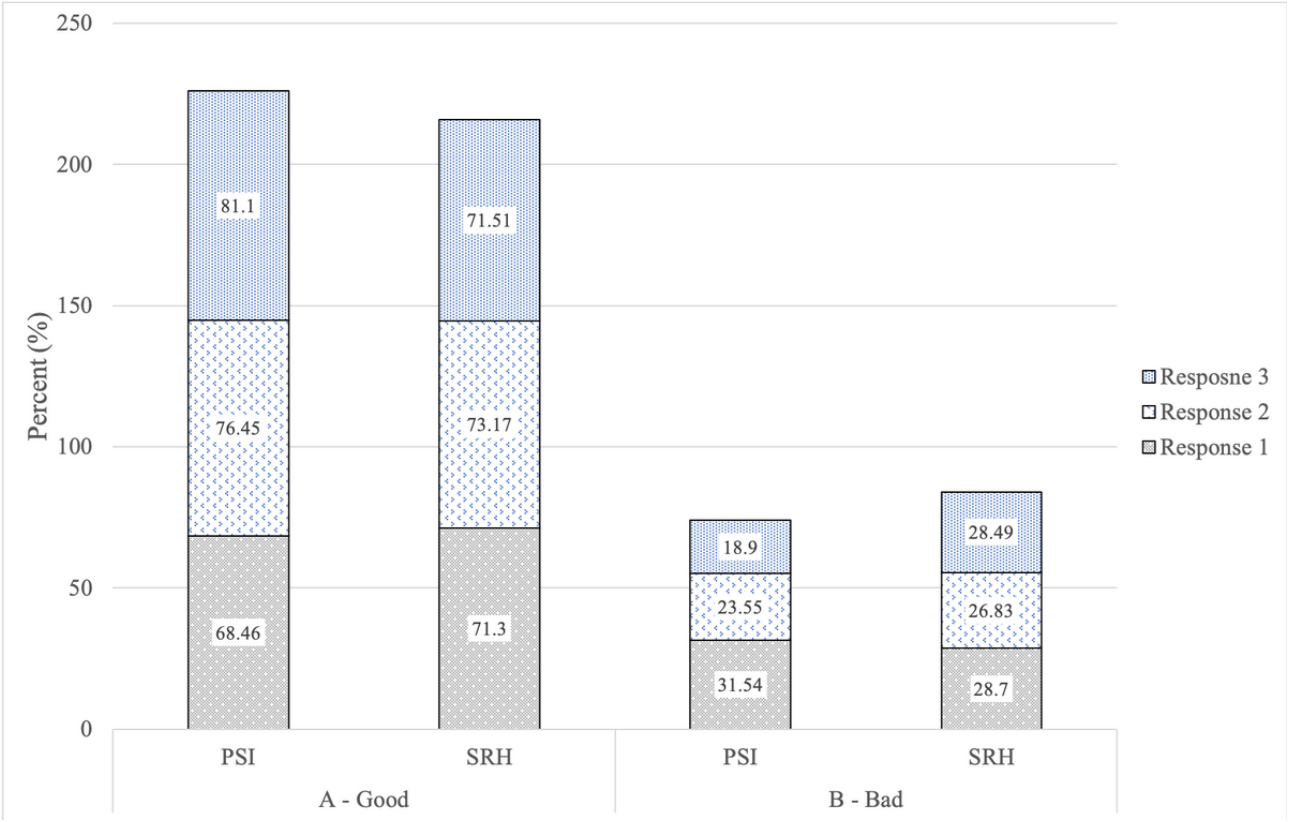
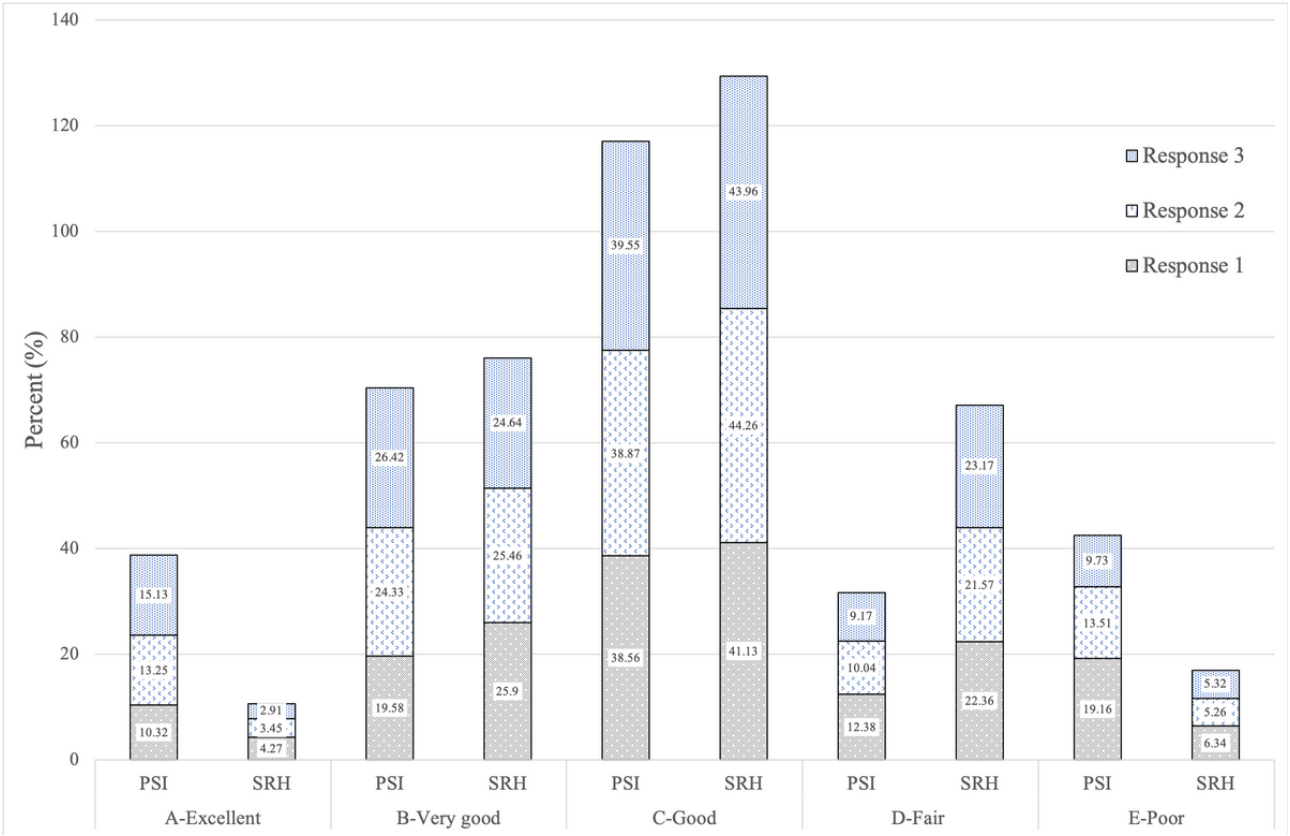


Figure 2. PSI by SRH responses across time: five categories. PSI: Parsley Symptom Index; SRH: self-rated health.



Discussion

Principal Findings

This study investigated the concordant validity of the PSI, a digital-first e-PROM, by comparing it to the SRH in a large adult population. We found moderate association and agreement (ie, relative concordance) between the PSI and SRH. When the PSI was scored as an ordinal, it did not perfectly match the 5 health categories in the SRH; however, they were consistent in terms of their position as good health (excellent, very good, good) versus bad health (fair, poor). In other words, the PSI and SRH generally point in the same directions for self-reported health categorization.

Various analyses were performed to explore association and agreement. First, we analyzed whether collapsing PSI scores into ordinal categorical variables (vs continuous) would change the association with SRH. The results were similar between continuous (polyserial correlation) and categorical (polychoric correlation) when compared to the categorical SRH. *t* tests showed no significant difference between these correlations. We also explored whether agreement between PSI and SRH were different between the first patient visit versus subsequent visits. Agreement between PSI and SRH for patients with repeated assessments remained consistent over time, suggesting consistency for the PSI from first visit to follow-up visits.

We noted that patients tended to report better health on the PSI than on the SRH. In this study, the SRH question was asked at the end of the PSI. It is possible that while answering the PSI questions, patients were reminded of their health symptoms leading them to be more likely to rate their health poorly in the SRH. The order of administration may play a role in the agreement level [32]. Future studies should incorporate A/B testing to explore whether the order of administration impacts the self-reported perception of well-being.

Although the focus of this study was not to assess or describe longitudinal changes between the PSI and SRH, we did observe that the PSI captured improvement in symptoms over time with treatment (Figure 1 and Figure 2). In comparison, the SRH remained relatively static over time. This implies that the PSI, with its greater degree of granularity, can capture symptom changes in a way that we would not expect from a single-item question like the SRH [33]. We did observe a broad range of PSI and SRH responses that fell into a normal distribution, indicating the full spectrum of perceived health statuses. This normal distribution persisted over time for both measures, but the PSI as reported was more sensitive to detecting changes over time. Further research should investigate potential moderators and mediators that influence PSI response change over time, such as baseline health status, age, sex, race, and pre-existing conditions.

Beyond the effects of administration order, there are conceptual differences between the PSI and SRH that may contribute to the degree of agreement. Although they address the same broad clinical concepts, the 45-item PSI captures more information than does the single-item SRH [34]. We would expect a general trend of agreement or relative concordance between the two, but not to such a high degree that it would match perfectly (literal concordance).

The PSI was created because a short-form e-PROM to capture a review of systems did not exist. Other PROMs like the

PROMIS [11,12], SF-36 [13,14], and the MSQ [15] are powerful assessment tools in their own right, but none were created to be a digital first in this new era of telehealth-centric care delivery. Although the PROMIS has many useful short forms, the most general ones were not designed to replace the ROS in the clinical encounter.

However, these results suggest that further validation of the PSI would benefit from comparing it to a PROM with similar granularity (eg, bodily system level) even if this PROM would not be a perfect conceptual match. The PROMIS, SF-36, and MSQ are similar enough that we hope to compare these tools to the PSI in future studies to better understand the PSI as a conceptually valid yet distinctly useful tool.

Limitations

The majority of Parsley Health members are White and female, so the study population was skewed in that direction, limiting the ecological validity of our results. Additionally, there was no randomization of PSI and SRH item presentation to address response biases. As the SRH was nested within the existing PSI, the infrastructure of the electronic health record could not support randomization. Future studies should consider randomization or A/B testing. There was also a lack of conceptually and operationally similar PROMs which we could use to validate the PSI. This is the reason that we created the PSI. In this study, we chose to compare the PSI to the SRH, a single-item questionnaire, to demonstrate convergent validity. Future studies will compare the PSI to PROMs that are similar in item length if not perfect matches in their design and intent.

Conclusions

This convergent validation study compared the best available questionnaire (SRH) to the PSI. Although the SRH and PSI fall under the same conceptual umbrella, they are different in their level of granularity. Further validation studies should compare the PSI to other multi-item, short-form PROMs of similar scope to continue the validation process. As telehealth will inevitably continue to grow, PROMs will be increasingly used and built as exclusively digital tools. Therefore, PROMs being used in the digital space must be researched and validated within the telehealth environment. This is a paradigm shift in the world of PROM development and validation. As this field evolves, we will need to assess what it means to validate a tool that is no longer administered to a captive audience in a physical waiting room, but rather, one that is engaged with remotely. Measures of engagement and “stickiness” will need to be considered as we build tools that can be completed anywhere and at any time of the day. These digital PROMs will need to be validated against previously validated tools while also being able to stand up to the test of our modern, all-access world.

Acknowledgments

The authors would like to thank Dylan Ray, Elena Céspedes, Patrick Hanaway, and Susan Silva for their support in this study.

This material was fully supported by Parsley Health. The funder had the following involvement with the study: study design, research, and preparation of the manuscript.

Authors' Contributions

HW, SS, RB, and HH contributed to the conception of the study design, article preparation, and data collection. KL and CO contributed to the article preparation. All authors read and approved the final manuscript. This statement confirms that this article has been submitted solely to this journal and is not published, in press, or submitted elsewhere.

Conflicts of Interest

All authors are either employees or consultants to Parsley Health at the time of analysis. All authors declare no other competing interests.

Multimedia Appendix 1
Checklist for reporting results of internet e-surveys.
[PDF File (Adobe PDF File), 72 KB - formative_v6i11e40063_app1.pdf]

References

1. Fair Health Study analyzes telehealth. FAIR Health. 2019. URL: https://www.fairhealth.org/article/fair-health-study-analyzes-telehealth [accessed 2022-02-03]
2. Peden CJ, Mohan S, Pagán V. Telemedicine and COVID-19: an observational study of rapid scale up in a US academic medical system. J GEN INTERN MED 2020 Jun 04;35(9):2823-2825. [doi: 10.1007/s11606-020-05917-9]
3. Koonin L, Hoots B, Tsang C, Leroy Z, Farris K, Jolly T, et al. Trends in the use of telehealth during the emergence of the COVID-19 pandemic - United States, January-March 2020. MMWR Morb Mortal Wkly Rep 2020 Oct 30;69(43):1595-1599 [FREE Full text] [doi: 10.15585/mmwr.mm6943a3] [Medline: 33119561]
4. Barbosa W, Zhou K, Waddell E, Myers T, Dorsey ER. Improving access to care: telemedicine across medical domains. Annu Rev Public Health 2021 Apr 01;42:463-481 [FREE Full text] [doi: 10.1146/annurev-publhealth-090519-093711] [Medline: 33798406]

5. Snoswell CL, Taylor ML, Comans TA, Smith AC, Gray LC, Caffery LJ. Determining if telehealth can reduce health system costs: scoping review. J Med Internet Res 2020 Oct 19;22(10):e17298 [FREE Full text] [doi: 10.2196/17298] [Medline: 33074157]
6. Buttorff C, Ruder T, Bauman M. Multiple chronic conditions in America. RAND Corporation. URL: https://www.rand.org/pubs/tools/TL221.html [accessed 2022-02-03]
7. Perez J, Niburski K, Stoopler M, Ingelmo P. Telehealth and chronic pain management from rapid adaptation to long-term implementation in pain medicine: A narrative review. PR9 2021 Jan 9;6(1):e912. [doi: 10.1097/pr9.0000000000000912]
8. Theodore BR, Whittington J, Towle C, Tauben DJ, Endicott-Popovsky B, Cahana A, et al. Transaction cost analysis of in-clinic versus telehealth consultations for chronic pain: preliminary evidence for rapid and affordable access to interdisciplinary collaborative consultation. Pain Med 2015 Jun;16(6):1045-1056 [FREE Full text] [doi: 10.1111/pme.12688] [Medline: 25616057]
9. Wicklund E. Hospital's telehealth program reduces ER visits, treatment costs. mHealth Intelligence. URL: https://mhealthintelligence.com/news/hospitals-telehealth-program-reduces-er-visits-treatment-costs [accessed 2022-02-10]
10. Greenhalgh J, Gooding K, Gibbons E, Dalkin S, Wright J, Valderas J, et al. How do patient reported outcome measures (PROMs) support clinician-patient communication and patient care? A realist synthesis. J Patient Rep Outcomes 2018 Dec;2:42 [FREE Full text] [doi: 10.1186/s41687-018-0061-6] [Medline: 30294712]
11. Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, et al. 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]
12. DeWalt DA, Rothrock N, Yount S, Stone AA, PROMIS Cooperative Group. Evaluation of item candidates: the PROMIS qualitative item review. Med Care 2007 May;45(5 Suppl 1):S12-S21 [FREE Full text] [doi: 10.1097/01.mlr.0000254567.79743.e2] [Medline: 17443114]
13. LoMartire R, Äng BO, Gerdle B, Vixner L. Psychometric properties of Short Form-36 Health Survey, EuroQol 5-dimensions, and Hospital Anxiety and Depression Scale in patients with chronic pain. Pain 2020 Jan;161(1):83-95 [FREE Full text] [doi: 10.1097/j.pain.0000000000001700] [Medline: 31568237]
14. Cordier R, Brown T, Clemson L, Byles J. Evaluating the Longitudinal Item and Category Stability of the SF-36 Full and Summary Scales Using Rasch Analysis. Biomed Res Int 2018;2018:1013453 [FREE Full text] [doi: 10.1155/2018/1013453] [Medline: 30519566]
15. Hyman MA. Symptom/Toxicity Questionnaire. MSQ-Medical. URL: http://drhyman.com/downloads/MSQ_Fillable.pdf [accessed 2015-01-01]
16. Chung AE, Basch EM. Incorporating the patient's voice into electronic health records through patient-reported outcomes as the "review of systems". J Am Med Inform Assoc 2015 Jul;22(4):914-916 [FREE Full text] [doi: 10.1093/jamia/ocu007] [Medline: 25614143]
17. Franklin P, Chenok K, Lavalee D, Love R, Paxton L, Segal C, et al. Framework to guide the collection and use of patient-reported outcome measures in the learning healthcare system. EGEMS (Wash DC) 2017 Sep 04;5(1):17 [FREE Full text] [doi: 10.5334/egems.227] [Medline: 29881737]
18. Williams H, Steinberg S, Berzin R. The development of a digital patient-reported outcome measurement for adults with chronic disease (The Parsley Symptom Index): prospective cohort study. JMIR Form Res 2021 Jun 11;5(6):e29122 [FREE Full text] [doi: 10.2196/29122] [Medline: 33999007]
19. Cullati S, Mukhopadhyay S, Sieber S, Chakraborty A, Burton-Jeangros C. Is the single self-rated health item reliable in India? A construct validity study. BMJ Glob Health 2018;3(6):e000856 [FREE Full text] [doi: 10.1136/bmjgh-2018-000856] [Medline: 30483411]
20. Schnittker J, Bacak V. The increasing predictive validity of self-rated health. PLoS One 2014;9(1):e84933 [FREE Full text] [doi: 10.1371/journal.pone.0084933] [Medline: 24465452]
21. Kaplan MS, Berthelot J, Feeny D, McFarland BH, Khan S, Orpana H. The predictive validity of health-related quality of life measures: mortality in a longitudinal population-based study. Qual Life Res 2007 Nov;16(9):1539-1546. [doi: 10.1007/s11136-007-9256-7] [Medline: 17899447]
22. U.S. DOHS. Code of Federal Regulations–Public Welfare: Title 45(Sections 46-101 to 46-409): Protection of human subjects. US Department of Health and Human Services. 1995. URL: https://www.hhs.gov/ohrp/sites/default/files/ohrp/policy/ohrpregulations.pdf [accessed 2022-02-10]
23. Federal Drug Administration (FDA) 2019. Patient-reported outcome measures: use in medical product development to support labeling claims. URL: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/patient-reported-outcome-measures-use-medical-product-development-support-labeling-claims [accessed 2022-02-10]
24. Zajacova A, Dowd JB. Reliability of self-rated health in US adults. American Journal of Epidemiology 2011 Sep 02;174(8):977-983. [doi: 10.1093/aje/kwr204] [Medline: 21890836]
25. Ranganathan P, Pramesh CS, Aggarwal R. Common pitfalls in statistical analysis: Measures of agreement. Perspect Clin Res 2017;8(4):187-191 [FREE Full text] [doi: 10.4103/picr.PICR_123_17] [Medline: 29109937]
26. Julius S, Chris CW, The KSIRSU. The Kappa Statistic in reliability studies: use, interpretation, and sample size requirements. Physical Therapy 2005;85(3):257-268 [FREE Full text] [doi: 10.1093/ptj/85.3.257]

27. Jürges H, Avendano M, Mackenbach JP. Are different measures of self-rated health comparable? An assessment in five European countries. Eur J Epidemiol 2008 Sep;23(12):773-781. [doi: 10.1007/s10654-008-9287-6] [Medline: 18814040]
28. Khan MAB, Hashim MJ, King JK, Govender RD, Mustafa H, Al Kaabi J. Epidemiology of type 2 diabetes - global burden of disease and forecasted trends. J Epidemiol Glob Health 2020 Mar;10(1):107-111 [FREE Full text] [doi: 10.2991/jegh.k.191028.001] [Medline: 32175717]
29. Zaki R, Bulgiba A, Ismail R, Ismail NA. Statistical methods used to test for agreement of medical instruments measuring continuous variables in method comparison studies: a systematic review. PLoS One 2012;7(5):e37908 [FREE Full text] [doi: 10.1371/journal.pone.0037908] [Medline: 22662248]
30. R: A language and environment for statistical computing. R Foundation for Statistical Computing. 2016. URL: https://www.r-project.org/ [accessed 2016-11-01]
31. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
32. Zini, Banfi G. A narrative literature review of bias in collecting patient reported outcomes measures (PROMs). Int J Environ Res Public Health 2021 Nov 26;18(23):124-145 [FREE Full text] [doi: 10.3390/ijerph182312445] [Medline: 34886170]
33. K. Van Ginneken J. A single- vs. multi-item self-rated health status measure: a 21-country study. TOPHJ 2012 Apr 12;5(1):1-9 [FREE Full text] [doi: 10.2174/1874944501205010001]
34. Bowling A. Just one question: If one question works, why ask several? J Epidemiol Community Health 2005 May;59(5):342-345. [doi: 10.1136/jech.2004.021204] [Medline: 15831678]

Abbreviations

CHERRIES: Checklist for Reporting Results of Internet E-Surveys
e-PROM: electronic patient-reported outcomes measure
FDA: Federal Drug Agency
MSQ: Medical Symptom Toxicity Questionnaire
PROMIS: Patient-Reported Outcomes Measurement Information System
PSI: Parsley Symptom Index
ROS: review of systems
SF-36: 36-Item Short Form Health Survey
SRH: self-rated health

Edited by A Mavragani; submitted 03.06.22; peer-reviewed by KL Mauco, A Logan; comments to author 29.07.22; revised version received 19.08.22; accepted 02.09.22; published 03.11.22.

Please cite as:
Williams H, Steinberg S, Leon K, O'Shea C, Berzin R, Hagg H
Validity of the Parsley Symptom Index—an Electronic Patient-Reported Outcomes Measure Designed for Telehealth: Prospective Cohort Study
JMIR Form Res 2022;6(11):e40063
URL: https://formative.jmir.org/2022/11/e40063
doi:10.2196/40063
PMID:36326802

Original Paper

A Short Digital Food Frequency Questionnaire (DIGIKOST-FFQ) Assessing Dietary Intake and Other Lifestyle Factors Among Norwegians: Qualitative Evaluation With Focus Group Interviews and Usability Testing

Hege Berg Henriksen[1], PhD; Markus Dines Knudsen[2,3,4], PhD; Monica Hauger Carlsen[1], PhD; Anette Hjartåker[1], PhD; Rune Blomhoff[1,5], PhD

[1]Department of Nutrition, Institute of Basic Medical Sciences, University of Oslo, Oslo, Norway

[2]Section for Colorectal Cancer Screening, Cancer Registry of Norway, Oslo, Norway

[3]Department of Transplantation Medicine, Division of Surgery, Inflammatory Diseases and Transplantation, Oslo University Hospital, Rikshospitalet, Oslo, Norway

[4]Department of Epidemiology, Harvard TH Chan School of Public Health, Boston, MA, United States

[5]Division of Cancer Medicine, Oslo University Hospital, Oslo, Norway

Corresponding Author:
Hege Berg Henriksen, PhD
Department of Nutrition
Institute of Basic Medical Sciences
University of Oslo
Sognsvannsveien 9
Oslo, 0372
Norway
Phone: 47 99459673
Email: h.b.henriksen@medisin.uio.no

Abstract

Background: In-person dietary counseling and interventions have shown promising results in changing habits toward healthier lifestyles, but they are costly to implement in large populations. Developing digital tools to assess individual dietary intake and lifestyle with integrated personalized feedback systems may help overcome this challenge. We developed a short digital food frequency questionnaire, known as the DIGIKOST-FFQ, to assess diet and other lifestyle factors based on the Norwegian Food-Based Dietary Guidelines. The DIGIKOST-FFQ includes a personalized feedback system, the DIGIKOST report, that benchmarks diet and lifestyle habits. We used qualitative focus group interviews and usability tests to test the feasibility and usability of the DIGIKOST application.

Objective: We aimed to explore attitudes, perceptions, and challenges in completing the DIGIKOST-FFQ. We also investigated perceptions and understanding of the personalized feedback in the DIGIKOST report and the technical flow and usability of the DIGIKOST-FFQ and the DIGIKOST report.

Methods: Healthy individuals and cancer survivors were invited to participate in the focus group interviews. The transcripts were analyzed using thematic analysis. Another group of healthy individuals completed the usability testing, which was administered individually by a moderator and 2 observers. The results were analyzed based on predefined assignments and discussion with the participants about the interpretation of the DIGIKOST report and technical flow of the DIGIKOST-FFQ.

Results: A total of 20 individuals participated in the focus group interviews, divided into 3 groups of healthy individuals and 3 groups of cancer survivors. Each group consisted of 3 to 4 individuals. Five main themes were investigated: (1) completion time (on average 19.1, SD 8.3, minutes, an acceptable duration), (2) layout (participants reported the DIGIKOST-FFQ was easy to navigate and had clear questions but presented challenges in reporting dietary intake, sedentary time, and physical activity in the last year), (3) questions (the introductory questions on habitual intake worked well), (4) pictures (the pictures were very helpful, but some portion sizes were difficult to differentiate and adding weight in grams would have been helpful), and (5) motivation (users were motivated to obtain personalized feedback). Four individuals participated in the usability testing. The

results showed that the users could seamlessly log in, give consent, fill in the DIGIKOST-FFQ, and receive, print, and read the DIGIKOST report. However, parts of the report were perceived as difficult to interpret.

Conclusions: The DIGIKOST-FFQ was overall well received by participants, who found it feasible to use; however, some adjustments with regard to reporting dietary intake and lifestyle habits were suggested. The DIGIKOST report with personalized feedback was the main motivation to complete the questionnaire. The results from the usability testing revealed a need for adjustments and updates to make the report easier to read.

(JMIR Form Res 2022;6(11):e35933) doi:10.2196/35933

KEYWORDS

digital assessment tool; assessment tool; food frequency questionnaire; food; diet; nutrition; questionnaire; focus group; interview; usability; physical activity; lifestyle factor; dietary intake; digital health; chronic disease; chronic condition; health promotion; cancer; survivor; usability; thematic analysis; research tool; measurement tool

Introduction

The Norwegian Food-Based Dietary Guidelines (Norwegian FBDG), published by the Norwegian health authorities, aim to reduce the risk of lifestyle-related chronic diseases and promote overall health in the general Norwegian population [1]. There is a lack of, and therefore a need for, easily accessible dietary and lifestyle assessment tools with a low respondent burden, (eg, by being feasible, quick to complete, motivational to use, and able to collect accurate, precise lifestyle data) [2-5]. Motivation is a crucial factor when recruiting people to undertake the challenge of filling in questionnaires. Personal feedback has been used in other studies to motivate study participants [6-8]. Thus, a digital diet and lifestyle questionnaire tool that automatically gives personalized feedback, including health and diet-related advice, would theoretically increase completion in future studies.

Digital applications assessing diet and lifestyle behaviors for use in epidemiological and clinical studies are emerging [2,6,8-12]. For instance, a web-based, semiquantitative food frequency questionnaire (FFQ) assessing habitual diet over the last year was found to be feasible for use among healthy adults living in Norway [13] and served as a valuable tool to be used in epidemiological studies. However, this web FFQ was solely developed to collect data on dietary intake, without any digital application for individual feedback reports. Forster et al [3] developed a dietary feedback system based on dietary intake, Food4Me FFQ, which was found to be well accepted by participants and feasible for use [8]. Moreover, the MyFood decision-support system was developed to assess symptoms and dietary intake among hospitalized patients at risk of malnutrition and to generate reports on personalized nutritional treatment for use by nurses or other health care professionals [9,14,15]. Another digital application for clinical use is eCHANGE, developed as a personalized digital intervention aiming at providing self-management support for long-term weight maintenance [12].

To the best of our knowledge, no digital application has been developed to assess dietary intake according to the Norwegian FBDG with integrated personalized feedback reports. We developed the DIGIKOST-FFQ, a short, digital, semiquantitative food and lifestyle frequency questionnaire, designed to assess adherence to the Norwegian FBDG. The DIGIKOST-FFQ is applicable in a number of settings where information on diet and lifestyle is needed. Based on a respondent's answers to the DIGIKOST-FFQ, a report, known as the DIGIKOST report, is automatically generated and immediately made available to the respondent after completion. It gives individual feedback on the respondent's adherence to the Norwegian FBDG and on other lifestyle factors, as well as advice on how to fulfill the recommendations.

The process of creating new digital dietary and lifestyle assessment tools involves several developmental stages, from defining the different constructs in the questionnaires to making it feasible for use. In addition, new research tools must be evaluated to explore their validity and reproducibility [16]. To evaluate digital tools that assess diet or lifestyle, qualitative methods, such as focus group interviews, are increasingly being used, constituting mobile health (mHealth) [16-18]. Focus groups are particularly useful for exploring people's knowledge and experiences, and can be used to examine not only what people think, but also how they think and why they think the way they do [19,20]. Results from focus group interviews are used in the further development of the tools in question.

During a focus group interview, the participants are invited to share their views, comments, and perspectives, phrased in their own words and in synergy with the other participants in the group. A moderator ensures that the structure and framework of the interview follow the focus group interview guide [19,20].

Usability testing is also a key component in the development of digital applications [21,22] and is critical in the development and improvement of the design, function, and understanding of the tool being developed. It is performed by real users trying to accomplish typical goals and tasks in a test version of the digital tool under controlled conditions, allowing researchers and the development team to observe and take notes [21,22]. Results from these observations are then used in the development of the tool.

As part of the development of the DIGIKOST-FFQ, we performed several focus group interviews with healthy individuals and cancer survivors, because both of these groups are expected to be important study populations in future research studies using the DIGIKOST-FFQ. Furthermore, an independent group of healthy individuals was invited to the usability testing of both the DIGIKOST-FFQ and the DIGIKOST report. In the current paper, we present the results from the focus group

interviews and the usability testing of the DIGIKOST-FFQ and the DIGIKOST report.

Methods

The DIGIKOST-FFQ

The DIGIKOST-FFQ is derived from a paper-based, validated, short, semiquantitative food frequency questionnaire called the NORDIET-FFQ [23,24], which was designed to measure adherence to the Norwegian FBDG. The first draft of the DIGIKOST-FFQ that underwent evaluation in the current study consisted of 80 questions on diet; 5 on physical activity, time being sedentary, and sleeping; 10 on tobacco use; and 9 on demographic data. In addition, the questionnaire included introductory questions about the usual intake of specific food groups. If a participant indicated no intake of any food item in a food group, the participant was redirected to the next food group. Moreover, when reporting no intake for a specific food item, the associated question on amount disappeared, due to an automatic function in the questionnaire, and the participant was redirected to the next food group.

The DIGIKOST Report

The DIGIKOST report presents adherence to the Norwegian FBDG in different ways (Multimedia Appendix 1). First, dietary intake and physical activity are estimated from the DIGIKOST-FFQ and presented in a table that compares the results with the Norwegian FBDG. Next, the same components are presented graphically, with columns presenting the degree of adherence to the Norwegian FBDG, measured as a percentage with traffic-light coloring. In the next section, adherence to the recommendations is presented as a health index consisting of 5 lifestyle components: diet, weight status (BMI), physical activity, smoking, and intake of alcohol. Each component is equally weighted. The degree of adherence is divided into a 3-level scoring system for diet, weight status, and physical activity, ranging from no adherence (0 points) to intermediate adherence (0.5 points) and full adherence (1 point). For alcohol and tobacco use, the degree of adherence is binary (0 or 1). The total health index ranges from 0 to 5 points. The participant's achievements in the health index score are presented and compared with the maximum score of the index. In addition, benchmarking of recorded individual diet and other lifestyle factors against the Norwegian FBDG is presented, along with individual advice on how to fulfill the recommendations. At the end of the report, the Healthy Eating Plate is presented (Multimedia Appendix 1). The Healthy Eating Plate is a commonly used model that focuses on diet quality. The plate features 3 individual sections: one-third should consist of vegetables, one-third of whole grains and starchy vegetables, and one-third of fish, meat, or legumes [25].

Subjects

Both healthy adult individuals and cancer survivors were recruited to the focus group interviews, while only healthy individuals were recruited to the usability testing. Healthy individuals were recruited using Facebook announcements, with separate recruitment processes for focus group interviews and the usability testing. Cancer survivors were recruited from an ongoing randomized controlled trial, the CRC-NORDIET study, and from the Norwegian Cancer Society user group panel [26,27]. The recruitment period for both studies was from April to June 2020. The focus group interviews and the usability testing were all conducted remotely by video meetings on Zoom due to the COVID-19 pandemic in 2020 and 2021. Group sizes of 5 to 10 and 4 to 5 individuals are recommended for focus group interviews and usability testing, respectively [28,29].

Focus Group Interviews

A moderator and 2 researchers from the Department of Nutrition at the University of Oslo (UiO) led the focus group interviews. Recording of the interviews was done with both Zoom and the Dictaphone app for smartphones, which sent all data directly to a secure server, the Services for Sensitive Data ("Tjenester for sensitive data," abbreviated "TSD" in Norwegian), at the University Center for Information Technology (USIT) at UiO [28]. In addition, the moderator and the assistants recorded feedback with written notes. All recorded data were safely stored at TSD. The focus group interviews were conducted according to the interview guide for focus groups that is included as part of the DIGIKOST-FFQ (Multimedia Appendix 2). The DIGIKOST reports were not ready for use at the time of the focus group interviews; therefore, the focus group interviews only tested the DIGIKOST FFQ and not the DIGIKOST reports. However, the focus group participants were asked whether a personal report on their individual lifestyles would motivate them to complete the DIGIKOST-FFQ.

All transcripts of the recorded focus group interviews were made with f4transkript software (version 6.2.5 Pro; Dr. Dresing & Pehl GmbH) [30]. The transcripts were analyzed by taking notes and coming to an overall understanding of the basic responses (ie, themes) and creating codes and constructs according to the DIGIKOST-FFQ interview guide for focus groups (Multimedia Appendix 2). The results were stratified into groups comprising healthy individuals and cancer survivors. Analysis of the transcripts was done manually and independently by 2 researchers, HBH and MDK, to identify the main themes and constructs in the responses. Afterwards, a thorough evaluation of the results was performed collaboratively by the same researchers. The most frequent responses were noted, as well as the single response evaluated as most improving the feasibility of the DIGIKOST-FFQ; these were included in the revised version.

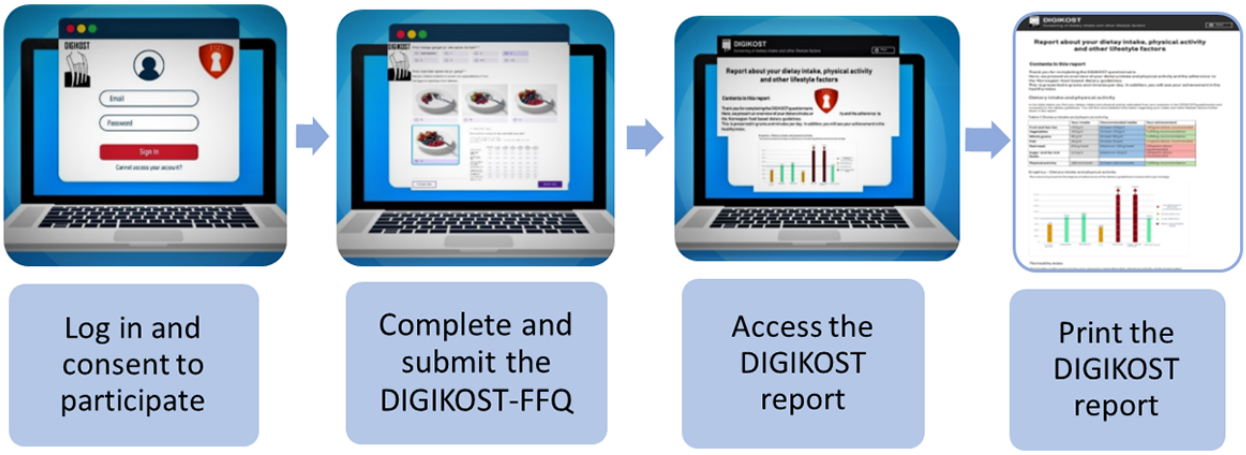
Usability Testing of the DIGIKOST Report

The usability testing was completed individually and conducted on Zoom with each participant. At this time, the DIGIKOST report was ready for using and testing. A moderator from USIT at UiO led the usability testing, with 2 observers from the Department of Nutrition at UiO taking notes on how the participants completed the planned tasks. The purpose was to test the technical flow of the DIGIKOST-FFQ, from consent and completion of the DIGIKOST-FFQ to opening and printing the individual DIGIKOST reports (Figure 1).

Furthermore, after printing the DIGIKOST report, the participants shared their understanding and views of the report (Multimedia Appendix 1) by answering questions about (1)

their understanding of the graphs in the report, (2) their comprehension of the specific recommendations for diet and health improvements, (3) their conception of the table, and (4) their conception of the health index included in the report. The usability testing was performed according to a standard protocol developed by USIT [31,32].

Figure 1. Technical flow of usability testing.



Ethics

The current study was carried out in accordance with the Helsinki Declaration; informed consent was obtained from all participants. The Norwegian Centre for Research Data approved the focus group protocol, the usability testing, and the informed consent (277679). All participants signed the informed consent form before completing the DIGIKOST-FFQ, focus group interviews, and usability testing. After accepting the study invitation, the participants were asked to fill out the DIGIKOST-FFQ before the focus group interviews. For the usability testing, the participants did not have access to the DIGIKOST-FFQ or the DIGIKOST report before attending the testing.

Results

Focus Group Interviews

A total of 20 adults, including 11 women and 9 men, participated in 6 focus group interviews, with 3 to 4 participants in each group. Each interview lasted for 1 to 1.5 hours. The focus group discussions had a natural flow, but were guided by the motivator using the interview guide. Five main themes and subthemes were identified through analysis of the transcripts: (1) the time it took to complete the questionnaire, (2) the layout of the questionnaire, (3) the questions in the questionnaire, (4) the pictures of portion sizes in the questionnaire, and (5) motivations for the participant to fill out the questionnaire (Table 1).

In general, the responses from the 2 groups of participants were similar and addressed the same themes and topics. However, cancer survivors had a harder time reporting physical activity and time being sedentary than healthy individuals. Healthy individuals asked for more questions about plant-based food and a third gender option.

Table 1. Summary of responses from the focus groups on main themes obtained from a discussion that was based on the interview guide included in the DIGIKOST-FFQ. Results are stratified by group.

Main themes	Responses from healthy individuals (7 females, 3 males)	Responses from cancer survivors (4 females, 6 males)
Completion time		
Completion time (minutes), average (SD, range)	17.2 (4.5, 10-25)	21.1 (11.1, 15-45)
Main comments about completion time	"Got tired at the end"; "the questionnaire cannot be longer"	"I have no belief in completing in 15 minutes"
Layout		
What themes worked well	Differentiation of portion sizes, ease of navigation, clarity of questions, reporting physical activity	Having to answer every question, automatic calculation of bread slices
What themes did not work well	Questions on money spent on snuff; reporting intakes in the previous year, over several seasons, or of foods you eat less than once a week when the answer is equal to never; and about duration of residence in Norway when you have answered that you were born in Norway	Questions about money spent last year or season on snuff, eggs, or jam; sedentary time; physical activity; and about duration of residence in Norway when you have answered that you were born in Norway. Pictures of glasses filled with alcoholic beverages would have been helpful
Questions		
Yes/no as introduction to a food group	Worked well, could have been implemented for all thematic question groups	Worked well
Missing questions/options	Questions were missing on plant-based cold cuts, alternative milk products (eg, soy, oat, or rice), and potatoes; there were only 2 categories for answers on marital status; add student as a separate category from education; Norway should be at the top of the pull-down choice list of countries; there were no answer options for a third category for gender or separate categories for full-time and part-time work	Questions were missing on legumes, eggs, and potatoes; Norway should be at the top of the pull-down choice list, "transport" was not included in the answer options for sedentariness; no option for power naps
Pictures		
Portion size	Informative text should have been added to the pictures and some portions, such as for A, B, C, and D, which were difficult to tell apart	Informative text should have been added to the pictures and some portions, such as for A, B, C, and D, which were difficult to tell apart
Were the pictures helpful or could they have been text only?	Pictures that included weight measurements in grams would have been helpful; references to the pictures would improve the information, as would more text before new food items and pictures	Pictures of better quality would help reporting quantitative intake; weight measurements in grams should have been included as references in the pictures; more text should have been included before new food items and pictures
Motivation	The individual reports were motivational	The individual reports were motivational

Completion Time

All participants completed the DIGIKOST-FFQ within a scheduled time of about 15 to 20 minutes. However, many respondents stressed that the questionnaire should not take more time than this to fill out:

> *I see that I got tired at the end...* [Healthy female participant; June 22, 2020]
>
> *I noticed that in the end it became more skim reading.* [Healthy male individual; June 22, 2020]

However, the participants also stated that the digital format was easier than a paper questionnaire:

> *It is much simpler compared to the whole paper mill where you have to sit for hours to complete.* [Female cancer survivor; May 8, 2020]

Layout

Both groups agreed that the DIGIKOST-FFQ was easy to navigate and the questions were easy to understand:

> *I think it was easy to navigate and there was a good progression and so, you were not surprised by any of the questions. The questions appeared clear and concise and easy to understand.* [Healthy male individual; June 22, 2020]
>
> *...and then the question and portion sizes are simple and straight forward. Period.* [Male cancer survivor; May 8, 2020]

The advantage of including different kinds of questions in the questionnaire to prevent it from becoming boring and tiring toward the end was also emphasized:

> *I believe in general that it is good if you change the way you ask the questions, so you do keep yourself*

> *awake and not just "click" your way down.* [Healthy female individual; March 23, 2020]
>
> *A questionnaire to be completed with a lot of fun...* [Healthy male individual; June 23, 2020]

We observed 2 challenges with the questionnaire: reporting seasonal variation in dietary intake and registering time being sedentary:

> *We are living in a society with big variation according to season. Especially in the north where the access to fruit and vegetables is determined by season.* [Healthy female individual; June 12, 2020]
>
> *I believe that it is very season dependent what we eat now and what we eat later. In the winter you eat paprika, whereas in the summer you eat strawberries, which has just started now.* [Healthy male individual; June 23, 2020]
>
> *I found it (sedentary time) seriously complicated to answer, I didn't have a chance.* [Female cancer survivor; June 22, 2020]
>
> *I agree (sedentary time), I was very... I have no idea, however then I tried to picture an ordinary day, but I think there were a huge difference between what I pictured and what I noted for an ordinary day, to say it that way.* [Healthy male individual; June 22, 2020]

Questions

The introductory questions were well accepted by the participants:

> *It worked very well.* [Female and male healthy individuals; June 12, 2020]
>
> *I think... for my part it fitted very well, you know... I can say "no" to things that I for sure never eat, and when I say "yes" to something, I pretty much found what I eat. However, I am probably more average when it comes to diet [laughter].* [Male cancer survivor; May 12, 2020]
>
> *It worked very well.* [Female cancer survivor; April 14, 2020]
>
> *I think it was okay to answer yes or no, I had no problems with that.* [Healthy female individual; June 22, 2020]

There were a few alternative answer options that the participants felt were missing from the questionnaire, such as categories for gender:

> *I noticed that there were only two variables [for gender].* [Healthy male individual; June 22, 2020]
>
> *Yes, in modern society you probably should have a third alternative for gender. And I, working at a health clinic for rare diseases, know that this exists.* [Healthy female individual; June 12, 2020]

Moreover, the cancer survivors emphasized the importance of being able to include power naps and transportation time in the definition of sedentary time. It was suggested that a function for summing the different levels of physical activity (eg, time being sedentary, sleeping, or engaging in physical activity) throughout a 24-hour period should be included to increase the feasibility of reporting those questions correctly. The cancer survivors also felt that questions were missing about intake of legumes, potatoes, and eggs. Reporting intakes of berries used in homemade jam was also challenging for some, because they were unsure whether to report homemade jam as jam in the questions about spreads or in the questions about berries.

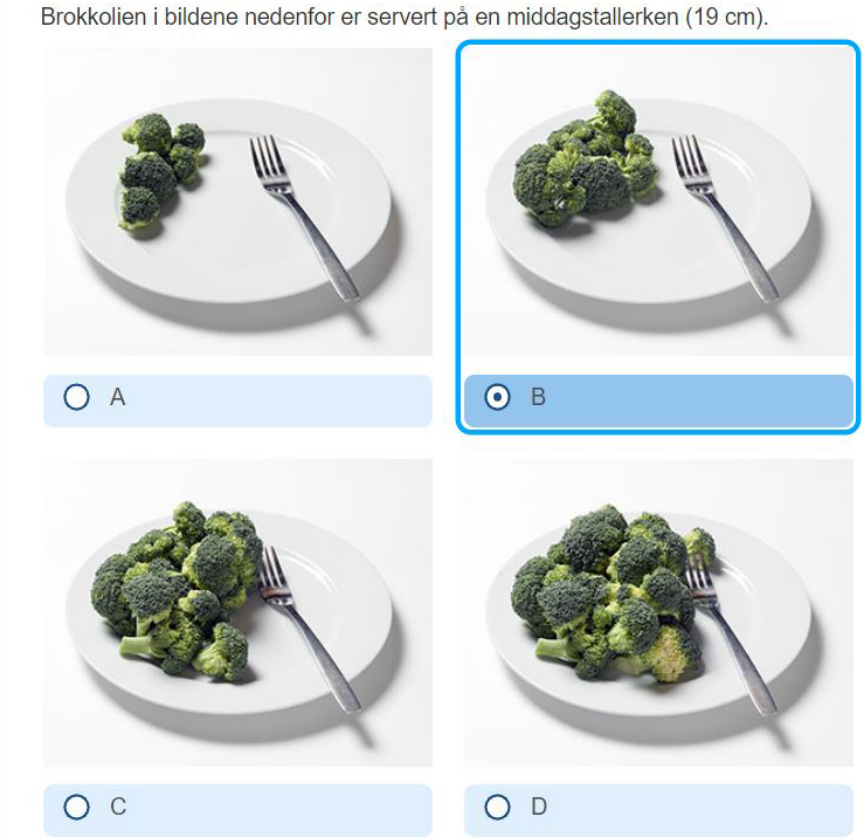
Pictures

There was a request for more text and the addition of grams to each image with portion sizes to increase the accuracy of the reporting. In particular, for some participants, it was difficult to tell some portion illustrations apart, especially portion sizes A and B and portion sizes C and D, as illustrated in Figure 2.

Some of the participant responses were as follows:

> *Now we are closing in on where I really missed grams for references.* [Male cancer survivor; May 8, 2020]
>
> *Beautiful pictures, particularly the berries.* [Female cancer survivor; April 14, 2020]

Figure 2. Example of a question from the DIGIKOST-FFQ regarding intake of broccoli, with 4 illustrations of portion sizes.



Motivation

A report about adherence to the Norwegian FBDG, along with a trending curve showing the change in intake over time (for studies where the participants would fill in the DIGIKOST-FFQ repeatedly over time), would have motivated participants in both groups to complete the DIGIKOST-FFQ:

> *I would like to find out more about my diet and lifestyle out of curiosity.* [Healthy female individual; June 23, 2020]
>
> *I think that some kind of report about your lifestyle, would be great fun.* [Healthy female individual; June 22, 2020]
>
> *Yes, it is obvious, it is fun to get feedback. See how you are placed... What kind of actions should I take.* [Healthy female individual; June 22, 2020]
>
> *It would be nice with feedback and compare your current to how you did in previous rounds... see if you have improved or worsened, preferably in tables or figures.* [Male cancer survivor; May 12, 2020]

Usability Testing

A total of 4 of 5 invited individuals, including 2 women and 2 men, participated in the usability testing. All attendees showed good technical skills. There was wide variation in their educational backgrounds, ages, and residential locations in Norway (data not shown). It took approximately 1 hour to conduct the usability testing for each attendee. All participants completed the 4 test tasks (Figure 1). However, due to technical issues with the ID portal, 1 of the participants could not test the first 3 steps in the technical flow. Therefore, that participant completed the questionnaire without the log-in function. Moreover, due to a technical issue, the DIGIKOST reports were emailed to the participants after completion of the DIGIKOST-FFQ, instead of appearing automatically online. This technical issue has now been resolved.

Technical Flow

The 3 participants who completed the technical flow testing (ie, the first 3 steps of the test) of the DIGIKOST-FFQ performed well, but 1 of the participants had minor challenges completing and submitting the DIGIKOST-FFQ (Table 2).

Table 2. Summary of participant performance in the usability and technical flow testing, based on the usability protocol.

Technical flow steps	Participant 1	Participant 2	Participant 3	Participant 4
1. Log in and consent to participate	Passed	Passed	Passed	Did not pass due to technical issues
2. Complete and submit the DIGIKOST-FFQ	Passed	Passed, but with minor issues	Passed	Did not pass due to technical issues
3. Access the personal DIGIKOST report via email	Passed	Passed	Passed	Did not pass due to technical issues
4. Print the DIGIKOST report	Passed	Passed	Passed	Passed

Interpretation of the DIGIKOST Report

The participants were, in general, positive about the DIGIKOST report (Multimedia Appendix 1). The coloring used to indicate adherence to the recommendations was well accepted, because it illustrated the degree of adherence in a clear manner. Three of the participants found the first table informative, whereas 1 preferred to look at the graphic with adherence shown as a percentage rather than read the table (because there was too much text). The report also included a graph showing adherence to the recommendations as a percentage. When a respondent had a higher intake of a food group than the minimum recommended amount, such as if they ate more than 5 fruits or vegetables a day, they would get a score of above 100%. All participants understood the graphic showing adherence to the recommendations as a percentage. However, 1 participant pointed out that the concept of percentages >100% could be challenging to understand.

The aim of the health index included in the report appeared to be unclear and difficult to understand for most participants. One participant asked whether the aim was to increase your BMI to achieve a full score by increasing dietary intake. Another participant found it difficult to see the difference between the 2 indices (ie, BMI and health index). Some found it difficult to understand the total sum score of 5 points when each component in the score reached only a maximum of 1 point. The participants suggested either improving the presentation of the health index by making it simpler or removing it from the report entirely. The participants liked the use of traffic-light coloring; however, it was pointed out that the colors used could be challenging for individuals who are color blind.

All participants reacted positively to the immediate individual response with advice on how to improve adherence to the recommendations, and 1 pointed out that it was very helpful. Another participant suggested reorganizing the responses by presenting the advice on improving adherence to the recommendations first and presenting the recommendations that were fulfilled second.

Discussion

Principal Findings

The DIGIKOST-FFQ, with the DIGIKOST report, is the first short digital FFQ and personal report that has been benchmarked against the Norwegian FBDG. The findings from both the focus group interviews and the usability testing showed that the DIGIKOST-FFQ and the DIGIKOST report were overall well accepted and easy to use. However, the study also revealed challenges for users and a need for some improvements. Particularly for the DIGIKOST-FFQ, there were challenges related to reporting seasonal variation in dietary intake over the last year, reporting physical activity, and differentiating images illustrating different portion sizes. FFQs rely on memory and a participant's conceptualization of portion size and frequency of intake, and these are frequently mentioned as challenges in the literature [33-35]. Seasonal variation in dietary intake is another well-known challenge when using FFQs that collect data from several seasons or ask about intake of foods that are specifically seasonal [33-35]. For the DIGIKOST report, interpretation of the graphics and percentages in the histograms was the main challenge.

To understand the willingness to respond and the motivation to participate in surveys and thereby improve survey effectiveness, de Leeuw et al [36] used an attitude scale to measure survey attitudes. Results from this study revealed 3 dimensions representing important contributors to participation in surveys: survey enjoyment (this reflects the individual perception of surveys as a positive experience), survey value (ie, salience, relevance, and usefulness), and survey burden (ie, if the survey is perceived as a burden on the individual, it has a negative influence on motivation and participation). All participants in the current study acknowledged the value of the DIGIKOST-FFQ and the importance of monitoring diet and lifestyle factors in the prevention of disease and improvement of quality of life. They also experienced survey enjoyment through the pleasant design of the questionnaire. Pictures and other visual aids (eg, reference objects, household measures, and food packaging) have been shown to be preferable and beneficial when assessing portion sizes and helpful in improving the accuracy of food quantification [5,37,38]. This is in accord with the current study, where pictures of portions were perceived as helpful, although adding more information, such as amounts in grams or household measures, was suggested as a way to increase the usability even more.

All participants agreed that the questionnaire should be short and take no more than 15 to 20 minutes to complete to reduce the survey burden and maintain enjoyment and motivation to participate. Previous studies have found that the use of digital dietary assessment tools is perceived as more fun, more motivational to use, and preferable to paper-based dietary assessment tools [2,4,39,40]. This was also supported by the participants in the current study.

Overall, the results from the focus group interviews showed that there were no large differences in the feedback from the healthy individuals and the cancer survivors, indicating that completion of the DIGIKOST-FFQ was equally feasible for both groups. However, some differences in their feedback should be pointed out. The cancer survivors found it most challenging to report daily activities and intake of traditional foods, whereas the challenges identified by the healthy individuals were related more to social status and the lack of questions about novel food products available on the market today. We assume that the differences in the feedback from healthy individuals and cancer survivors might be due to age; however, we do not have information on the age of the participants.

The results from the usability testing showed that the technical flow of the questionnaire was good. Most participants found the DIGIKOST reports easy to understand, and all enjoyed the individual advice and recommendations presented in the text at the end of the report. However, some preferred a more visual presentation of the results, such as percentage adherence, rather than the textual information in the table. We speculate that it would be difficult to please all individual preferences on how to visualize the results, and that the solution might therefore be to include both tables and graphics in the report to suit both preferences. The interpretation of the health index varied to a

great extent, and most participants found it difficult to understand. Difficulties in perceiving health risk factors presented as percentages or other statistical terms have been documented in patients with low numeracy, whereas interactive graphics may be more easily perceived [41-43]. Thus, when communicating health risk factors, it is important to be aware that different formats generate different risk perceptions among patients with different levels of numeracy [42].

Strengths and Limitations

DIGIKOST-FFQ is accessible from multiple electronic devices, such as personal computers, phones, and tablets, allowing for high flexibility in future use of the tool, minimal respondent burden, and potentially reduced selection bias. Further advantages are a low demand for personnel and economic resources and easy implementation in research settings, including both observational and interventional studies; the literature has reported similar advantages for previous digital questionnaires [2]. Moreover, the questions in the digital platform are easy to change and adapt to follow future updates in dietary and lifestyle guidelines. A strength of our study is the inclusion of both healthy individuals and cancer survivors in the focus group interviews. Moreover, we included new participants in the usability testing to make sure it was the first time the participants tested the DIGIKOST-FFQ and DIGIKOST reports so that we could obtain their first-impression feedback of its usability. Another strength is that the participants had a good variety of backgrounds (in gender, work and education, and location in Norway).

A limitation of the current study could be that the focus group interviews were carried out by video call on Zoom due to the COVID-19 pandemic, which could have affected the group dynamics and the ability of the participants to freely discuss their ideas. Some participants found it difficult to participate on Zoom due to technical issues and low-quality sound or video, and we speculate that the study population may have been biased toward people with access to high-quality digital equipment and greater technological skills. Nevertheless, the video-call format made it possible to include individuals living across Norway.

Future Improvements in DIGIKOST-FFQ and the DIGIKOST Report

The focus group interviews contributed valuable knowledge about users' challenges and led to suggestions for improvements. For instance, in the future, we will include more text about each food item and present food quantities by weight or volume, in addition to pictures representing different portion sizes. Moreover, as most of the participants found it challenging to report intake and activities over the previous year due to seasonal variations, the time frame of reporting will be revised and reduced to the previous 2 months.

Several aspects of the DIGIKOST reports were challenging to understand for the participants. Therefore, we will remove the graphics showing achievements as percentages, the health index, and the Healthy Eating Plate recommendations. Moreover, we will add more lifestyle factors to the "individual advice for you" section, such as alcohol intake, smoking, physical activity, and body weight.

Conclusions

The DIGIKOST-FFQ and the DIGIKOST report were well received by the participants, who found it easy to log in to and navigate the system and understand the questions. The completion time was acceptable. Changes in the questionnaire and report to address difficulties in recalling dietary intake over the previous year and due to seasonal variation will be implemented. Also, text with additional information on weight or volume will be added to the portion-size pictures. All participants found it motivational to receive personalized feedback reports with dietary advice. The usability testing showed that the log-in system worked well, but that some adjustments were needed to the reports in order to make the personalized feedback more understandable.

Acknowledgments

The Throne Holst Foundation for Nutrition Research and Aktieselskabet Freia Chocolade Fabriks Medisinske Fond funded this study. The authors would like to acknowledge the University Center for Information Technology at the University of Oslo for its contribution to the design and development of the DIGIKOST-FFQ, and particularly information technology designer Mette Sundal at the University Center for Information Technology for administering and performing the usability testing. We also acknowledge Nina Norberg Gjerrestad and Anne Lene Nordengen for the management of all participants, performance of the focus group interviews, and data handling.

Authors' Contributions

HBH had the main responsibility for writing the manuscript. HBH, MDK, MHC, AH, and RB contributed to the conception and the design of the study and drafting of the manuscript. HBH, MDK, and MHC contributed to the acquisition, analysis, and interpretation of the data. All authors contributed to writing and approval of the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
DIGIKOST report, first draft.

[PDF File (Adobe PDF File), 8466 KB - formative_v6i11e35933_app1.pdf]

Multimedia Appendix 2
Interview guide for the focus groups on DIGIKOST-FFQ.
[DOCX File , 18 KB - formative_v6i11e35933_app2.docx]

References

1. Kostråd for å fremme folkehelsen og forebygge kroniske sykdommer: metodologi og vitenskapelig kunnskapsgrunnlag. Oslo, Norway: Nasjonalt råd for ernæring, Helsedirektoratet; 2011.
2. 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]
3. Forster H, Fallaize R, Gallagher C, O'Donovan CB, Woolhead C, Walsh MC, et al. Online dietary intake estimation: the Food4Me food frequency questionnaire. J Med Internet Res 2014 Jun 09;16(6):e150 [FREE Full text] [doi: 10.2196/jmir.3105] [Medline: 24911957]
4. Rusin M, Arsand E, Hartvigsen G. Functionalities and input methods for recording food intake: a systematic review. Int J Med Inform 2013 Aug;82(8):653-664. [doi: 10.1016/j.ijmedinf.2013.01.007] [Medline: 23415822]
5. Kristal AR, Kolar AS, Fisher JL, Plascak JJ, Stumbo PJ, Weiss R, et al. Evaluation of web-based, self-administered, graphical food frequency questionnaire. J Acad Nutr Diet 2014 Apr;114(4):613-621 [FREE Full text] [doi: 10.1016/j.jand.2013.11.017] [Medline: 24462267]
6. McDermott MS, Oliver M, Iverson D, Sharma R. Effective techniques for changing physical activity and healthy eating intentions and behaviour: A systematic review and meta-analysis. Br J Health Psychol 2016 Nov;21(4):827-841. [doi: 10.1111/bjhp.12199] [Medline: 27193530]
7. O'Connor S, Hanlon P, O'Donnell CA, Garcia S, Glanville J, Mair FS. Understanding factors affecting patient and public engagement and recruitment to digital health interventions: a systematic review of qualitative studies. BMC Med Inform Decis Mak 2016 Sep 15;16(1):120 [FREE Full text] [doi: 10.1186/s12911-016-0359-3] [Medline: 27630020]
8. Forster H, Walsh MC, O'Donovan CB, Woolhead C, McGirr C, Daly EJ, et al. A dietary feedback system for the delivery of consistent personalized dietary advice in the web-based multicenter Food4Me study. J Med Internet Res 2016 Jun 30;18(6):e150 [FREE Full text] [doi: 10.2196/jmir.5620] [Medline: 27363307]
9. Paulsen MM, Varsi C, Paur I, Tangvik RJ, Andersen LF. Barriers and facilitators for implementing a decision support system to prevent and treat disease-related malnutrition in a hospital setting: qualitative study. JMIR Form Res 2019 May 09;3(2):e11890 [FREE Full text] [doi: 10.2196/11890] [Medline: 31094333]
10. Triantafyllidis AK, Tsanas A. Applications of machine learning in real-life digital health interventions: review of the literature. J Med Internet Res 2019 Apr 05;21(4):e12286 [FREE Full text] [doi: 10.2196/12286] [Medline: 30950797]
11. Lewkowicz D, Slosarek T, Wernicke S, Winne A, Wohlbrandt AM, Bottinger E. Digital therapeutic care and decision support interventions for people with low back pain: systematic review. JMIR Rehabil Assist Technol 2021 Nov 19;8(4):e26612 [FREE Full text] [doi: 10.2196/26612] [Medline: 34807837]
12. Asbjørnsen RA, Hjelmesæth J, Smedsrød ML, Wentzel J, Ollivier M, Clark MM, et al. Combining persuasive system design principles and behavior change techniques in digital interventions supporting long-term weight loss maintenance: design and development of eCHANGE. JMIR Hum Factors 2022 May 27;9(2):e37372 [FREE Full text] [doi: 10.2196/37372] [Medline: 35622394]
13. Carlsen MH, Andersen LF, Hjartåker A. Reproducibility and feasibility of an online self-administered food frequency questionnaire for use among adult Norwegians. Food Nutr Res 2021;65:10.29219/fnr.v65.7561 [FREE Full text] [doi: 10.29219/fnr.v65.7561] [Medline: 34908922]
14. Paulsen MM, Paur I, Gjestland J, Henriksen C, Varsi C, Tangvik RJ, et al. Effects of using the MyFood decision support system on hospitalized patients' nutritional status and treatment: A randomized controlled trial. Clin Nutr 2020 Dec;39(12):3607-3617 [FREE Full text] [doi: 10.1016/j.clnu.2020.03.012] [Medline: 32241711]
15. Paulsen M, Varsi C, Andersen L. Process evaluation of the implementation of a decision support system to prevent and treat disease-related malnutrition in a hospital setting. BMC Health Serv Res 2021 Mar 25;21(1):281 [FREE Full text] [doi: 10.1186/s12913-021-06236-3] [Medline: 33766017]
16. Barbour RS, Morgan DL. A New Era in Focus Group Research: Challenges, Innovation and Practice. London, UK: Palgrave Macmillan; 2017.
17. Cleland JA. The qualitative orientation in medical education research. Korean J Med Educ 2017 Jun;29(2):61-71 [FREE Full text] [doi: 10.3946/kjme.2017.53] [Medline: 28597869]
18. Trakman GL, Forsyth A, Hoye R, Belski R. Developing and validating a nutrition knowledge questionnaire: key methods and considerations. Public Health Nutr 2017 Oct;20(15):2670-2679. [doi: 10.1017/S1368980017001471] [Medline: 28735598]
19. Kitzinger J. Qualitative research. Introducing focus groups. BMJ 1995 Jul 29;311(7000):299-302 [FREE Full text] [doi: 10.1136/bmj.311.7000.299] [Medline: 7633241]

20. Tausch AP, Menold N. Methodological aspects of focus groups in health research: results of qualitative interviews with focus group moderators. Glob Qual Nurs Res 2016;3:2333393616630466 [FREE Full text] [doi: 10.1177/2333393616630466] [Medline: 28462326]
21. Zapata BC, Fernández-Alemán JL, Idri A, Toval A. Empirical studies on usability of mHealth apps: a systematic literature review. J Med Syst 2015 Feb;39(2):1. [doi: 10.1007/s10916-014-0182-2] [Medline: 25600193]
22. 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]
23. Henriksen HB, Carlsen MH, Paur I, Berntsen S, Bøhn SK, Skjetne AJ, et al. Relative validity of a short food frequency questionnaire assessing adherence to the Norwegian dietary guidelines among colorectal cancer patients. Food Nutr Res 2018;62:1 [FREE Full text] [doi: 10.29219/fnr.v62.1306] [Medline: 29545734]
24. Henriksen HB, Berntsen S, Paur I, Zucknick M, Skjetne AJ, Bøhn SK, et al. Validation of two short questionnaires assessing physical activity in colorectal cancer patients. BMC Sports Sci Med Rehabil 2018;10:8 [FREE Full text] [doi: 10.1186/s13102-018-0096-2] [Medline: 29854408]
25. Slik kan du sette sammen et sunt måltid. Norwegian Directorate of Health. URL: https://www.helsenorge.no/kosthold-og-ernaring/sma-grep-for-et-sunt-kosthold/dagens-maltider/#tallerkenmodellen [accessed 2022-10-20]
26. Norwegian Cancer Society. URL: https://kreftforeningen.no/en/ [accessed 2022-10-20]
27. Henriksen HB, Ræder H, Bøhn SK, Paur I, Kværner AS, Billington, et al. The Norwegian dietary guidelines and colorectal cancer survival (CRC-NORDIET) study: a food-based multicentre randomized controlled trial. BMC Cancer 2017 Jan 30;17(1):83 [FREE Full text] [doi: 10.1186/s12885-017-3072-4] [Medline: 28137255]
28. University Center for Information Technology. URL: https://www.usit.uio.no/english/ [accessed 2022-10-20]
29. Krueger R. Designing and Conducting Focus Group Interviews. St. Paul, MN: University of Minnesota; 1954.
30. f4transkript. Universitetet i Oslo. URL: https://www.uio.no/tjenester/it/forskning/datafangst-og-analyse/f4-transcribe/ [accessed 2022-10-20]
31. First Rule of Usability? Don't Listen to Users. Nielsen Norman Group. 2001. URL: https://www.nngroup.com/articles/first-rule-of-usability-dont-listen-to-users/ [accessed 2022-10-20]
32. Usability testing 2017. University Center for Information Technology. URL: https://www.usit.uio.no/om/organisasjon/bnt/web/ux/blogg/2017/brukertest.html [accessed 2022-10-20]
33. Cade J, Thompson R, Burley V, Warm D. Development, validation and utilisation of food-frequency questionnaires - a review. Public Health Nutr 2002 Aug;5(4):567-587. [doi: 10.1079/PHN2001318] [Medline: 12186666]
34. Margetts B, Nelson M. Design Concepts in Nutritional Epidemiology. New York, NY: Oxford University Press; 1997.
35. Willett W. Nutritional Epidemiology. Oxford, UK: Oxford University Press; 2013.
36. de Leeuw E, Hox J, Silber H, Struminskaya B, Vis C. Development of an international survey attitude scale: measurement equivalence, reliability, and predictive validity. Meas Instrum Soc Sci 2019 Dec 02;1(1):1-9 [FREE Full text] [doi: 10.1186/s42409-019-0012-x]
37. Faulkner GP, Livingstone MBE, Pourshahidi LK, Spence M, Dean M, O'Brien S, et al. An evaluation of portion size estimation aids: Consumer perspectives on their effectiveness. Appetite 2017 Jul 01;114:200-208. [doi: 10.1016/j.appet.2017.03.027] [Medline: 28336467]
38. Bouchoucha M, Akrout M, Bellali H, Bouchoucha R, Tarhouni F, Mansour AB, et al. Development and validation of a food photography manual, as a tool for estimation of food portion size in epidemiological dietary surveys in Tunisia. Libyan J Med 2016;11:32676 [FREE Full text] [doi: 10.3402/ljm.v11.32676] [Medline: 27585631]
39. Boushey CJ, Kerr DA, Wright J, Lutes KD, Ebert DS, Delp EJ. Use of technology in children's dietary assessment. Eur J Clin Nutr 2009 Feb;63 Suppl 1:S50-S57 [FREE Full text] [doi: 10.1038/ejcn.2008.65] [Medline: 19190645]
40. Vaportzis E, Clausen MG, Gow AJ. Older adults perceptions of technology and barriers to interacting with tablet computers: a focus group study. Front Psychol 2017 Oct 04;8:1687 [FREE Full text] [doi: 10.3389/fpsyg.2017.01687] [Medline: 29071004]
41. Ancker JS, Chan C, Kukafka R. Interactive graphics for expressing health risks: development and qualitative evaluation. J Health Commun 2009;14(5):461-475 [FREE Full text] [doi: 10.1080/10810730903032960] [Medline: 19657926]
42. Peters E, Hart PS, Fraenkel L. Informing patients: the influence of numeracy, framing, and format of side effect information on risk perceptions. Med Decis Making 2011;31(3):432-436. [doi: 10.1177/0272989X10391672] [Medline: 21191122]
43. Ancker JS, Weber EU, Kukafka R. Effects of game-like interactive graphics on risk perceptions and decisions. Med Decis Making 2011;31(1):130-142 [FREE Full text] [doi: 10.1177/0272989X10364847] [Medline: 20393103]

Abbreviations

CRC-NORDIET study: The Norwegian dietary guidelines and colorectal cancer survival study: a food-based multicenter randomized controlled study
FBDG: Food-Based Dietary Guidelines
FFQ: food frequency questionnaire
ID: Identification data

TSD: Tjenester for sensitive data (Services for Sensitive Data)
UiO: University of Oslo
USIT: University Center for Information Technology

Edited by A Mavragani; submitted 22.12.21; peer-reviewed by V Kaufman-Shriqui; comments to author 07.09.22; revised version received 27.09.22; accepted 27.09.22; published 08.11.22.

Please cite as:
Henriksen HB, Knudsen MD, Carlsen MH, Hjartåker A, Blomhoff R
A Short Digital Food Frequency Questionnaire (DIGIKOST-FFQ) Assessing Dietary Intake and Other Lifestyle Factors Among Norwegians: Qualitative Evaluation With Focus Group Interviews and Usability Testing
JMIR Form Res 2022;6(11):e35933
URL: https://formative.jmir.org/2022/11/e35933
doi:10.2196/35933
PMID:36346647

Viewpoint

How to Use the Six-Step Digital Ethnography Framework to Develop Buyer Personas: The Case of Fan Fit

Alex Fenton[1], BSc, MA, PhD; Aleksej Heinze[2], BSc, PhD; McVal Osborne[3], BA; Wasim Ahmed[4], BA, MSc, PhD

[1]Business School, University of Chester, Chester, United Kingdom

[2]Marketing Department, KEDGE Business School, Marseille, France

[3]Flight Story, London, United Kingdom

[4]Management School, University of Stirling, Stirling, United Kingdom

Corresponding Author:
Wasim Ahmed, BA, MSc, PhD
Management School
University of Stirling
Stirling, FK9 4LA
United Kingdom
Phone: 44 1786 467
Email: Wasim.Ahmed@Stirling.ac.uk

Abstract

Background: One of the key features of digital marketing is customer centricity, which can be applied to the domain of health. This is expressed through the ability to target specific customer segments with relevant content using appropriate channels and having data to track and understand each interaction. In order to do this, marketers create buyer personas based on a wide spectrum of quantitative and qualitative data. Digital ethnography is another established method for studying web-based communities. However, for practitioners, the complexity, rigor, and time associated with ethnographical work are sometimes out of reach.

Objective: This paper responds to the gaps in the practically focused method of using social media for digital ethnography to develop buyer personas. This paper aims to demonstrate how digital ethnography can be used as a way to create and refine buyer personas.

Methods: Using a case study of the Fan Fit smartphone app, which aimed to increase physical activity, a digital ethnography was applied to create a better understanding of customers and to create and refine buyer personas.

Results: We propose two buyer personas, and we develop a 6-step digital ethnography framework designed for the development of buyer personas.

Conclusions: The key contribution of this work is the proposal of a 6-step digital ethnography framework designed for the development of buyer personas. We highlight that the 6-step digital ethnography could be a robust tool for practitioners and academicians to analyze digital communications for the process of creating and updating data-driven buyer personas to create deeper insights into digital and health marketing efforts.

(JMIR Form Res 2022;6(11):e41489) doi:10.2196/41489

KEYWORDS

health tracking; digital; ethnography; apps; mobile app; customer; physical activity

Introduction

Digital communication channels offer great benefits to marketers for understanding customer behavior to bring customer centricity to the core of digital marketing strategies. The buyer persona concept was designed specifically for digital marketing strategies and is linked with the concept of the buyer persona spring, as outlined by Heinze et al [1]. Buyer personas can also be applied within the health domain and are useful for marketing health-related initiatives. The focus of this case study is in the health domain, which is increasingly becoming an important research area [2], more specifically in the segment of digital fitness and well-being devices, which is projected to reach over US $58 billion in 2022 and US $99 billion by 2027 [3].

Heinze et al [1] present the buyer persona concept as a fictitious representation of the customer segment, which has a number of elements allowing marketers to understand customer needs and how they would like to be interacted with. The core of the buyer

persona insights is the pain points—the main issues that the buyer persona faces and attempts to alleviate through the use of the products or services of a particular organization. For each pain point, a corresponding trust point is also established, which helps to understand the perceived solutions to these problems.

The reasons why these solutions are *perceived* and not necessarily the definitive correct solutions is because buyer personas are based on aggregate knowledge from multiple sources, and buyers might be misinformed or simply unaware of the "correct" solution that is currently available on the market. Realizing these points also helps marketers to identify keyword phrases associated with these areas of problems and solutions and, as a consequence, estimate the size of the audience and the seasonality of these terms.

While the buyer persona concept is driven by practitioners, it is not discussed in the academic literature—only textbooks [4] that focus on how to use buyer personas. Other studies [5] have limitations in that they only use a simplistic approach and state that certain questions in relation to a buyer persona have to be answered without going in-depth on the process of answering these questions. A more comprehensive look at the buyer persona development was given by Revella [6], who acknowledges the need for multiple data sources, including interviews, customer surveys, and big data, in particular the use of loyalty cards and web analytics. While it is a comprehensive book that outlines buyer persona understanding and development, it misses two key data sources available to digital marketers—past keyword phrases searched and a methodical approach to studying social media.

Formulating and refining personas based on an ever-increasing volume of social media channels and content is an ongoing challenge for digital marketers because the digital landscape changes constantly. This calls for a flexible method that can adapt to understand customers now and in the coming years. This knowledge allows marketers to refine products and services and keep customer centricity at the center of digital and social media marketing activity.

Two key sources of data for buyer persona development are keyword research based on search engine use and social media analysis based on publicly available posts and engagement data. The process of refining and developing insights from search engine use data benefits from search engine keyword research tools such as Google Ads and the Baidu Index. Despite an established set of tools, there is still a need for interpretation of those search terms. The selection of buyer persona strategic terms does still require human analysis.

However, marketing automation tools that manage paid advertising campaigns do exist, and there are various options for larger-scale campaigns. The keyword research data is relatively well structured, showing when, where, and how often a particular search term was used when compared to social media. The use of social media targeting is more unstructured, on the other hand. No one particular method is apparent, but methods such as netnography [7] and digital ethnography [8] are well-established ways of understanding more about people and social media communities.

While these methods are well established for academic and practitioner research, they can be considered complex, time-consuming, and challenging to operationalize. Furthermore, the General Data Protection Regulation and ethical issues such as obtaining informed consent from each one of the observed individuals are complex and shifting landscapes. Netnography [2,7] provides the clearest and most established set of guidelines for conducting social media ethnographical research, and following these guidelines would enable buyer personas to be created and refined [1]. In this paper, however, we explore an alternative approach to digital ethnography. Ethnography as an approach is methodologically flexible; therefore, in this paper, we explore digital ethnography as a flexible approach to the practical creation of buyer personas.

A lot of brands might talk about "Net Zero," but what does it mean in reality for a brand? There is a need to go beyond the numbers, starting with quantitative data but then flexibly returning to qualitative data and back again; this responsive adaptiveness is required to formulate and maintain buyer personas.

Therefore, this formative paper aims to propose and illustrate a 6-step digital ethnography (SSDE) as one of the methods for data analysis to create buyer personas. The paper is structured as follows. First, the concept of a buyer persona is reviewed and discussed. Second, the use of digital ethnography as a core method and its challenges are discussed based on the literature. Third, the application of digital ethnography is illustrated using the health case study of Fan Fit as an example.

Buyer Personas

The traditional marketing literature has been concerned with the segmentation of the markets for decades. The idea that the same message cannot be shared with all prospective customers and customer segmentations is decades old [9]; in his article, Marcus [9] shares a method for dividing existing customers into groups based on past purchase frequency and discusses other statistical methods to divide customers into groups. While this method is useful for understanding customer behavior, it gives us a single dimension to understanding customers based on their past purchases. These kinds of segmentations suffer from the survivorship bias [10], where we only get to know the needs of those who are already our customers and not those who chose not to use our products and services. Further studies on customer segmentation tend to also focus on statistical analysis methods that allow for a more detailed understanding of usually existing customers [11]. Furthermore, these approaches are not suitable for small- and medium-sized organizations and start-ups that do not usually possess statistical know-how nor do they have large enough data sets of customers for querying. This is why the concept of a practical but methodical approach is necessary in digital marketing, and SSDE is therefore relevant and practical.

The buyer persona and buyer persona spring concept are outlined in the book *Digital and Social Media Marketing: A Results-Driven Approach* [1]. The buyer persona has much in common with other customer personas because it is a fictitious representation of a customer segment. Customers and potential

customers come in various ages, races, classes, genders, and other demographics. Digital technologies and data capture are improving all the time in understanding and delivering personalized and timely information to people. Even so, marketers cannot understand everything about their audiences individually, so personas are of great value to marketers. While offering us a number of examples and suggestions on how to develop a buyer persona template, there is no methodical approach that helps in understanding the transparency of how these questions should be answered.

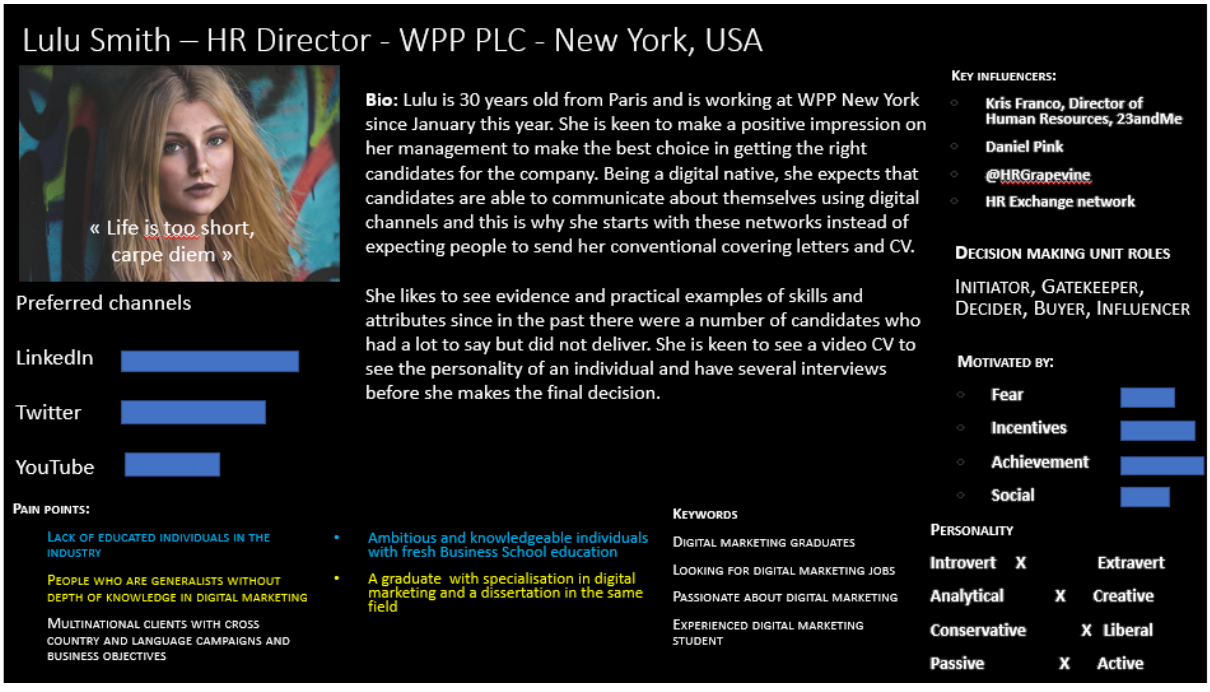
Buyer personas are similar to user personas, but the main difference is buyer personas are representative of the ones who will be "buying" into your ideas, products, or services as opposed to simply using something. For example, for children's books, the user persona will consider how the child will use a book, whereas the buyer persona for the same child would focus on how to encourage the same child to use a book. The purchase process does not always need to be financial—it could be simply the "buying of a political idea" or understanding a concept. Buyer personas also state the role that this individual has in the purchase process—for example, a child could have a role of an influencer, while the parent has the role of a gatekeeper, decision maker, or purchaser. The buyer persona, therefore, is an invaluable way of better understanding audiences in order to create targeted products and service communications, and for health marketing purposes.

Two potential approaches for the creation of buyer personas are science fiction prototyping and data-driven development. Science fiction prototyping is usually used by organizations that want to become leaders in the market through disruptive innovation. These organizations often develop products or services for which there is no current customer segment, which makes creating a real-life buyer persona impossible. For example, when developing software products for the management of yet-to-be-invented robotics systems, AutoDesk used science fiction prototyping [12] in its buyer persona development. This is because there was not yet anything like this product on the market and, therefore, a lack of data as to what drives customer purchase decisions before the product even exists.

Alternatively, the data-driven buyer persona is based in areas where an existing product or service does exist on the market. The content of a buyer persona generally includes demographics, average order value, locations, national and regional culture, socioeconomic background, and decision-making patterns. Central to buyer persona information are the "pain touch points" and corresponding "trust touch points." It should also include the keywords they might use on search engines, the social networks that they might engage with, and other information that helps marketers to focus content based on buyers' needs. The ability to target and select specific and unique audience segments is being made possible through the proliferation of digital communications and social media channels. The purpose of developing a detailed buyer persona is, therefore, to better know this target audience through this technique of digital personification. Figure 1 provides an example of a buyer persona.

There are several ways to develop a buyer persona using a data-driven strategy. The data-driven process uses real data to understand the pain points and corresponding hot points or trust points of the buyer persona. Data from surveys and publicly available data can be used to develop the profile. Social media is also a critical source in order to understand the customer. For the social media marketer, interacting with customers is generally a very regular occurrence, and this participation is critical to developing buyer personas. A key aspect of ethnography is participant observation and interaction with people. Digital ethnography is therefore particularly suited to engaging with an audience and for observing, understanding, formulating, and refining buyer personas.

Figure 1. An example buyer persona.



Digital Ethnography

Ethnography as a research method has a long history in studying people and storytelling [8]. Some consider it the most in-depth of all possible research methods because it "enables a researcher to see what people are doing as well as what they say they are doing" [8]. Digital ethnography is a newer branch of ethnography that takes into account digital communities as a key focus [4]. Various forms of ethnography are dedicated to studying technology and digitally enabled communities. Digital ethnography was selected for this paper because it has long been established [6] as well as its focus on "yielding new insights for our understanding of data" [4] and its practical flexibility. The *Routledge Handbook of Digital Ethnography* [4] is a useful guide to applying this method and states that "Drawing from the empirical, ethnography can provide insight into motivations and practices that in turn shape future directions for digital media."

The application of digital ethnography to create buyer personas shapes the direction of digital marketing practice and campaigns. Therefore, it is an important concept for developing and refining buyer personas. This is critical to the formulation and execution of digital marketing strategies.

Table 1 outlines the 6 steps of digital ethnography used over the past 5 years, from 2017 to 2022, to teach digital marketing students in higher education institutions in the United Kingdom and France.

This teaching was done at specialized master's programs with digital marketing courses as well as for MBA students. About 300 students per year have taken these courses at internationally recognized business schools. Students were tasked to create a buyer persona for a particular case study, whereby the group was recommended to focus on different buyer personas for the same brand. This allowed participants to undertake the same process for a range of different markets—both business-to-business as well as business-to-consumer. One of these live case studies was the Fan Fit case, where students had the option to expand the application of the market currently relevant to the Fan Fit white-label app. The steps and tasks for students were refined over the years to facilitate a clearer process that works for all cases.

The table was cross-referenced with practices used by one of the largest digital marketing agencies that used buyer personas on a regular basis for all of their communications.

Table 1. Six-step digital ethnography framework.

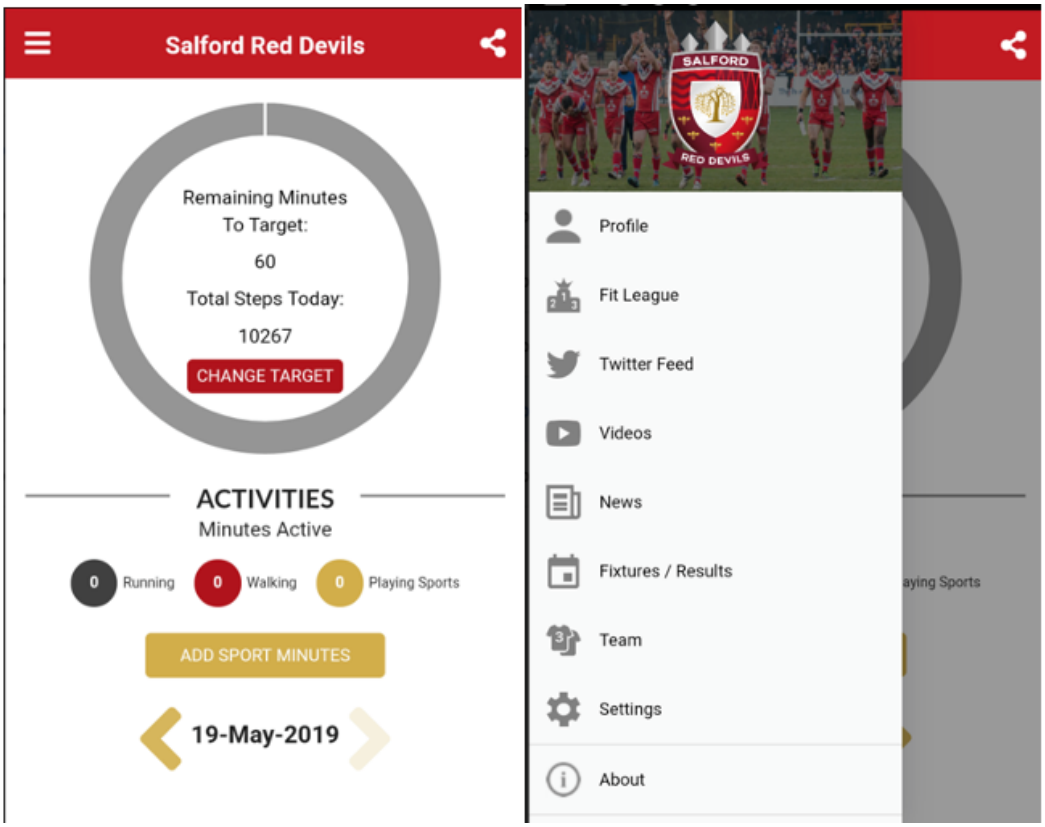
Stage	Tools and techniques	Outputs for buyer persona
(1) Ideation: documenting views of the potential buyer persona	Brainstorming of buyer persona elements—there are no data, just ideas of individuals who undertake this work	All elements of buyer persona based on the perceptions and preconceptions of the initial idea
(2) Social proofing: first impressions of real data existence for relevant topics	Initial observations of potential online communities and reviews of products or services; hashtag analysis, online reviews for related products and services among your own and competitor communications such as social media channels	Confirmation, disproof, and identification of new themes for pain and trust points; pictures and quotes of real issues expressed online by potential buyer persona representatives
(3) Horizon scanning: industry reports, analysis, and wider statistics	Reports done by market analysis firms such as Mintel and MarketLine, and wider statistics such as those offered by governments or panel data such as Statista	Confirmation, disproof, and identification of new themes for pain and trust points, adding more contextual data through wider reports and statistical publications, if available
(4) Keyword research: using search data to see evidence	Using search engine past search behavior trends as an indication and prediction of interest; keyword research tools such as Google Trends and Google Keyword Planner or Baidu Index for the Chinese market	Identification of keywords with potential demand, which could be integrated into the content development for the buyer persona; understanding their terms' use and potential seasonality
(5) Content audit: skyscraper content and influencer identification	Using identified keyword terms confirming most visible content, places, and individuals/organizations who are engaging existing buyer personas; various tools can be used in this stage	List of channels, content examples (screenshots, links, and evidence of its popularity); list of influencers who regularly engage on these topics using identified keywords
(6) Proofing: finding evidence that the updated buyer persona exists	Using focus groups, surveys, and social concept tests to identify if the issues highlighted in the buyer persona research reflect the reality of the target audience	Confirmation and update of the buyer persona understanding; refinement of the buyer persona as necessary

Fan Fit Health Case Study

The main case study used in this section is "Fan Fit" [13]. This case study is used as an example, and an overview of the app is shown in Figure 2.

Fan Fit began in 2015 as a project started at the University of Salford in the United Kingdom, which aimed to use technologies to engage fans and customers around fitness and well-being. The first major implementation of the project was a smartphone app created for Android and iPhone platforms, which eventually became the first and only official club app of the Salford Red Devils Rugby Club. Red Devils are a top-division Super League club that, like many sports clubs, had a website and social media channels but no smartphone app to engage with their fans.

Figure 2. Fan Fit Red Devils smartphone app.



The app offered the usual sports club app functionality of pulling in news, fixtures, and social media. A key differentiating feature between this and other similar apps was the addition of an activity tracker for measuring walking and running automatically through the phone or a wristband. This enabled the app to create fitness leagues between fans (global, monthly, or personal), gamifying user engagement. Using gamification principles, fans could win digital badges and physical prizes such as season tickets, shirts, or fitness wristbands through the app, creating a social media buzz and health-focused dialog between the club and fans. The project used a human-centered design approach, involving fans at each stage of development through surveys, focus groups, and online and offline participant observation.

Ultimately, the project aimed to:

- engage fans through interactive technology to build a new kind of brand community and capture new types of user data, and
- raise awareness around major public health challenges related to heart disease and obesity while encouraging fans to be more active and develop healthier lifestyles.

To develop a minimum viable product that sports clubs could test, the Fan Fit team integrated fitness-tracking technology into a brandable (white labeled) official club app. This allowed them to obtain a "first mover advantage" and bring a new product to market before other competitors could do the same. This minimum viable product opened up doors to other opportunities—discussions with sports clubs and the first major sign-up with the Salford Red Devils. This success created opportunities at the national level to get further funding and the signing of a major football club (Rangers FC).

A key aspect of the Fan Fit project was the creation of a digital marketing strategy for the project to reach new sports clubs and other organizations interested in the project (business-to-business). In addition, for the version of the app that existed, a business-to-consumer digital marketing strategy was used to reach fans of those clubs.

In all cases, the research team used a combination of surveys, focus groups, and ongoing digital ethnography to understand more about the audience and formulate and refine buyer personas. Having outlined the importance of the buyer persona, we now illustrate how this can be applied to our case of Fan Fit. Two buyer persona examples are shown, one for the fan (business-to-consumer) and another for a sports club (business-to-business).

Buyer Persona for Fan Fit: Jim Smith (Fan)

Jim Watson is a 52-year-old man from Salford. He is a season ticket holder with the Salford Red Devils and attends most home games. He is also an avid football fan, attending some Manchester United matches and following them on TV. He also attends some matches of the Salford FC. Jim is a van driver in Salford, his wife is Marie, and he has three grown-up children and two grandchildren.

Channels

Jim uses television, Twitter, Facebook, and YouTube to find the latest news on his favorite teams, keep in touch with family, and interact with other fans on Facebook and Twitter.

Frustrations and Pain Points

Jim has thought about buying a fitness wristband and is trying to be more active and develop healthier habits, but he does not know which one to buy or how to set up or use it. His son has a FitBit but is always very busy with his work and children, so Jim does not want to bother his son by asking for a recommendation.

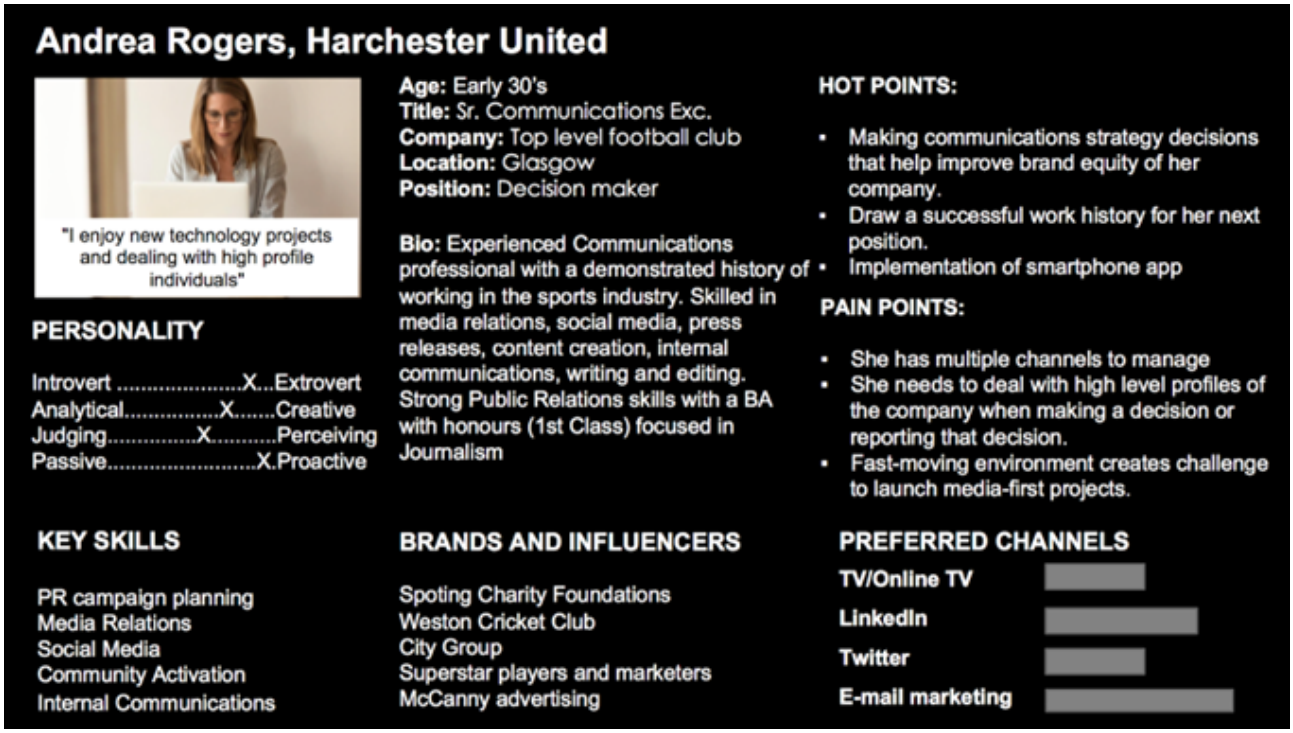
Hot or Trust Points

Jim has become more aware of the importance of fitness after his friend had a minor heart attack last year. He thinks a mobile exercise tracking app can help him prevent ill health and improve his quality of life.

Buyer Persona for Fan Fit: Andrea Rogers (Club)

Andrea is a 33-year-old woman originally from Manchester (Figure 3). She is a senior marketing and communications professional with a major football club. She manages a small team who runs the club website and social media channels, and a wider pool of journalists and content producers. She is recently married but does not have any children.

Figure 3. Buyer persona for Andrea Rogers.



Channels

Andrea uses television, Twitter, Facebook, YouTube, Instagram, and LinkedIn. She uses Facebook and Instagram primarily to keep in touch with her friends and family. She uses Twitter and LinkedIn primarily as her professional network and to seek new opportunities and promote initiatives.

Frustrations and Pain Points

Andrea is acutely aware that her club has no plans for a smartphone app and has started to fall behind other rival clubs with their use of innovative technologies and digital fan engagement.

Hot and Trust Points

Andrea is interested in new technologies such as apps, the Internet of Things, virtual reality, augmented reality, eSports, and anything that would potentially bring the club to a new audience of young fans. She would also like to engage more female fans through digital channels, as they are not proportionately represented on social media.

How These Buyer Personas Were Developed: Data

Digital and social media channels like Google, Twitter, LinkedIn, Instagram, Facebook, and Reddit provide marketers and health researchers with treasure troves of information on their target audiences. The content published across these channels can be exhaustive in terms of scope and detail, and the data captured from this content can be incredibly valuable in producing accurate buyer personas. Data come in all kinds of shapes and forms. In terms of the buyer persona spring and as part of the digital entrepreneurship strategy, we are most interested in the data sources that help to better understand the buyer persona, the channels, and the content. Data sources, including things like websites and social media analytics, can be invaluable sources of information and insight. Data in the raw format must be cleaned and visualized in such a way that they can be used to add meaning to the strategy in an ongoing way.

Data-gathering strategies can also be devised to create new primary data sources. For example, making a survey whose

results can be visualized is a simple way to gather some useful quantitative data to inform the strategy and answer many questions; in order to delve deeper and answer questions relating to your digital strategy, it may be necessary to also gather some qualitative data derived from your audience, experts, or the general public. This kind of qualitative data can be derived from interviews, focus groups, and even through social media analysis of comments and posts, including netnography and digital ethnography.

The ultimate goal of a data monitoring strategy is to keep track of achieving strategic business objectives. Strategic in this sense means that they are long-term—perhaps 1 year, 3 years, or 5 years long. The longer the time span, the more it is necessary to have key performance indicators along the way to see if these objectives are going to be reached. Using the SMART objective setting helps to make them more transparent.

Conclusions

Fan Fit, our health-based case study, has at least two distinct buyer personas. The first one is fans (eg, Jim Smith), and the second is marketing and communications staff and decision makers within organizations that may want to adopt a version of Fan Fit for themselves (Andrea Rogers).

In the buyer personas outlined above, we demonstrated which channels Jim and Andrea are using. In the case of Fan Fit, for example, Twitter may be an appropriate channel to reach Jim. In terms of content, combining official messages from the brand he follows (Red Devils) and the conversational approach among fans, Fan Fit founders, and the club would help build the social capital and community to reach Jim.

A competition to win a season ticket and the competition with his friends and fellow fans are enough to get him to download the Fan Fit Red Devils app and get more active and engaged. In Andrea's case, LinkedIn may be an appropriate channel to identify who she is and send her a personal connection request and a personalized message about Fan Fit. It is also possible to produce press releases and share these more widely using LinkedIn. Our findings are likely to be of interest to health marketing researchers and practitioners looking to promote positive health behaviors.

This paper has outlined a key methodological area that allows students, practitioners, and researchers to follow an approach to finding the answers to the questions around buyer persona development. The challenges highlighted in previous studies (eg, [9]) are reduced in the current process in that we are exploring data sets that are mainly openly available. The previous studies had survival bias, where classic customer segmentation approaches use the statistics of only existing customers. By reaching out to multiple data sources and cross-referencing these, we are supporting the ideas advocated by Revella [6] and developing a combination of data sources in a more digital ethnography-focused way.

The limitation of this study is that the method originated from discussions with teachers looking for a systemic, practical, and effective way to help students develop digital marketing strategies. The traditional methods of segmentations and buyer persona development were not accessible to learners as well as smaller organizations that do not have large internal databases to interrogate and develop their buyer personas. The practitioner perspectives are also only taken into account by one organization that has offices in Germany, the United Kingdom, and the United States. Other geographic areas, including the growing African digital ecosystems and Chinese perspectives, can add further richness to the application and further generalization of the framework.

Researchers are invited to explore and test the use of the SSDE in other settings. Practitioners can use the 6 steps of ideation, social proofing, horizon scanning, keyword research, content audit, and proofing for structuring their activities and offer transparency to the way in which their buyer personas are presented.

Acknowledgments

We would like to acknowledge all organizations and funding bodies that have supported the development of our understanding of digital entrepreneurship, in particular, the European Union–funded projects such as Passport to Trade 2.0 and Joint European Masters in Search and Social Media; national initiatives such as CreativeHive and Creative Entrepreneur; and finally, the Fan Fit project.

Conflicts of Interest

None declared.

References

1. Heinze A, Fletcher G, Rashid T, Cruz A, editors. Digital and Social Media Marketing: A Results-Driven Approach. England, UK: Routledge, Taylor & Francis Group; 2020.
2. Quévat A, Heinze A. The digital transformation of preventive telemedicine in France based on the use of connected wearable devices. Glob Bus Org Exc 2020 Jun 29;39(6):17-27. [doi: 10.1002/joe.22054]
3. Statista. 2022. URL: https://www.statista.com/outlook/dmo/digital-health/digital-fitness-well-being/digital-fitness-well-being-devices/worldwide [accessed 2022-10-26]
4. Cruz A, Karatzas S. Understanding your buyer persona. In: Heinze A, Fletcher G, Rashid T, Cruz A, editors. Digital and Social Media Marketing: A Results-Driven Approach. London: Routledge, Taylor & Francis Group; 2020.

5. Akre V, Rajan A, Ahamed J, Al Amri A, Al Daisi S. Smart Digital Marketing of Financial Services to Millennial Generation using emerging technological tools and buyer persona. 2019 Presented at: 2019 Sixth HCT Information Technology Trends; November 20-21, 2019; Ras Al Khaimah, United Arab Emirates p. 120-125. [doi: 10.1109/itt48889.2019.9075106]
6. Revella A. Buyer Personas: How to Gain Insight Into Your Customer's Expectations, Align Your Marketing Strategies, and Win More Business. Hoboken, NJ: John Wiley & Sons; 2015.
7. Kozinets R. Netnography: The Essential Guide to Qualitative Social Media Research. Newbury Park, CA: SAGE; 2019.
8. Hjorth L, Horst H, Galloway A, Bell G. The Routledge Companion to Digital Ethnography. London, UK: Routledge; 2017.
9. Marcus C. A practical yet meaningful approach to customer segmentation. J Consum Market 1998;15(5):504. [doi: 10.1108/07363769810235974]
10. Garcia C, Gould F. Survivorship bias. J Portfolio Manag 1993 Apr 30;19(3):52-56. [doi: 10.3905/jpm.1993.409445]
11. Cooil B, Aksoy L, Keiningham TL. Approaches to customer segmentation. J Relatsh Mark 2008 Jan 14;6(3-4):9-39. [doi: 10.1300/j366v06n03_02]
12. Atherton E, Johnson BD. Science fiction prototyping at work. Computer 2016 Aug;49(8):109-111. [doi: 10.1109/mc.2016.229]
13. Fenton A, Cooper-Ryan A, Hardey M, Ahmed W. Football fandom as a platform for digital health promotion and behaviour change: a mobile app case study. Int J Environ Res Public Health Interne 2022;19(14):A. [Medline: 35886270]

Abbreviations

SSDE: six-step digital ethnography

Edited by G Eysenbach; submitted 28.07.22; peer-reviewed by M McEachern; comments to author 18.10.22; revised version received 27.10.22; accepted 31.10.22; published 25.11.22.

Please cite as:
Fenton A, Heinze A, Osborne M, Ahmed W
How to Use the Six-Step Digital Ethnography Framework to Develop Buyer Personas: The Case of Fan Fit
JMIR Form Res 2022;6(11):e41489
URL: https://formative.jmir.org/2022/11/e41489
doi:10.2196/41489
PMID:36427232

Original Paper

Brief App-Based Cognitive Behavioral Therapy for Anxiety Symptoms in Psychiatric Inpatients: Feasibility Randomized Controlled Trial

Gaurav Sharma[1,2], MD; Lisa Schlosser[2]; Brett D M Jones[1,2], MD; Daniel M Blumberger[1,2], MSc, MD; David Gratzer[1,2], MD; M Omair Husain[1,2], MBBS; Benoit H Mulsant[1,2], MS, MD; Lance Rappaport[3], PhD; Vicky Stergiopoulos[1,2], MHSc, MD; Muhammad Ishrat Husain[1,2], MBBS, MD (Res)

[1]Department of Psychiatry, University of Toronto, Toronto, ON, Canada

[2]Centre for Addiction and Mental Health, Toronto, ON, Canada

[3]Department of Psychology, University of Windsor, Windsor, ON, Canada

Corresponding Author:
Gaurav Sharma, MD
Department of Psychiatry
University of Toronto
250 College Street, 8th Floor
Toronto, ON, M5T 1R8
Canada
Phone: 1 416 979 6948
Email: mr.sharma@mail.utoronto.ca

Abstract

Background: Psychiatric inpatients often have limited access to psychotherapeutic education or skills for managing anxiety, a common transdiagnostic concern in severe and acute mental illness. COVID-19–related restrictions further limited access to therapy groups on inpatient psychiatric units. App-based interventions may improve access, but evidence supporting the feasibility of their use, acceptability, and effectiveness in psychiatric inpatient settings is limited. MindShift CBT is a free app based on cognitive behavioral therapy principles with evidence for alleviating anxiety symptoms in the outpatient setting.

Objective: We aimed to recruit 24 participants from an acute general psychiatric inpatient ward to a 1-month randomized control study assessing the feasibility and acceptability of providing patients with severe and acute mental illness access to the MindShift CBT app for help with managing anxiety symptoms.

Methods: Recruitment, data collection, analysis, and interpretation were completed collaboratively by clinician and peer researchers. Inpatients were randomized to two conditions: treatment as usual (TAU) versus TAU plus use of the MindShift CBT app over 6 days. We collected demographic and quantitative data on acceptability and usability of the intervention. Symptoms of depression, anxiety, and psychological distress were measured in pre- and poststudy surveys for preliminary signals of efficacy. We conducted individual semistructured interviews with participants in the MindShift CBT app group at the end of their trial period, which were interpreted using a standardized protocol for thematic analysis.

Results: Over 4 weeks, 33 inpatients were referred to the study, 24 consented to participate, 20 were randomized, and 11 completed the study. Of the 9 randomized participants who did not complete the study, 7 were withdrawn because they were discharged or transferred prior to study completion, with a similar distribution among both conditions. Among the enrolled patients, 65% (13/20) were admitted for a psychotic disorder and no patient was admitted primarily for an anxiety disorder. The average length of stay was 20 days (SD 4.4; range 3-21) and 35% (7/20) of patients were involuntarily admitted to hospital. Small sample sizes limited accurate interpretation of the efficacy data. Themes emerging from qualitative interviews included acceptability and usability of the app, and patient agency associated with voluntary participation in research while admitted to hospital.

Conclusions: Our study benefitted from collaboration between peer and clinician researchers. Due to rapid patient turnover in the acute inpatient setting, additional flexibility in recruitment and enrollment is needed to determine the efficacy of using app-based psychotherapy on an acute psychiatric ward. Despite the limited sample size, our study suggests that similar interventions may be feasible and acceptable for acutely unwell inpatients. Further study is needed to compare the efficacy of psychotherapeutic apps with existing standards of care in this setting.

Trial Registration: ClinicalTrials.gov NCT04841603; https://clinicaltrials.gov/ct2/show/NCT04841603

(JMIR Form Res 2022;6(11):e38460) doi:10.2196/38460

KEYWORDS

inpatient; mental health; mental disorder; psychiatry; psychiatric; smartphone app; cognitive behavioral therapy; CBT; anxiety; mobile app; mobile health; mHealth; health app; digital health; eHealth; feasibility study; randomized controlled trial; RCT; feasibility; acceptability

Introduction

Inpatient admission plays a crucial role in the treatment of patients with serious mental illness. Patients requiring hospitalization typically present with severe symptoms that require a combination of pharmacological, psychological, and social interventions [1]. While pharmacotherapy is the mainstay of psychiatric inpatient care, most inpatients report that their psychosocial needs are not adequately addressed [2]. For example, prior to the COVID-19 pandemic, only 4% of psychiatric inpatients participated in group psychotherapy and 84% reported being socially disengaged and inactive [3]. This is despite evidence from meta-analyses that inpatient psychotherapeutic interventions are both effective and highly valued by patients [4-6]. COVID-19 pandemic–related social distancing measures have further limited inpatient access to traditional psychosocial interventions.

Anxiety symptoms are frequently targeted in inpatient psychosocial interventions owing to their transdiagnostic ubiquity and relevance [7-9]. Cognitive behavioral therapy (CBT) is an evidence-based intervention for anxiety symptoms and provides a broadly applicable psychotherapy framework suitable for most psychiatric conditions [10,11]. The core principles of CBT can be communicated to people with a range of cognitive abilities, such as children [12] or people with cognitive impairment [13], which is important for inpatients for whom severe symptoms may impede more complex psychotherapeutic modalities. In addition, during inpatient hospitalizations, inpatients have a significant amount of time in a controlled environment with their care team nearby, providing a structure to assist with the acquisition, deployment, and generalization of skills learned in CBT. Group CBT was commonly endorsed in inpatient settings prior to the COVID-19 pandemic to treat a range of symptoms and conditions [3,6]. However, financial and logistical challenges to implementing in-person inpatient group programming have been magnified by the pandemic [14,15]. Innovative solutions to deliver CBT to psychiatric inpatients were being investigated prepandemic and need further investigation urgently [16,17].

Some governments (eg, in Australia) are now providing free access to digital CBT [18]. There is robust and growing evidence for digital CBT for outpatients with mental illness; however, its use on psychiatric inpatient units has not been sufficiently studied [19]. MindShift CBT is a freely available, smartphone-based app developed by Anxiety Canada. There is some preliminary evidence suggesting that this app can be helpful for outpatients with moderate or severe anxiety [20]. While some inpatient programs have started to make personal-use electronics available to the inpatients they serve, studies of specific app-based psychosocial interventions are limited. A study that examined the use of Headspace, a commercially available meditation app, on an inpatient unit had promising feasibility results [17]. However, our review of the literature identified no trials of a CBT app on a psychiatric inpatient unit. Thus, we assessed the feasibility and acceptability of the MindShift CBT app on an inpatient unit and compared its impact to usual inpatient care.

Methods

Setting and Participants

This pilot randomized controlled trial was performed at one of the acute general psychiatry units of a large academic psychiatric hospital, the Centre for Addiction and Mental Health (CAMH), which provides secondary and tertiary inpatient care in Toronto, Canada. The study was registered on ClinicalTrials.gov (identifier: NCT04841603). In addition to the clinician researchers, a peer researcher (ie, a researcher with lived experience of mental illness) was involved in recruitment of patients and in the collection and analyses of qualitative data as per best practices for patient involvement in mental health research [21]. This report was prepared according to the CONSORT (Consolidated Standards of Reporting Trials) statement (Multimedia Appendix 1).

All patients admitted to the unit between April 5 and May 5, 2021, were screened by the admitting inpatient team for eligibility and referred to the research team if they assented to hear more about the study. Inclusion criteria were age between 18 and 65 years, fluency in English, a Dynamic Appraisal for Situational Aggression (DASA) score < 3 at the time of referral (which indicates lower acute risk of violence/aggression [22]), and capacity to consent to participation as assessed by the treating team. Exclusion criteria were diagnoses of moderate or severe intellectual disability, learning disability, or neurocognitive disorder, as these participants may have had difficulty navigating the app.

Ethical Considerations

The study was performed in accordance with the Ethical Principles of Psychologists and Code of Conduct as set out by the British Association for Behavioural and Cognitive Psychotherapies and British Psychological Society, and with the Declaration of Helsinki. All participants provided written informed consent after being provided verbal and written information about the study and before the initiation of any study procedures. The CAMH Research Ethics Review Board approved the protocol, all supplementary documents, and the informed consent form (#116/2020-01). Participants could be withdrawn from the study at their request, the treatment team's

request, or if they experienced significant worsening of symptoms as determined by the research team or treating physicians. Study data were deidentified and anonymized, and participant records were stored in a secure locked drawer in a locked room. Participants were not compensated for their participation in the study. After the trial period, participants in the control group were provided access to the MindShift CBT app.

Randomization

After provision of informed consent and confirmation of eligibility, participants were randomized without stratification, using the open-source randomizer randomizer.org, to MindShift+treatment as usual (TAU) or TAU. By design, neither the research team nor participants were blinded to group assignment upon commencement of the study.

Intervention and Control Conditions

After being randomized, participants were involved in the study for 6 days (ie, participants whose length of stay was longer than 6 days were involved for only 6 days). The 6-day study period started on each Tuesday of 4 consecutive weeks. On the first day of the study, each participant met individually with a research team member uninvolved with clinical care on the unit. Participants in both conditions were provided with a tablet to complete the baseline questionnaires (see below). Participants randomized to the intervention (ie, MindShift+TAU) completed an individual introductory session to the app, comprising a review of several sections: the "Home" section, including CBT-based tools such as a fillable thought record and exposure ladder; the "Learn" section containing basic information about anxiety and CBT; the "Quick Relief" section containing meditation exercises; and the "Goals" section, which allows users to set and review progress toward specific and measurable goals. Patients were asked to open the app and use their preferred features for at least 10 minutes and to use the "check in" function daily. This feature requests that users: (1) rate their current mood on a 1-10 Likert scale calibrated with qualitative labels and emoticons; (2) type a response to an open question stating "What's going on? Describe what's going on in your life right now and/or what's on your mind"; and (3) indicate active symptoms of anxiety selected from a checklist. Due to concerns about confidential patient information being shared with others on the unit inadvertently, the tablets used were programmed to require login to the MindShift app each time they were opened. On the third day, each participant was offered a 15-minute session with a clinician or peer member of the research team to address any questions about the study. For participants randomized to the intervention, the use of the app and its core features were also reviewed during these sessions. On the sixth and final day, all participants completed poststudy questionnaires and a debriefing session. Participants in the intervention condition also completed a 30-minute semistructured qualitative interview with a peer or clinician researcher.

All tablets were stored in the nursing station when not in use by patients or when requiring charging. The tablet given to the participants randomized to the control (TAU) condition did not include the MindShift app but had otherwise identical apps installed and accessible, including a video-streaming app and internet browser. TAU on the inpatient unit comprised daily assessments by a psychiatrist, meetings with a social worker as determined by the clinical team, 24/7 access to nursing staff, and pharmacotherapy management. No group psychosocial activities were conducted as part of TAU during the study period due to COVID-19 pandemic–related restrictions. After completing the exit questionnaires on the sixth day of the study, control group participants were offered access to tablets with the MindShift app.

Measures and Outcomes

The primary outcomes of this trial were indicators of feasibility and acceptability related to the use of the MindShift CBT app in an inpatient setting. Feasibility was assessed with the rates of consent, study completion, and withdrawal, and by completeness of data in pre- and poststudy questionnaires. Acceptability was assessed quantitatively with the Client Satisfaction Questionnaire (CSQ-8), a self-reported scale with 8 items that describes satisfaction with a health service [23]; participants in the intervention condition also completed a user-experience questionnaire for app-based interventions (Multimedia Appendix 2). Given the feasibility-sized sample, changes in symptoms from baseline to the end of the study were classified as secondary quantitative outcomes. The Generalized Anxiety Disorder (GAD-7) scale was used to assess anxiety [24], the 9-item Patient Health Questionnaire (PHQ-9) was used for depressive symptoms [25], and Kessler Psychological Distress Scale (K10) was used as a global assessment of psychological function [26].

Qualitative Data Analysis

To assess qualitative data, a thematic analysis was applied to transcribed records of the semistructured qualitative interviews using the 15-point checklist of criteria for good thematic analysis published by Braun and Clarke [27]. Transcripts were reviewed and coded independently by the lead clinical researcher and peer researcher, and themes arising from the codes were identified independently. The researchers then met to establish a consensus regarding themes, and items were recorded in alignment with the themes identified by consensus.

Sample Size, Power, and Statistical Analysis

One of the main goals of this feasibility study was to obtain data that could be used to calculate sample sizes for a future larger, confirmatory trial. Prior research has suggested that a sample size of 12 in each arm suffices for a pilot feasibility study [28]. Given the patient flow on the inpatient unit on which the study was conducted, we planned to conduct the study over 4 weeks with the expectation that this would be long enough to enable randomization of 24 participants.

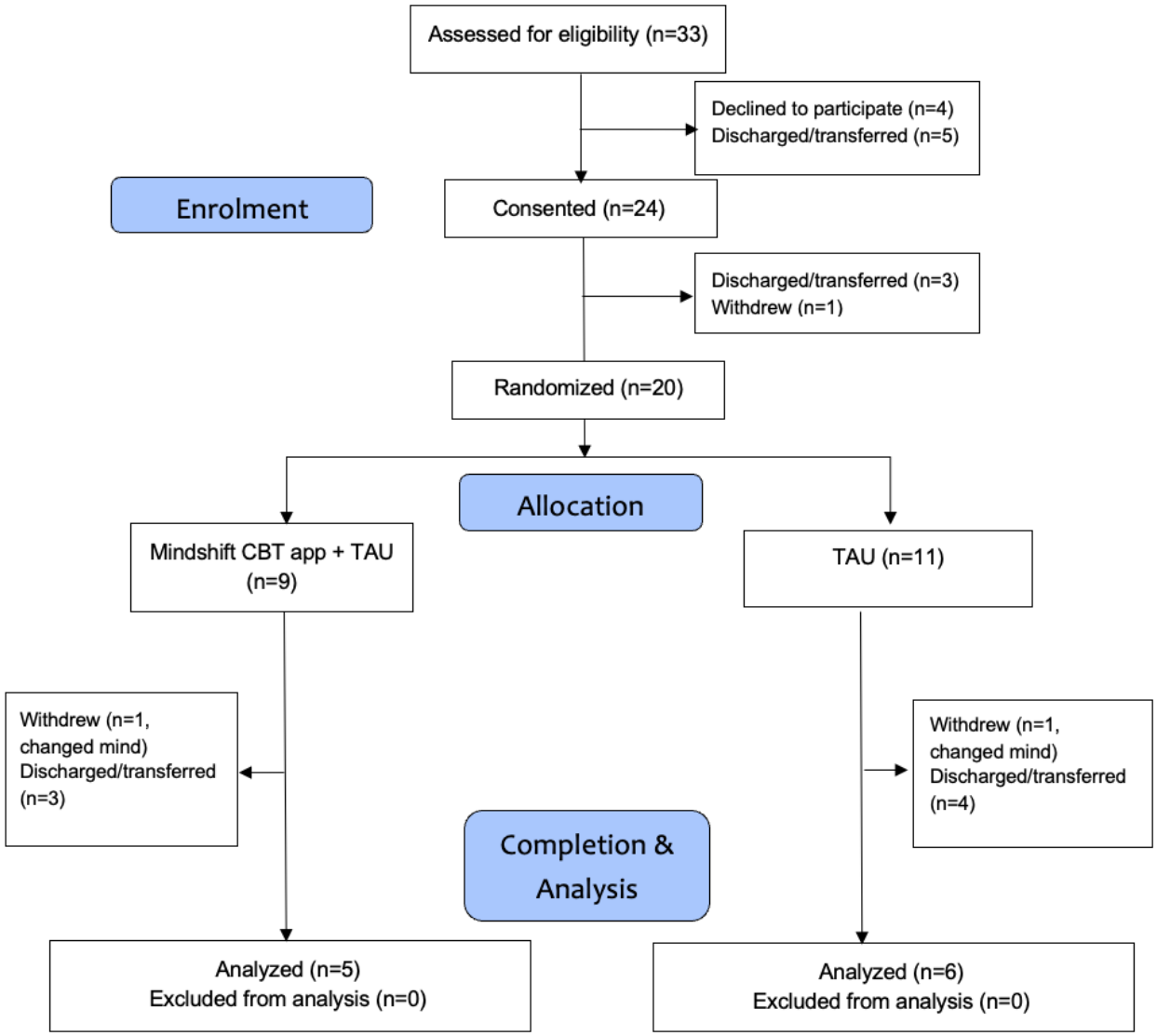
Statistical analysis was completed using SPSS 28 software. Descriptive statistics, including mean, SD, and range, were calculated for the entire study sample and for each group. The baseline characteristics were compared between the two groups using Kruskal-Wallis tests.

Results

Feasibility Measures

Figure 1 presents the flow of referrals and participation. During the study period (April 5 to May 5, 2021), 33 patients were referred to the study, 5 of whom were discharged or transferred prior to being invited to participate. Of the 28 patients invited to participate, 4 declined, resulting in a consent rate of 24/28 (86%). Of the 24 consented participants, 1 withdrew and 3 were transferred or discharged before being randomized, resulting in a randomization rate of 20/24 (75%). Of the 20 randomized participants, 11 (55%) completed the study and provided pre- and poststudy data that were analyzable. Of the 9 participants who did not complete the study, 2 withdrew (1 in each condition) and 7 were withdrawn because of being discharged or transferred prior to study completion (3 in the MindShift+TAU condition and 4 in the TAU condition).

Figure 1. Flow of participant inclusion. CBT: cognitive behavioral therapy; TAU: treatment as usual.



Characteristics and Clinical Outcomes of Randomized Participants

Table 1 summarizes the demographic and baseline clinical characteristics of the participants in the two study conditions. Overall, intervention and control condition participants did not show significant differences in baseline characteristics, except for a higher proportion of women in the control condition. Of note, 65% of the participants were admitted with a primary diagnosis of psychotic disorder, and none of the patients who participated was admitted for treatment of a primary anxiety disorder. Overall, 35% of the patients were involuntarily admitted under the Mental Health Act. The mean length of stay of the 20 participants was 20 (SD 4.4, range 3-21) days from the time of admission to the first of either discharge, study completion, withdrawal from the study, or transfer to another unit. Table 1 also presents the clinical outcomes of the 11 participants who completed the study and for whom pre- and poststudy clinical measures were available. The differences in GAD-7, PHQ-9, or K10 scores were not interpretable given the small sample size.

Table 1. Demographic and clinical characteristics of the 20 randomized participants.

Characteristic	MindShift+TAU[a] (n=9)	TAU (n=11)
Age (years), mean (SD)	34.8 (8.0)	29.4 (9.8)
Self-reported sex as female, n (%)	1 (11)	4 (36)
Self-reported racial/ethnic group, n (%)		
White	3 (33)	4 (36)
Black	2 (22)	2 (18)
Middle Eastern	1 (11)	1 (9)
South Asian	1 (11)	1 (9)
Southeast Asian	1 (11)	1 (9)
Not reported	1 (11)	2 (18)
Some postsecondary education, n (%)	3 (33)	4 (36)
Annual income below poverty line, n (%)	3 (33)	4 (36)
Primary diagnosis, n (%)		
Schizophrenia-spectrum disorder	6 (67)	6 (55)
Borderline personality disorder	0 (0)	2 (18)
Major depressive disorder	0 (0)	2 (18)
Amphetamine-induced psychotic disorder	1 (11)	0 (0)
Bipolar disorder	0 (0)	1 (9)
Opioid use disorder	1 (11)	0 (0)
Posttraumatic stress disorder	1 (11)	0 (0)
Admitted voluntarily, n (%)	6 (67)	7 (64)
GAD-7[b] score, mean (SD)		
Prestudy	10.1 (7.3), n=9	8.2 (4.8), n=11
Poststudy	10.8 (3.3), n=5	9.2 (3.3), n=6
PHQ-9[c] score, mean (SD)		
Prestudy	11.0 (8.8), n=9	11.1 (5.9), n=11
Poststudy	13.6 (3.7), n=5	10.8 (3.7), n=6
K10[d] score, mean (SD)		
Prestudy	27.3 (8.2), n=9	27.3 (8.2), n=11
Poststudy	27.6 (6.0), n=5	25.8 (6.7), n=6

[a]TAU: treatment as usual.

[b]GAD-7: Generalized Anxiety Disorder Assessment.

[c]PHQ-9: Patient Health Questionnaire.

[d]K10: Kessler Psychological Distress Questionnaire.

Acceptability Measures

The scores on the CSQ-8 questionnaire (possible range 8-32, higher scores indicate higher satisfaction) completed by the 5 participants in the MindShift+TAU condition (mean 20.2, SD 3.4; range 17-25) indicated overall positive satisfaction with the intervention. Similarly, the scores on the user-experience questionnaire (possible range 0-110, higher scores indicate better usability), available for 4 participants (mean 80.5, SD 27.4; range 46-108), reflected moderate to high user engagement and satisfaction.

Semistructured Exit Interviews

Four of five participants in the MindShift+TAU condition completed exit interviews, providing information that was grouped into five major themes: feasibility of the intervention in the inpatient setting, usability and patterns of app use, acceptability of the features of the app, a desire to provide feedback and to be seen as active participants during inpatient admissions (which we characterized as agency), and impact of the app on symptoms (see Textbox 1). The benefit of having access to the tablets and app on the inpatient unit was a recurring theme. For instance, one participant commented, “It helped a

lot…It gave me access to a tablet and the programs on it, so indirectly I had access to the programs…and do your work as well, so I think it was beneficial." Users found the app simple to navigate. Frequency of use varied from multiple times per day to just a few times over the week. The "chill zone" feature of the app was the most commented upon and received consistently positive feedback. The participants spontaneously provided suggestions for improvement or additional features, and talked about their participation in research and how their feedback could impact future programming. For instance, one participant commented, "I've been asked to participate in the research. This means it's not just inpatient and that's it...you could put input in it, you feel that you're worth something." There were mixed comments about the impact the app had on symptoms, with some believing use was beneficial, while one participant commented that the app may trigger negative thoughts: "I didn't like to read too much into it because the info could scare me. It's like phobias, 'I'm scared I'm scared!'" Two participants described being confused about the purpose of the pre- and poststudy questionnaires, with one participant conflating the prestudy questionnaires with the CBT app itself.

Textbox 1. Themes and representative quotes emerging from the exit interviews.

Theme 1: Feasibility (18 corresponding statements)

"It helped a lot…It gave me access to, a tablet and the programs on it, so indirectly I had access to the programs…and do your work as well, so I think it was beneficial"

"It was difficult because it took a lot of mental concentration to think through these questions…They're not like very generic questions like hey, how are you?... so I found myself getting flustered" (in reference to pre- and poststudy surveys)

Theme 2: Usability (16 corresponding statements)

"It was very easy to use and click through and enter different…get different information"

"I worked on the app in a different mood and every time I was able to connect"

Theme 3: Acceptability (21 corresponding statements)

"It would provide different avenues to have audio abilities, capabilities so you didn't have to read if you were lethargic, or couldn't read"

"The chill zone I found was useful, that like mindfulness meditation was great...It's also very simple. You just sit there and you listen to directions and you chill out"

Theme 4: Symptoms (11 corresponding statements)

"I haven't been able to notice too much of a difference, but I'm sure the app has helped me"

"I didn't like to read too much into it because the info could scare me. It's like phobias, 'I'm scared I'm scared!'"

"The more information I have, the more it helps"

Theme 5: Agency (22 corresponding statements)

"It should come back and say you actually did a good job. This is a reward for doing a good job" (in regard to the journal entry capacity of the app)

"I've been asked to participate in the research. This mean it's not just inpatient and that's it...you could put input in it, you feel that you're worth something"

"Is me speaking, its it going to the originator who made the program...and will they like to run with it? Or what is the end result from all of this?"

Discussion

Principal Findings

This study provides informative data on the feasibility of studying an app-based intervention on an acute psychiatric inpatient unit. A survey of health care provider attitudes toward digital interventions in this setting raised important questions about feasibility, including implementation and suitability for acutely ill patients [29]. Our study supports that the use of digital psychotherapeutic interventions on the inpatient unit requires careful planning and considerations, but could have benefits for an acutely unwell, transdiagnostic population. The high consent rate (24/28, 86%) among inpatients referred to the study demonstrates interest in and willingness to participate in such an intervention. However, the relatively low completion rate (11/20, 55%) emphasizes the challenges of conducting research in acute inpatient settings, where rapid patient turnover can impact the participation and retention of participants. Our goal of obtaining data to calculate sample sizes for a larger confirmatory trial was not achieved because of limitations in our study design and time allotted for data collection, which will inform the design of future trials moving forward. We gathered pre- and poststudy clinical measures of anxiety, depressive symptoms, and psychological distress, but could not accurately interpret these data owing to the small sample size.

While the limited sample size also restricts our ability to draw broad conclusions from the thematic analysis, the quotes provide rich context to reflect patient experiences of the intervention. An unexpected finding was that participants felt empowered by contributing to research and were eager to offer feedback. Voluntary participation in inpatient research projects may stand in psychological contrast to the disempowerment experienced by at least one-third of our participants who were admitted involuntarily. Some participants indicated an intent to download the app on their personal devices after completion of the study and planned to continue to use it after discharge from the

hospital. If included in discharge planning, such an intervention could serve as a "transitional object" between inpatient and outpatient settings [30].

A core feature of our study was the collaboration between clinician researchers and a peer researcher for recruitment, data collection, and analysis. However, collaborative efforts must occur earlier in the research process to guide research question development and study design [21]. Compiling research teams of diverse stakeholders, including people with lived experience of mental illness, will be key to designing patient-centered research objectives in the historically coercive inpatient setting.

Limitations

Of the 20 randomized participants who withdrew from the study, 7 (35%) had to be withdrawn due to being transferred to another more specialized unit or being discharged. Due to limited research personnel, the study operated on a fixed schedule with a 1-week intervention that started on each Tuesday of 4 consecutive weeks, which contributed to low completion rates. A future study would need to use a rolling-entry design in which enrolled participants would be randomized and start the intervention on the day they provide consent. Although the focus of our study was transdiagnostic use of the MindShift app for anxiety symptoms, no participants in the intervention condition had a primary anxiety disorder, which may have impacted the acceptability of the app that is tailored to people with a primary anxiety disorder diagnosis. The requirement to log on to the app at each use presented an additional barrier compared to typical use on personal devices, where users are only required to log in once upon first use of the app and then remain logged in. Additionally, there were unanticipated logistical challenges, which can be generalized to researching other inpatient technology-based interventions, including requiring a system for tablet sign-out, technical troubleshooting, and a need for policies on the inpatient unit about acceptable use of tablets. In addition to the challenges posed to research, these limitations are also relevant to implementation of app-based therapy for clinical use.

Conclusions

Despite the limitations of this study, our findings support the overall feasibility and acceptability of use of psychotherapeutic apps by inpatients admitted with a variety of psychiatric diagnoses. Specifically, the interactive features of guided meditation and mindfulness in the MindShift CBT app were well-received among our participants. An unintended but important finding was the empowerment described by inpatients from being included in a research study run collaboratively by peer and clinician researchers. Our feasibility trial adds to the limited literature on the use of digital psychosocial interventions on acute psychiatric wards. These interventions could be particularly useful in acute transdiagnostic inpatient settings where tablets or mobile devices are already available, or when traditional group-based psychosocial interventions are not practical (ie, when patients are on infection control precautions). However, it is unclear whether investing in technology-based psychosocial interventions for psychiatric inpatient units would be universally appropriate, or whether these are reasonable substitutes for existing psychosocial treatments. Potential future directions for research include exploring gaps in knowledge about the efficacy of this type of inpatient intervention; exploration of the use of other apps with psychotherapeutic features that are interactive, such as guided meditation and mindfulness; and head-to-head comparison with standard-of-care inpatient group psychotherapy interventions.

Acknowledgments

We would like to thank the Centre for Addiction and Mental Health Empowerment Council for their support in connecting our peer and clinician researchers.

Conflicts of Interest

LR sits on the Scientific Advisory Committee for Anxiety Canada and receives funding from Anxiety Canada to study the MindShift CBT app. He reviewed and provided feedback on the draft manuscript but did not provide input toward study design, data collection, or analysis. BHM holds and receives support from the Labatt Family Chair in Biology of Depression in Late-Life Adults at the University of Toronto. He also receives compensation from the Department of Psychiatry, University of Toronto, and the Centre for Addiction and Mental Health (CAMH), Toronto, Canada.

Multimedia Appendix 1
CONSORT (Consolidated Standards of Reporting Trials) checklist.
[PDF File (Adobe PDF File), 1218 KB - formative_v6i11e38460_app1.pdf]

Multimedia Appendix 2
User experience questionnaire.
[DOCX File , 16 KB - formative_v6i11e38460_app2.docx]

References

1. Bowers L, Chaplin R, Quirk A, Lelliott P. A conceptual model of the aims and functions of acute inpatient psychiatry. J Mental Health 2009 Aug 13;18(4):316-325. [doi: 10.1080/09638230802053359]

2. Hopkins J, Loeb S, Fick D. Beyond satisfaction, what service users expect of inpatient mental health care: a literature review. J Psychiatr Ment Health Nurs 2009 Dec;16(10):927-937. [doi: 10.1111/j.1365-2850.2009.01501.x] [Medline: 19930367]
3. Radcliffe J, Smith R. Acute in-patient psychiatry: how patients spend their time on acute psychiatric wards. Psychiatr Bull 2018 Jan 02;31(5):167-170. [doi: 10.1192/pb.bp.105.009290]
4. Cuijpers P, Ciharova M, Miguel C, Harrer M, Ebert DD, Brakemeier E, et al. Psychological treatment of depression in institutional settings: a meta-analytic review. J Affect Disord 2021 May 01;286:340-350 [FREE Full text] [doi: 10.1016/j.jad.2021.03.017] [Medline: 33773217]
5. Paterson C, Karatzias T, Dickson A, Harper S, Dougall N, Hutton P. Psychological therapy for inpatients receiving acute mental health care: a systematic review and meta-analysis of controlled trials. Br J Clin Psychol 2018 Nov 16;57(4):453-472. [doi: 10.1111/bjc.12182] [Medline: 29660770]
6. Xia J, Merinder L, Belgamwar M. Psychoeducation for schizophrenia. Cochrane Database Syst Rev 2011 Jun 15(6):CD002831 [FREE Full text] [doi: 10.1002/14651858.CD002831.pub2] [Medline: 21678337]
7. Kösters M, Burlingame GM, Nachtigall C, Strauss B. A meta-analytic review of the effectiveness of inpatient group psychotherapy. Group Dyn Theory Res Pract 2006 Jun;10(2):146-163. [doi: 10.1037/1089-2699.10.2.146]
8. Lavey R, Sherman T, Mueser KT, Osborne DD, Currier M, Wolfe R. The effects of yoga on mood in psychiatric inpatients. Psychiatr Rehabil J 2005;28(4):399-402. [doi: 10.2975/28.2005.399.402] [Medline: 15895926]
9. Zwerenz R, Becker J, Knickenberg RJ, Siepmann M, Hagen K, Beutel ME. Online self-help as an add-on to inpatient psychotherapy: efficacy of a new blended treatment approach. Psychother Psychosom 2017;86(6):341-350. [doi: 10.1159/000481177] [Medline: 29131090]
10. Berking M, Ebert D, Cuijpers P, Hofmann SG. Emotion regulation skills training enhances the efficacy of inpatient cognitive behavioral therapy for major depressive disorder: a randomized controlled trial. Psychother Psychosom 2013 May 22;82(4):234-245. [doi: 10.1159/000348448] [Medline: 23712210]
11. Bloom JM, Woodward EN, Susmaras T, Pantalone DW. Use of dialectical behavior therapy in inpatient treatment of borderline personality disorder: a systematic review. Psychiatr Serv 2012 Sep 01;63(9):881-888. [doi: 10.1176/appi.ps.201100311] [Medline: 22751905]
12. Oar EL, Johnco C, Ollendick TH. Cognitive behavioral therapy for anxiety and depression in children and adolescents. Psychiatr Clin North Am 2017 Dec;40(4):661-674. [doi: 10.1016/j.psc.2017.08.002] [Medline: 29080592]
13. Tay K, Subramaniam P, Oei TP. Cognitive behavioural therapy can be effective in treating anxiety and depression in persons with dementia: a systematic review. Psychogeriatrics 2019 May 13;19(3):264-275. [doi: 10.1111/psyg.12391] [Medline: 30548731]
14. Bojdani E, Rajagopalan A, Chen A, Gearin P, Olcott W, Shankar V, et al. COVID-19 pandemic: impact on psychiatric care in the United States. Psychiatry Res 2020 Jul;289:113069 [FREE Full text] [doi: 10.1016/j.psychres.2020.113069] [Medline: 32413707]
15. Li S, Zhang Y. Mental healthcare for psychiatric inpatients during the COVID-19 epidemic. Gen Psychiatr 2020 Apr 01;33(2):e100216 [FREE Full text] [doi: 10.1136/gpsych-2020-100216] [Medline: 32363326]
16. Ben-Zeev D. Technology-based interventions for psychiatric illnesses: improving care, one patient at a time. Epidemiol Psychiatr Sci 2014 Jul 21;23(4):317-321. [doi: 10.1017/s2045796014000432]
17. Mistler LA, Ben-Zeev D, Carpenter-Song E, Brunette MF, Friedman MJ. Mobile mindfulness intervention on an acute psychiatric unit: feasibility and acceptability study. JMIR Ment Health 2017 Aug 21;4(3):e34 [FREE Full text] [doi: 10.2196/mental.7717] [Medline: 28827214]
18. Andersson G, Carlbring P, Titov N, Lindefors N. Internet interventions for adults with anxiety and mood disorders: a narrative umbrella review of recent meta-analyses. Can J Psychiatry 2019 Jul 16;64(7):465-470 [FREE Full text] [doi: 10.1177/0706743719839381] [Medline: 31096757]
19. Dülsen P, Bendig E, Küchler AM, Christensen H, Baumeister H. Digital interventions in adult mental healthcare settings: recent evidence and future directions. Curr Opin Psychiatry 2020 Jul;33(4):422-431. [doi: 10.1097/YCO.0000000000000614] [Medline: 32427592]
20. Paul AM, Eubanks Fleming CJ. Anxiety management on campus: an evaluation of a mobile health intervention. J Technol Behav Sci 2018 Sep 19;4(1):58-61. [doi: 10.1007/s41347-018-0074-2]
21. Jennings H, Slade M, Bates P, Munday E, Toney R. Best practice framework for Patient and Public Involvement (PPI) in collaborative data analysis of qualitative mental health research: methodology development and refinement. BMC Psychiatry 2018 Jun 28;18(1):213 [FREE Full text] [doi: 10.1186/s12888-018-1794-8] [Medline: 29954373]
22. Ogloff JRP, Daffern M. The dynamic appraisal of situational aggression: an instrument to assess risk for imminent aggression in psychiatric inpatients. Behav Sci Law 2006;24(6):799-813. [doi: 10.1002/bsl.741] [Medline: 17171770]
23. Attkisson C, 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]
24. 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]

25. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
26. Andrews G, Slade T. Interpreting scores on the Kessler Psychological Distress Scale (K10). Aust N Z J Public Health 2001 Dec;25(6):494-497. [doi: 10.1111/j.1467-842x.2001.tb00310.x] [Medline: 11824981]
27. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
28. Julious SA. Sample size of 12 per group rule of thumb for a pilot study. Pharmaceut Statist 2005 Oct;4(4):287-291. [doi: 10.1002/pst.185]
29. Sander J, Bolinski F, Diekmann S, Gaebel W, Günther K, Hauth I, et al. Online therapy: an added value for inpatient routine care? Perspectives from mental health care professionals. Eur Arch Psychiatry Clin Neurosci 2022 Feb 16;272(1):107-118 [FREE Full text] [doi: 10.1007/s00406-021-01251-1] [Medline: 33725165]
30. Casher M. "There's no such thing as a patient": reflections on the significance of the work of D. W. Winnicott for modern inpatient psychiatric treatment. Harv Rev Psychiatry 2013;21(4):181-187. [doi: 10.1097/HRP.0b013e31828ea604] [Medline: 24651506]

Abbreviations

CAMH: Centre for Addiction and Mental Health
CBT: cognitive behavioral therapy
CONSORT: Consolidated Standards of Reporting Trials
CSQ-8: Client Satisfaction Questionnaire
DASA: Dynamic Appraisal for Situational Aggression
GAD-7: Generalized Anxiety Disorder scale
K10: Kessler Psychological Distress Scale
PHQ-9: 9-item Patient Health Questionnaire
TAU: treatment as usual

Edited by A Mavragani; submitted 02.05.22; peer-reviewed by I Choi, A Pencer; comments to author 23.08.22; revised version received 13.09.22; accepted 29.09.22; published 02.11.22.

Please cite as:
Sharma G, Schlosser L, Jones BDM, Blumberger DM, Gratzer D, Husain MO, Mulsant BH, Rappaport L, Stergiopoulos V, Husain MI
Brief App-Based Cognitive Behavioral Therapy for Anxiety Symptoms in Psychiatric Inpatients: Feasibility Randomized Controlled Trial
JMIR Form Res 2022;6(11):e38460
URL: https://formative.jmir.org/2022/11/e38460
doi:10.2196/38460
PMID:36322113

Original Paper

Internet-Based Intervention Compared to Brief Intervention for Smoking Cessation in Brazil: Pilot Study

Nathalia Munck Machado[1,2*], PhD; Henrique Pinto Gomide[3*], PhD; Heder Soares Bernardino[4*], PhD; Telmo Mota Ronzani[2*], PhD

[1]Department of Population Health, University of Kansas Medical Center, Kansas City, KS, United States
[2]Department of Psychology, Universidade Federal de Juiz de Fora, Juiz de Fora, Brazil
[3]Department of Education, Universidade Federal de Viçosa, Viçosa, Brazil
[4]Department of Computer Science, Universidade Federal de Juiz de Fora, Juiz de Fora, Brazil
[*]all authors contributed equally

Corresponding Author:
Nathalia Munck Machado, PhD
Department of Population Health
University of Kansas Medical Center
3901 Rainbow Blvd
Kansas City, KS, 66160
United States
Phone: 1 8163287381
Email: nmachado@kumc.edu

Abstract

Background: Smoking is still the leading cause of preventable death. Governments and health care providers should make available more accessible resources to help tobacco users stop.

Objective: This study describes a pilot longitudinal study that evaluated the efficacy of an internet-based intervention compared to the brief intervention for smoking cessation among Brazilians.

Methods: Eligible participants were recruited and randomly allocated to one of the two interventions. Measures were drawn by comparing cessation rates, motivation scores, and sought treatment between groups, assessed 1 and 3 months after the intervention. Inferential analysis was performed to compare the participants' characteristics, and the intention to treat was calculated.

Results: A total of 49 smokers were enrolled in this study (n=25, 51% in the brief intervention group; n=24, 49% in the internet-based intervention group). Mean age was 44.5 (SD 13.3) years; most were male (n=29, 59.2%), had elementary school (n=22, 44.9%), smoked 14.5 cigarettes per day on average (SD 8.6), and had a mean score of 4.65 for nicotine dependence and 5.7 for motivation to quit. Moreover, 35 (71%) participants answered follow-up 1, and 19 (39%) answered follow-up 2. The results showed similar rates of cessation and reduction for both intervention groups.

Conclusions: The internet-based intervention was slightly more effective for smoking cessation, while the brief intervention was more effective in reducing the number of cigarettes smoked per day. This difference was small and had no statistical significance even after adjusting for intention-to-treat analysis. These results should be interpreted with caution, especially due to the small sample size.

(JMIR Form Res 2022;6(11):e30327) doi:10.2196/30327

KEYWORDS

smoking cessation; internet-based intervention; digital intervention; mobile health; mHealth; tobacco; addiction; public health; digital intervention; substance use

Introduction

Smoking is the leading cause of cancer, preventable death, and disability worldwide, causing around 8 million deaths per year [1]. Overall mortality among smokers is about 3 times higher than never smokers [2]. Despite the severe health risks, 1.3 billion people are still smokers worldwide [1]. Of these, about 80% live in low- and middle-income countries where the burden

of tobacco-related illness and death is even more significant [1]. In Brazil, 12.6% of the adult population smokes—15.9% men and 9.6% women [3]. Smoking is responsible for more than 161.853 deaths per year in Brazil, equivalent to 443 deaths per day and 13% of all deaths in people older than 35 years [4].

Several health promotion methods are used for smoking prevention and cessation [5]. However, it is well reported that professional counseling combined with pharmacotherapy is the most effective treatment for smoking cessation [6]. However, due to the high cost of face-to-face treatments, less costly and effective treatment forms can help address the treatment demand. Moreover, the governments and health care providers should make available more accessible resources to help tobacco users stop, as suggested by the World Health Organization Framework Convention on Tobacco Control, Article 14 [7].

Brief advice can significantly increase the odds of quitting [8]. The brief intervention (BI) based on the motivational interview integrates different strategies to increase motivation to change behaviors [9]. BI has better results than simple counseling, especially among those not ready to quit smoking [10], and it is considered a cost-effective strategy for smoking cessation [8,11,12]. There is good evidence of BI effectiveness provided by a therapist, such as advice from a doctor, and it yields a quit rate of 13.4% [6].

The internet can reach many people and has great potential to provide behavior change interventions to them at a low cost [13]. Internet-based interventions are attractive due to their low cost, convenience, and confidentiality [14,15]. These interventions can also reach smokers who might not access in-person support due to limited health care availability or stigmatization. They can provide an opportunity for psychological help to those who could not receive it otherwise [16-18].

Tobacco users have increasingly used web-based resources to search ways for smoking cessation, with data showing that more than 12 million smokers have used the internet to seek help for quitting in 2017 [19]. The effectiveness of internet-based interventions is well established, with reported quit rates ranging from 12.8% to 14.3% [20,21], and the available evidence is already enough to justify its use for smoking cessation [13,22,23].

Digital technology has been evolving rapidly, requiring it to be updated and refined after evaluative trials to not become obsolete and less attractive when available outside research contexts [24]. People can, however, use digital health interventions differently in real-world contexts compared to the conditions of the studies [25]. For these reasons, in addition to evaluating the viability and effectiveness of the interventions in clinical contexts, it is essential to examine the use of digital interventions in real-world contexts. In this matter, this study aimed to evaluate the efficacy of a computerized intervention compared to the brief intervention (face-to-face intervention) for smoking cessation among Brazilians in a real-world context.

Methods

Participants

Participants were recruited from the Federal University of Juiz de Fora and a city-owned company in Juiz de Fora (Brazil). Potential participants were identified through flyers and company meetings and were contacted and invited individually to participate in a smoking cessation program. Eligibility criteria were currently smoking and aged 18 years. Exclusion criteria were participating in smoking cessation treatment at the time of the study. The participants completed an eligibility screening and provided written informed consent.

Ethics Approval

This study was approved by the Human Research Ethics Committee of the Federal University of Juiz de Fora (CAEE: 84446218.4.0000.5147).

Sample Size

Studies suggest a minimum sample size of 12 subjects per treatment arm [26] or at least 30 [27] to 70 participants [28] in pilot trials based on the rules of thumb. In this study, recruitment was terminated when there was no more flow of participants to be recruited.

Study Design and Interventions

Eligibility criteria were checked at baseline, and participants who met the criteria were randomized on a 1:1 ratio using a uniformly distributed random number generator. Participants were allocated to one of the following two intervention arms: (1) the life without tobacco (LWT) web-based intervention [29] or (2) the face-to-face BI. These two interventions are described below.

LWT Intervention

This is an open-source web intervention available in 7 languages. It was developed based on scientific research and treatment protocols to offer psychoeducation to smokers [29]. Information about smoking management is based on the "Treating tobacco use and dependence - 2008 update" guidelines [6]. The intervention is divided into the following three stages: (1) "Is it worth stopping?"—intended for smokers who are not yet confident about attempting to quit; (2) "Ready to quit?"—intended for smokers confident in attempting to quit; and (3) "Have you stopped?"—intended for smokers who have gone through the previous phase or relapsed.

Educational content includes information about the consequences of tobacco use and the benefits of quitting, effective cessation methods and medications, nicotine dependence, and comorbidities related to smoking. The main objective of the intervention is to develop a personalized plan to stop smoking, which focuses on preparing to choose a quit date, coping with slips, and preventing relapse. After selecting the stop date, the user receives a follow-up by email for 12 months

Face-to-face BI

BI involves opportunistic advice, discussion, negotiation, and encouragement. It is a structured, focal, and objective

intervention strategy focused on behavior change [30]. A protocol was developed including the following essential elements of the BI process aimed at users of psychoactive substances [31]: screening, feedback, setting goals, discussing the pros and cons of use, counseling, and development of the patient's self-efficacy. The intervention was based on the Stages of Change Model, considering the stage of change in which the participant was [32]. The protocol was printed to be followed during the intervention to make the BIs as similar as possible. The intervention was performed in a single session of approximately 20 minutes. The objectives were the same as those of the web-based intervention—developing a personalized plan to stop, setting up a quit date, helping smokers cope with slips, and preventing relapses.

Measures

At baseline, a questionnaire was performed before the intervention. The questionnaire consisted of the following measures:

Measures of demographic characteristics, which include age, sex, level of education, and health insurance. Measures related to smoking history, which include questions about the type of tobacco product used, the number of cigarettes smoked per day, use frequency, age of use initiation, attempts to quit, and methods to quit.

The Fagerström Test for Nicotine Dependence [33] is a standard instrument for assessing nicotine dependence. The test consists of 6 items with scores ranging from 0 to 10, which permit the classification of nicotine dependence into the following five levels: very low (0 to 2 points); low (3 to 4 points); moderate (5 points); high (6 to 7 points); and very high (8 to 10 points).

The Contemplation Ladder assesses the readiness to consider smoking cessation based on the individual's motivational stages for change [34]. It consists of a single question with a response range from 1 to 10; higher scores mean higher motivation.

The Patient Health Questionnaire-9 (PHQ-9) assesses the degree of depression severity through nine items directly based on the nine diagnostic criteria for major depressive disorder in the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition [35]. The final score ranges from 0 to 27 and can be classified into the following five levels: minimal depression (1 to 4 points); mild depression (5 to 9 points); moderate depression (10 to 14 points); moderately severe depression (15 to 19 points); and severe depression (20 to 27 points).

The Alcohol Use Disorders Identification Test is a 3-question screen that can help identify hazardous drinkers or those who have alcohol use disorders [36]. It is scored on a scale of 0-12 points. In men, a score of 4 points or more is considered positive for alcohol misuse; in women, a score of 3 points or more is considered positive.

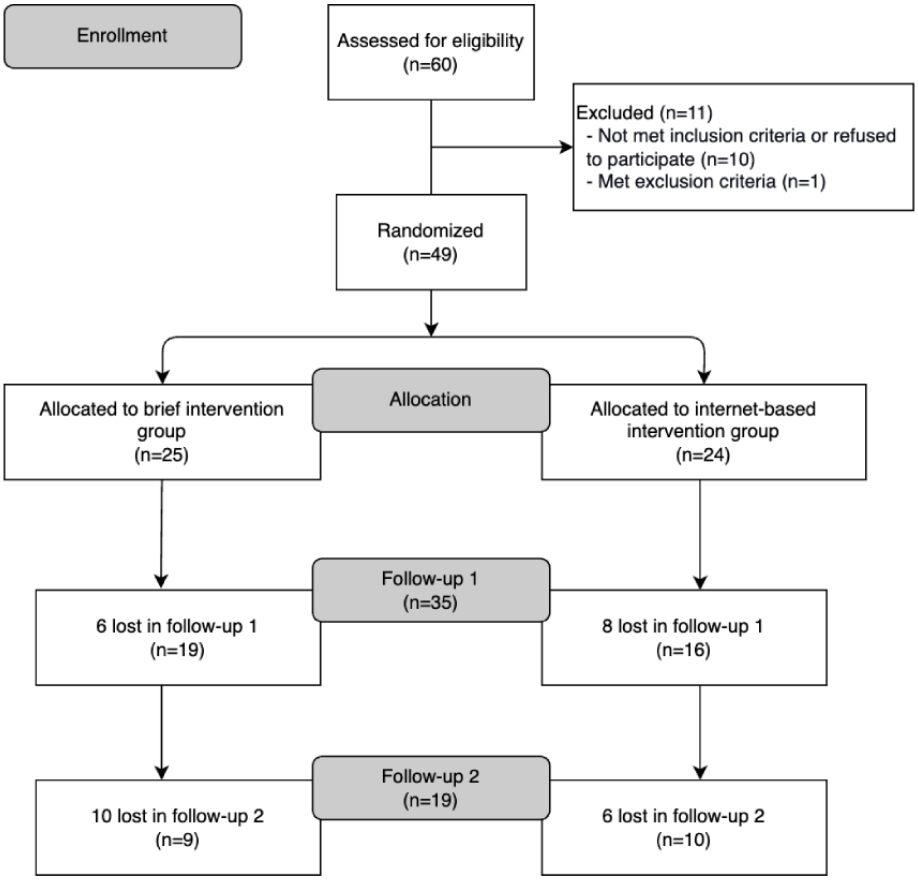
At the 30-day follow-up, participants were contacted by phone to know if their smoking status had changed after the intervention. Specifically, they were asked whether they stopped smoking or decreased the number of cigarettes smoked per day, and the Ladder scale was reassessed to compare motivation with the baseline. As a secondary outcome, it was also accessed if they sought intensive treatment for smoking cessation, which was recommended after the intervention. The participants were contacted again at 90 days after the intervention date to assess their abstinence, smoking status, and whether they had sought intensive treatment.

Procedures

After being assigned to one of the two groups, the participants received a brief intervention or were given a tablet to access the life without tobacco website. The researcher was present during access, and the intervention usage and adherence were similar.

Data collection and follow-ups occurred between August 2018 and May 2019. The participants were contacted by phone to fill in a follow-up questionnaire 1 and 3 months after intervention. This process and the final sample size are presented in the flowchart (Figure 1).

Figure 1. Flowchart of sample size and study procedures.



Data Analysis

The primary outcome was to assess the efficacy of the intervention through the cessation rate between groups, assessed 1 and 3 months after the intervention. Secondary outcomes involved comparing motivation between baseline and follow-ups and seeking an intensive smoking cessation treatment recommended after the intervention. Because of the possible bias the treatment-seeking individual could have in cessation, we also assessed the association between treatment seeking and cessation.

First, the normality Shapiro-Wilk test was performed, which found that the data distribution was not normal (P=.01). Consequently, nonparametric tests were performed for inferential analysis to compare the participants' characteristics between the two groups. The P value was generated using the Wilcoxon signed-rank test for continuous data and the Fisher exact 2-sided test for nominal variables. For the intention-to-treat analysis, a chi-square test was performed. All participants were included in the analysis. Those who did not respond to the follow-up were treated as smokers. A chi-square test was also used to determine the association between treatment seeking and abstinence. All analyses were performed through the 2020 R software (R Core Team) [37].

Results

Participants' Characteristics at Baseline

A total of 49 smokers were allocated to one of the two intervention groups, 25 (51%) in the BI group and 24 (49%) in the LWT. The participants were between 21-65 years old (mean age of 44.5, SD 13.3 years). Most of them were male (n=29, 59.2%) and had elementary school as the highest level of education (n=22, 44.9%). Regarding tobacco use history, they smoked an average of 14.5 (SD 8.6) cigarettes per day, and the mean score for nicotine dependence was 4.53, which means low to moderate dependence. Their motivation to quit was 5.75 on average. The number of cigarettes smoked per day (P=.002) and the dependence (P=.02) were statistically different between groups. The results are presented in Table 1.

Table 1. Differences between participants' characteristics in the two interventions groups at baseline (N=49).

Characteristics	BI[a] (n=25)	LWT[b] (n=24)	Total (n=49)	*P* value
Age, mean (SD)	42.9 (14.1)	46.1 (12.5)	44.5 (13.3)	.77
Sex, n (%)				.99
Male	15 (60)	14 (58)	29 (59)	
Female	10 (40)	10 (42)	20 (41)	
Education, n (%)				.28
Elementary	12 (48)	10 (42)	22 (45)	
High school	5 (20)	9 (38)	14 (29)	
College	3 (12)	4 (17)	7 (14)	
Graduation	5 (20)	1 (4)	6 (12)	
Health insurance, n (%)				.99
Public	14 (29)	13 (27)	27 (55)	
Private	11 (22)	11 (23)	22 (45)	
Type of tobacco product (cigarettes), n (%)	20 (41)	22 (45)	42 (86)	.47
Age of use initiation, mean (SD)	17.4 (3.1)	15.7 (3.0)	16.6 (3.1)	.44
Cigarettes per day, mean (SD)	12.4 (8.2)	16.7 (8.7)	14.5 (8.6)	.002
Frequency of use (daily), n (%)	23 (47)	23 (47)	46 (94)	.99
Attempt to quit (yes), n (%)	13 (27)	12 (25)	25 (51)	.92
Methods to quit, n (%)				.60
Counseling	3 (0.06)	2 (0.04)	5 (0.1)	
Nicotine replacement therapy	2 (0.04)	0 (0)	2 (0.04)	
Non-nicotine medications	1 (0.02)	0 (0)	1 (0.01)	
Combination of methods	2 (0.04)	3 (0.06)	5 (0.10)	
Dependence, mean (SD)	4.1 (2.3)	5.0 (1.8)	4.5 (2.1)	.02
Motivation to quit, mean (SD)	5.7 (1.7)	5.8 (2.1)	5.6 (1.9)	.79
Depression (PHQ-9[c]), mean (SD)	15.4 (5.9)	16.3 (5.3)	15.9 (5.64)	.67
Alcohol (AUDIT-C[d]), mean (SD)	4.2 (3.1)	4.9 (3.4)	4.6 (3.2)	.36

[a]BI: brief intervention.

[b]LWT: life without tobacco.

[c]PHQ-9: Patient Health Questionnaire-9.

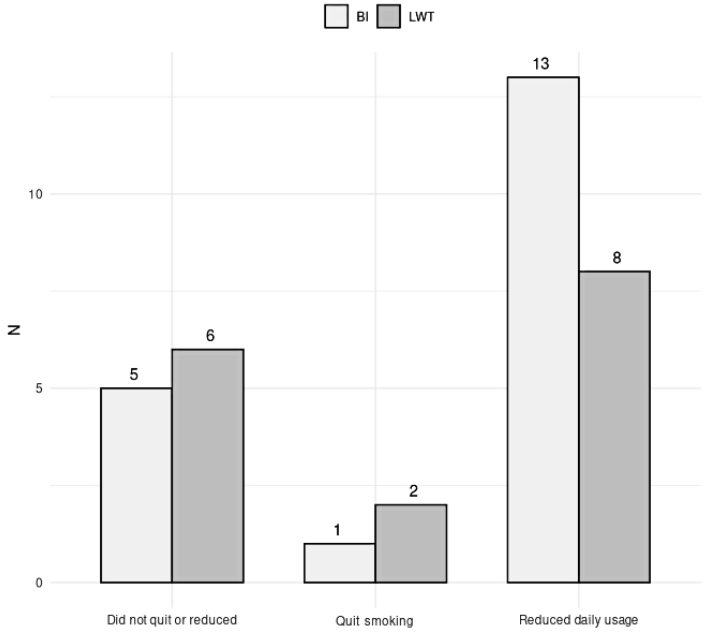
[d]AUDIT-C: Alcohol Use Disorders Identification Test.

Follow-up 1

A total of 35 (n=16, 46% LWT vs n=19, 54% BI) answered the first follow-up questionnaire (35/49, 71%). Of these 35 participants, 3 (9%) had stopped smoking (n=2, 6% LWT vs n=1, 3% BI), and 21 (63%) had reduced the number of cigarettes per day (n=8, 23% LWT vs n=13, 37% BI) by 55.5% on average. Moreover, 11 (31%) participants had not quit smoking (n=6, 17% LWT vs n=5, 14% BI). This difference was not statistically significant in both complete case analysis (χ^2_2=1.367, *P*=.50) and intention to treat analysis (χ^2_2=1.864, *P*=.39). The results of the first follow-up are presented in Figure 2.

There was a slight increase in the baseline average of 5.75 to 6.14 (SD 2.11) in the 30-day follow-up regarding motivation to quit smoking. Separated by group, the average score was 6.18 (SD 2.28) for the LWT group and 6.10 (SD 2.02) for BI (χ^2_9=9.479, *P*=.39). A total of 6 people reported seeking intensive treatment for smoking cessation after the intervention, 5 (83%) in the BI group versus 1 (17%) in the LWT group. However, this difference was not statistically significant (χ^2_1=1.221, *P*=.27). Regarding the association between seeking treatment and cessation, of the 3 people who reported quitting smoking, 1 (33%) had sought treatment, with no significant difference between groups (χ^2_2=0.462, *P*=.79).

Figure 2. Results from follow-up 1. BI: brief intervention; LWT: life without tobacco.

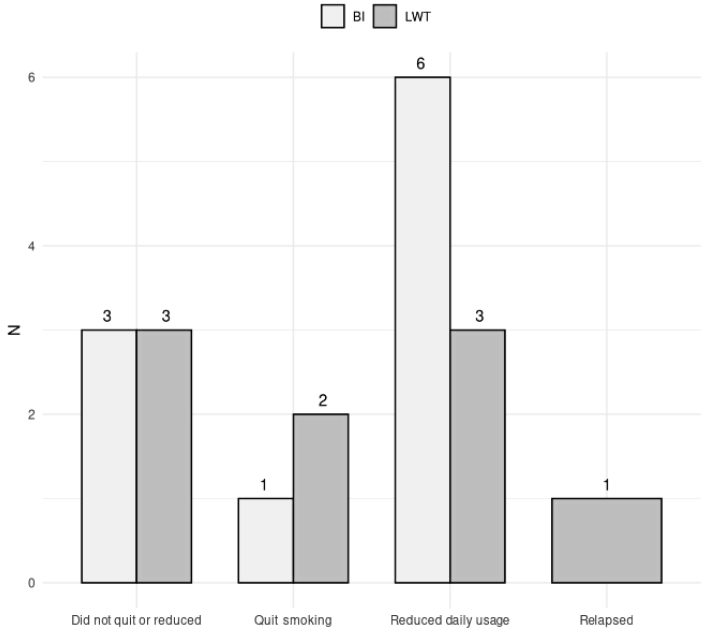


Follow-up 2

Three months after the intervention, 19 (19/49, 39%) participants (n=10, 53% LWT vs n=9, 47% BI) completed the second follow-up. A total of 9 people continued to decrease the number of cigarettes smoked daily (22% additional reduction on average; n=3, 33% LWT vs n=6, 67% BI); 6 people did not quit (n=3, 50% LWT vs n=3, 50% BI), and 1 participant relapsed and returned to smoking (LWT). Additionally, 3 people had quit smoking (n=2, 67% LWT vs n=1, 33% BI). This difference was not statistically significant (χ^2_3=2.287, *P*=.52). The results of the intention-to-treat analysis regarding cessation measures also did not demonstrate statistical significance between the intervention groups (χ^2_2=1.340, *P*=.51). Moreover, 3 participants reported seeking for intensive smoking cessation treatment (n=1, 33% LWT vs n=2, 67% BI). Of the 3 people who reported quitting smoking, no one had sought treatment, with no significant difference between groups (χ^2_3=2.1, *P*=.55). The results of the second follow-up are presented in Figure 3.

The results of the intention-to-treat analysis regarding cessation measures also did not demonstrate statistical significance between the intervention groups in both follow-ups (χ^2_2=1.864, *P*=.39 for follow-up 1; χ^2_2=1.340, *P*=.51 for follow-up 2).

Figure 3. Results from follow-up 2. BI: brief intervention; LWT: life without tobacco.



Discussion

This study showed similar rates of cessation and reduction for both BI and internet-based groups. According to the follow-up results, the internet-based intervention seems slightly more effective for smoking cessation. By contrast, the brief intervention was more effective in reducing the number of cigarettes smoked per day. However, because this difference was not statistically significant and the sample size is small, these results should be interpreted cautiously.

Characteristics of Sample

One result that differs from the literature [38,39] is that our sample included more men than women. Although the smoking prevalence is higher among men [40], women are more likely to seek smoking cessation approaches [41]. However, this result may be due to the higher prevalence of male workers where the intervention was carried out.

Compared to another study that assessed demographic characteristics of Brazilian smokers [42], smokers in this study were slightly younger (average 44 years old versus 49 years old). However, the age was consistent with another study that pointed out that younger smokers are more likely than older smokers to try to quit smoking [43]. They also smoked fewer cigarettes per day (an average of 14.5 cigarettes per day versus 20 cigarettes per day). As about half of the smokers in this sample have already tried to quit smoking, this lower average number of cigarettes per day may reflect these attempts, resulting in a decrease in daily consumption.

Most of the smokers in this sample have elementary school as the highest level of education, which is consistent with the smoking literature that points out the relation between lower levels of education and higher cigarette smoking rates [44,45].

Although almost all the variables were similar between the two groups, the number of cigarettes smoked per day, and the nicotine dependence differs, with the LWT group having a higher average of cigarettes per day compared with the BI group. The LWT group had a higher average of cigarettes per day than the BI group. The intensity of consumption is strongly associated with the level of nicotine addiction [46], which explains why the level of dependence was also higher in the LWT group. However, due to the small sample size, we could not control this difference in the baseline. As higher levels of nicotine dependence are associated with difficulties in quitting [47], participants in the LWT group may have encountered greater challenges in quitting smoking compared to the BI group. Although the difference in the cessation rate was not significant between the groups, this fact may have an advantage over the internet-based intervention, implying that the web-based intervention can be effective for heavy smokers with a high level of nicotine dependence.

Cessation Rates

This study has found cessation rates of 12% for the internet intervention and 5% through the brief advice after a 1-month follow-up. This was similar to other studies that showed quit rates ranging from 12.8% [21] to 14.3% [20] for interactive and tailored internet-based interventions. A literature review on internet-based interventions [48] also found high rates of smoking cessation when compared to the control group, with studies reporting cessation rates ranging from 1% to 42.8%, depending on the follow-up period. The internet-based intervention has demonstrated effectiveness comparable to other recommended forms of cessation treatment [19]. Given the low rate of smokers using traditional cessation methods [49], these results may imply an increase in the reach of smokers in quitting attempts and in the success rates.

Motivation to Quit Smoking

Regarding the secondary outcome of this study, the motivation score increased slightly in both groups after the intervention was received. Studies also reported that both BI [50] and internet-based intervention [51] increased the motivation score to stop smoking. Motivation is a fundamental prerequisite for a quit attempt [52], and to the contrary, lack of motivation is a fundamental barrier to engagement [53]; thus, both interventions are important tools to increase motivation, and both methods elucidate an attempt to stop smoking.

Seeking for Traditional Treatment to Stop Smoking

Finally, BI was more effective than the internet-based intervention in getting participants who seek intensive treatment to stop smoking, but this difference was not statistically significant. This is consistent with previous studies that found the brief intervention effective in achieving treatment referral for problem drinkers [54]. We found no association between seeking treatment and cessation. Because of the small frequency of participants, more robust analyzes could not be performed. Thus, future studies are necessary to confirm this finding.

This study has some strengths. First, this study evaluated two interventions for smoking cessation among Brazilians in a real-world context. People can use digital health interventions differently in real-world contexts compared to the conditions of the studies [25]. In this way, it is important to examine the use of digital interventions in real-world contexts. Furthermore, despite the small sample size, participation in this research was voluntary, and the participants did not receive any incentive. Besides that, the interventions had reasonable cessation rates, which is also a good indicator that these interventions can be effective in the real world.

This study also has some limitations. First, the sample size is relatively small; therefore, definitive conclusions about the effectiveness of such interventions cannot be made. Although the results are promising on the efficacy of these interventions, future studies should include a larger number of participants for more generalizable conclusions. Another common limitation in longitudinal studies is the decrease in the response rate over the follow-up. Although we made several attempts to contact the participants in this study, some patients were lost to follow-up, which biases the conclusion of the results.

Conclusion

Both interventions were effective in the cessation and reduction of cigarette consumption. This conclusion was based on the cessation rate results, 12% for the internet intervention and 5% through the brief advice after a 1-month follow-up. Although

the results need to be interpreted with caution as it is a pilot study, they point out that it is feasible to carry out a clinical study to measure the real impact of such interventions. In this matter, a larger trial will be necessary to better understand the effectiveness of these interventions for smoking cessation. Future investigations should also include longer follow-up periods to determine the long-term impact of internet-based interventions on smoking cessation.

Implications

Because smokers are not using traditional forms of smoking cessation, new and effective forms to address tobacco treatment are needed. This is the first study to evaluate a web-based intervention for smoking cessation in Brazil. Results showed good evidence of efficacy and pointed out that this intervention may help this population quit smoking. Future research is needed to evaluate long-term abstinence in this population.

Acknowledgments

This work was supported by the National Counsel of Technological and Scientific Development (CNPq); Coordination for the Improvement of Higher Education Personnel (CAPES); Foundation of Support Research of the State of Minas Gerais (FAPEMIG); and Federal University of Juiz de Fora (UFJF). The article processing charges related to the publication of this article were supported by The University of Kansas (KU) One University Open Access Author Fund sponsored jointly by the KU Provost, KU Vice Chancellor for Research & Graduate Studies, and KUMC Vice Chancellor for Research and managed jointly by the Libraries at the Medical Center and KU - Lawrence.AZ.

Conflicts of Interest

None declared.

References

1. Tobacco fact sheet. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/tobacco [accessed 2022-10-12]
2. National Center for Chronic Disease Prevention and Health Promotion (US) Office on Smoking and Health. The health consequences of smoking—50 years of progress: A report of the Surgeon General. Atlanta (GA): Centers for Disease Control and Prevention (US); 2014. URL: https://www.ncbi.nlm.nih.gov/books/NBK179276/
3. Observatório da Política Nacional de Controle do Tabaco. Instituto Nacional do Câncer (INCA). 2022. URL: https://www.inca.gov.br/observatorio-da-politica-nacional-de-controle-do-tabaco/dados-e-numeros-prevalencia-tabagismo [accessed 2022-10-12]
4. Palacios V, Pinto M, Barros L, Bardach A, Casarini A, Rodríguez Cairoli F, et al. A importância de aumentar os impostos do tabaco na Brasil. Instituto de Efectividad Clínica y Sanitaria 2020 12:epub [FREE Full text] [doi: 10.1590/0102-311x00129118]
5. Golechha M. Health Promotion Methods for Smoking Prevention and Cessation: A Comprehensive Review of Effectiveness and the Way Forward. Int J Prev Med 2016;7:7 [FREE Full text] [doi: 10.4103/2008-7802.173797] [Medline: 26941908]
6. Fiore FM, Jaén C, Baker T. Tobacco Use and Dependence Guideline Panel. Treating Tobacco Use and Dependence: 2008 Update. Rockville (MD), US: US Department of Health and Human Services; 2008.
7. WHO report on the global tobacco epidemic, 2008: The MPOWER package. World Health Organization. 2008. URL: https://apps.who.int/iris/bitstream/handle/10665/43818/9789241596282_eng.pdf?sequence=1&isAllowed=y [accessed 2022-10-11]
8. Stead LF, Bergson G, Lancaster T. Physician advice for smoking cessation. Cochrane Database Syst Rev 2008 Apr 16(2):CD000165. [doi: 10.1002/14651858.CD000165.pub3] [Medline: 18425860]
9. Miller WR, Rollnick S. Entrevista motivacional: preparando as pessoas para a mudança de comportamentos adictivos. Porto Alegre, Brasil: Artmed; 2001.
10. Butler CC, Rollnick S, Cohen D, Bachmann M, Russell I, Stott N. Motivational consulting versus brief advice for smokers in general practice: a randomized trial. Br J Gen Pract 1999;49(445):611-616.
11. Aveyard P, Begh R, Parsons A, West R. Brief opportunistic smoking cessation interventions: a systematic review and meta-analysis to compare advice to quit and offer of assistance. Addiction 2012 Jun;107(6):1066-1073. [doi: 10.1111/j.1360-0443.2011.03770.x] [Medline: 22175545]
12. Smoking: brief interventions and referrals. National Institute for Health and Clinical Excellence. URL: https://www.nice.org.uk/guidance/ph1 [accessed 2022-08-04]
13. Swartz LHG, Noell JW, Schroeder SW, Ary DV. A randomised control study of a fully automated internet based smoking cessation programme. Tob Control 2006 Feb;15(1):7-12 [FREE Full text] [doi: 10.1136/tc.2003.006189] [Medline: 16436397]
14. Marks IM, Cavanagh K, Gega L. Computer-aided psychotherapy: revolution or bubble? Br J Psychiatry 2007 Dec;191:471-473. [doi: 10.1192/bjp.bp.107.041152] [Medline: 18055948]
15. Murray E. Internet-delivered treatments for long-term conditions: strategies, efficiency and cost-effectiveness. Expert Rev Pharmacoecon Outcomes Res 2008 Jun;8(3):261-272. [doi: 10.1586/14737167.8.3.261] [Medline: 20528378]

16. Fleming TM, de Beurs D, Khazaal Y, Gaggioli A, Riva G, Botella C, et al. Maximizing the Impact of e-Therapy and Serious Gaming: Time for a Paradigm Shift. Front Psychiatry 2016;7:65. [doi: 10.3389/fpsyt.2016.00065] [Medline: 27148094]
17. Richards D, Richardson T. Computer-based psychological treatments for depression: a systematic review and meta-analysis. Clin Psychol Rev 2012 Jun;32(4):329-342. [doi: 10.1016/j.cpr.2012.02.004] [Medline: 22466510]
18. Rogers MA, Lemmen K, Kramer R, Mann J, Chopra V. Internet-Delivered Health Interventions That Work: Systematic Review of Meta-Analyses and Evaluation of Website Availability. J Med Internet Res 2017 Mar 24;19(3):e90 [FREE Full text] [doi: 10.2196/jmir.7111] [Medline: 28341617]
19. Graham AL, Amato MS. Twelve Million Smokers Look Online for Smoking Cessation Help Annually: Health Information National Trends Survey Data, 2005-2017. Nicotine Tob Res 2019 Jan 04;21(2):249-252 [FREE Full text] [doi: 10.1093/ntr/nty043] [Medline: 29660037]
20. Graham A, Carpenter K, Cha S, Cole S, Jacobs M, Raskob M, et al. Systematic review and meta-analysis of Internet interventions for smoking cessation among adults. SAR 2016 May:55. [doi: 10.2147/sar.s101660]
21. Taylor GMJ, Dalili MN, Semwal M, Civljak M, Sheikh A, Car J. Internet-based interventions for smoking cessation. Cochrane Database Syst Rev 2017 Sep 04;9:CD007078 [FREE Full text] [doi: 10.1002/14651858.CD007078.pub5] [Medline: 28869775]
22. Graham AL, Papandonatos GD, Cha S, Erar B, Amato MS, Cobb NK, et al. Improving Adherence to Smoking Cessation Treatment: Intervention Effects in a Web-Based Randomized Trial. Nicotine Tob Res 2017 Mar 01;19(3):324-332. [doi: 10.1093/ntr/ntw282] [Medline: 27836982]
23. Myung S, McDonnell DD, Kazinets G, Seo HG, Moskowitz JM. Effects of Web- and computer-based smoking cessation programs: meta-analysis of randomized controlled trials. Arch Intern Med 2009 May 25;169(10):929-937. [doi: 10.1001/archinternmed.2009.109] [Medline: 19468084]
24. Poppelaars M, Tak YR, Lichtwarck-Aschoff A, Engels RCME, Lobel A, Merry SN, et al. A randomized controlled trial comparing two cognitive-behavioral programs for adolescent girls with subclinical depression: A school-based program (Op Volle Kracht) and a computerized program (SPARX). Behav Res Ther 2016 May;80:33-42. [doi: 10.1016/j.brat.2016.03.005] [Medline: 27019280]
25. Fleming T, Bavin L, Lucassen M, Stasiak K, Hopkins S, Merry S. Beyond the Trial: Systematic Review of Real-World Uptake and Engagement With Digital Self-Help Interventions for Depression, Low Mood, or Anxiety. J Med Internet Res 2018 Jun 06;20(6):e199 [FREE Full text] [doi: 10.2196/jmir.9275] [Medline: 29875089]
26. Julious SA. Sample size of 12 per group rule of thumb for a pilot study. Pharmaceut. Statist 2005 Oct;4(4):287-291. [doi: 10.1002/pst.185]
27. Browne RH. On the use of a pilot sample for sample size determination. Stat Med 1995 Sep 15;14(17):1933-1940. [doi: 10.1002/sim.4780141709] [Medline: 8532986]
28. Teare MD, Dimairo M, Shephard N, Hayman A, Whitehead A, Walters SJ. Sample size requirements to estimate key design parameters from external pilot randomised controlled trials: a simulation study. Trials 2014 Jul 03;15:264 [FREE Full text] [doi: 10.1186/1745-6215-15-264] [Medline: 24993581]
29. Gomide HP, Bernardino HS, Richter K, Martins LF, Ronzani TM. Development of an open-source web-based intervention for Brazilian smokers - Viva sem Tabaco. BMC Med Inform Decis Mak 2016 Aug 02;16(1):103 [FREE Full text] [doi: 10.1186/s12911-016-0339-7] [Medline: 27485018]
30. De Micheli D, Formigoni MOS, Carneiro APL. SUPERA Sistema para detecção do Uso abusivo e dependência de substâncias Psicoativas: Encaminhamento, Intervenção breve, Reinserção social e Acompanhamento. Secretaria Nacional de Políticas sobre Drogas. 2017. URL: https://www.supera.org.br/wp-content/uploads/2016/06/SUP9_Guia.pdf [accessed 2022-10-12]
31. Miller WR, Sanches VC. Motivating young adults for treatment and lifestyle change. In: Issues in Alcohol Use and Misuse in Young Adults. Notre Dame, IN, US: University of Notre Dame Press; 1994:55-81.
32. 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]
33. Heatherton TF, Kozlowski LT, Frecker RC, Fagerström KO. The Fagerström Test for Nicotine Dependence: a revision of the Fagerström Tolerance Questionnaire. Br J Addict 1991 Sep;86(9):1119-1127. [doi: 10.1111/j.1360-0443.1991.tb01879.x] [Medline: 1932883]
34. Biener L, Abrams DB. The Contemplation Ladder: validation of a measure of readiness to consider smoking cessation. Health Psychol 1991;10(5):360-365. [doi: 10.1037//0278-6133.10.5.360] [Medline: 1935872]
35. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
36. Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. Arch Intern Med 1998 Sep 14;158(16):1789-1795. [doi: 10.1001/archinte.158.16.1789] [Medline: 9738608]
37. The R Project for Statistical Computing. URL: https://www.r-project.org/ [accessed 2022-10-12]
38. Balmford J, Borland R, Benda P. Patterns of use of an automated interactive personalized coaching program for smoking cessation. J Med Internet Res 2008 Dec 17;10(5):e54 [FREE Full text] [doi: 10.2196/jmir.1016] [Medline: 19097975]

39. Prout MN, Martinez O, Ballas J, Geller AC, Lash TL, Brooks D, et al. Who uses the Smoker's Quitline in Massachusetts? Tob Control 2002 Jun;11 Suppl 2:ii74-ii75 [FREE Full text] [doi: 10.1136/tc.11.suppl_2.ii74] [Medline: 12034987]
40. Drope J, Liber AC, Cahn Z, Stoklosa M, Kennedy R, Douglas CE, et al. Who's still smoking? Disparities in adult cigarette smoking prevalence in the United States. CA Cancer J Clin 2018 Mar;68(2):106-115 [FREE Full text] [doi: 10.3322/caac.21444] [Medline: 29384589]
41. Zhu S, Melcer T, Sun J, Rosbrook B, Pierce JP. Smoking cessation with and without assistance. American Journal of Preventive Medicine 2000 May;18(4):305-311. [doi: 10.1016/s0749-3797(00)00124-0]
42. Rodrigues NC, Neves RD, Mendes DDS, Mendes CP, Martins IH, Reis IN, et al. Profile of Brazilian smokers in the National Program for Tobacco Control. Braz J Psychiatry 2015;37(2):150-154. [doi: 10.1590/1516-4446-2014-1611] [Medline: 26083814]
43. Curry SJ, Sporer AK, Pugach O, Campbell RT, Emery S. Use of Tobacco Cessation Treatments Among Young Adult Smokers: 2005 National Health Interview Survey. Am J Public Health 2007 Aug;97(8):1464-1469. [doi: 10.2105/ajph.2006.103788]
44. 2016 National Survey on Drug Use and Health. Center for Behavioral Health Statistics and Quality. URL: https://www.samhsa.gov/data/sites/default/files/NSDUH-DetTabs-2016/NSDUH-DetTabs-2016.pdf [accessed 2022-10-12]
45. Warren GW, Alberg AJ, Kraft AS, Cummings KM. The 2014 Surgeon General's report: "The health consequences of smoking--50 years of progress": a paradigm shift in cancer care. Cancer 2014 Jul 01;120(13):1914-1916 [FREE Full text] [doi: 10.1002/cncr.28695] [Medline: 24687615]
46. Shiffman S, Ferguson SG, Dunbar MS, Scholl SM. Tobacco dependence among intermittent smokers. Nicotine Tob Res 2012 Nov 22;14(11):1372-1381 [FREE Full text] [doi: 10.1093/ntr/nts097] [Medline: 22529224]
47. Tanner JA, Chenoweth M, Tyndale R. Pharmacogenetics of nicotine and associated smoking behaviors. Curr Top Behav Neurosci 2015;23:37-86. [doi: 10.1007/978-3-319-13665-3_3] [Medline: 25655887]
48. Park E, Drake E. Systematic review: internet-based program for youth smoking prevention and cessation. J Nurs Scholarsh 2015 Jan;47(1):43-50. [doi: 10.1111/jnu.12104] [Medline: 25130325]
49. Caraballo RS, Shafer PR, Patel D, Davis KC, McAfee TA. Quit Methods Used by US Adult Cigarette Smokers, 2014-2016. Prev Chronic Dis 2017 Apr 13;14:E32 [FREE Full text] [doi: 10.5888/pcd14.160600] [Medline: 28409740]
50. Rohsenow DJ, Martin RA, Monti PM, Colby SM, Day AM, Abrams DB, et al. Motivational interviewing versus brief advice for cigarette smokers in residential alcohol treatment. J Subst Abuse Treat 2014 Mar;46(3):346-355 [FREE Full text] [doi: 10.1016/j.jsat.2013.10.002] [Medline: 24210533]
51. Mavrot C, Stucki I, Sager F, Etter J. Efficacy of an Internet-based, individually tailored smoking cessation program: A randomized-controlled trial. J Telemed Telecare 2016 Jun 30;23(5):521-528. [doi: 10.1177/1357633x16655476]
52. Buczkowski K, Marcinowicz L, Czachowski S, Piszczek E. Motivations toward smoking cessation, reasons for relapse, and modes of quitting: results from a qualitative study among former and current smokers. PPA 2014 Oct:1353. [doi: 10.2147/ppa.s67767]
53. Simblett S, Greer B, Matcham F, Curtis H, Polhemus A, Ferrão J, et al. Barriers to and Facilitators of Engagement With Remote Measurement Technology for Managing Health: Systematic Review and Content Analysis of Findings. J Med Internet Res 2018 Jul 12;20(7):e10480 [FREE Full text] [doi: 10.2196/10480] [Medline: 30001997]
54. Bien TH, Miller WR, Tonigan JS. Brief interventions for alcohol problems: a review. Addiction 1993 Mar;88(3):315-335. [doi: 10.1111/j.1360-0443.1993.tb00820.x] [Medline: 8461850]

Abbreviations

BI: brief intervention
LWT: life without tobacco

Edited by A Mavragani; submitted 10.05.21; peer-reviewed by S Luo, F Tzelepis; comments to author 07.06.21; revised version received 03.08.21; accepted 14.05.22; published 03.11.22.

Please cite as:
Machado NM, Gomide HP, Bernardino HS, Ronzani TM
Internet-Based Intervention Compared to Brief Intervention for Smoking Cessation in Brazil: Pilot Study
JMIR Form Res 2022;6(11):e30327
URL: https://formative.jmir.org/2022/11/e30327
doi:10.2196/30327
PMID:36326817

Original Paper

Use of a Rapid Qualitative Method to Inform the Development of a Text Messaging Intervention for People With Serious Mental Illness Who Smoke: Formative Research Study

Catherine S Nagawa[1], PhD; Ian A Lane[2], BA, RN; Colleen E McKay[3], MA, CAGS; Ariana Kamberi[1], MBA; Lisa L Shenette[1], MSc, RD, LDN; Megan M Kelly[4], PhD; Maryann Davis[3], PhD; Rajani S Sadasivam[1,3], PhD

[1]Department of Population and Quantitative Health Sciences, UMass Chan Medical School, Worcester, MA, United States

[2]Tan Chingfen Graduate School of Nursing, UMass Chan Medical School, Worcester, MA, United States

[3]Department of Psychiatry, UMass Chan Medical School, Worcester, MA, United States

[4]VISN 1 New England Mental Illness Research, Education, and Clinical Center, VA Bedford Healthcare System, Bedford, MA, United States

Corresponding Author:
Catherine S Nagawa, PhD
Department of Population and Quantitative Health Sciences
UMass Chan Medical School
55 Lake Avenue North
Worcester, MA, 01655
United States
Phone: 1 508 856 8999
Email: cnagawa@mgh.harvard.edu

Abstract

Background: People with serious mental illness are disproportionately affected by smoking and face barriers to accessing smoking cessation treatments in mental health treatment settings. Text-based interventions are cost-effective and represent a widely accessible approach to providing smoking cessation support.

Objective: We aimed to identify key factors for adapting text-based cessation interventions for people with serious mental illness who smoke.

Methods: We recruited 24 adults from mental health programs who had a serious mental illness and currently smoked cigarettes or had quit smoking within the past 5 years. We then conducted virtual qualitative interviews between November 2020 and August 2021. Data were analyzed using the rapid thematic analytic approach.

Results: We identified the following 3 major themes: (1) interplay between smoking and having a serious mental illness, (2) social contextual factors of smoking in adults with serious mental illness, and (3) smoking and quitting behaviors similar to the general population. Participants reported barriers and facilitators to quitting across the 3 themes. Within the "interplay between smoking and having a serious mental illness" theme, barriers included smoking to manage stress and mental health symptoms, and facilitators to quitting included the awareness of the harm of smoking on mental health and patient-provider discussions on smoking and mental health. In the "social contextual factors of smoking in adults with serious mental illness" theme, barriers included high social acceptability of smoking among peers. Positive support and the combined social stigma of smoking and having a mental health condition outside of peer groups motivated individuals to quit. Some participants indicated that low exposure to other smokers during the COVID-19 pandemic helped them to engage in cessation efforts. In the "smoking and quitting behaviors similar to the general population" theme, barriers included smoking after eating, having coffee, drinking alcohol, and experiencing negative social support, and facilitators included health concerns, improvement in the general quality of life, and use of evidence-based tobacco treatments when available.

Conclusions: People with serious mental illness often smoke to cope with intense emotional states, manage mental health symptoms, or maintain social bonds. Text message content emphasizing equally effective and less harmful ways for stress reduction and mental health symptom management may improve quit rates in individuals with serious mental illness.

(JMIR Form Res 2022;6(11):e40907) doi:10.2196/40907

KEYWORDS

serious mental illness; mental disorder; psychiatric disorder; tobacco use; smoking cessation; text messaging; intervention; smoking; mental health; virtual; COVID-19; pandemic; symptom

Introduction

People with serious mental illness report a higher smoking prevalence (41%) than people without serious mental illness (14%) [1,2]. Unfortunately, most people with serious mental illness who smoke have little or no access to pharmacotherapeutic cessation aids and behavioral approaches, which are the recommended treatments for smoking cessation [3]. While mental health treatment settings offer a possibility for engaging people with serious mental illness in smoking cessation, only a minority of them are currently receiving treatment [4]. In 2019, 46% of the 52.9 million US adults with any mental illness had received mental health services in the past year [4]. Further, many mental health settings do not routinely offer tobacco treatment [5]. Alternative approaches to engage and treat people with serious mental illness for smoking cessation are needed.

Text-based interventions are cost-effective and eliminate many barriers to accessing traditional treatments [6,7]. Ninety percent of individuals with a mental health diagnosis own more than one mobile device, including a mobile phone, and many report high use of multiple devices [8]. In the general population, text-based cessation interventions significantly improve quit rates and use of telephone counseling (ie, Quitline) services [9]. Studies have shown the feasibility of text-based cessation interventions in people with serious mental illness who smoke [10-12]. However, text-based cessation interventions need to address barriers to quitting that are specific to people with serious mental illness.

In the literature, many factors have been noted for increased tobacco use rates among people with serious mental illness. For instance, people with mental health conditions report using cigarettes to alleviate emotional problems, such as depression and anxiety, stabilize their mood, and relieve stress [13-15]. Many of these factors are shared with the general population of people who smoke. However, studies have not fully explored how these factors and mental health conditions intersect to increase tobacco use rates and quitting challenges in people with mental illness. Adapting text-based cessation interventions to address this intersection directly may increase their relevance and effectiveness, and improve engagement among people with serious mental illness who smoke [16].

The effectiveness of the text-based messaging intervention being adapted has been demonstrated in a large randomized trial conducted among individuals without serious mental illness [17,18]. Our intervention included use of expert and peer-written messages [19]. We developed the expert-written messages through an iterative group review process guided by theoretical frameworks and existing smoking cessation guidelines [3]. Peer-written messages were written by current and former smokers responding to an online survey. The content of expert messages was more "biomedical" in nature (avoidance, behavioral strategies, and health), while the content of peer messages focused on "social" and "real-life" aspects of smoking (expectations, money, quality of life, attitudes, and friends) [19]. The intervention group received motivational text messages weekly for 6 months, and compared with the control group, receiving text messages increased the odds of 7-day point prevalence cessation at 6 months (odds ratio 1.69, 95% CI 1.03-2.8) [17,18]. However, these messages do not address barriers associated with serious mental illness.

In this study, we explored tobacco use behaviors and barriers along with facilitators of quitting among 24 people with serious mental illness to identify key factors for adapting a text messaging intervention that addresses their cessation needs. We also detailed a 4-step approach to identify themes from the qualitative interviews that could be a structured process for future text-based intervention developers.

Methods

Study Overview

We conducted an in-depth qualitative exploration of smoking behaviors, barriers, and facilitators among adults with serious mental illness who currently smoke or quit within the past 5 years.

We used a 4-step approach to identify key themes that will inform the development of text messages targeted at people with mental health conditions. Briefly, the 4-step approach included identifying domains that aligned with key interview questions, summarizing transcript data under each domain, consolidating the data into a matrix, and interactively identifying themes under each domain. We provide details in the data analysis section.

Study Setting and Participant Recruitment

Participants were recruited nationally from 3 mental health agencies, most of which are publicly funded (Clubhouses accredited by the International Center for Clubhouse Development [20], Thresholds Agency, and Massachusetts Department of Mental Health agencies such as Open Sky). These mental health programs provide services to individuals with various mental health conditions. Services include the opportunity for friendship, employment, housing, education, and access to medical and psychiatric services. Clubhouses are accessed by about 100,000 people with mental health conditions [21]. Thresholds is a large community mental health center in Chicago that serves 12,500 adults and youth each year. The Massachusetts Department of Mental Health programs and Clubhouses affiliated with Clubhouse International provide various clinical and nonclinical services to adults with serious mental illness.

We recruited 24 adults with serious mental illness through virtual information sessions administered by a research coordinator. Individuals were eligible to participate if they (1) were 18 years or older, (2) were currently smoking cigarettes

or had quit within the past 5 years, (3) were willing and able to provide consent, and (4) were currently receiving services from mental health agencies (ie, Clubhouses, Thresholds Agency, and Massachusetts Department of Mental Health agencies such as Open Sky). We excluded pregnant individuals or those attending correctional programs (ie, prisoners).

Ethical Considerations

Potential participants provided contact information to the research coordinator, who set up one-on-one virtual meetings to provide more information about the study. Those interested in participating provided informed electronic consent (e-consent) before enrollment into the study. Electronic signatures were captured using the Research Electronic Data Capture e-consent framework. Participants received a US $15 Amazon electronic gift card for the study. This project was approved by the ethics review boards at the Department of Mental Health in Massachusetts and the University of Massachusetts Chan Medical School (reference number: H00019687_3).

Interview Guide Development

The interview guide was developed by researchers with expertise in mental health and tobacco use based on prior literature [22-24]. The interview guide was structured using domains that aligned with the interview questions and the overall study objective of developing text messages for people with serious mental illness. We chose the approach outlined in *Rapid qualitative inquiry: A field guide to team-based assessment* by James Beebe [25] to generate the domains explored in our study. Informed by the literature, the research team identified the key factors that would be the most meaningful in informing the implementation aspects of a smoking cessation study. Such factors included the context in which individuals with serious mental illness smoked, barriers and facilitators to quitting, and how one's mental health diagnosis may influence their smoking habits. These factors guided the formulation of the key research questions/interview guide, which shaped our study domains.

The final interview questions broadly focused on motivations, beliefs, and barriers and facilitators to smoking or quitting. We also asked about smoking triggers or cues; symptoms of tobacco dependence; previous experiences reducing or stopping smoking, including use of and barriers to evidence-based cessation; and the role of interpersonal relationships in smoking cessation. During the interviews, we prompted participants to reflect on how their diagnosis has influenced their smoking behaviors and quitting experiences. The specified 9 domains that aligned with these questions included the context in which participants smoked, smoking triggers, reasons for smoking, barriers to active quitting, facilitators of active quitting, mental health–specific quitting strategies, mental health–specific facilitators of quitting maintenance, and the influence of the COVID-19 pandemic on smoking.

We refined the interview guide based on feedback from a stakeholder advisory panel of 9 adults with serious mental illness who also currently smoked or had quit within the past 5 years. Panel members were recruited from the same mental health agencies as study participants to ensure the representativeness of the study population.

Data Collection

Two trained researchers conducted the qualitative interviews using a semistructured interview guide via Zoom video conference software (Zoom Video Communications Inc). Interviews were conducted between November 2020 and August 2021, and lasted for 30 to 40 minutes. All interviews were audio recorded, deidentified, and professionally transcribed. In addition to the qualitative interview, participants completed a baseline questionnaire that captured demographic information (eg, age, sex, race/ethnicity, and educational level) and smoking behavior or history of those who recently quit.

Data Analysis

We used the rapid thematic qualitative analytic approach to analyze qualitative data. This rapid approach is a team-based qualitative inquiry that uses triangulation and iterative data analysis to quickly develop a preliminary understanding of a situation from the insider's perspective [25-27]. This approach to data analysis provided insights into the context in which participants smoked and their attitudes and beliefs toward quitting smoking.

The 4-step process is as follows:

1. We used the domains developed in the interview development step to create a transcript summary template in Microsoft Word (Microsoft Corp). The summary template was structured so that each section on the template corresponded to a domain. The summary template guided the analysis team in mapping each participant's responses to the corresponding domain.
2. We performed a "test drive" to assess the domains' usability, relevance, and consistency. In this step, 2 coders (CSN and IAL) used the template to summarize the same transcript to evaluate whether the specified domains were identifiable in the data and to check for consistency across coders in capturing the domains. We modified the template based on feedback from the rest of the research team (RSS, MD, AK, and CEM) before testing with a second transcript. Once consistency was established, we divided the transcripts equally across the 2 coders and summarized the transcripts using the modified template.
3. We transferred transcript summaries into a matrix. We placed information from the 24 transcript summaries into a Microsoft Excel matrix. The transcript summaries were put into a matrix to analyze each domain's depth and breadth of data [25]. Each column in the Excel document represented a prespecified domain, and the rows represented study participants. The cells in the Excel document contained summarized responses from each study participant that corresponded to a specific domain.
4. We identified study themes and subthemes using the matrix. Themes were identified within each domain. The research team collaboratively and iteratively reviewed, discussed, and sorted the data to refine the initial themes and subthemes, and highlight the most salient quotes. The research team members provided multiple perspectives; a triangulation (investigator triangulation [28]) approach was designed to increase the findings' reliability. In addition to investigator triangulation, we used the peer-review approach

[29], in which a researcher (MMK) who was not involved in data collection reviewed the evidence that supported the interpretation of the data and conclusions. The final themes highlight the views and experiences with using family or peer support for smoking cessation and considerations for support of cessation interventions. A brief descriptor (subtheme) of what participants reported is provided within each theme. Each of these themes will inform the text message content.

Results

Participant Characteristics

Participant characteristics are presented in Table 1. Most participants were between 36 and 54 years old (11/24, 46%) and male (15/24, 63%). About half self-identified as non-Hispanic White (13/24, 54%) and had a high school education (13/24, 54%). More than half (10/17, 59%) of the participants who were currently smoking had made a quit attempt in the past year. A high proportion (20/24, 83%) of study participants owned a cell phone or smartphone, and a majority (19/20, 95%) of those who owned a cell phone or smartphone often used it to send and receive text messages (Table 1).

Table 1. Sociodemographic characteristics of the study participants.

Participant characteristic	Value (N=24), n (%)
Age (years)	
18-34	9 (37%)
36-54	11 (46%)
55-65	4 (17%)
Gender	
Male	15 (63%)
Female	8 (33%)
Nonbinary	1 (4%)
Race and ethnicity	
Non-Hispanic White	13 (54%)
Non-Hispanic Black	5 (21%)
Other	6 (25%)
Education level	
Some high school/high school graduate	13 (54%)
How hard is it for you (and your family) to pay for medical care?	
Very hard	3 (12%)
Somewhat hard	4 (17%)
Not very hard	16 (67%)
Don’t know	1 (4%)
Do you take medication for your mental health?	
Yes	23 (96%)
Do you currently smoke?	
Yes	17 (71%)
Past 12 months quit attempts among those who currently smoke (N=17)	10 (59%)
Past 30-day electronic cigarette use	
Every day or some days	6 (25%)
Not at all	18 (75%)
Is smoking allowed where you live?	
Yes	20 (83%)
Besides yourself, does anyone who lives with you smoke cigarettes now?[a]	
No	16 (70%)
Yes	7 (30%)
Do you own a cell phone or smartphone?	
Yes	20 (83%)
Do you ever use your cell phone to send or receive text messages? (N=20)	
Yes	19 (95%)

[a]Information is missing for 1 participant.

Factors for Adapting Text-Based Interventions for People With Serious Mental Illness Who Smoke

We found 3 major themes that provided insights into the experiences of quitting smoking in people with serious mental illness. The themes included interplay between smoking and having a serious mental illness, social context of smoking, and similarities in smoking and quitting behaviors between the participants and the general population (Table 2). There were barriers and facilitators within each theme, as presented below.

Table 2. Factors for adapting text-based interventions for people with serious mental illness who smoke.

Variable	Theme 1: Interplay between smoking and having a serious mental illness	Theme 2: Social contextual factors of smoking in adults with serious mental illness	Theme 3: Smoking and quitting behaviors similar to the general population
Smoking patterns	Smoking occurs when anxiety, stress, depression, or nervousness is heightened	Individuals tend to smoke with friends and at social events (or parties)	Smoking is habitual; usually done after eating or having coffee, or when drinking alcohol
Reasons for smoking	Smoking is perceived to improve mood and reduce anxiety	Smoking is perceived to help with social interactions or be part of social activities	Stress management
Motivations to quit smoking	Negative impact on mental health	The social stigma associated with being a smoker and having a mental health condition	Motivated to quit for esthetic reasons Negative impact on physical health Financial cost of smoking
Challenges of quitting	Ways to manage mental health when quitting	Socially acceptable or rewarding to smoke	Negative or counter-productive support Boredom and cigarette sharing
Facilitators of quitting	Mental health care providers	Positive social support	Use of evidence-based cessation approaches
Impact of the pandemic on quitting	Smoking continued in part as a coping response to the stress/anxiety of the pandemic	Lower levels of exposure to other individuals who smoke facilitated cessation	The pandemic did not add to the motivation to quit for those who perceived their risk of COVID-19 to be low

Theme 1: Interplay Between Smoking and Having a Serious Mental Illness

Barriers

Individuals often smoked when mental health symptoms (such as feelings of anxiety, stress, or nervousness) were heightened, and the perception was that smoking helped with these symptoms.

Participants said:

> *Because I know that when I'm having a bipolar episode, and I started having voice issues and more paranoia and stuff like that, something about having a cigarette just calms it down.* [Female, 45-54 years old, currently smokes]

> *Anytime, like anytime someone raised their voice with me. Anytime I thought I might be in trouble. Anytime I had to get something done, but time was short. Anytime I had to make a phone call. Phones always used to cause me anxiety. Anytime I had to… Gosh, the exhaustion of having to make dinner. Oh, my gosh. I better just go smoke, you know?* [Female, 35-44 years old, recently quit]

One participant who found smoking helpful in managing mental health symptoms said:

> *I mean any time that I get depressed, I'm able to like to go out and have a cigarette. And it just seems like as I'm smoking, like all my worries just go away because it's letting me calm down rather than getting worse.* [Female, 21-23 years old, currently smokes]

Smoking continued in part as a coping response to the stress and anxiety induced by the COVID-19 pandemic.

> *When COVID hit, it just made it worse. I wanted to drink. I'm trying to drown my sorrows away. Like I'm depressed. I'm sad for the people that's like literally dying. They didn't… You know what I'm saying? They didn't think it was going to be their time yet. It was just so much death. Everybody was getting sick. It made it intensify. Like, oh, my God. I gotta smoke. I gotta have a drink. I gotta smoke. You know what I'm saying? Like I gotta have a cigarette. I'm stressed out. The world's stressed out. The world's goin' crazy. The world's ending! Like it has affected me. I'm not gonna lie.* [Female, 30-34 years old, currently smokes]

Facilitators

Participants were motivated to stop smoking due to the awareness of the negative impact of smoking on their mental health. Specifically, 2 participants said:

> *So, it wrecks my health, because my mental health and my physical health are very, very intertwined. So, I turn into an anxious monster because the moment I extinguish a cigarette, my body goes into withdrawal, and I just want another cigarette. And then I'm looking and hunting for that, which for me produces a lot of anxiety.* [Female, 35-44 years old, recently quit]

> *I think sometimes it makes [mental health diagnosis] worse. Yeah, the smoking makes like my anxiety worse sometimes, also. cause it gets my heart goin'. And then if I can't like to control my breathing, but then sometimes it's also the opposite of where it just takes me away from everything.* [Male, 24-26 years old, currently smokes]

Although managing mental health symptoms during cessation was a significant barrier to successful quitting, participants found that speaking with their mental health care provider helped them navigate the quitting process better.

> *I even tried Chantix, which my psychiatrist was very wary about prescribing me because of my history with suicidality. But I was, like, in a really good place now. And you know, I'm always honest about my feelings. So, can we give it a try? And then if it does make my symptoms worse, we'll just take me off of it right away. And I didn't even last seven days on it before I started feeling suicidal. So, I couldn't even take Chantix.* [Female, 45-54 years old, currently smokes]

> *I know at one point we tried Wellbutrin for both mental health and smoking, and that was one I don't think had any effect on me. I am currently waiting to see if quitting smoking has an effect on my sleep, which would then have an effect on my psychiatric sleep medication. But we haven't quite got there yet because I have to get off some of my Chantix to be able to see that. So that's the big one that we're waiting to see if my sleep gets better now that I'm not smoking because that smoking causes some bad sleep apnea.* [Female, 35-44 years old, recently quit]

Theme 2: Social Contextual Factors of Smoking in Adults With Serious Mental Illness

The social context played a vital role in the smoking and quitting behaviors of people with serious mental illness.

Barriers

Smoking seemed socially acceptable among peers and was viewed as a social bonding activity.

Individuals often smoked during social gatherings. Participants said:

> *Yeah, when you go out there, there's like two or three picnic tables in the winter or the summer, and there's like maybe 10 smokers or 15 smokers. So, you join in with the crowd and you don't feel left out. I feel more blended in with them because they smoke.* [Male, 55-65 years old, recently quit]

> *If I'm trying to have a conversation, I feel a little uneasy without a cigarette, but I'm getting much better about that. But it used to be, if I am, you know, out on a patio with friends, it doesn't feel right unless I'm smoking. I'm getting better with that. But yeah, social situations. A cigarette always makes you feel like you've got your best friend there.* [Female, 35-44 years old, currently smokes]

Smoking seemed socially rewarding. Some individuals used smoking to feel socially included or make friends. Participants said:

> *Obviously, smoking cigarettes can start friendships or bring people together in the first place because you might be out somewhere. Say you're going to a bar or something and you go outside, and everybody's had a few, and then you smoke a cigarette. You might get to know somebody you don't.* [Male, 35-44 years old, recently quit]

> *I know there were certain times I met people because we'd go… I'd go to smoke, and they were there, and it was like oh, I didn't know you smoked. Oh, hi. Hi. And then suddenly, you're friends. But there are better ways to meet people. So, see everything I say is really over time I've learned that there really is no good reason to smoke. Not that I shame anyone for doing it.* [Female, 35-44 years old, recently quit]

Facilitators

Participants were motivated to stop smoking due to the existing social stigma of smoking outside their peer networks and viewed positive support as a facilitator to quitting smoking. For instance, 1 smoker indicated that the social stigma of smoking, when added to the stigma of having a mental illness, was a motivator to stop smoking.

> *There's a stigma with smokers, you know? And there's a stigma with mental illness. I don't like stigma. I don't like labels. I don't like boxes, you know? I'd rather be a nonsmoker in peoples' eyes if I had to choose.* [Female, 45-54 years old, recently quit]

Social support also facilitated successful quitting. One participant said:

> *I think as I just got older, there were just more people in my life, who would say, you know, you should really stop smoking. You know, doctors, family members, friends who didn't smoke. It's just like there were more of those types of people in my life. And so, even though I didn't necessarily listen to them right then and there. In a way, I really did listen to them. I just kinda put it into the back of my mind. And I just eventually was like, you know, you don't really need this.* [Male, 45-54 years old, recently quit]

While some participants responded that social isolation increased their smoking behaviors, those who had low exposure to other smokers due to social distancing were able to engage in cessation efforts.

> *I wasn't necessarily even really enjoying it as much. Maybe because I was alone, and there wasn't that like, because of the pandemic, there wasn't that like social component to it, that made it even easier, I guess, to maybe quit I guess.* [Male, 45-54 years old, recently quit]

Theme 3: Smoking and Quitting Behaviors Similar to the General Population

There were similarities in smoking and quitting behaviors between people with serious mental illness and the general population.

Barriers (Smoking Behavior)

Similar to the general population, participants tended to smoke out of habit, such as after eating or having coffee, or when drinking alcohol.

> *It's like clockwork. You have to smoke right after you eat. I don't even know why. I don't even think I did it because of any real craving or anything. It's like something you just become used to doing.* [Male, 35-44 years old, recently quit]

> *Like something about coffee and cigarettes, like I don't know. For me, they go together well. Like you know, you're trying to boost your energy up.* [Female, 30-34 years old, currently smokes]

Participants also smoked due to boredom or having easy access to cigarettes through friends.

> *Time, having time. I've had to take a lot of time to find other routines and other things to have in my hands. Especially as someone who is unemployed and disabled, it's like I don't need to be sitting around smoking. We have to do other things.* [Female, 35-44 years old, recently quit]

> *Because everybody here is smoking, you know, cigarettes. Like there's a guy here, and he buys, you know, a pound of tobacco at a time. And he'll roll a bunch, you know, with the tubes. And so he just hands them out. It's like very communal.* [Male, 24-26 years old, currently smokes]

Similar to the general population, negative social support made it challenging to quit. One smoker said:

> *And then I didn't really have anybody supporting me in it. I have family telling me that they didn't think I was gonna succeed. I didn't have anybody like supporting me, and saying, yeah, you can do it. I didn't have anyone like that. Just people saying, when are you gonna start smoking again?* [Female, 45-54 years old, currently smokes]

Facilitators (Quitting Experiences)

However, individuals had concerns about the negative impact of smoking on their physical health and general quality of life, which motivated them to quit smoking.

> *You know, again, I guess I was thinking down the line like I can't keep doing this until I'm 40 or 50. And so it was more, again, like thinking of my health down the road, as opposed to right now. And so, it was basically thinking long-term, I can't keep this up. I don't want to, you know, have any cancer or risk of heart attack, stroke, or heart disease.* [Male, 30-34 years old, recently quit]

> *In your clothing, in your hair. Everywhere around you, you know. Everything smells like… well, especially you come back home from fresh air from outside, it's just… or when you came out of shower and you walk in the room, it's just disgusting.* [Male, 55-65 years old, currently smokes]

Participants used evidence-based tobacco treatments and behavioral approaches to quit smoking.

> *Yeah, I've tried Chantix a couple of times. I used to not really believe in the patch or the gum up until, you know, recently.* [Male, 24-26 years old, currently smokes]

> *I have the Nicotrol inhaler. It's a prescription quit-smoking thing that's not available over the counter yet for some reason. I've tried it with the gum. I've tried with the patch. I think the Nicotrol inhaler is more effective because you're still puffing on something. So the Nicotrol inhaler's more helpful because you're puffing on something. So that was the only thing I was ever able to quit with at all.* [Female, 45-54 years old, currently smokes]

Discussion

Main Study Findings

Using the 4-step analytical approach, we identified the following 3 key aspects that can inform the development of text-based cessation interventions for people with serious mental illness who smoke: (1) interplay between smoking and having a serious mental illness, (2) social contextual factors of smoking, and (3) smoking and quitting behaviors similar to the general population. Across the 3 themes, participants reported barriers, motivators, and facilitators to quitting smoking. The variations in the content provided by the participants can facilitate the development of messages that can be implemented in text-based interventions targeted at people with serious mental illness who smoke.

Participants strongly believed that smoking helped manage mental health symptoms and was often used to cope with stress. Stress is often cited as a reason for smoking in individuals with and without serious mental illness [30,31]. However, because stress has a strong positive correlation with mental illness [32], people with serious mental illness may not progress through the stages of change at a rate comparable to that in those without serious mental illness. In addition, the term “stress” could be used as a proxy for mental health symptoms in this patient group. In such a case, smoking could be used as self-medication to alleviate feelings of intense emotions or to feel an immediate sense of relaxation even though individuals are aware of the harmful effects of smoking on their mental and physical health.

Individuals found that speaking with their mental health care provider when quitting helped them navigate the quitting process better. In discussions with their providers, they expressed the impact of tobacco medication (ie, Chantix) on their mental health symptoms. This finding implies that patient-provider discussions regarding smoking and mental health may facilitate successful quitting. Therefore, engaging health care providers when making a quit attempt is essential in this population. Patient-provider discussions regarding smoking also provide a critical opportunity for health care providers to monitor for smoking/quitting-drug interactions [33] as patients may not be aware of these interactions.

Adapting text-based cessation interventions to address barriers to smoking among people with serious mental illness may increase the intervention’s relevance, engagement, and effectiveness. Past text-based cessation interventions have not included content specific to mental health conditions [34]. A systematic review of text-based cessation interventions found that message content focused on increasing self-efficacy and encouraging smokers to quit or maintain their quit status by providing quit tips [34]. This content is very similar to the one we used in our own past studies [35,36]. Currently, text messages used in cessation interventions lack content that highlights mental health symptoms and stress management, and content that emphasizes the need to engage mental health care

providers when quitting. Therefore, text messages that offer tips on evidence-based strategies for managing stress, regulating emotions, managing negative affect, and dealing with unhelpful thoughts/anxiety can be a meaningful inclusion in text-based interventions. Messages that serve as reminders to engage mental health care providers when quitting may increase the relevance of text-based cessation interventions in this patient population.

We found a strong tendency to smoke in people with mental health conditions and in the mental health environments in which they were embedded. From our observations, smoking seemed ingrained in all aspects of their social lives. For instance, smoking was perceived as a social bonding activity so much that individuals often perceived smoking as a way to be socially included or form friendships. Given the high level of social acceptance observed in this patient population [24,37], these barriers to quitting smoking extend beyond those experienced by people in the general population who smoke. Recognizing the broader factors that drive smoking rates within the living and social environments of people with serious mental illness can complement existing efforts aimed at helping them quit successfully. A few past text messaging interventions have included messages with content relevant to the social context [38]. The messages are about seeking social support or how to deal with a partner who smokes. Additional content could include messages that denormalize smoking [39], provide alternatives to smoking as a group bonding activity that could counteract the perceived social rewards of smoking, and provide information on how to identify cessation role models.

In addition to the distinct challenges identified by people with serious mental illness, they also face challenges commonly observed among people in the general population who smoke. Participants often smoked out of habit, such as after a meal or while drinking alcohol, due to boredom and having access to cigarettes. They reported similar reasons for quitting, including health benefits, the financial cost of cigarettes, and the general improvement in quality of life. Participants also used evidence-based strategies to quit smoking, including nicotine replacement therapy and tobacco treatment medications. These findings are similar to past study findings [40], which indicate that people with serious mental illness are motivated to quit and can engage in cessation using evidence-based resources, yet they often cannot quit. It is important to modify existing text-based interventions to provide information on mental health service access and to include content that increases antismoking attitudes.

Other researchers have discussed the need for a structured process to support the development of text messages [41]. The approach we used in this paper provided an efficient way to identify key factors that can inform text message development from formative interviews. We chose the rapid qualitative approach over other qualitative data analysis approaches because it provided a systematic way to describe the lived experiences of people with serious mental illness who smoke within each domain (eg, the context in which participants smoked, smoking triggers, reasons for smoking, barriers to active quitting, and facilitators of active quitting).

Limitations

This study had limitations. Our study recruited participants from mental health agencies, who had access to mental health services. In many ways, individuals with access to mental health services could differ from those without access to mental health services regarding smoking behaviors, attitudes toward quitting, and cessation experiences. Targeting people with serious mental illness without access to health services could provide additional insights into these patients' challenges. Additionally, we relied on self-reported information when assessing smoking status.

Conclusion

People with serious mental illness may experience life stressors more intensely than those without serious mental illness. As a result, they often use smoking to cope with intense emotional states, manage mental health symptoms, or maintain social bonds. Based on our study findings, text message content should be modified to (1) provide information on the effect of smoking on mental health and alternative approaches to managing mental health symptoms, (2) incorporate smoking denormalization strategies that increase antismoking attitudes, and (3) provide healthier alternatives to smoking for social bonding activities. Adapted text-based interventions can ensure that people with serious mental health illness receive the appropriate support to stop smoking.

Data Availability

This article's data can be shared on reasonable request to the project principal investigators (MD and RSS).

Conflicts of Interest

None declared.

References

1. Jamal A, Dube SR, Malarcher AM, Shaw L, Engstrom MC, Centers for Disease ControlPrevention (CDC). Tobacco use screening and counseling during physician office visits among adults--National Ambulatory Medical Care Survey and National Health Interview Survey, United States, 2005-2009. MMWR Suppl 2012 Jun 15;61(2):38-45. [Medline: 22695462]
2. Jamal A, King BA, Neff LJ, Whitmill J, Babb SD, Graffunder CM. Current cigarette smoking among adults - United States, 2005-2015. MMWR Morb Mortal Wkly Rep 2016 Nov 11;65(44):1205-1211 [FREE Full text] [doi: 10.15585/mmwr.mm6544a2] [Medline: 27832052]

3. Tobacco Use and Dependence Guideline Panel. Treating Tobacco Use and Dependence: 2008 Update. Rockville, MD: US Department of Health and Human Services; 2008.
4. Key substance use and mental health indicators in the United States: Results from the 2019 National Survey on Drug Use and Health. Substance Abuse and Mental Health Services Administration. Rockville, MD: Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration; 2020. URL: https://www.samhsa.gov/data/sites/default/files/reports/rpt29393/2019NSDUHFFRPDFWHTML/2019NSDUHFFR1PDFW090120.pdf [accessed 2022-10-31]
5. Sheals K, Tombor I, McNeill A, Shahab L. A mixed-method systematic review and meta-analysis of mental health professionals' attitudes toward smoking and smoking cessation among people with mental illnesses. Addiction 2016 Sep 03;111(9):1536-1553 [FREE Full text] [doi: 10.1111/add.13387] [Medline: 27003925]
6. Stanczyk NE, Smit ES, Schulz DN, de Vries H, Bolman C, Muris JWM, et al. An economic evaluation of a video- and text-based computer-tailored intervention for smoking cessation: a cost-effectiveness and cost-utility analysis of a randomized controlled trial. PLoS One 2014 Oct 13;9(10):e110117 [FREE Full text] [doi: 10.1371/journal.pone.0110117] [Medline: 25310007]
7. Hoermann S, McCabe KL, Milne DN, Calvo RA. Application of synchronous text-based dialogue systems in mental health interventions: Systematic review. J Med Internet Res 2017 Jul 21;19(8):e267 [FREE Full text] [doi: 10.2196/jmir.7023] [Medline: 28784594]
8. Gay K, Torous J, Joseph A, Pandya A, Duckworth K. Digital technology use among individuals with schizophrenia: Results of an online survey. JMIR Ment Health 2016 May 04;3(2):e15 [FREE Full text] [doi: 10.2196/mental.5379] [Medline: 27146094]
9. Spohr SA, Nandy R, Gandhiraj D, Vemulapalli A, Anne S, Walters ST. Efficacy of SMS text message interventions for smoking cessation: A meta-analysis. J Subst Abuse Treat 2015 Sep;56:1-10. [doi: 10.1016/j.jsat.2015.01.011] [Medline: 25720333]
10. Heffner JL, Kelly MM, Waxmonsky J, Mattocks K, Serfozo E, Bricker JB, et al. Pilot randomized controlled trial of web-delivered acceptance and commitment therapy versus smokefree.gov for smokers with bipolar disorder. Nicotine Tob Res 2020 Aug 24;22(9):1543-1552 [FREE Full text] [doi: 10.1093/ntr/ntz242] [Medline: 31883336]
11. Brown RA, Minami H, Hecht J, Kahler CW, Price LH, Kjome KL, et al. Sustained care smoking cessation intervention for individuals hospitalized for psychiatric disorders: The helping HAND 3 randomized clinical trial. JAMA Psychiatry 2021 Aug 01;78(8):839-847 [FREE Full text] [doi: 10.1001/jamapsychiatry.2021.0707] [Medline: 33950156]
12. Brunette MF, Ferron JC, McGurk SR, Williams JM, Harrington A, Devitt T, et al. Brief, web-based interventions to motivate smokers with schizophrenia: Randomized trial. JMIR Ment Health 2020 Mar 08;7(2):e16524 [FREE Full text] [doi: 10.2196/16524] [Medline: 32039811]
13. Lerman C, Audrain J, Orleans C, Boyd R, Gold K, Main D, et al. Investigation of mechanisms linking depressed mood to nicotine dependence. Addictive Behaviors 1996 Jan;21(1):9-19. [doi: 10.1016/0306-4603(95)00032-1]
14. Lawn SJ, Pols RG, Barber JG. Smoking and quitting: a qualitative study with community-living psychiatric clients. Social Science & Medicine 2002 Jan;54(1):93-104. [doi: 10.1016/s0277-9536(01)00008-9]
15. Clancy N, Zwar N, Richmond R. Depression, smoking and smoking cessation: a qualitative study. Fam Pract 2013 Oct 08;30(5):587-592. [doi: 10.1093/fampra/cmt032] [Medline: 23836095]
16. Wiltsey Stirman S, Baumann AA, Miller CJ. The FRAME: an expanded framework for reporting adaptations and modifications to evidence-based interventions. Implement Sci 2019 Jun 06;14(1):58 [FREE Full text] [doi: 10.1186/s13012-019-0898-y] [Medline: 31171014]
17. Houston TK, Sadasivam RS, Ford DE, Richman J, Ray MN, Allison JJ. The QUIT-PRIMO provider-patient Internet-delivered smoking cessation referral intervention: a cluster-randomized comparative effectiveness trial: study protocol. Implement Sci 2010 Nov 17;5(1):87 [FREE Full text] [doi: 10.1186/1748-5908-5-87] [Medline: 21080972]
18. Houston TK, Sadasivam RS, Allison JJ, Ash AS, Ray MN, English TM, et al. Evaluating the QUIT-PRIMO clinical practice ePortal to increase smoker engagement with online cessation interventions: a national hybrid type 2 implementation study. Implement Sci 2015 Nov 02;10(1):154 [FREE Full text] [doi: 10.1186/s13012-015-0336-8] [Medline: 26525410]
19. Coley HL, Sadasivam RS, Williams JH, Volkman JE, Schoenberger Y, Kohler CL, National Dental PBRNQUITPRIMO Collaborative Group. Crowdsourced peer- versus expert-written smoking-cessation messages. Am J Prev Med 2013 Nov;45(5):543-550 [FREE Full text] [doi: 10.1016/j.amepre.2013.07.004] [Medline: 24139766]
20. Macias C, Jackson R, Schroeder C, Wang Q. What is a clubhouse? Report on the ICCD 1996 survey of USA clubhouses. Community Ment Health J 1999 May;35(2):181-190. [doi: 10.1023/a:1018776815886] [Medline: 10412626]
21. McKay CE, Yates BT, Johnsen M. Costs of clubhouses: an international perspective. Adm Policy Ment Health 2007 Jan 12;34(1):62-72. [doi: 10.1007/s10488-005-0008-0] [Medline: 16220235]
22. de Leon J, Diaz FJ. A meta-analysis of worldwide studies demonstrates an association between schizophrenia and tobacco smoking behaviors. Schizophr Res 2005 Jul 15;76(2-3):135-157. [doi: 10.1016/j.schres.2005.02.010] [Medline: 15949648]
23. Young-Wolff KC, Henriksen L, Delucchi K, Prochaska JJ. Tobacco retailer proximity and density and nicotine dependence among smokers with serious mental illness. Am J Public Health 2014 Aug;104(8):1454-1463. [doi: 10.2105/ajph.2014.301917]

24. Kerr S, Woods C, Knussen C, Watson H, Hunter R. Breaking the habit: a qualitative exploration of barriers and facilitators to smoking cessation in people with enduring mental health problems. BMC Public Health 2013 Mar 12;13(1):221 [FREE Full text] [doi: 10.1186/1471-2458-13-221] [Medline: 23497231]
25. Beebe J. Rapid Qualitative Inquiry: A Field Guide to Team-Based Assessment. Lanham, MD: Rowman & Littlefield Publishers; 2014.
26. Hamilton AB, Finley EP. Qualitative methods in implementation research: An introduction. Psychiatry Res 2019 Oct;280:112516 [FREE Full text] [doi: 10.1016/j.psychres.2019.112516] [Medline: 31437661]
27. Gale RC, Wu J, Erhardt T, Bounthavong M, Reardon CM, Damschroder LJ, et al. Comparison of rapid vs in-depth qualitative analytic methods from a process evaluation of academic detailing in the Veterans Health Administration. Implement Sci 2019 Feb 01;14(1):11 [FREE Full text] [doi: 10.1186/s13012-019-0853-y] [Medline: 30709368]
28. Carter N, Bryant-Lukosius D, DiCenso A, Blythe J, Neville AJ. The use of triangulation in qualitative research. Oncology Nursing Forum 2014 Aug 26;41(5):545-547. [doi: 10.1188/14.onf.545-547]
29. Spall S. Peer debriefing in qualitative research: Emerging operational models. Qualitative Inquiry 2016 Jun 29;4(2):280-292. [doi: 10.1177/107780049800400208]
30. Fatani BZ, Al-Yahyawi H, Raggam A, Al-Ahdal M, Alzyoud S, Hassan AN. Perceived stress and willingness to quit smoking among patients with depressive and anxiety disorders seeking treatment. Health Sci Rep 2022 Mar 24;5(2):e503 [FREE Full text] [doi: 10.1002/hsr2.503] [Medline: 35229052]
31. Villanti AC, Bover Manderski MT, Gundersen DA, Steinberg MB, Delnevo CD. Reasons to quit and barriers to quitting smoking in US young adults. Fam Pract 2016 Apr 04;33(2):133-139 [FREE Full text] [doi: 10.1093/fampra/cmv103] [Medline: 26733658]
32. Beiter R, Nash R, McCrady M, Rhoades D, Linscomb M, Clarahan M, et al. The prevalence and correlates of depression, anxiety, and stress in a sample of college students. J Affect Disord 2015 Mar 01;173:90-96. [doi: 10.1016/j.jad.2014.10.054] [Medline: 25462401]
33. Juan D Molina FM. CYP450 and its implications in the clinical use of antipsychotic drugs. Clin Exp Pharmacol 2015;05(03). [doi: 10.4172/2161-1459.1000176]
34. Kong G, Ells DM, Camenga DR, Krishnan-Sarin S. Text messaging-based smoking cessation intervention: a narrative review. Addict Behav 2014 May;39(5):907-917 [FREE Full text] [doi: 10.1016/j.addbeh.2013.11.024] [Medline: 24462528]
35. Nagawa CS, Faro JM, Menon AJ, Ito Fukunaga M, Williams JH, Mourao D, et al. Written advice given by African American smokers to their peers: Qualitative study of motivational messages. JMIR Form Res 2021 Apr 30;5(4):e21481 [FREE Full text] [doi: 10.2196/21481] [Medline: 33929332]
36. Coley HL, Sadasivam RS, Williams JH, Volkman JE, Schoenberger Y, Kohler CL, National Dental PBRNQUITPRIMO Collaborative Group. Crowdsourced peer- versus expert-written smoking-cessation messages. Am J Prev Med 2013 Nov;45(5):543-550 [FREE Full text] [doi: 10.1016/j.amepre.2013.07.004] [Medline: 24139766]
37. Ziedonis D, Hitsman B, Beckham JC, Zvolensky M, Adler LE, Audrain-McGovern J, et al. Tobacco use and cessation in psychiatric disorders: National Institute of Mental Health report. Nicotine Tob Res 2008 Dec 01;10(12):1691-1715. [doi: 10.1080/14622200802443569] [Medline: 19023823]
38. Abroms LC, Ahuja M, Kodl Y, Thaweethai L, Sims J, Winickoff JP, et al. Text2Quit: results from a pilot test of a personalized, interactive mobile health smoking cessation program. J Health Commun 2012 May 02;17 Suppl 1(sup1):44-53 [FREE Full text] [doi: 10.1080/10810730.2011.649159] [Medline: 22548598]
39. Bayer R, Stuber J. Tobacco control, stigma, and public health: Rethinking the relations. Am J Public Health 2006 Jan;96(1):47-50. [doi: 10.2105/ajph.2005.071886]
40. Centers for Disease Control and Prevention (CDC). Vital signs: current cigarette smoking among adults aged ≥18 years--United States, 2005-2010. MMWR Morb Mortal Wkly Rep 2011 Oct 09;60(35):1207-1212 [FREE Full text] [Medline: 21900875]
41. Abroms LC, Whittaker R, Free C, Mendel Van Alstyne J, Schindler-Ruwisch JM. Developing and pretesting a text messaging program for health behavior change: Recommended steps. JMIR Mhealth Uhealth 2015 Dec 21;3(4):e107 [FREE Full text] [doi: 10.2196/mhealth.4917] [Medline: 26690917]

Edited by A Mavragani; submitted 08.07.22; peer-reviewed by K Wiseman, S Veldheer; comments to author 20.09.22; revised version received 11.10.22; accepted 26.10.22; published 07.11.22.

Please cite as:
Nagawa CS, Lane IA, McKay CE, Kamberi A, Shenette LL, Kelly MM, Davis M, Sadasivam RS
Use of a Rapid Qualitative Method to Inform the Development of a Text Messaging Intervention for People With Serious Mental Illness Who Smoke: Formative Research Study
JMIR Form Res 2022;6(11):e40907
URL: https://formative.jmir.org/2022/11/e40907
doi:10.2196/40907
PMID:36342765

Original Paper

Remotely Assessing Mechanisms of Behavioral Change in Community Substance Use Disorder Treatment to Facilitate Measurement-Informed Care: Pilot Longitudinal Questionnaire Study

Kevin A Hallgren[1], PhD

Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, United States

Corresponding Author:
Kevin A Hallgren, PhD
Department of Psychiatry and Behavioral Sciences
University of Washington
University of Washington Medical Center, Box 356560
1959 NE Pacific Street
Seattle, WA, 98195
United States
Phone: 1 206 616 2906
Email: khallgre@uw.edu

Abstract

Background: Research shows that improvements in coping strategies, abstinence self-efficacy, craving, and depression are potential mechanisms of behavioral change (MOBC) in treatments for substance use disorders (SUDs). However, little is known about how these insights regarding MOBC can be applied to SUD treatment settings. One way to facilitate MOBC-informed care in frontline settings could be to measure and monitor changes in MOBC throughout treatment using brief, frequent questionnaires that patients complete by using mobile technologies (eg, smartphones). The results derived from these questionnaires could potentially be used for clinical monitoring (ie, measurement-based care) to better understand whether individual patients are experiencing treatment-related improvements on key clinical targets.

Objective: This study evaluated whether brief, weekly MOBC questionnaires completed by patients remotely can potentially provide clinically meaningful information about changes in MOBC in the context of real-world, community-based SUD treatment.

Methods: A total of 30 patients (14/30, 47% female; 13/30, 43% racial or ethnic minority) in a community SUD treatment clinic participated in a pilot study where they were invited to complete brief, weekly questionnaires that assessed various MOBC, including coping strategies, abstinence self-efficacy, craving, depression, and therapeutic alliance. Questionnaires were typically completed remotely via smartphone for up to 6 months; 618 questionnaires were completed in total. Participants also completed longer, psychometrically validated measures of the same MOBC at baseline and 6-month research appointments. Statistical analyses tested whether brief, weekly, remotely completed MOBC questionnaires exhibited characteristics that would be desirable for real-world longitudinal clinical monitoring, including a tendency to detect within-person changes in MOBC over time; cross-sectional and longitudinal associations with longer, psychometrically validated measures completed at research appointments; and similar patterns of associations with 6-month percentage of days abstinent as longer, psychometrically validated MOBC measures completed at research appointments.

Results: The results of this study indicated that the brief, weekly, remotely completed MOBC measures exhibited characteristics that are desirable for clinical monitoring, including a tendency to vary longitudinally (within patients over time) more often than measures of alcohol and drug consumption, generally having medium to large cross-sectional and longitudinal correlations with longer psychometrically validated measures of MOBC completed at research appointments, and generally having similar patterns of association with 6-month percentage of days abstinent from alcohol and drugs as longer psychometrically validated MOBC measures completed at research appointments.

Conclusions: The results of this pilot study provide initial evidence that incorporating brief, weekly, and remotely completed MOBC questionnaires into community SUD treatment may be a viable approach for facilitating MOBC-informed care. Such questionnaires can potentially support measurement-based care by providing meaningful information about within-patient changes in clinical domains that are often directly targeted in SUD treatments and predict long-term substance use outcomes.

(JMIR Form Res 2022;6(11):e42376) doi:10.2196/42376

KEYWORDS

addiction; clinical pilot; measurement-based care; mechanisms of change; mobile health; mHealth; mobile phone

Introduction

At least 2 million people receive treatment for substance use disorders (SUDs) annually in the United States [1]. Research has identified several efficacious treatments for SUDs, including behavioral and medication-based treatments [2,3]. To help improve the effectiveness and efficiency of SUD treatments, research has become increasingly focused on identifying the specific mechanisms of behavior change (MOBC) that may be most responsible for driving treatment outcomes [4,5].

Studies show that multiple MOBC—including (but not limited to) increases in substance-related coping skills and abstinence self-efficacy, reductions in craving and depression symptoms, and a strong therapeutic alliance—typically improve during a variety of SUD treatments, and these improvements often predict long-term substance use outcomes [6-19]. Moreover, these MOBC often reflect highly meaningful clinical targets for patients and clinicians; for example, patients may feel distressed by cravings and depression symptoms and often want to improve their coping skills and abstinence self-efficacy.

What remains less understood is how the growing body of MOBC-related findings will be used to improve the treatments that are offered in community-based SUD treatment settings; that is, how can MOBC science be used to help frontline clinicians offer better treatment? Because MOBC are potentially observable indicators of whether SUD treatments are working for individual patients in treatment, it is possible that measuring and monitoring MOBC longitudinally throughout treatment could help clinicians and patients detect whether a treatment is affecting meaningful proximal treatment targets that are associated with long-term outcomes (ie, MOBC). For example, clinicians could routinely administer brief questionnaires to assess different MOBC and then score, graph, and review changes in these measures over time throughout treatment as a form of treatment progress monitoring or measurement-based care [20]. These self-monitoring data could help draw clinical attention toward MOBC and facilitate clinical discussions about factors that may be contributing to changes in MOBC; for example, how specific behavior changes helped a patient increase their abstinence self-efficacy or reduce their depression symptoms. However, many community treatment settings are currently unable to systematically measure and monitor MOBC as a part of routine care. As a result, there are significant limitations to the extent that MOBC-related research findings affect the SUD treatments that patients receive in real-world settings.

To help reduce this gap, we developed a tool designed to facilitate MOBC-informed measurement-based care and pilot-tested its feasibility and usability in a community-based treatment setting [21,22]. The measurement-based care system included a patient-facing brief questionnaire that patients could complete remotely (eg, via smartphone), called *the weekly check-in*, which included questions assessing MOBC and other domains that clinicians reported as potentially helpful to measure frequently during routine care. The system also included a web-based dashboard for clinicians to review the longitudinal results of their patients' weekly check-ins. During the clinical pilot, clinicians and patients had high rates of engagement with the system and provided favorable subjective ratings regarding its usability and clinical helpfulness [21].

Despite the high rates of engagement and favorable usability ratings of the weekly check-in [21], it is currently unclear whether brief and remotely completed weekly MOBC assessments could provide quantitatively meaningful data about changes in MOBC during real-world SUD treatment. Therefore, this study provides a preliminary evaluation of such MOBC assessments. If the results of this study indicate that such questionnaires provide quantitatively meaningful information, it would suggest that incorporating brief, remote, and frequent MOBC assessments during community SUD treatment could be a viable approach to facilitating MOBC-informed measurement-based care in real-world settings, which could be further tested in large-scale implementation studies. Specifically, this study evaluates whether brief, weekly, remotely completed MOBC questionnaires completed by patients in community-based treatment exhibited characteristics that would be desirable for real-world longitudinal clinical monitoring, including (1) a tendency to detect longitudinal within-person changes in MOBC over time; (2) cross-sectional and longitudinal associations with longer, psychometrically validated measures completed at research appointments; and (3) similar patterns of associations with 6-month percentage of days abstinent as longer, psychometrically validated MOBC measures completed at research appointments. It was hypothesized that the brief and remotely completed MOBC measures completed on the weekly check-in would vary longitudinally for most patients and that within-patient variability would be more common for measures of MOBC than for measures of alcohol and drug consumption, in part because many patients initiate abstinence from alcohol and drugs before starting SUD treatment [23,24]. It was also hypothesized that the remotely completed MOBC measures would have high cross-sectional and longitudinal associations with longer, psychometrically validated measures completed at baseline and 6-month research appointments, and that MOBC measures completed in both modalities—brief, remotely completed measures on the weekly check-in and longer, psychometrically validated measures completed at research appointments—would have similar patterns of association with 6-month percentage of days abstinent.

Methods

Setting and Participants for Clinical Pilot

In total, 30 participants were recruited from an addiction and mental health treatment clinic in an urban setting in Washington

State, United States. The clinic offered treatment in various formats, including counseling (individual and group), peer recovery support, case management, pharmacotherapy for psychiatric, alcohol, and opioid use disorders, and harm reduction–oriented treatment (ie, for patients with nonabstinent goals). Participants were recruited via handouts that clinicians could give to patients during treatment sessions and flyers that were posted in public spaces within the clinic. Eligibility criteria for the clinical pilot included being enrolled in SUD treatment, receiving care from a clinician on 1 of 2 participating treatment teams, having an iPhone or Android smartphone, the ability to read and speak English, being at least 18 years old, and having an Alcohol Use Disorders Identification Test-Consumption version score >4 for men or 3 for women [25] or self-reporting past-year use of nonprescribed or illicit drugs, including alcohol, cannabis, hallucinogens, inhalants, opioids, sedatives, hypnotics, anxiolytics, or stimulants [26]. Eligibility screening was completed by phone by a member of the research team. Eligible patients were invited to complete a baseline appointment, where they provided informed consent to participate in the pilot. Participants completed additional follow-up research appointments with a research coordinator at 6, 12, and 24 weeks after baseline. Participants were recruited between October 2019 and June 2021. A sample size of 30 patients was recruited during this period, which was determined to be adequate for the formative aims of this research. Additional details regarding the setting and procedures are provided in a previous study [21].

Measures

Weekly Check-in

A brief questionnaire, called *the weekly check-in*, was sent to patients every week via SMS text message or email through a REDCap server [27]. Patients completed the first weekly check-in at the baseline research appointment, and subsequent weekly check-ins were completed independently for up to 24 weeks. The weekly check-in contained 2 questions that assessed past-week substance use and 10 questions that assessed 5 MOBC domains described subsequently, with 1 to 3 questions per domain included from previously validated measures. The 5 MOBC domains were selected based on clinician input during formative research [22] and empirical evidence supporting their role as MOBC in SUD treatment.

The weekly check-in also included 10 other questions that were not evaluated here but were included in the weekly check-in based on clinician input [22]; these included 2 questions assessing positive life outlook, 6 questions assessing goals for the coming week, and 2 open-ended questions where patients could write additional goals or provide other notes they wanted to communicate to their clinicians. The specific questions included in the weekly check-in are provided in Multimedia Appendix 1 [28-32].

Research Assessments

Participants were also invited to complete longer, psychometrically validated measures of substance use and MOBC at research appointments completed at the baseline and at the 6-, 12-, and 24-week follow-ups. These appointments initially occurred in person but were later completed as telephone-based appointments to facilitate social distancing during the COVID-19 pandemic.

Substance Use and MOBC Domains

Overview

The measures used to assess substance use and MOBC domains are described in the following sections. Each section describes the brief, remotely completed measures that were included in the weekly check-in and the corresponding longer, psychometrically validated measures that were completed at the baseline and at the 6-, 12-, and 24-week research appointments.

Substance Use

On the weekly check-in, drinking and drug use were measured using 2 questions that asked participants how many days they "drank too much" and "used drugs" over the past week. The response options for both questions were on a 5-point scale ranging from "not at all" to "every day." The 2 questions used in the weekly check-in were derived from the Substance Use Recovery Evaluator [32] with minor modifications to question wording based on earlier usability testing with patients that indicated some confusion with the original question phrasing [33].

At research appointments, drinking and drug use were assessed over a 30-day period via structured interviews with the Addiction Severity Index-Lite [34]. The specific indices used in this study reflected (1) the percentage of days (out of the past 30) that participants drank to intoxication, (2) the percentage of days (out of the past 30) that participants used any illicit or nonprescribed psychoactive drugs, excluding alcohol and tobacco, and (3) the percentage of days (out of the past 30) that participants reported complete abstinence from alcohol and illicit or nonprescribed psychoactive drugs, excluding tobacco.

Coping Strategies

The weekly check-in included 3 questions asking patients how often they used coping strategies that could be helpful for preventing alcohol and drug use. The 3 items were derived from the Coping Strategies Scale [28] and were selected for the weekly check-in because they reflected broad cognitive behavioral strategies that are applicable to a range of scenarios, including stimulus control (ie, avoiding people, places, and things that can lead to substance use), alternative reinforcement (ie, engaging in activities that can replace substance use), and high-risk planning (eg, planning ahead for situations that pose a high-risk for substance use). The wording of the questions and responses was modified to make the question structure more consistent with other items in the weekly check-in, as informed by preliminary usability testing with patients. Response options were on a 5-point scale ranging from "not at all" to "always."

At research appointments, coping strategies were assessed using the 59-item Coping Strategies Scale [28], modified to reflect coping strategies relevant to preventing any type of substance use rather than alcohol specifically. The 59-item Coping Strategies Scale was initially excluded from the research appointment questionnaire battery owing to concerns about potential assessment fatigue to participants, as the questionnaire

was considerably longer than all other measures. However, most patients were able to complete the initial questionnaire battery without assessment fatigue, so the questionnaire was added to the battery midway through the study. Therefore, it was only available at baseline for the last 14 enrolled patients and at the 6-month follow-up for the last 16 enrolled patients, one of whom declined to complete it because of assessment fatigue.

Abstinence Self-efficacy

The weekly check-in included two items asking patients how confident they felt in their ability to not drink or use drugs (ie, abstinence self-efficacy), including (1) when they were emotionally upset or in physical pain and (2) when they felt an urge or craving. The questions were derived from the Brief Situational Confidence Questionnaire-8 [29] with modifications to make phrasing more consistent with other items in the weekly check-in, as informed by preliminary usability testing with patients. Response options were on a 5-point scale ranging from "not at all confident" to "extremely confident."

At research appointments, confidence in avoiding alcohol or drug use was assessed using the full Brief Situational Confidence Questionnaire-8 questionnaire [29].

Craving

The weekly check-in included one question that asked how many days the patient experienced alcohol and drug cravings over the past week using a question included in the Substance Use Recovery Evaluator [32]. Response options were on a 5-point scale ranging from "not at all" to "every day."

At research appointments, craving was assessed over the past week using the 5-item Penn Alcohol Craving Scale questionnaire [35], which was modified to assess craving for drugs in addition to alcohol.

Depression

The weekly check-in included the Patient Health Questionnaire-2 (PHQ-2) to briefly assess depression symptoms [30]. The PHQ-2 is a 2-item measure derived from the first 2 items of the 9-item PHQ [36]. Response options are on a 4-point scale ranging from "not at all" to "nearly every day."

At the research appointments, depression symptoms were measured using the full PHQ-9 [36].

Therapeutic Alliance

Therapeutic alliance was assessed on the weekly check-in using 2 questions asking patients to report how much their clinicians agreed with them about what was important to work on in treatment and how often their clinicians gave them new ways of looking at their problems. The 2 questions were derived from the Working Alliance Inventory-Short Revised [31] with modifications to make the questions and responses more consistent with the other questions on the weekly check-in. Response options were on a 5-point scale ranging from "seldom" to "always."

At research appointments, the therapeutic alliance was assessed using the full Working Alliance Inventory-Short Revised [31].

Ethical Considerations

All procedures used in this study were approved by the University of Washington Institutional Review Board (approval number: STUDY00007996). All participants provided written documentation of informed consent before participating in the study. All study data were deidentified before analysis. Research participants were paid US $50 in prepaid debit cards for attending each research appointment (up to 4 appointments or US $200 in total).

Analysis Plan

Statistical analyses aimed to evaluate whether the MOBC measures obtained from the brief, remotely completed weekly check-in conveyed information that would be desirable for longitudinal monitoring of changes in MOBC during SUD treatment, including (1) a tendency to vary longitudinally within persons over time; (2) high cross-sectional and longitudinal associations with longer, psychometrically validated measures completed at research appointments; and (3) similar patterns of association with end-of-pilot percentage of days abstinent as the longer, psychometrically validated MOBC measures completed at research appointments.

Longitudinal, within-patient variability in weekly check-in domains was characterized descriptively to understand how much information each domain on the weekly check-in could potentially provide about *changes in* MOBC within patients over time (ie, measures with limited variability within patients over time may have less utility for longitudinal clinical monitoring and measurement-based care). Descriptive analyses were performed using data visualization methods that illustrated the variability and ranges of scores for weekly substance use and MOBC measures within patients across all time points. The number of patients whose weekly check-in responses indicated no variability over the 24-week period (eg, number of patients reporting no changes in substance use or no changes in coping strategies) was also identified to evaluate how often the weekly measures had no within-person variability.

Cross-sectional and longitudinal associations between the weekly check-in measures (assessed using 1-3 questions per domain) and their corresponding longer, psychometrically validated measures completed at research appointments (5-59 questions per domain) were examined. Cross-sectional associations were examined by computing Pearson correlations between the psychometrically validated measures and the corresponding weekly check-in measures when both were completed at baseline. Longitudinal associations were evaluated using Pearson correlations of *change scores* on the psychometrically validated measures between the baseline and the 6-month research assessments and *change scores* on the corresponding weekly check-in measures over a matched time period (ie, the "matched" weekly check-ins were those that were temporally closest to each research appointment, allowing a gap of up to 14 days between the research appointment and the weekly check-in).

Patterns of cross-sectional and longitudinal associations between percentages of days abstinent and MOBC domains assessed using both modalities (ie, weekly check-in and the longer,

psychometrically validated measures completed at research appointments) were also examined. Cross-sectional analyses examined whether the pattern of Pearson correlations between the 6-month percentage of days abstinent and 6-month MOBC domains measured using both modalities were similar in magnitude and direction. Longitudinal analyses similarly examined whether the pattern of Pearson correlations between *changes in* percentage of days abstinent from baseline to 6 months and *changes in* MOBC over a matched period were similar in magnitude and direction for the MOBC measures obtained using both modalities.

Although the statistical power in Pearson correlation analyses was limited owing to the small sample size of this clinical pilot, the ability to test the aims of the study relied less on testing whether the correlation coefficients were statistically significant and more on examining whether the patterns of associations (effect sizes) observed for the weekly check-in measures were similar to the patterns of associations observed for the longer, psychometrically validated measures.

Results

Descriptive Statistics

A total of 30 patients were enrolled in this clinical pilot study. Of these, 8 (27%) patients were aged 18 to 34 years, 19 (63%) patients were aged 35 to 54 years, and 3 patients (10%) were aged ≥55 years. Furthermore, 53% (16/30) of the patients were male; 57% (17/30) of the patients were White, 13% (4/30) were Black, 7% (2/30) were American Indian or Alaska Native, 7% (2/30) were Hispanic or Latinx, and 17% (5/30) were from other racial or ethnic groups. The sample demographics were similar to those of the clinic's patient population. Of the 30 patients, 90% (27/30) of the patients reported that they were not employed, and approximately half (16/30, 53%) reported that they were unhoused, had transitional or temporary or other housing, or were living in a house that someone else owned or leased. Patients self-reported that they were receiving treatment for the use of stimulants (18/30, 60%), opioids (16/30, 53%), alcohol (15/30, 50%), cannabis (5/30, 17%), sedatives (4/30, 13%), and hallucinogens (1/30, 3%). Additional patient descriptive statistics are available in a previous study [21].

Only 1 patient elected to receive the weekly check-in reminder via email; the remaining 29 patients received it via an SMS text message. Patients completed a mean of 20.60 weekly check-ins (SD 5.54; 85.8% of the maximum 24 weekly check-ins available to each patient; range 4-24; total number of weekly check-ins=618). Internal timing mechanisms on the weekly check-in indicated that, on average, the weekly check-in took <5 minutes to complete. All participants (30/30, 100%) completed the baseline research appointments. All but 1 patient (29/30, 97%) completed the 24-week research appointment. The participant who did not complete the 24-week research appointment also did not complete the 6- or 12-week appointments and was excluded from the analyses involving change scores.

Within-Patient Variability in Weekly Substance Use and MOBC

Within-patient variability over the 24 weeks of completing the weekly check-in is depicted in Figures 1 and 2. Each graph shows a separate weekly check-in domain, with each of the 30 patients represented by a different letter on the x-axis. Each weekly check-in score is represented as a dot, and the ranges of the scores reported for each patient are represented as vertical lines. For example, on the "drank too much" domain (Figure 1, top panel), patient A provided responses ranging from "not at all" to "every day" over the 24-week period (higher within-patient variability), patient B provided responses ranging from "not at all" to "1 or 2 days" (lower within-patient variability), and patient C only provided responses of "not at all" (no within-patient variability). As shown in Figure 1, a total of 50% (15/30) of the patients reported no variability on the drinking domain, 47% (14/30) of the patients reported no variability on the drug use domain, and 33% (10/30) of the patients reported no variability in both the drinking and drug use domains. In contrast, all patients displayed at least some variability in MOBC domains reflecting craving, coping strategies, and abstinence self-efficacy; 90% (27/30) of patients displayed at least some variability in depression and 93% (28/30) of patients displayed at least some variability in therapeutic alliance.

Figure 1. Within-patient variability in weekly check-in domains reflecting alcohol use, drug use, coping strategies, and substance refusal self-efficacy. Higher scores indicate more alcohol use, drug use, coping strategy use, and abstinence self-efficacy. Each letter on the x-axis reflects an individual patient and each dot reflects a score from a single weekly check-in. The vertical lines reflect each patient's range of scores across repeated weekly check-ins.

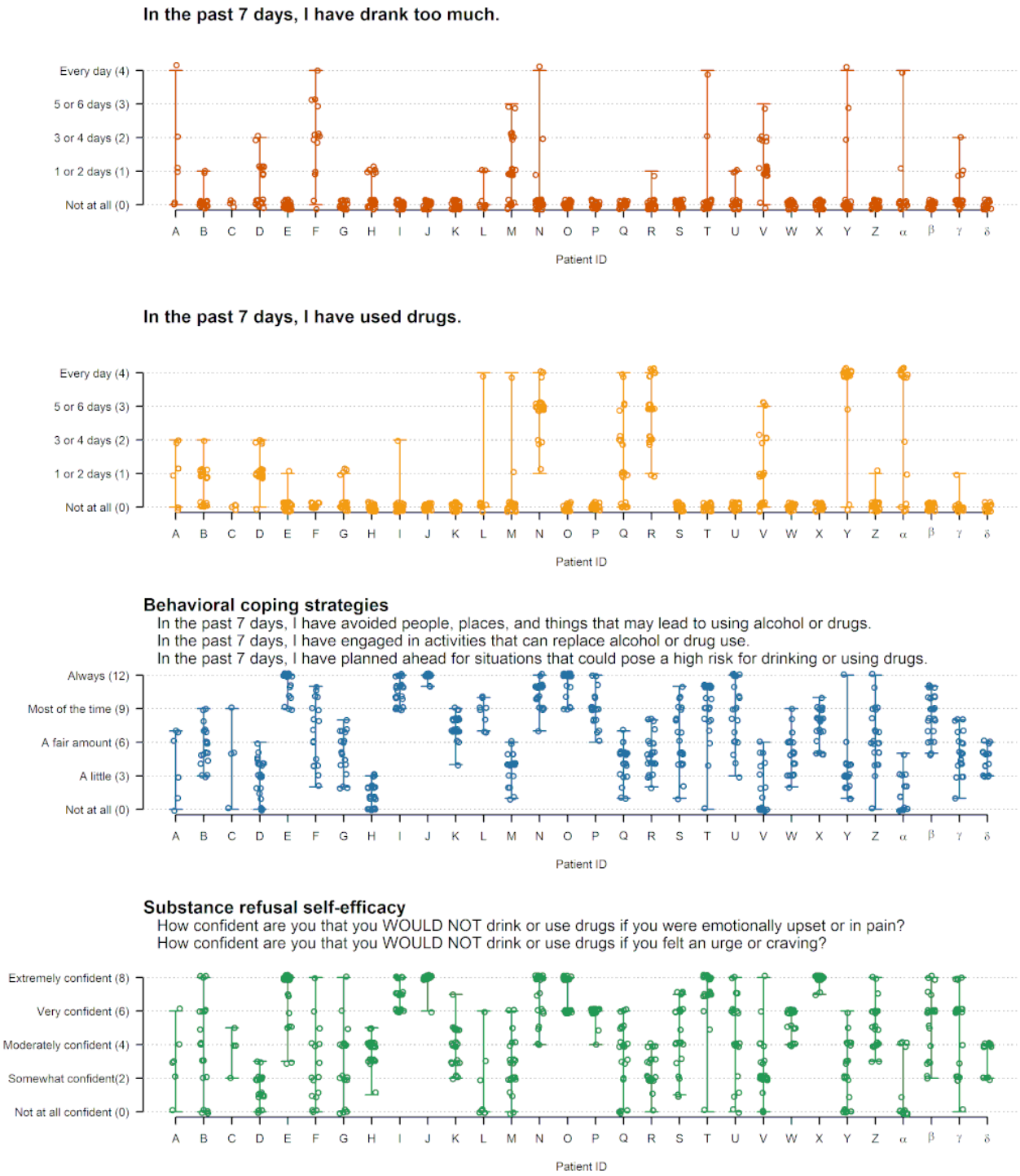
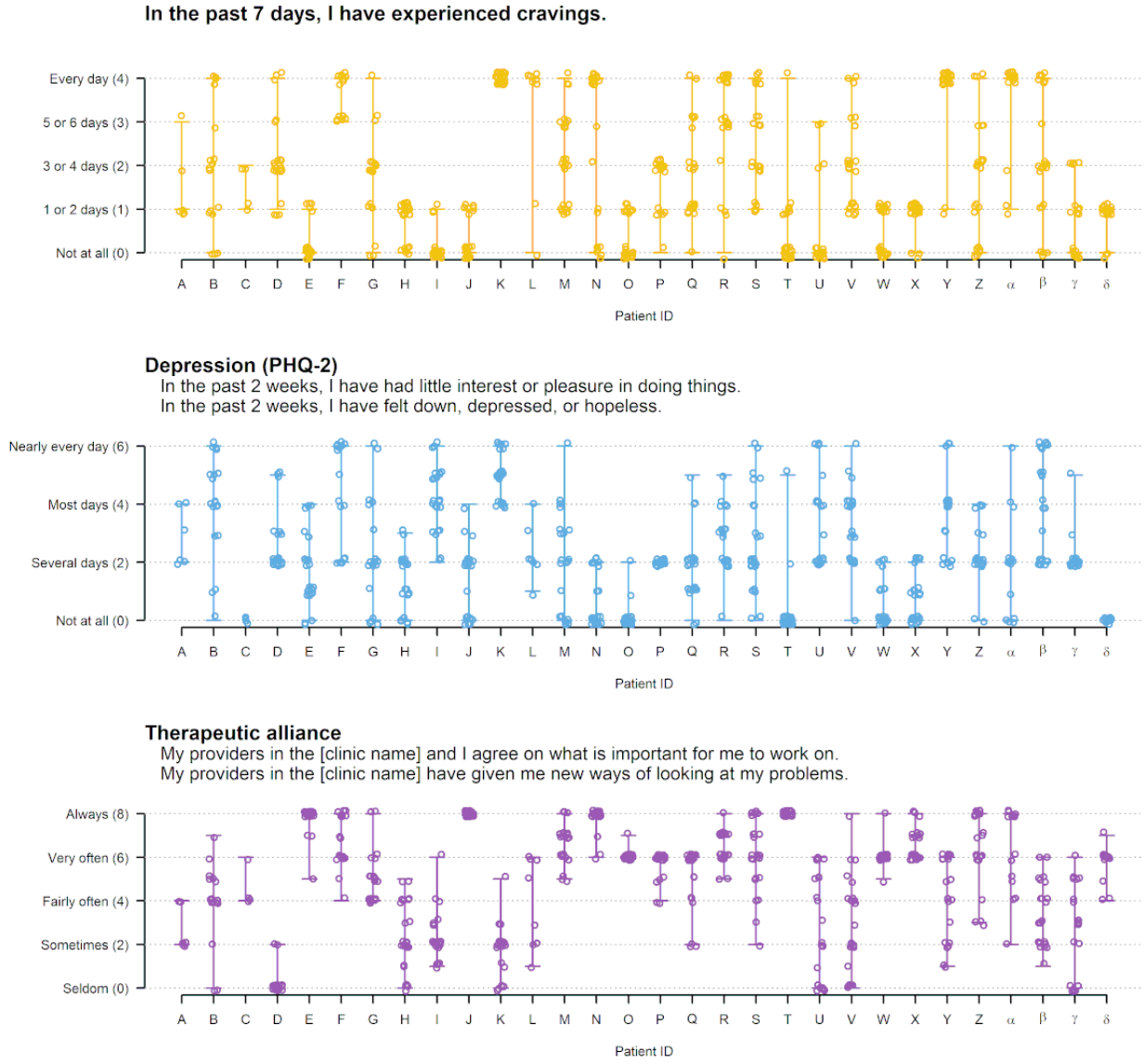


Figure 2. Within-patient variability in weekly check-in domains reflecting craving, depression, and therapeutic alliance. Higher scores indicate more frequent craving, higher self-efficacy, and higher therapeutic alliance. Each letter on the x-axis reflects an individual patient and each dot reflects a score from a single weekly check-in. The vertical lines reflect each patient's range of scores across repeated weekly check-ins. PHQ-2: Patient Health Questionnaire-2



Associations Between Weekly Check-in and Longer Psychometrically Validated Instruments

Scores on the first weekly check-in had significant cross-sectional correlations with scores obtained from longer, psychometrically validated instruments completed at baseline research appointments, with large effect sizes for domains reflecting substance use (r=0.51-0.64) and MOBC (r=0.57-0.85; Table 1). Correlations between *changes in* the weekly check-in measures and *changes in* the corresponding longer, psychometrically validated measures from baseline to 6 months were also significant, with large effect sizes (r=0.53-0.68) for domains reflecting *changes in* alcohol use, drug use, craving, depression, and abstinence self-efficacy. Correlations were significant, positive, and medium in size for *changes in* therapeutic alliance (r=0.42) but nonsignificant and smaller for *changes in* coping strategies (r=0.33).

Table 1. Correlations of substance use and mechanisms of behavioral change (MOBC) domains measured on brief, remotely completed weekly check-ins and substance use and MOBC domains measured on longer, psychometrically validated instruments completed at research appointments.

Weekly check-in domain	Cross-sectional correlations between domains measured on the weekly check-in and corresponding domains measured at research appointments at baseline			Longitudinal correlations between *changes in* domains measured on the weekly check-in and *changes in* corresponding domains measured at research appointments, from baseline to 6 months		
	r (95% CI)	Number of observations included in the analysis	*P* value	*r* (95% CI)	Number of observations included in the analysis	*P* value
Substance use						
Alcohol use	0.51 (0.19 to 0.74)	30	.004	0.63 (0.33 to 0.82)	27	<.001
Drug use	0.64 (0.37 to 0.82)	30	<.001	0.53 (0.19 to 0.75)	28	.004
MOBC						
Coping strategies[a]	0.57 (0.02 to 0.85)	13	.04	0.33 (−0.30 to 0.76)	12	.29
Abstinence self-efficacy	0.73 (0.51 to 0.87)	30	<.001	0.62 (0.33 to 0.81)	28	<.001
Craving	0.81 (0.63 to 0.91)	30	<.001	0.57 (0.24 to 0.79)	26	.002
Depression	0.85 (0.70 to 0.93)	30	<.001	0.68 (0.42 to 0.84)	28	<.001
Therapeutic alliance	0.72 (0.48 to 0.86)	30	<.001	0.42 (0.06 to 0.69)	28	.03

[a]A standardized questionnaire for coping strategies was only added to the research appointment assessment battery midway through the study because of initial concerns about potential assessment fatigue to patients (ie, the Coping Strategies Scale contains 59 items). Therefore, it was only available at baseline for the last 14 enrolled patients and at the 6-month follow-up for the last 16 enrolled patients, one of whom declined to complete it because of assessment fatigue at the research appointment.

Associations Between MOBC Measures and 6-Month Percentage of Days Abstinent

The 6-month percentage of days abstinent generally had similar patterns of association with MOBC measured via the weekly check-in and MOBC measured using longer, psychometrically validated measures at research appointments. In cross-sectional analyses at 6 months (Table 2: cross-sectional correlation between MOBC measure and percentage of days abstinent, both measured at 6 months), effect sizes of the associations between percentages of days abstinent and MOBC were similar in magnitude and direction for MOBC measured via the weekly check-in versus MOBC measured via longer, psychometrically validated measures completed at research appointments (ie, absolute differences in Pearson correlation effect sizes, *r*, were always <0.10) for coping strategies, abstinence self-efficacy, depression, and therapeutic alliance. However, for craving, the cross-sectional effect size was larger in magnitude for the modified Penn Alcohol Craving Scale completed at the research appointments compared with the single-item craving measure completed remotely on the weekly check-in. Although detecting statistical significance was not a primary objective of this analysis, 6-month percentages of days abstinent had significant cross-sectional associations with 6-month coping strategies measured on the weekly check-in, 6-month abstinence self-efficacy measured on the weekly check-in and at research appointments, and 6-month craving measured at research appointments.

For the analyses of change scores (Table 2: longitudinal correlation between changes in MOBC measure and changes in percentage of days abstinent, both examined from baseline to 6 months), the effect sizes of the associations between *changes in* percentages of days abstinent and *changes in* MOBC domains were similar when MOBC were measured via the weekly check-in and when MOBC were measured via longer, psychometrically validated measures completed at research appointments for all domains (ie, absolute difference in Pearson correlation effect size, *r*, was always <0.10). Although detecting statistical significance was not a primary objective of this analysis, *changes in* the percentages of days abstinent were significantly associated with *changes in* coping strategies measured on the weekly check-in and *changes in* therapeutic alliance measured on the weekly check-in and at research appointments.

Table 2. Correlations between percentages of days abstinent and mechanisms of behavioral change (MOBC) domains measured on brief, remotely completed weekly check-ins and MOBC domains measured on longer, psychometrically validated instruments completed at research appointments.

MOBC domain	Cross-sectional correlation between MOBC measure and percentage of days abstinent, both measured at 6 months			Longitudinal correlation between *changes in* MOBC measure and *changes in* percentage of days abstinent, both examined from baseline to 6 months		
	r (95% CI)	Number of observations included in the analysis	*P* value	*r* (95% CI)	Number of observations included in the analysis	*P* value
Coping strategies						
Weekly check-in	0.45 (0.09 to 0.70)	28	.02	0.42 (0.05 to 0.68)	28	.03
Research appointments[a]	0.40 (−0.14 to 0.76)	15	.14	0.36 (−0.27 to 0.77)	12	.25
Abstinence self-efficacy						
Weekly check-in	0.62 (0.33 to 0.81)	28	<.001	0.26 (−0.13 to 0.57)	28	.19
Research appointments	0.58 (0.27 to 0.79)	28	.001	0.26 (−0.12 to 0.58)	28	.18
Craving						
Weekly check-in	−0.25 (−0.57 to 0.13)	28	.19	−0.29 (−0.60 to 0.10)	27	.14
Research appointments	−0.58 (−0.78 to −0.25)	27	.002	−0.37 (−0.66 to 0.01)	27	.06
Depression						
Weekly check-in	−0.02 (−0.39 to 0.36)	28	.94	−0.26 (−0.58 to 0.12)	28	.18
Research appointments	−0.11 (−0.46 to 0.28)	28	.58	−0.29 (−0.60 to 0.09)	28	.13
Therapeutic alliance						
Weekly check-in	0.26 (−0.13 to 0.58)	28	.18	0.49 (0.14 to 0.73)[a]	28	.008
Research appointments	0.31 (−0.07 to 0.61)	28	.11	0.54 (0.21 to 0.76)[a]	28	.003

[a]A standardized questionnaire for coping strategies was only added to the research appointment assessment battery midway through the study because of initial concerns about potential assessment fatigue to patients (ie, the Coping Strategies Scale contains 59 items). Therefore, it was only available at baseline for the last 14 enrolled patients and at the 6-month follow-up for the last 16 enrolled patients, one of whom declined to complete the study due to assessment fatigue.

Discussion

Principal Findings

This study provided a preliminary evaluation of whether brief, weekly, remotely completed, patient-reported MOBC assessments could provide quantitatively meaningful information about changes in MOBC during community-based SUD treatment. The results of this pilot study provide preliminary support for our hypotheses; the weekly MOBC measures (1) varied considerably within patients over time and typically varied more than measures of substance use; (2) had large cross-sectional and longitudinal associations with longer, psychometrically validated MOBC measures completed at research appointments; and (3) had patterns of cross-sectional and longitudinal associations with percentages of days abstinent that were generally similar in magnitude and direction as longer, psychometrically validated MOBC measures completed at research appointments. The results of this pilot study provide promising preliminary support for the feasibility of remotely measuring multiple MOBC domains via a brief weekly patient self-report questionnaire for providing clinically meaningful information during SUD treatment as usual as it is offered in a community treatment setting.

There were 2 findings that did not align with the study's hypotheses; one involved the 3-item coping strategies measure, which had a small longitudinal association with the 59-item Coping Strategies Scale despite there being a high cross-sectional association between these 2 measures and similar patterns of association with the percentage of days abstinent. One potential reason for this discrepancy is that the Coping Strategies Scale's assessment of 59 specific cognitive behavioral coping strategies may capture more nuanced changes in coping behavior because it captures 59 specific behavioral changes that patients could make, in contrast to the 3 coping strategies measured on the weekly check-in that were intended to capture more general coping strategies that are applicable for a range of situations. Nonetheless, the similar cross-sectional and longitudinal associations of both coping measures with percentages of days abstinent suggest that both may capture important information about coping strategies that may help monitor patients' use of strategies that can help with avoiding substance use. The second finding that was inconsistent with the hypotheses involved the weekly check-in measure of craving, which had a nominally smaller cross-sectional association with the percentage of days abstinent at 6 months than the 5-item Penn Alcohol Craving Scale. A broader assessment of craving across multiple dimensions (eg, frequency and intensity) using the Penn Alcohol Craving Scale may provide a more complete

picture of patients' experiences with craving, particularly as craving relates to current substance use. However, notably, the single-item craving measure on the weekly check-in had large cross-sectional and longitudinal associations with the Penn Alcohol Craving Scale, suggesting that a single-item question on the weekly check-in may capture much of the same information as the Penn Alcohol Craving Scale, potentially warranting the use of a single-item craving measure when the benefits of a briefer assessment outweigh the potential costs of a longer measure (eg, if a 5-item measure is potentially burdensome to patients or impractical for a given clinical setting).

Previous studies have shown that abstinence self-efficacy, depression symptoms, and craving can be reliably measured using brief questionnaires [30,37,38]. Previous work has also shown that a brief, clinician-administered assessment measure that includes questions about MOBC can predict substance use outcomes when administered every 3 months in the context of an effectiveness trial [39]. This study builds on these findings and shows that brief, weekly, remotely completed, patient-reported MOBC assessments can provide quantitatively meaningful information about changes in MOBC when embedded in a measurement-based care system that is added to community-based SUD treatment. Measuring MOBC briefly, remotely, and using patient self-reports may increase the feasibility of longitudinal, multidimensional measurement-based care in SUD treatment. The resulting information may help clinicians monitor multiple dimensions that could help indicate whether treatment is affecting the MOBC, which are expected to improve during evidence-based treatments and predict long-term substance use outcomes. Monitoring in such a manner could help clinicians obtain frequent information to support measurement-based care and could complement less frequent clinician-administered assessments [39] and patients' narrative reports about their treatment progress.

Monitoring MOBC could also potentially help guide clinical attention toward the importance of MOBC as pertinent treatment targets; for example, by reminding patients that treatment can address coping strategies, self-efficacy, craving, and depression symptoms—not just alcohol and drug consumption [40,41]. The results from this study suggest that there may be a particular utility in measuring hypothesized MOBC in SUD treatment rather than focusing measurement specifically on alcohol and drug *consumption*, given the tendency for MOBC to vary over time more often than measures of substance use. Measures of *substance use* are typically specified as primary treatment outcome measures in SUD clinical trials [42,43] and are among the most common outcomes for patients and clinicians to focus their attention on during SUD treatment; for example, clinicians often check in with their patients whether they are currently using substances, how often they are using them, or how long they have gone without using substances. In contrast, MOBC (eg, coping strategies, abstinence self-efficacy, craving, depression symptoms, and therapeutic alliance) may be discussed less frequently in routine care despite representing highly pertinent domains that can cause distress (eg, craving, depression), motivate treatment-seeking (eg, low self-efficacy or limited coping strategies), signal risk for treatment dropout (eg, low therapeutic alliance [44]), and are often directly addressable with interventions that clinicians can offer during sessions (eg, practicing new coping skills, helping patients obtain medications for craving or depression, and clarifying reasons for poor therapeutic alliance).

Limitations and Strengths

This study had several limitations. Because this was a pilot study, the sample was intentionally small, which limited the statistical power to detect statistical significance in some correlational analyses. In addition, the 59-item Coping Strategies Scale was only added to the assessment batteries for the last 14 enrolled patients owing to concerns about it creating assessment fatigue; therefore, the statistical power was further limited for analyses involving this measure. Despite the small sample size, some associations were found to be statistically significant. Moreover, the overarching goal of this pilot study was to compare patterns of association between the weekly check-in and longer, psychometrically validated measures rather than to test whether any given association was statistically significant. This study was conducted in a single urban addiction and mental health clinic, and future studies are needed to evaluate the performance of tools that briefly and longitudinally assess MOBC in additional community-based treatment settings, including those with varying service models, workflows, and patient populations. The brief questionnaires used in this study were not previously validated, and this study lacked an adequate sample size to perform analyses that focused on psychometric validation. However, the questions were derived from longer, validated measures, and this study is still able to show proof-of-concept that measuring MOBC using brief, weekly, remotely completed, patient-report questionnaires can potentially provide clinically meaningful information.

This study also had noteworthy strengths. The remote monitoring system tested here was developed with multiple rounds of patient and clinician inputs. It was tested as an add-on to SUD treatment as it is usually offered in a community setting, which provides high external validity, and it was found to be engaging, usable, and clinically helpful by patients and clinicians [21]. Thus, the tested system may be feasible to incorporate into other community settings. Follow-up rates were high. Despite the smaller sample size for the pilot study, the frequency of measurement yielded a large number of data points (over 600 weekly check-ins completed, over 20 repeated measures per patient on average). The sample was also diverse in terms of gender and race.

Conclusions

Measuring MOBC frequently and remotely using brief patient-report measures may be a viable approach for obtaining clinically meaningful information about changes in MOBC during treatment. Implementing systems that facilitate measurement and monitoring MOBC longitudinally during SUD treatment could be one approach to support the delivery of MOBC-informed care in community treatment settings.

Acknowledgments

This research was supported by the National Institute on Alcohol Abuse and Alcoholism of the National Institutes of Health under award number . The content is solely the responsibility of the authors and does not necessarily represent the official views of the Alcohol Abuse and Alcoholism or the National Institutes of Health.

Data Availability

The data sets generated during this study are not publicly available because consent was not obtained from participants to share the research data publicly; however, the data are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Questions included in the weekly check-in.
[DOCX File , 500 KB - formative_v6i11e42376_app1.docx]

References

1. Park-Lee E, Lipari RN, Hedden SL, Kroutil LA, Porter JD. Receipt of Services for Substance Use and Mental Health Issues Among Adults: Results from the 2016 National Survey on Drug Use and Health. Rockville, MD, USA: Substance Abuse and Mental Health Services Administration (US); Sep 2017.
2. Klein JW. Pharmacotherapy for substance use disorders. Med Clin North Am 2016 Jul;100(4):891-910. [doi: 10.1016/j.mcna.2016.03.011] [Medline: 27235620]
3. Magill M, Ray LA. Cognitive-behavioral treatment with adult alcohol and illicit drug users: a meta-analysis of randomized controlled trials. J Stud Alcohol Drugs 2009 Jul;70(4):516-527 [FREE Full text] [doi: 10.15288/jsad.2009.70.516] [Medline: 19515291]
4. Hallgren KA, Wilson AD, Witkiewitz K. Advancing analytic approaches to address key questions in mechanisms of behavior change research. J Stud Alcohol Drugs 2018 Mar;79(2):182-189 [FREE Full text] [doi: 10.15288/jsad.2018.79.182] [Medline: 29553344]
5. Longabaugh R, Magill M. Recent advances in behavioral addiction treatments: focusing on mechanisms of change. Curr Psychiatry Rep 2011 Oct;13(5):382-389 [FREE Full text] [doi: 10.1007/s11920-011-0220-4] [Medline: 21750958]
6. Connors GJ, Maisto SA, Schlauch RC, Dearing RL, Prince MA, Duerr MR. Therapeutic alliances predict session by session drinking behavior in the treatment of alcohol use disorders. J Consult Clin Psychol 2016 Nov;84(11):972-982 [FREE Full text] [doi: 10.1037/ccp0000130] [Medline: 27548032]
7. Epstein EE, McCrady BS, Hallgren KA, Gaba A, Cook S, Jensen N, et al. Individual versus group female-specific cognitive behavior therapy for alcohol use disorder. J Subst Abuse Treat 2018 May;88:27-43 [FREE Full text] [doi: 10.1016/j.jsat.2018.02.003] [Medline: 29606224]
8. Glasner-Edwards S, Tate SR, McQuaid JR, Cummins K, Granholm E, Brown SA. Mechanisms of action in integrated cognitive-behavioral treatment versus twelve-step facilitation for substance-dependent adults with comorbid major depression. J Stud Alcohol Drugs 2007 Sep;68(5):663-672. [doi: 10.15288/jsad.2007.68.663] [Medline: 17690799]
9. Hallgren KA, Holzhauer CG, Epstein EE, McCrady BS, Cook S. Optimizing the length and reliability of measures of mechanisms of change to support measurement-based care in alcohol use disorder treatment. J Consult Clin Psychol 2021 Apr;89(4):277-287 [FREE Full text] [doi: 10.1037/ccp0000643] [Medline: 34014690]
10. Helstrom AW, Blow FC, Slaymaker V, Kranzler HR, Leong S, Oslin D. Reductions in alcohol craving following naltrexone treatment for heavy drinking. Alcohol Alcohol 2016 Sep;51(5):562-566. [doi: 10.1093/alcalc/agw038] [Medline: 27402770]
11. Johnson JE, Finney JW, Moos RH. End-of-treatment outcomes in cognitive-behavioral treatment and 12-step substance use treatment programs: do they differ and do they predict 1-year outcomes? J Subst Abuse Treat 2006 Jul;31(1):41-50. [doi: 10.1016/j.jsat.2006.03.008] [Medline: 16814009]
12. Kruger ES, Serier KN, Pfund RA, McKay JR, Witkiewitz K. Integrative data analysis of self-efficacy in 4 clinical trials for alcohol use disorder. Alcohol Clin Exp Res 2021 Nov;45(11):2347-2356. [doi: 10.1111/acer.14713] [Medline: 34523721]
13. Prince MA, Connors GJ, Maisto SA, Dearing RL. Within treatment therapeutic alliance ratings profiles predict posttreatment frequency of alcohol use. Psychol Addict Behav 2016 Mar;30(2):184-193 [FREE Full text] [doi: 10.1037/adb0000161] [Medline: 26999350]
14. Roos CR, Carroll KM, Nich C, Frankforter T, Kiluk BD. Short- and long-term changes in substance-related coping as mediators of in-person and computerized CBT for alcohol and drug use disorders. Drug Alcohol Depend 2020 Jul 01;212:108044 [FREE Full text] [doi: 10.1016/j.drugalcdep.2020.108044] [Medline: 32422538]
15. Roos CR, Witkiewitz K. A contextual model of self-regulation change mechanisms among individuals with addictive disorders. Clin Psychol Rev 2017 Nov;57:117-128 [FREE Full text] [doi: 10.1016/j.cpr.2017.08.008] [Medline: 28866435]

16. Wilcox CE, Pearson MR, Tonigan JS. Effects of long-term AA attendance and spirituality on the course of depressive symptoms in individuals with alcohol use disorder. Psychol Addict Behav 2015 Jun;29(2):382-391 [FREE Full text] [doi: 10.1037/adb0000053] [Medline: 26076099]
17. Wilcox CE, Tonigan JS. Changes in depression mediate the effects of AA attendance on alcohol use outcomes. Am J Drug Alcohol Abuse 2018;44(1):103-112 [FREE Full text] [doi: 10.1080/00952990.2016.1249283] [Medline: 27892692]
18. Witkiewitz K, Bowen S, Douglas H, Hsu SH. Mindfulness-based relapse prevention for substance craving. Addict Behav 2013 Feb;38(2):1563-1571 [FREE Full text] [doi: 10.1016/j.addbeh.2012.04.001] [Medline: 22534451]
19. Wolitzky-Taylor K, Schiffman J. Predictive associations among the repeated measurements of anxiety, depression, and craving in a dual diagnosis program. J Dual Diagn 2019;15(3):140-146. [doi: 10.1080/15504263.2019.1589660] [Medline: 30982462]
20. Epstein EE, McCrady BS. Overcoming Alcohol Use Problems: A Cognitive-Behavioral Treatment Program. Oxford, UK: Oxford University Press; 2009.
21. Hallgren KA, Cohn EB, Ries RK, Atkins DC. Delivering remote measurement-based care in community addiction treatment: engagement and usability over a 6-month clinical pilot. Front Psychiatry 2022 Apr 7;13:840409 [FREE Full text] [doi: 10.3389/fpsyt.2022.840409] [Medline: 35463505]
22. Tauscher JS, Cohn EB, Johnson TR, Diteman KD, Ries RK, Atkins DC, et al. What do clinicians want? Understanding frontline addiction treatment clinicians' preferences and priorities to improve the design of measurement-based care technology. Addict Sci Clin Pract 2021 Jun 15;16(1):38 [FREE Full text] [doi: 10.1186/s13722-021-00247-5] [Medline: 34130724]
23. Epstein EE, Drapkin ML, Yusko DA, Cook SM, McCrady BS, Jensen NK. Is alcohol assessment therapeutic? Pretreatment change in drinking among alcohol-dependent women. J Stud Alcohol 2005 May;66(3):369-378. [doi: 10.15288/jsa.2005.66.369] [Medline: 16047526]
24. Hallgren KA, McCrady BS, Epstein EE. Trajectories of drinking urges and the initiation of abstinence during cognitive-behavioral alcohol treatment. Addiction 2016 May;111(5):854-865 [FREE Full text] [doi: 10.1111/add.13291] [Medline: 26709608]
25. Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. Arch Intern Med 1998 Sep 14;158(16):1789-1795. [doi: 10.1001/archinte.158.16.1789] [Medline: 9738608]
26. Hearon BA, Pierce CL, Björgvinsson T, Fitzmaurice GM, Greenfield SF, Weiss RD, et al. Improving the efficiency of drug use disorder screening in psychiatric settings: validation of a single-item screen. Am J Drug Alcohol Abuse 2015 Mar;41(2):173-176. [doi: 10.3109/00952990.2015.1005309] [Medline: 25700005]
27. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
28. Litt MD, Kadden RM, Cooney NL, Kabela E. Coping skills and treatment outcomes in cognitive-behavioral and interactional group therapy for alcoholism. J Consult Clin Psychol 2003 Feb;71(1):118-128. [doi: 10.1037//0022-006x.71.1.118] [Medline: 12602432]
29. Breslin FC, Sobell LC, Sobell MB, Agrawal S. A comparison of a brief and long version of the Situational Confidence Questionnaire. Behav Res Ther 2000 Dec;38(12):1211-1220. [doi: 10.1016/s0005-7967(99)00152-7] [Medline: 11104185]
30. Kroenke K, Spitzer RL, Williams JB. The Patient Health Questionnaire-2: validity of a two-item depression screener. Med Care 2003 Nov;41(11):1284-1292. [doi: 10.1097/01.MLR.0000093487.78664.3C] [Medline: 14583691]
31. Munder T, Wilmers F, Leonhart R, Linster HW, Barth J. Working Alliance Inventory-Short Revised (WAI-SR): psychometric properties in outpatients and inpatients. Clin Psychol Psychother 2010;17(3):231-239. [doi: 10.1002/cpp.658] [Medline: 20013760]
32. Neale J, Vitoratou S, Finch E, Lennon P, Mitcheson L, Panebianco D, et al. Development and validation of 'sure': a patient reported outcome measure (PROM) for recovery from drug and alcohol dependence. Drug Alcohol Depend 2016 Aug 01;165:159-167 [FREE Full text] [doi: 10.1016/j.drugalcdep.2016.06.006] [Medline: 27344196]
33. Hallgren KA, Diteman KD, Atkins DA. Designing measurement-based care technology for frontline addiction treatment settings: usability testing and iterative design improvement. Alcohol Clin Exp Res 2020 Jun 1;44:154 [FREE Full text] [doi: 10.1111/acer.14358]
34. Cacciola JS, Alterman AI, McLellan AT, Lin YT, Lynch KG. Initial evidence for the reliability and validity of a "Lite" version of the Addiction Severity Index. Drug Alcohol Depend 2007 Mar 16;87(2-3):297-302. [doi: 10.1016/j.drugalcdep.2006.09.002] [Medline: 17045423]
35. Flannery BA, Volpicelli JR, Pettinati HM. Psychometric properties of the Penn Alcohol Craving Scale. Alcohol Clin Exp Res 1999 Aug;23(8):1289-1295. [doi: 10.1111/j.1530-0277.1999.tb04349.x] [Medline: 10470970]
36. 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] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]

37. Hoeppner BB, Kelly JF, Urbanoski KA, Slaymaker V. Comparative utility of a single-item versus multiple-item measure of self-efficacy in predicting relapse among young adults. J Subst Abuse Treat 2011 Oct;41(3):305-312 [FREE Full text] [doi: 10.1016/j.jsat.2011.04.005] [Medline: 21700411]
38. McHugh RK, Fitzmaurice GM, Griffin ML, Anton RF, Weiss RD. Association between a brief alcohol craving measure and drinking in the following week. Addiction 2016 Jun;111(6):1004-1010 [FREE Full text] [doi: 10.1111/add.13311] [Medline: 26780476]
39. Van Horn DH, Goodman J, Lynch KG, Bonn-Miller MO, Thomas T, Del Re AC, et al. The predictive validity of the progress assessment, a clinician administered instrument for use in measurement-based care for substance use disorders. Psychiatry Res 2020 Oct;292:113282 [FREE Full text] [doi: 10.1016/j.psychres.2020.113282] [Medline: 32711168]
40. Hallgren KA, Bauer AM, Atkins DC. Digital technology and clinical decision making in depression treatment: current findings and future opportunities. Depress Anxiety 2017 Jun;34(6):494-501 [FREE Full text] [doi: 10.1002/da.22640] [Medline: 28453916]
41. Scott K, Lewis CC. Using measurement-based care to enhance any treatment. Cogn Behav Pract 2015 Feb;22(1):49-59 [FREE Full text] [doi: 10.1016/j.cbpra.2014.01.010] [Medline: 27330267]
42. Robinson SM, Sobell LC, Sobell MB, Arcidiacono S, Tzall D. Alcohol and drug treatment outcome studies: new methodological review (2005-2010) and comparison with past reviews. Addict Behav 2014 Jan;39(1):39-47. [doi: 10.1016/j.addbeh.2013.09.029] [Medline: 24140306]
43. Witkiewitz K, Finney JW, Harris AH, Kivlahan DR, Kranzler HR. Recommendations for the design and analysis of treatment trials for alcohol use disorders. Alcohol Clin Exp Res 2015 Sep;39(9):1557-1570 [FREE Full text] [doi: 10.1111/acer.12800] [Medline: 26250333]
44. Meier PS, Barrowclough C, Donmall MC. The role of the therapeutic alliance in the treatment of substance misuse: a critical review of the literature. Addiction 2005 Mar;100(3):304-316. [doi: 10.1111/j.1360-0443.2004.00935.x] [Medline: 15733244]

Abbreviations

MOBC: mechanisms of behavioral change
PHQ: Patient Health Questionnaire
SUD: substance use disorder

Edited by A Mavragani; submitted 01.09.22; peer-reviewed by M Kapsetaki, V Palzes; comments to author 27.09.22; revised version received 10.10.22; accepted 11.10.22; published 07.11.22.

Please cite as:
Hallgren KA
Remotely Assessing Mechanisms of Behavioral Change in Community Substance Use Disorder Treatment to Facilitate Measurement-Informed Care: Pilot Longitudinal Questionnaire Study
JMIR Form Res 2022;6(11):e42376
URL: https://formative.jmir.org/2022/11/e42376
doi:10.2196/42376
PMID:36342773

Original Paper

Behavior Change Training for Health Professionals: Evaluation of a 2-Hour Workshop

Barbara Mullan[1,2,3*], PhD; Caitlin Liddelow[4*], PhD; Darren Haywood[1,3,5*], BPsych (Hons); Hayley Breare[1,3*], BPsych (Hons)

[1]enAble Institute, Faculty of Health Sciences, Curtin University, Bentley, Australia

[2]Western Australian Cancer Prevention Research Unit, Curtin University, Bentley, Australia

[3]School of Population Health, Curtin University, Bentley, Australia

[4]School of Psychology, University of Wollongong, Wollongong, Australia

[5]Department of Mental Health, St Vincent's Hospital Melbourne, Fitzroy, Australia

[*]all authors contributed equally

Corresponding Author:
Barbara Mullan, PhD
enAble Institute
Faculty of Health Sciences
Curtin University
Kent Street
Bentley, 6102
Australia
Phone: 61 08 9266 2468
Email: barbara.mullan@curtin.edu.au

Abstract

Background: Rates of noncommunicable diseases continue to rise worldwide. Many of these diseases are a result of engaging in risk behaviors. Without lifestyle and behavioral intervention, noncommunicable diseases can worsen and develop into more debilitating diseases. Behavioral interventions are an effective strategy to reduce the burden of disease. Behavior change techniques can be described as the “active ingredients” in behavior change and address the components that need to be altered in order for the target behavior to change. Health professionals, such as pharmacists and nurses, can engage in opportunistic behavior change with their patients, to encourage positive health behaviors.

Objective: We aimed to develop, implement, and evaluate a behavior change workshop targeted at health professionals in Australia, with the goal of increasing knowledge of behavior change techniques and psychological variables.

Methods: A prospective study design was used to develop and evaluate a 2-hour behavior change workshop targeted at health professionals. The workshop was developed based on the Capability, Opportunity, Motivation, and Behavior Model and had five core objectives: (1) to detail the role of health professionals in delivering optimal care, (2) to demonstrate opportunities to change behavior, (3) to describe principles of behavior change, (4) to explain behavior change techniques, and (5) to determine the most appropriate behavior change techniques to use and when to use them. A total of 10 workshops were conducted. To evaluate the workshops and identify any potential long-term changes in behavior, we collected pre- and postworkshop data on knowledge and psychological constructs from the attendees.

Results: A final sample of 41 health professionals comprising general practitioners, nurses, and pharmacists completed the pre- and postworkshop surveys. Following the workshops, there were significant improvements in knowledge of behavior change techniques (t_{40}=–5.27, P<.001), subjective norms (t_{40}=–3.49, P=.001), descriptive norms (t_{40}=–3.65, P<.001), perceived behavioral control (t_{40}=–3.30, P=.002), and intention (t_{36}=–3.32, P=.002); each had a large effect size. There was no significant difference in postworkshop attitude (t_{40}=0.78, P=.44). The participants also found the workshops to be highly acceptable.

Conclusions: A 2-hour, theoretically informed workshop designed to facilitate the use of behavior change techniques by health professionals was shown to be largely effective. The workshops resulted in increases in knowledge, descriptive and subjective norms, perceived behavioral control, and intention, but not in attitude. The intervention was also shown to be highly acceptable, with the large majority of participants deeming the intervention to be needed, useful, appropriate, and applicable, as well as

interesting and worth their time. Future research should examine the lasting impacts of the workshop on health professionals' practices.

(JMIR Form Res 2022;6(11):e42010) doi:10.2196/42010

KEYWORDS

behaviour change; psychology; psychological; BCT; health professional; health care professional; medical education; health care provider; continuing education; professional development; theory of planned behaviour; COM-B; workshop; intervention; clinical practice

Introduction

Rates of noncommunicable diseases such as cancer, coronary heart disease, and diabetes continue to rise worldwide and account for 71% of deaths globally [1]. Many of these diseases are a result of engaging in unhealthy behaviors such as smoking, excessive alcohol consumption, and being sedentary. Cancer is one of many major chronic diseases that impacts individuals, families, communities, and economies. In Australia, 42% of the cancer burden is attributable to personal attributes and engagement with risk behaviors [2]. In 2021, it was estimated that more than 1 million individuals were impacted by either living with or having lived with cancer in Australia alone. Without lifestyle and behavioral intervention, noncommunicable diseases can worsen and develop into more debilitating diseases. For many chronic diseases, using behavioral interventions can be an effective strategy for reducing the burden of disease in Australia [3,4].

There are a range of theoretical frameworks that have been used to understand and predict behavior and inform the development of behavioral interventions. One popular framework is the Capability, Opportunity, Motivation, and Behavior (COM-B) model [5]. The COM-B model provides a framework to inform the choice and use of behavioral interventions. The COM-B model proposes that 3 components are integral in the production, and thereby changing, of behavior: capability (ie, a person's perception of whether they are physically and psychologically able to complete the behavior), opportunity (ie, the physical and social opportunity to complete the behavior), and motivation (ie, the desire or need to complete the behavior over competing behaviors) [5]. A range of interventions have successfully used behavior change approaches targeted at improving capability, opportunity, or motivation with the goal of changing behavior [6]. The application of behavior change approaches can be implemented at the population level, through large-scale interventions, or on the individual level, through opportunistic implementation.

Recently, there have been calls for health professionals, such as pharmacists and nurses, to engage in opportunistic behavior change [7]. Public health policies from the National Health Service of the United Kingdom, such as Make Every Contact Count, promote the use of behavior change strategies and interventions by health professionals to engage with patients who may need additional assistance managing their health [8]. Keyworth et al [7] found that in practice, health professionals did not deliver opportunistic behavior change interventions on 50% of occasions when they were perceived as necessary. Health professionals reported a number of barriers to delivery of these interventions, such as beliefs about their capability and the consequences of these interactions; views on their professional role and identity; and more discipline-specific barriers, such as prioritization, time pressures, and workload pressures [9].

Among approaches that have a strong evidence base for successfully promoting healthy behaviors among patients and changing their behavior [10,11] are behavior change techniques [12]. Behavior change techniques can be described as the "active ingredients" in behavior change and often focus on the components that need to be altered in order for the target behavior to change. Behavior change techniques have been successfully used in previous interventions in both patients and health professionals [13]. A 2009 meta-analysis showed that interventions aimed at patients were more effective in improving medication adherence when active behavior change techniques were used [14]. Similarly, health care provider–led interventions that incorporated behavior change techniques, such as educating patients on the consequences of nonadherence, were successful in increasing the odds of medication adherence in patients with acute coronary syndrome by 54% [15]. Finally, a systematic review of reviews showed that behavior change interventions aimed at changing the practice behavior of health professionals were successful when they were interactive and multifaceted [16].

The aim of this study was to develop, implement, and evaluate a 2-hour behavior change workshop based on the COM-B model, targeted at health professionals in Western Australia. More specifically, the workshop was targeted at health professionals who treat or encounter patients experiencing various chronic diseases. The Theory of Planned Behavior variables [17] was used to obtain a deeper understanding of health professionals' attitudes (ie, how positively or negatively they view engaging in a behavior), subjective norms (ie, how likely it is that others similar to them engage or believe they should engage in the behavior), perceived behavioral control (ie, how much control they have over engaging in the behavior), and intention (ie, whether they intend to change their behavior) regarding behavior change techniques and using them in practice. This theory is widely used in behavior research and suggests that the combination of attitudes, subjective norms, and perceived behavioral control predicts the intention to engage in a behavior. Intention and perceived behavioral control are then said to significantly predict engagement in the target behavior [17]. Through the piloting of the workshops, we hypothesized that health professionals would report a pre- to postworkshop increase in knowledge of behavior change techniques and the Theory of Planned Behavior variables. These

variables are commonly used to assess and inform the likelihood of people engaging in target behaviors following participation in an intervention [18-21]. The acceptability of the workshop from the point of view of the health professionals was also assessed as part of the evaluation.

Methods

Workshop

The interactive workshop was developed based on the COM-B model [5] and provided information on the importance and use of behavior change techniques that could be used in everyday practice for opportunistic behavior change. There were five core objectives of the workshop: (1) to detail the role of health professionals in delivering optimal care; (2) to demonstrate opportunities to change behavior; (3) to describe the principles of behavior change; (4) to explain behavior change techniques; and (5) to determine the most appropriate behavior change techniques to use and when to use them.

The workshops were conducted in collaboration with the Pharmaceutical Society of Australia (PSA), Western Australia Division. Health professionals involved in the care of chronic disease patients, specifically those working with patients with heart failure, were invited to attend a professional development day organized by the PSA. The first component of the professional development day was run by a cardiologist and focused on educating the attendees on the medical aspects of heart failure. This was presented separately from our workshop and was not part of the intervention evaluation. The second and final component of the professional development day was the 2-hour behavior change workshop, which was facilitated by 3 of the authors. If attendees provided informed consent to participate in the research evaluation of the workshop, they completed a baseline survey before the workshop and a follow-up survey immediately after the workshop.

The workshop was presented in 3 overall sections (see Multimedia Appendix 1 for the workshop schedule). The first section introduced the COM-B model and how it can be applied in practice. An interactive case-study activity was then introduced. The case study provided an example of a woman who did not want to adhere to hypertension medication and showed how the COM-B model could be used to identify barriers and facilitators to behavior change. Each attendee spent approximately 10 minutes working on the case study, with the group reconvening after this time to debrief and share responses. The second section focused on why behavior is hard to change and highlighted some of the difficulties health professionals encounter with their own behavior and attitudes when trying to change the behavior of their patients. The third and final part of the workshop introduced behavior change techniques and provided a more in-depth overview of certain groupings of behavior change techniques that may be appropriate in practice. Seven individual groups of behavior change techniques were discussed: social support, self-monitoring of behavior, verbal persuasion for capability, focus on past success, planning, attitude change, and automaticity. These groups were selected using Cards for Change, a toolkit for behavior change developed by researchers at the University of Manchester and Manchester Metropolitan University [22]. These cards provide clear definitions of each technique and supporting example activities that can assist educators and trainers in teaching behavior change techniques.

After discussing the behavior change groupings, attendees were randomly placed into small groups and instructed to “choose two of the behavior change techniques that have just been introduced, explore when they might work best, and what the barriers to implementation may be.” After approximately 10 minutes, the attendees reconvened, and the workshop presenters facilitated debriefing and sharing of responses. The final activity was another group discussion activity, which asked attendees, in small groups, to complete the following task: “using the 12 behavior change techniques that have been introduced, explore what combinations might work best and in what circumstances.” A final debriefing was then facilitated by the workshop presenters.

Study Design and Procedure

A prospective study design was used to implement and evaluate the behavior change workshop. A total of 10 workshops were conducted between September and November 2021, with 6 held in person in metropolitan Perth and 4 held online for those in regional areas. All workshop attendees were contacted by the PSA through its database of health professionals and were invited to attend the free professional development day. When signing up for the professional development day, all attendees provided basic demographic information to the PSA.

Prior to participating in the 2-hour workshop, all attendees were provided with a link to the survey, hosted on Qualtrics. The first page of the survey provided a participant information sheet and a consent form. Attendees that were interested in participating in the evaluation of the workshop were asked to provide informed consent for both time points after reading the participant information sheet. After providing consent, participants were asked to provide their first name and email address, so that their baseline and postworkshop surveys could be linked. No other demographic information was collected from participants. After this, participants completed questions related to their current knowledge of behavior change techniques, current use of behavior change techniques in practice, and the psychological variables. All attendees then participated in the workshop, and at the end they were provided another Qualtrics link to the follow-up survey. This survey was the same as the preworkshop survey, but had additional items related to the acceptability of the workshop.

Ethical Considerations

Ethics approval was obtained from the Human Research Ethics Committee of the Curtin University, Australia (HRE2021-0567) prior to any data collection or facilitation of the workshops. Participants provided informed consent by marking a checkbox at the start of the preworkshop survey. Survey data were deidentified following the merging of participants’ responses from the pre- and postworkshop surveys. Participants were not provided compensation for their time, as the professional development workshop was provided to participants free of charge through the PSA.

Measures

All psychosocial measures were based on the standardized procedures for measure development outlined by Ajzen [17]. This included defining the behavior and research population and formulating reflective and direct measures to address each of the main constructs of the Theory of Planned Behavior [23].

Perceived Knowledge

Participants' perceived knowledge of behavior change techniques was assessed and measured both before and after the workshop using a single item: "On a scale of no understanding to perfect understanding, how would you rate your knowledge of behavior change techniques?" Participants rated their level of understanding on a 7-point Likert scale. Scores were summed to yield a total perceived knowledge score. Higher scores indicated greater perceived knowledge.

Attitudes

To measure attitudes toward behavior change techniques in practice, participants were provided with the single phrase "For me, changing my professional practice to reduce the effects of heart failure would be..." and were asked to complete this item for 2 attitudes, wisdom and usefulness, with responses on a sliding 7-point Likert scale, ranging from 1 (very wise/very useful) to (7 very unwise/very useless). Items were reverse scored and responses to the 2 items were averaged. Higher scores indicated positive attitudes toward behavior change techniques. The attitude measure demonstrated excellent internal consistency before and after the workshop (α=.98 and α=.95, respectively).

Social and Descriptive Norms

Two items were developed to assess social and descriptive norms related to changes in professional practice to incorporate behavior change techniques. These items were provided both pre- and postworkshop. Participants rated their agreement with each statement on a 7-point Likert scale, ranging from 1 (strongly disagree) to 7 (strongly agree). For descriptive norms, the item was "The people who are important to me think I should change my professional practice to reduce the effects of heart failure." For subjective norms, the item was "People like me think I should change my professional practice to reduce the effects of heart failure." Agreement with each norm was represented by a higher score.

Perceived Behavioral Control

Perceived behavioral control was measured using 1 item developed for this study. Participants rated the item "I am confident I can change my professional practice to reduce the effects of heart failure" on a 7-point Likert scale, ranging from 1 (strongly disagree) to 7 (strongly agree). Perceived behavioral control was measured both before and after the workshop. Higher perceived behavioral control to change behavior in professional practice was indicated by a higher score.

Intention

A single item was used to assess participants' intention to use behavior change techniques over the next 4 weeks: "I intend to provide behavior change techniques to my heart failure patients over the next four weeks." Participants rated how much they agreed with the statement on a 7-point Likert scale, ranging from 1 (strongly agree) to 7 (strongly disagree). Intention was measured both before and after the workshop. A higher score indicated greater intention to use behavior change techniques in practice.

Workshop Acceptability

Two items were used to assess the acceptability of the workshops. The items were constructed based on the feasibility and acceptability questionnaire developed by Kothe and Mullan [24]. The first set of items asked participants to rate their agreement with each of 7 statements on their feelings about whether the workshops were needed, useful, appropriate for the profession, applicable to their current practices, interesting, exciting, and worth their time. The statements were rated on a 5-point Likert scale, ranging from 1 (completely disagree) to 5 (completely agree). The internal consistency was good, with Cronbach α=.94. The second set of items asked participants to indicate how satisfied they were with the workshop on a 5-point Likert scale, ranging from 1 (completely dissatisfied) to 5 (completely satisfied). The statements were summed to create a total score. A higher score represented greater overall acceptance of the behavior change workshop. Higher scores on item 2 represented greater satisfaction with the workshop.

Data Analysis

Pre- and postworkshop survey responses were matched using the participants' email addresses, which were then removed and replaced with an anonymous participant ID. Data were screened for errors and missing values. Missing values were imputed using expectation maximization. Differences between pre- and postworkshop scores for knowledge, attitude, social and descriptive norms, perceived behavioral control, and intention were assessed using 2-tailed paired-samples *t* tests. We adjusted our α level for multiple comparisons with Bonferroni correction. Therefore, our α level for the paired-samples *t* tests was α=.008. Pearson correlations were also used to determine the association between pre- and postworkshop scores.

Results

Workshop Attendees

A total of 127 health professionals from Western Australia attended 1 of 10 workshops on the professional development day organized by the PSA. This sample of workshop attendees included general practitioners (n=4, 3.1%), nurses (n=16, 12.6%), and pharmacists (n=107, 84.3%). The majority of the workshop attendees were women (n=108, 85%), and just over half attended the in-person workshops (n=75, 59.1%).

Survey Participants

Of the 127 attendees, 71 completed the preworkshop survey. Four participants did not provide consent, 10 participants did not complete any part of the survey, and 1 response was a duplicate. Removing these responses left data from 56 participants. The postworkshop survey was completed by 58 attendees; only 1 participant was removed due to not completing any items on the survey. The final data combined pre- and

postworkshop survey responses. Seventy-two completed surveys were screened, and email addresses were matched. Of these completed surveys, 6 duplicates were removed. Twelve participants completed only the preworkshop survey and 13 completed only the postworkshop survey. A final sample of 41 participants who completed both the pre- and postworkshop surveys was used for the analyses.

The demographic information of the 41 survey participants was not collected. Although the PSA collected attendee demographics when they signed up for the professional development day, due to the anonymization of the survey responses, the demographics of the survey participants could not be linked to the PSA data. However, given the large proportion of workshop attendees who were pharmacists and the large proportion of women, it is likely that most survey participants were pharmacists and that most were women.

Correlations Between Pre- and Postworkshop Psychosocial Variables

Table 1 shows the Pearson correlations between pre- and postworkshop psychosocial variables.

Table 1. Pearson correlations for pre- and postworkshop psychosocial variables.

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Pre workshop																
1 Knowledge																
r	1	–0.10	–0.13	–0.11	0.15	0.02	0.23	0.13	0.04	–0.04	–0.04	–0.04	0.25	0.09	0.28	0.33
P value	—[a]	.60	.48	.54	.36	.90	.15	.41	.82	.80	.80	.80	.11	.51	.70	.05
2 Attitude (wise/unwise)																
r	–0.10	1	0.98	0.99	–0.16	–0.19	0.25	0.19	0.17	0.59	0.56	0.59	–0.07	0.04	0.10	0.04
P value	.60	—	<.001	<.001	.30	.23	.12	.23	.29	<.001	<.001	<.001	.66	.79	.53	.79
3 Attitude (useful/useless)																
r	–0.13	0.98	1	0.99	–0.20	–0.22	0.20	0.16	0.18	0.53	0.51	0.53	–0.11	0.02	0.09	0.06
P value	.48	<.001	—	<.001	.21	.17	.21	.33	.26	<.001	<.001	<.001	.50	.90	.58	.70
4 Attitude average																
r	–0.11	0.99	0.99	1	–0.18	–0.21	0.23	0.17	0.18	0.56	0.54	0.56	–0.09	0.03	0.10	0.05
P value	.54	<.001	<.001	—	.25	.20	.16	.28	.28	<.001	<.001	<.001	.57	.84	.53	.75
5 Descriptive norms																
r	0.15	–0.16	–0.20	–0.18	1	0.65	0.29	0.16	–0.13	–0.19	–0.27	–0.24	0.70	0.47	0.10	0.17
P value	.36	.30	.21	.25	—	<.001	.07	.32	.43	.23	.08	.13	<.001	.002	.53	.28
6 Subjective norms																
r	0.02	–0.19	–0.22	–0.21	0.65	1	–0.02	0.03	0.08	–0.22	–0.24	–0.23	0.69	0.51	0.06	0.13
P value	.90	.23	.17	.20	<.001	—	.88	.85	.63	.18	.13	.14	<.001	<.001	.69	.43
7 Perceived behavioral control																
r	0.23	0.25	0.20	0.23	0.29	–0.02	1	0.78	0.17	0.15	0.10	0.13	0.07	0.05	0.54	0.50
P value	.15	.12	.21	.16	.07	.88	—	<.001	.28	.34	.54	.43	.68	.76	<.001	<.001
8 Intention																
r	0.13	0.19	0.16	0.17	0.16	0.03	0.78	1	0.38	0.13	0.14	0.14	0.04	–0.02	0.52	0.54
P value	.41	.23	.33	.28	.32	.85	<.001	—	.02	.43	.39	.40	.82	.93	<.001	<.001
Postworkshop																
9 Knowledge																
r	0.04	0.17	0.18	0.18	–0.13	0.08	0.17	0.38	1	0	0.04	0.02	–0.04	–0.12	0.31	0.31
P value	.82	.29	.26	.28	.43	.63	.28	.02	—	1	.83	.91	.81	.44	.052	.053
10 Attitude (wise/unwise)																
r	–0.04	0.59	0.53	0.56	–0.19	–0.22	0.15	0.13	0	1	0.92	0.98	–0.04	0.09	0.32	0.15
P value	.80	<.001	<.001	<.001	.23	.18	.34	.43	1	—	<.001	<.001	.81	.58	.05	.35
11 Attitude (useful/useless)																
r	–0.04	0.56	0.51	0.54	–0.27	–0.24	0.10	0.14	0.04	0.92	1	0.98	–0.08	0.05	0.27	0.13
P value	.80	<.001	<.001	<.001	.08	.13	.54	.39	.83	<.001	—	<.001	.61	.74	.08	.43
12 Attitude average																
r	–0.04	0.59	0.53	0.56	–0.24	–0.23	0.13	0.14	0.02	0.98	0.98	1	–0.06	0.07	0.30	0.14
P value	.80	<.001	<.001	<.001	.13	.14	.43	.40	.91	<.001	<.001	—	.70	.66	.06	.38
13 Descriptive norms																
r	0.25	–0.07	–0.11	–0.09	0.70	0.69	0.07	0.04	–0.04	–0.04	–0.08	–0.06	1	0.69	0.21	0.26
P value	.11	.66	.50	.57	<.001	<.001	.68	.82	.81	.81	.61	.70	—	<.001	.18	.10
14 Subjective norms																

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
r	0.09	0.04	0.02	0.03	0.47	0.51	0.05	–0.02	–0.12	0.09	0.05	0.07	0.69	1	0.15	0.26
P value	.51	.79	.90	.84	.002	<.001	.76	.93	.44	.58	.74	.66	<.001	—	.37	.10
15 Perceived behavioral control																
r	0.28	0.10	0.09	0.10	0.10	0.06	0.54	0.52	0.31	0.32	0.27	0.30	0.21	0.15	1	0.84
P value	.70	.53	.58	.53	.53	.69	<.001		.052	.05	.08	.06	.18	.37	—	<.001
16 Intention																
r	0.33	0.04	0.06	0.05	0.17	0.13	0.50	0.54	0.31	0.15	0.13	0.14	0.26	0.26	0.84	1
P value	.05	.79	.70	.75	.28	.43	<.001	<.001	.053	.35	.43	.38	.10	.10	<.001	—

[a]Not applicable.

Knowledge of Behavior Change Techniques

A paired-samples test was used to evaluate the impact of the behavior change workshop on participants' perceived knowledge of behavior change techniques before (mean score 3.73, SD 1.55) and after (mean score 5.20, SD 0.93) the workshop. There was a difference in mean knowledge scores of –1.46 (95% CI –2.02 to –0.90). This difference was significant (t_{40}=–5.27, *P*<.001) and had a large effect size (Cohen *d*=1.78).

Attitude

A paired-samples *t* test revealed no significant increase in attitude scores from before to after the workshop (t_{40}=0.78, *P*=.44). There was a small difference in scores before (mean score 6.16, SD 1.27) and after (mean score 5.99, SD 1.63) the workshop, with a change in mean score of 0.17 (95% CI –0.27 to 0.61), but this was not a significant change.

Social Norms

A paired-samples *t* test comparing social norms before (mean score 4.78, SD 1.29) and after (mean score 5.46, SD 1.23) the workshop revealed a change in mean score of –0.68 (95% CI –1.08 to –0.29). This change was significant (t_{40}=–3.49, *P*=.001) and had a large effect size (Cohen *d*=1.25).

Descriptive Norms

A paired-samples *t* test was conducted and showed a significant change in descriptive norms from before to after the workshop (t_{40}=–3.65, *P*<.001). The mean score for descriptive norms changed by –0.61 (95% CI –0.95 to –0.27) from before (mean score 4.49, SD 1.47) to after (mean score 5.10, SD 1.22) the workshop. The effect size for this test was large (Cohen *d*=1.07).

Perceived Behavioral Control

A paired-samples *t* test revealed a significant change in perceived behavioral control scores from before (mean score 5.33, SD 1.15) to after (mean score 5.83, SD 0.74) the workshop (t_{40}=–3.30, *P*=.002). This represented a change in mean score of –0.50 (95% CI –0.81 to –0.20). There was a large effect size (Cohen *d*=.98).

Intention

A paired-samples *t* test was conducted to examine if there was a significant change in intention from before (mean score 5.30, SD 0.98) to after (mean score 5.75, SD 0.84) the workshop. There was a significant difference in intention (t_{36}=–3.32, *P*=.002), with a mean score increase of –0.46 (95% CI –0.74 to –0.18) and a large effect size (Cohen *d*=0.88).

Workshop Acceptability

Following the workshop, 78% (32/41) of participants agreed or strongly agreed that the workshop was needed, and the remainder neither agreed nor disagreed. No participant felt that the workshop was not needed. Participants also felt that the workshop was useful, with 85% (35/41) of participants agreeing, while the rest neither agreed nor disagreed. There were no participants who felt that the workshop was not useful. The majority of the participants felt that the training was appropriate for their profession, with 90% (37/41) of participants agreeing, while 10% (4/41) were undecided.

Workshop participants completely agreed or agreed that the workshop was applicable to their current practices (41/47, 90%), while 10% (4/41) were neutral. Most participants indicated that the workshop was interesting (35/41, 85% agreed or completely agreed), while 14% (5/41) were neutral and 3% (1/41) participants completely disagreed. Most participants agreed or completely agreed (33/41, 80%) that the workshop was worth their time, while only a small proportion of participants completely disagreed or disagreed (2/41, 5%) and only 15% (6/41) were neutral. Overall, participants demonstrated a high acceptance of the workshop (mean score 28.84, SD 4.44). Further, most participants were satisfied or completely satisfied (36/41, 88%) with the 2-hour workshop.

Discussion

Principal Results

The 2-hour behavior change workshop was effective in changing knowledge, social and descriptive norms, perceived behavioral control, and intention, but was not effective at changing attitudes. Further, health professionals found that the workshop was acceptable, and they were satisfied overall with the content and delivery of the behavior change workshops.

Knowledge of behavior change techniques can lead to changes in important psychosocial predictors of behavior. While knowledge alone is generally insufficient to change behavior, behavior change is more likely to occur with improvements in knowledge [25]. Thus, other skills and techniques are required to ensure health professionals have the capability, opportunity,

and motivation to change behavior. As suggested by Ruppar et al [26], no intervention focused on changing behavior should focus solely on patient education. Rather, patient education works best when it is combined with more active behavioral approaches. Thus, in the context of training health professionals, increasing their knowledge of behavior change techniques and providing them opportunities to discuss case studies and real-life applications of the techniques increases the likelihood of sustained changes in behavior. Increasing the participants' perceived knowledge of behavior change techniques alone does not ensure they will apply these techniques; previous research focused on improving health care professionals' knowledge and practices regarding adverse drug reactions showed that providing an intervention targeting both knowledge and other psychosocial variables had positive effects on professional practice at the end of 12 months [27]. However, the long-term effectiveness of such interventions and the training of health professionals beyond a 12-month follow-up is uncertain; future longitudinal research is required [13,27].

Both social and descriptive norms also significantly improved after the workshop. Improvement in norms, both descriptive (referring to people important to them) and social (referring to people like them), is theorized to increase health care professionals' intention to complete behaviors, in this case, the use of behavior change techniques with their patients. The workshops increased norms with a large effect size. Norms may have been improved by a multitude of workshop components, in particular the increase in knowledge of behavior change techniques. Across various domains, educational interventions that increase knowledge often result in improved ratings for subjective norms [28-30]. In the workshops, an increase in knowledge of behavior change techniques, including their evidence base and why and how they are used, may have increased health care professionals' perception of whether others (descriptive norms) and people similar to them (social norms) believe they should change their professional behaviors to ensure optimal care.

Participants' attitudes toward changing their professional practice to provide optimal care did not change significantly. Preworkshop attitudes were high, with a mean score of 6.16 on a scale from 1 to 7. Therefore, it is likely that the lack of significant improvement in attitudes following the workshop was due to a ceiling effect, in that improvements to attitudes were unlikely due to high preintervention levels. This has previously been seen in studies of the Theory of Planned Behavior–based interventions [18,31] and suggests that improvements to other components of the modes (ie, norms and perceived behavioral control), rather than attitudes, are required to improve intention and thereby behavior.

The 2-hour workshop intervention also resulted in large improvements to participants' ratings of perceived behavioral control. The Theory of Planned Behavior posits that perceived behavioral control not only influences intention to perform a behavior, but also directly predicts behavior [17]. Therefore, the demonstrated postworkshop improvements in perceived behavioral control show particular promise for not only improving participants' intention to use behavior change techniques, but also to facilitate the actual use of behavior change techniques with their patients.

Pre- and postworkshop comparisons revealed a significant improvement in the participants' intention to use behavior change techniques with their patients. This is unsurprising, as the Theory of Planned Behavior posits that changes in attitudes, norms, and perceived behavioral control will result in increased intention, and within our sample, norms (both descriptive and social) and perceived behavioral control significantly improved, with large effect sizes. Further, the Theory of Planned Behavior theorizes that intention directly predicts behavior [17]. Therefore, in the context of the intervention, improvements in the participants' intention should lead to the use of behavior change techniques with their patients. Indeed, a great deal of previous research has demonstrated that intention is a statistically significant predictor of behavior [18,32]. However, it is also important to note that the body of literature points toward an intention-behavior gap, as intention often only accounts for a limited amount of variance in behavior [33,34]. Therefore, although our intervention shows promise in facilitating the participants' use of behavior change techniques with their patients, future interventions might also incorporate techniques that lead to habitual use of behavior change techniques with patients to ensure more consistent use of behavior change techniques in the health domain [35-38]. However, this was beyond the scope of this intervention.

Limitations and Directions for Future Research

A key limitation of the current study was the high attrition from before to after the workshop, with the final comparison sample including only 58% (74/127) of those that completed the preworkshop survey. A related limitation is that we had an inadequate sample size to use analyses, such as structural equation modeling, that are part of the Theory of Planned Behavior, thus limiting our analysis to pre- and postworkshop comparisons. However, it is also important to note that assessing the relationships between the components of the Theory of Planned Behavior was not an aim of this study. Future research should seek to replicate our study with larger sample sizes that allow for the assessment of the fit of our theoretical model, using analyses such as structural equation modeling and confirmatory factor analysis [18,39-42].

A further limitation of the study design was that some items on the pre- and postworkshop surveys only contained a single item. This was done to facilitate practical assessment and to limit the burden on participants, thereby increasing validity [43]. Further, this approach has been shown to be valid in previous work [44]. However, future research should further validate our findings by using measures with multiple items for each domain. In addition, due to not having collected any demographic variables of the participants, we were unable to evaluate if there were any differences based on participant demographics (eg, pharmacist vs nurse) or the workshop delivery mode (eg, online vs face-to-face). This limited our understanding of how and for whom the intervention was most effective. Future research should consider evaluating any demographic differences between workshop participants to inform more targeted future interventions.

An additional limitation to this study was that there was no subsequent follow-up. This makes it difficult to know if the changes in psychosocial factors were maintained over time or if there is a need to engage health professionals in regular training in behavior change to sustain long-term changes in these psychosocial factors. Further, we could not assess the frequency of the participants' application of behavior change techniques over time. Future research should explore the lasting impacts of the intervention and include a measure of health professionals' behavior to determine the translation of the improvements to knowledge, subjective norms, descriptive norms, perceived behavioral control, and intention in the health professionals' practices over time.

Lastly, the participants demonstrated high attitude scores in the preworkshop survey, which suggests that there may have been some self-selection bias. The design of the study did not assess if participants had prior training, knowledge, or interest in behavior change, which may have biased the results. Self-selection bias has been noted as a limitation in other health behavior interventions [45,46] and future research should consider strategies to mitigate self-selection bias by including control groups (eg, wait-list control groups) to improve the internal validity of the study design.

This study did, however, use a theoretically informed and evidence-based workshop design incorporating both active and passive components. The participants came from a range of disciplines and our results demonstrate the effectiveness and acceptability of our intervention across different health professions.

Conclusion

Ultimately, the 2-hour, theoretically informed, evidence-based workshop designed to facilitate the use of behavior change techniques by health professionals was shown to be largely effective. The workshops resulted in a significant increase in knowledge, descriptive norms, subjective norms, perceived behavioral control, and intention. The intervention was also shown to be highly acceptable for many participants, who deemed the intervention needed, useful, appropriate, and applicable, as well as interesting and worth their time. Future research should explore the lasting impacts of the workshop on health professionals' practices, as well as how changes in their practices may impact their patients in clinical and community settings.

Data Availability

The data sets generated during and analyzed during the current study are not publicly available due to privacy and ethical restrictions but are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Workshop Schedule.
[PDF File (Adobe PDF File), 136 KB - formative_v6i11e42010_app1.pdf]

References

1. Noncommunicable diseases. World Health Organization. 2022. URL: https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases [accessed 2022-09-18]
2. Cancer in Australia 2021. Australian Institute of Health and Welfare. URL: https://www.aihw.gov.au/getmedia/0ea708eb-dd6e-4499-9080-1cc7b5990e64/aihw-can-144.pdf [accessed 2022-11-10]
3. Fekete EM, Antoni MH, Schneiderman N. Psychosocial and behavioral interventions for chronic medical conditions. Current Opinion in Psychiatry 2007 Mar;20(2):152-157. [doi: 10.1097/YCO.0b013e3280147724] [Medline: 17278914]
4. Allegrante JP, Wells MT, Peterson JC. Interventions to Support Behavioral Self-Management of Chronic Diseases. Annual Review of Public Health 2019 Apr 01;40:127-146 [FREE Full text] [doi: 10.1146/annurev-publhealth-040218-044008] [Medline: 30601717]
5. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implementation Science 2011 Apr 23;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
6. Shoneye CL, Mullan B, Begley A, Pollard CM, Jancey J, Kerr DA. Design and Development of a Digital Weight Management Intervention (ToDAy): Qualitative Study. JMIR mHealth and uHealth 2020 Sep 09;8(9):e17919 [FREE Full text] [doi: 10.2196/17919] [Medline: 32641284]
7. Keyworth C, Epton T, Goldthorpe J, Calam R, Armitage CJ. Are healthcare professionals delivering opportunistic behaviour change interventions? A multi-professional survey of engagement with public health policy. Implementation Science 2018 Sep 21;13(1):122 [FREE Full text] [doi: 10.1186/s13012-018-0814-x] [Medline: 30241557]
8. Making Every Contact Count (MECC): Consensus statement. National Health Service England. 2016. URL: https://www.england.nhs.uk/publication/making-every-contact-count-mecc-consensus-statement/ [accessed 2022-08-18]

9. Keyworth C, Epton T, Goldthorpe J, Calam R, Armitage CJ. 'It's difficult, I think it's complicated': Health care professionals' barriers and enablers to providing opportunistic behaviour change interventions during routine medical consultations. British Journal of Health Psychology 2019 Sep;24(3):571-592 [FREE Full text] [doi: 10.1111/bjhp.12368] [Medline: 30977291]
10. Presseau J, Hawthorne G, Sniehotta FF, Steen N, Francis JJ, Johnston M, et al. Improving Diabetes care through Examining, Advising, and prescribing (IDEA): protocol for a theory-based cluster randomised controlled trial of a multiple behaviour change intervention aimed at primary healthcare professionals. Implementation Science 2014 May 24;9(1):61 [FREE Full text] [doi: 10.1186/1748-5908-9-61] [Medline: 24886606]
11. van Achterberg T, Huisman-de Waal GGJ, Ketelaar N, Oostendorp R, Jacobs J, Wollersheim H. How to promote healthy behaviours in patients? An overview of evidence for behaviour change techniques. Health Promotion International 2011 Jun;26(2):148-162 [FREE Full text] [doi: 10.1093/heapro/daq050] [Medline: 20739325]
12. 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. Annals of Behavioral Medicine 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
13. Hatfield TG, Withers TM, Greaves CJ. Systematic review of the effect of training interventions on the skills of health professionals in promoting health behaviour, with meta-analysis of subsequent effects on patient health behaviours. BMC Health Services Research 2020 Jun 29;20(1):593 [FREE Full text] [doi: 10.1186/s12913-020-05420-1] [Medline: 32600404]
14. Conn VS, Hafdahl AR, Cooper PS, Ruppar TM, Mehr DR, Russell CL. Interventions to improve medication adherence among older adults: Meta-analysis of adherence outcomes among randomized controlled trials. Gerontologist 2009 Aug;49(4):447-462. [doi: 10.1093/geront/gnp037] [Medline: 19460887]
15. Crawshaw J, Auyeung V, Ashworth L, Norton S, Weinman J. Healthcare provider-led interventions to support medication adherence following ACS: A meta-analysis. Open Heart 2017;4(2):e000685 [FREE Full text] [doi: 10.1136/openhrt-2017-000685] [Medline: 29344366]
16. Chauhan BF, Jeyaraman M, Mann AS, Lys J, Skidmore B, Sibley KM, et al. Behavior change interventions and policies influencing primary healthcare professionals' practice-an overview of reviews. Implementation Science 2017 Jan 05;12(1):3 [FREE Full text] [doi: 10.1186/s13012-016-0538-8] [Medline: 28057024]
17. Ajzen I. The Theory of Planned Behavior. Organizational Behavior and Human Decision Processes 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
18. Mergelsberg ELP, Haywood D, Mullan BV, Pereira R, Smith A, Zaccagnini M, et al. Extending the theory of planned behaviour to predict and increase PASS attendance and students' performance: can academic motivation improve efficacy? SN Social Sciences 2021 Jan 11;1(1):1-23. [doi: 10.1007/s43545-020-00041-1]
19. Mullan B, Wong C. Using the Theory of Planned Behaviour to design a food hygiene intervention. Food Control 2010 Nov;21(11):1524-1529. [doi: 10.1016/j.foodcont.2010.04.026]
20. Sainsbury K, Mullan B, Sharpe L. A randomized controlled trial of an online intervention to improve gluten-free diet adherence in celiac disease. American Journal of Gastroenterology 2013 May;108(5):811-817. [doi: 10.1038/ajg.2013.47] [Medline: 23458849]
21. Kothe E, Mullan B, Butow P. Promoting fruit and vegetable consumption. Testing an intervention based on the theory of planned behaviour. Appetite 2012 Jun;58(3):997-1004. [doi: 10.1016/j.appet.2012.02.012] [Medline: 22349778]
22. Pearson E, Byrne-Davis L, Bull E, Hart J. Behavior change techniques in health professional training: Developing a coding tool. Translational Behavioral Medicine 2020 Feb 03;10(1):96-102. [doi: 10.1093/tbm/iby125] [Medline: 30535181]
23. Ajzen I. Constructing a Theory of Planned Behavior questionnaire. ResearchGate. 2006. URL: https://www.researchgate.net/publication/235913732_Constructing_a_Theory_of_Planned_Behavior_Questionnaire [accessed 2022-11-10]
24. Kothe EJ, Mullan BA. Acceptability of a theory of planned behaviour email-based nutrition intervention. Health Promotion International 2014 Mar;29(1):81-90. [doi: 10.1093/heapro/das043] [Medline: 22942273]
25. van Rijen D, Mergelsberg E, Hoor GT, Mullan B. Improving safe food-handling practices by increasing self-efficacy. Food Control 2021 Dec;130:108361. [doi: 10.1016/j.foodcont.2021.108361]
26. Ruppar TM, Dobbels F, Lewek P, Matyjaszczyk M, Siebens K, De Geest SM. Systematic Review of Clinical Practice Guidelines for the Improvement of Medication Adherence. International Journal of Behavioral Medicine 2015 Dec;22(6):699-708. [doi: 10.1007/s12529-015-9479-x] [Medline: 25805550]
27. Opadeyi AO, Fourrier-Réglat A, Isah AO. Educational intervention to improve the knowledge, attitude and practice of healthcare professionals regarding pharmacovigilance in South-South Nigeria. Therapeutic Advances in Drug Safety 2019;10:2042098618816279 [FREE Full text] [doi: 10.1177/2042098618816279] [Medline: 30719280]
28. Askari A, Jeihooni AK, Kashfi SM, Marzban A, Khiyali Z. The effect of educational program based on belief, attitude, subjective norm, and enabling factors model on changing the metabolic indices in elderly patients with type II diabetes. International Journal of Preventative Medicine 2018;9:74 [FREE Full text] [doi: 10.4103/ijpvm.IJPVM_308_16] [Medline: 30210739]
29. Rahman W, Setyowati DL, Ifroh RH. Effect of health education of safety riding using audiovisual media on knowledge, attitudes and subjective norms of junior high school students in Samarinda Indonesia. Public Health of Indonesia 2019 Sep 30;5(3):54-61 [FREE Full text] [doi: 10.36685/phi.v5i3.264]

30. Pirzadeh A, Hazavei M, Entezari M, Hasanzadeh A. The effect of educational intervention on girl's behavior regarding nutrition: Applying the beliefs, attitudes, subjective norms, and enabling factors. Journal of Education and Health Promotion 2014;3:79 [FREE Full text] [doi: 10.4103/2277-9531.139244] [Medline: 25250345]
31. Liddelow C, Mullan B, Novoradovskaya E. Exploring medication adherence amongst Australian adults using an extended theory of planned behaviour. International Journal of Behavioral Medicine 2020 Aug;27(4):389-399. [doi: 10.1007/s12529-020-09862-z] [Medline: 32072454]
32. Liddelow C, Ferrier A, Mullan B. Understanding the predictors of hand hygiene using aspects of the theory of planned behaviour and temporal self-regulation theory. Psychology & Health 2021 Sep 07:1-18. [doi: 10.1080/08870446.2021.1974862] [Medline: 34491138]
33. Sheeran P, Webb TL. The Intention-Behavior Gap. Social and Personality Psychology Compass 2016 Sep 05;10(9):503-518. [doi: 10.1111/spc3.12265]
34. Sheeran P. Intention—Behavior Relations: A Conceptual and Empirical Review. European Review of Social Psychology 2002 Jan;12(1):1-36. [doi: 10.1080/14792772143000003]
35. Charlesworth J, Mullan B, Howell J, Tan H, Abbott B, Potter A. Exploring the role of perceived risk and habit in safe food-handling behaviour change. Food Control 2022 Apr;134:108754. [doi: 10.1016/j.foodcont.2021.108754]
36. Gardner B, Arden MA, Brown D, Eves FF, Green J, Hamilton K, et al. Developing habit-based health behaviour change interventions: Twenty-one questions to guide future research. Psychology & Health 2021 Nov 14:1-23. [doi: 10.1080/08870446.2021.2003362] [Medline: 34779335]
37. Haywood D, Lawrence BJ, Baughman FD, Mullan BA. A Conceptual Model of Long-Term Weight Loss Maintenance: The Importance of Cognitive, Empirical and Computational Approaches. International Journal of Environmental Research and Public Health 2021 Jan 13;18(2):635 [FREE Full text] [doi: 10.3390/ijerph18020635] [Medline: 33451110]
38. Judah G, Mullan B, Yee M, Johansson L, Allom V, Liddelow C. A habit-based randomised controlled trial to reduce sugar-sweetened beverage consumption: The impact of the substituted beverage on behaviour and habit strength. International Journal of Behavioral Medicine 2020 Dec 05;27(6):623-635. [doi: 10.1007/s12529-020-09906-4] [Medline: 32500393]
39. Haywood D, Baughman F, Mullan B, Heslop K. What accounts for the factors of psychopathology? an investigation of the neurocognitive correlates of internalising, externalising, and the p-factor. Brain Sci 2022 Mar 22;12(4):421 [FREE Full text] [doi: 10.3390/brainsci12040421] [Medline: 35447951]
40. Haywood D, Baughman FD, Mullan BA, Heslop KR. One-Factor for All? Exploring the Applicability of Structural Models of Psychopathology within Subgroups of a Population. International Journal of Environmental Research and Public Health 2021 Jul 02;18(13):7108 [FREE Full text] [doi: 10.3390/ijerph18137108] [Medline: 34281043]
41. Haywood D, Baughman FD, Mullan BA, Heslop KR. Going "Up" to Move Forward: S-1 Bifactor Models and the Study of Neurocognitive Abilities in Psychopathology. International Journal of Environmental Research and Public Health 2021 Jul 11;18(14):7413 [FREE Full text] [doi: 10.3390/ijerph18147413] [Medline: 34299862]
42. Haywood D, Baughman FD, Mullan BA, Heslop KR. Psychopathology and Neurocognition in the Era of the P-Factor: The Current Landscape and the Road Forward. Psychiatry International 2021 Jun 23;2(3):233-249 [FREE Full text] [doi: 10.3390/psychiatryint2030018]
43. Lavrakas PJ. Encyclopedia of Survey Research Methods. New York, NY: SAGE Publications; 2008.
44. Charlesworth J, Mullan B, Howell J, Tan H, Abbott B, Potter A. Evaluating the impact of a pilot safe food-handling media campaign among consumers in Western Australia: Implications for public health messaging. Food Control 2021 Aug;126:108070. [doi: 10.1016/j.foodcont.2021.108070]
45. Milton A, Mullan B. Consumer food safety education for the domestic environment: A systematic review. British Food Journal 2010;112(9):1003-1022. [doi: 10.1108/00070701011074363]
46. Sainsbury K, Mullan B, Sharpe L. Reduced quality of life in coeliac disease is more strongly associated with depression than gastrointestinal symptoms. Journal of Psychosomatic Research 2013 Aug;75(2):135-141. [doi: 10.1016/j.jpsychores.2013.05.011] [Medline: 23915769]

Abbreviations

COM-B: Capability, Opportunity, Motivation, and Behavior
PSA: Pharmaceutical Society of Australia

Edited by A Mavragani; submitted 18.08.22; peer-reviewed by I Wilson, J Gamble; comments to author 27.09.22; revised version received 11.10.22; accepted 12.10.22; published 18.11.22.

Please cite as:
Mullan B, Liddelow C, Haywood D, Breare H
Behavior Change Training for Health Professionals: Evaluation of a 2-Hour Workshop
JMIR Form Res 2022;6(11):e42010
URL: https://formative.jmir.org/2022/11/e42010
doi:10.2196/42010
PMID:36399382

Review

Digital mHealth and Virtual Care Use During COVID-19 in 4 Countries: Rapid Landscape Review

Alison Müller[1], MSc, PhD; Alessandro Cau[1], BSc; Semakula Muhammed[2], MSc; Osman Abdullahi[3], PhD; Andrew Hayward[4], MD, MSc, PhD; Sabin Nsanzimana[5], DPhil, MD; Richard Lester[1], MD

[1]Division of Infectious Diseases, Department of Medicine, University of British Columbia, Vancouver, BC, Canada

[2]Institute of Diseases Prevention and Control, Rwanda Biomedical Centre, Kigali, Rwanda

[3]Department of Public Health, School of Health and Human Sciences, Pwani University, Kilifi, Kenya

[4]Department of Epidemiology & Public Health, University College London, London, United Kingdom

[5]University Teaching Hospital, Butare, Rwanda

Corresponding Author:
Alison Müller, MSc, PhD
Division of Infectious Diseases
Department of Medicine
University of British Columbia
575 - Vancouver General Hospital Research Pavilion
828 W 10th Ave
Vancouver, BC, V5Z1M9
Canada
Phone: 1 604 875 4111 ext 66294
Email: amuller8@mail.ubc.ca

Abstract

Background: As a result of the COVID-19 pandemic, providing health care while maintaining social distancing has resulted in the need to provide care remotely, support quarantined or isolated individuals, monitor infected individuals and their close contacts, as well as disseminate accurate information regarding COVID-19 to the public. This has led to an unprecedented rapid expansion of digital tools to provide digitized virtual care globally, especially mobile phone–facilitated health interventions, called mHealth. To help keep abreast of different mHealth and virtual care technologies being used internationally to facilitate patient care and public health during the COVID-19 pandemic, we carried out a rapid investigation of solutions being deployed and considered in 4 countries.

Objective: The aim of this paper was to describe mHealth and the digital and contact tracing technologies being used in the health care management of the COVID-19 pandemic among 2 high-income and 2 low-middle income countries.

Methods: We compared virtual care interventions used for COVID-19 management among 2 high-income countries (the United Kingdom and Canada) and 2 low-middle income (Kenya and Rwanda) countries. We focused on interventions used to facilitate patient care and public health. Information regarding specific virtual care technologies was procured from a variety of resources including gray literature, government and health organization websites, and coauthors' personal experiences as implementers of COVID-19 virtual care strategies. Search engine queries were performed to find health information that would be easily accessible to the general public, with keywords including "COVID-19," "contact-tracing," "tool-kit," "telehealth," and "virtual care," in conjunction with corresponding national health authorities.

Results: We identified a variety of technologies in Canada, the United Kingdom, Rwanda, and Kenya being used for patient care and public health. These countries are using both video and text message–based platforms to facilitate communication with health care providers (eg, WelTel and Zoom). Nationally developed contact tracing apps are provided free to the public, with most of them using Bluetooth-based technology. We identified that often multiple complimentary technologies are being utilized for different aspects of patient care and public health with the common purpose to disseminate information safely. There was a negligible difference among the types of technologies used in both high-income and low-middle income countries, although the latter implemented virtual care interventions earlier during the pandemic's first wave, which may account for their effective response.

Conclusions: Virtual care and mHealth technologies have evolved rapidly as a tool for health care support for both patient care and public health. It is evident that, on an international level, a variety of mHealth and virtual care interventions, often in combination, are required to be able to address patient care and public health concerns during the COVID-19 pandemic, independent of a country's economic standing.

(JMIR Form Res 2022;6(11):e26041) doi:10.2196/26041

KEYWORDS

COVID-19; virtual care; public health; mHealth; contact tracing; telehealth; Canada; United Kingdom; Kenya; Rwanda; global health; apps

Introduction

Background

Pandemics pose a considerable threat to global health security and place an enormous strain on health care systems. The current COVID-19 pandemic has already claimed over 2.5 million lives, with over 112 million reported cases as of February 26, 2020 [1]. Delivering care remotely, supporting individuals on home isolation, monitoring infected individuals and their close contacts, and disseminating accurate information regarding COVID-19 to the public have been major challenges [2]. Although virtual care technologies have been used in a variety of health care settings prior to COVID-19, their utilization has significantly increased as they provide solutions to address the many challenges that result from maintaining physical distancing while providing essential health care. Virtual care is a rapidly growing area that is well-positioned to alleviate many issues within various health care systems including lengthy wait times, overcrowded emergency departments, and avoidable hospital readmittance [3-5]. Due to the COVID-19's ease of transmissibility, self-quarantining and physical distancing measures have been implemented worldwide to mitigate its spread, in addition to border closures and significantly limiting the sizes of social gatherings.

In this landscape review—discussing the technology developed and used for COVID-19 management in Canada, the United Kingdom, Rwanda, and Kenya—we define virtual care as the remote delivery of health care using technology. This includes technology that facilitates video communication and text messaging between patients and their health care providers (HCPs). This broad category encompasses eHealth, telehealth, and clinical decision support tools in the field. Virtual care tools may provide care for patients from the time that they decide to access the health care system up until the end of their care experience, facilitating HCP-patient interactions and care throughout the entirety of the patient journey with minimal in-person interactions. Thus, virtual care strategies may improve access to care and quality of care, reduce health care costs, and empower patients to care for themselves while providing a medium through which they can comfortably request reliable information or advice [6]. An article released by the Canadian Medical Association has found that patients are overwhelmingly satisfied with their virtual health care during the COVID-19 pandemic [7]. At the time of publication, evidence was unavailable regarding perceptions of general virtual health care in other countries discussed in this paper.

mHealth has been defined by the World Health Organization as "the use of mobile and wireless technologies to support the achievement of health objectives" [8]. It encompasses any strategy that utilizes mobile wireless technology to deliver health care, including health and wellness apps in addition to digital wearable devices. Health care delivery through phone or video, also known as telehealth, can be classified into virtual visits and remote patient monitoring. Virtual visits, usually referred to as telemedicine, are online real time interactions between care providers and patients through a virtual care platform, a videoconferencing service, an app, or over the phone. Remote patient monitoring is the remote monitoring and collection of patients' health data including vital signs and glucose levels depending on which specialized devices the patients are using. The current COVID-19 health crisis is an important opportunity to provide insight into the implementation, utilization, and efficacy of virtual care approaches during pandemics.

In this paper, we explore the current virtual care strategies for both patient care and case contact tracing in 4 focus countries: Canada, the United Kingdom, Rwanda, and Kenya. These countries were selected because our research group received an urgent Canadian federal research funding to deploy and study an mHealth intervention associated with our research group and established partners in those countries related to other pandemics (HIV and tuberculosis).

Purpose

The purpose of this review was to outline and summarize the landscape of mHealth and virtual care in 4 countries (2 low-middle income and 2 high income) that have been purported to be used or enhanced specifically due to the COVID-19 pandemic and to serve as examples from diverse regions: Canada, the United Kingdom, Rwanda, and Kenya.

Methods

The information for this landscape review was procured by accessing a variety of resources including peer-reviewed literature, gray literature (such as press releases), government and health organization websites, and firsthand experiences of this article's authors and collaborators in Canada, the United Kingdom, Rwanda, and Kenya. We evaluated 2 high-income countries, classified by the World Bank as having a gross national income of more than US $12,535 per capita (Canada and the United Kingdom), and 2 lower-middle income countries, classified as having a gross national income of less than US $4045 per capita (Rwanda and Kenya) [9]. We focused on presenting rapidly published information and information

procured from our collaborators to ensure that the information presented in this landscape review was up-to-date. It should be noted that there is extremely limited primary scientific literature regarding the development, implementation, and effectiveness of rapidly developed and deployed virtual care and public health interventions in the nations' efforts to quell the spread of COVID-19 [10]. The authors contributed additional information regarding their country's specific COVID-19 virtual care and public health measures, many of which were personally involved in implementing them. Communication with the authors and collaborators occurred via a combination of emailing, WhatsApp messaging, and videoconferencing using Zoom. To find press releases, information about mobile apps, and commentaries pertaining to COVID-19 technologies, search engines (predominantly Google) were used. Additional information regarding specific apps were found on their supporting app stores (eg, Google Play). Search engine queries were performed to find health information that would be easily accessible to the general public, with keywords including "COVID-19," "contact-tracing," "tool-kit," "telehealth," and "virtual care," in conjunction with the corresponding national health authorities. We further focused on provincial health authorities in Canada only because of the diversity of COVID-19 measures undertaken by each province. The United Kingdom, Kenya, and Rwanda had more cohesive national development and implementation of interventions, so separation by regions was not necessary. We focused on procuring information pertaining to mHealth and virtual care technologies being used for both patient care and public health in the COVID-19 response. This includes interventions used for patient communication, appointment making, information platforms, and case contact tracing, among others. We understand that this method will invariably leave gaps, especially as information surrounding COVID-19 mHealth resources continues to evolve as the pandemic continues. We therefore encourage readers to seek additional resources, primary and secondary, and to contact our team regarding any errors, omissions, or suggestions to consider.

Results

Current Virtual Care Strategies

Canada

In Canada, several virtual care platforms with videoconferencing capabilities have existed and new ones emerged to assist physicians seeing patients remotely with some platforms that are offering their technology for free to assist with patient care during the COVID-19 pandemic. In Canadian provinces such as Nova Scotia and Prince Edward Island, health care websites highly recommend using Zoom for health care while other provincial health care sites do not specify a platform. The list of virtual care technologies is updated regularly with new platforms and features being developed as the COVID-19 situation evolves. In each province, virtual care "toolkits" have been developed to assist physicians when providing care for patients remotely (Multimedia Appendix 1). These can be accessed from provincial health authority websites easily found using internet search engines.

The federal government developed a nationwide contact tracing app called COVID Alert. It was developed in partnership with Shopify, Blackberry, and the Canadian Digital Service and was launched on July 31, 2020 [11]. It is a free app using Bluetooth technology, which uses Google and Apple's exposure notification application programming interface to notify an individual who is in proximity of one who has self-identified as being COVID-positive on the app. The app specifications indicate that it does not use GPS technology to track the user's location. The specifications also indicate that COVID Alert does not access private phone information, including the phone's contacts, or access to health information of the user or other nearby users. As of October 19, the app has been downloaded 4.7 million times, which comprises approximately 16% of the Canadian population [12]. COVID Alert can also be used to report a COVID-19 diagnosis in most provinces, excluding British Columbia, Alberta, Nunavut, and the Yukon and Northwest Territories. In Alberta, ABTraceTogether, another Bluetooth-based "exposure notification app" for digital case contact tracing was released on May 2, 2020. This app also allows COVID-19–positive patients in Alberta to provide Alberta Health Services access to facilitate contact tracing [13]. An exposure notification app by the artificial intelligence group Mila in Quebec was released in May, which features additional information such as alerting potentially infected users and providing relevant, personalized recommendations to improve the understanding of infection risks [14]. These preliminary platforms have yet to demonstrate their superiority to traditional case contact tracing strategies when it comes to being effective while maintaining anonymity. Another app available for download by the Canadian public is the Canada COVID-19 app, powered by the Canadian company Thrive Health. It allows the user to track symptoms and contains a self-assessment tool, in addition to providing timely updates with important news and updates from Canada's Ministry of Health.

The Government of Canada has also provided a variety COVID-19 resources to assist in evidence-based decision-making on the Canadian Institutes of Health Research website [15]. Numerous Canadian universities have synthesized COVID-19 evidence in the efforts to support policy and decision-making. McMaster University has COVID-END, which provides resources for decision-makers—researchers, presentations and products, as well as working groups. The 7 working groups, made up of international experts, focus on different aspects of COVID-19 knowledge translation: scoping, engaging, digitizing, synthesizing, recommending, packaging, and sustaining [16]. Ryerson University has developed a COVID-19 misinformation site to consolidate many internet resources focused on monitoring COVID-19 misinformation. They provide tools for reporting misinformation, fact-checking resources, and a misinformation and bots-watching dashboards, as well as a slideshow presentation debunking common COVID-19 claims [17]. The SPOR Evidence Alliance focuses on promoting evidence-informed health policy, practice, and service to ensure that decision-makers have access to up-to-date scientific evidence. They provide an online form for those seeking specific evidence to inform decisions related to health policy, practice, and service for COVID-19 to facilitate the acquisition of relevant, accurate, and timely data [18]. Several

provincial health authorities have also engaged with companies that provide assistance with COVID-19 case contact monitoring; however, information on these was not yet made public or readily available.

The United Kingdom

In the United Kingdom, many common in-person medical communications have transitioned into virtual communications. For example, testing centers are booking COVID-19 tests online, and care homes are arranging all their regular resident and staff testing virtually. For those who are self-isolating, the national health service (NHS) has provided information to local authority helplines if self-isolating individuals require practical or social support, support for someone a patient cares for, and financial support. The NHS is also supporting a virtual clinic service powered by medio.link, which is video link–based [19]. They are working with Barts Health NHS Trust to accomplish their goal of forming a national network of sites providing virtual care. The video consultations have been able to be adapted for a variety of settings, including across the entire prison estate. Additionally, video-observed therapy is being substituted for in-person, directly observed therapy to ensure patients with tuberculosis take their medication [20,21]. Digital risk assessment has been established where the analysis of primary care has been linked to mortality data, permitting the risk of death during the first wave to be automatically calculated in providing appropriate advice to guide the patients. The United Kingdom also has a locally developed app called Babylon Health, which is paid for by the client in order to consult virtually with medical staff affiliated with Babylon Health, not necessarily their primary care provider. Additionally, various symptom checker apps such as COVID Symptom Study, Ask NHS, and Symptomate are available for the public's use to assess whether they could be exhibiting COVID-19 symptoms. There is also Zoe Symptom Tracker, an app that allows for syndromic surveillance and assessment of symptom profiles.

With regards to other tracking options, the United Kingdom is also working on an app to track COVID-19 patients, which has completed trials on the Isle of Wight. Initially, the release of the app saw low download rates, with The Guardian reporting only 10% of the population in both England and Wales [22]. In addition to using Bluetooth technology similar to Canada's contact tracing app, it enables QR code scanning at venues so that if an individual who recently attended the venue reports a COVID-19–positive test result, other patrons of the venue are automatically notified. In addition to the app, the NHS is providing NHS Test and Trace, a free service that helps trace recent contacts of those testing positive for COVID-19 and notifying them, advising them to self-quarantine. There has also been an increase in digital surveillance throughout the United Kingdom, especially with venues and large employers. For example, daily surveys of hostel managers were performed to facilitate rapid testing and University College London is launching a program, Connect to Protect. This program aids people to report symptoms and positive tests to identify clusters in class, buildings, and residences [23].

Rwanda

Rwanda utilizes several existing virtual and mHealth solutions in its ecosystem. To assist in monitoring and supporting patients (cases and contacts) directly under home-based isolation and quarantine, they rapidly deployed the WelTel mHealth platform, which they had previously been using to support patients attending HIV clinics for adherence support. WelTel is an integrated virtual care and 2-way patient engagement digital health intervention that acts as a hub for HCPs to communicate with their patients. As it is primarily used for SMS-based text messaging, patients do not need a smartphone or internet access in order to communicate with their HCPs. WelTel was launched in the Rwanda national emergency operations center for the COVID-19 response in mid-March, within a week of identifying the opportunity of using SMS to reach patients, as previous technologies to reach out to Ebola contacts that were online had limited uptake due to accessibility of the internet. Currently, the WelTel platform is being used for virtual home-based care of COVID-19 patients who are asymptomatic. Above 80% of people who tested COVID-19 positive in Rwanda are asymptomatic. This platform enables the daily follow-up of patients at their home as well as their contacts. The program was rolled out, and Rwandan COVID-19 response teams were trained. The WelTel platform offers several options to interact with the patient through SMS chatting, email, or video call. In addition, at the beginning of the outbreak, WelTel was used to communicate COVID-19 testing results, mainly with those who tested negative. Furthermore, Rwanda uses the DHIS2 (District Health Information Software 2) tracker to record everyone having a COVID-19 test, and it is linked to laboratory systems that push results to the patient's mobile number and email. The data gathered through Weltel are stored at the Rwanda National Data Center. Rwanda has implemented an additional mHealth technology to help its citizens access HCPs. For example, robots have been implemented in health care centers to collect important patient information including temperature screenings and vital readings, to deliver video messages, and to instruct people to put on a mask. Most importantly, robots play a key role in reducing the exposure time of health care professionals with COVID-19 patients [24]. Two robots have even been deployed at the Kigali International Airport for screening and informing security of issues [25]. The Government of Rwanda has recently (February 2020) signed a 10-year partnership with Babylon's Rwanda-focused virtual care subsidiary, Babyl. Appointments are being paid for by the government's health insurance scheme. The service includes physician consults, prescriptions, and lab test codes. There are currently 2 million users. Instructions for Rwandans who might suspect they have COVID-19 are advised to call 114, dial *114# for automated screening, send an email, or send a WhatsApp message [26]. WelTel is also being expanded for remote virtual care for HIV and maternal child health due to access issues under COVID-19.

Kenya

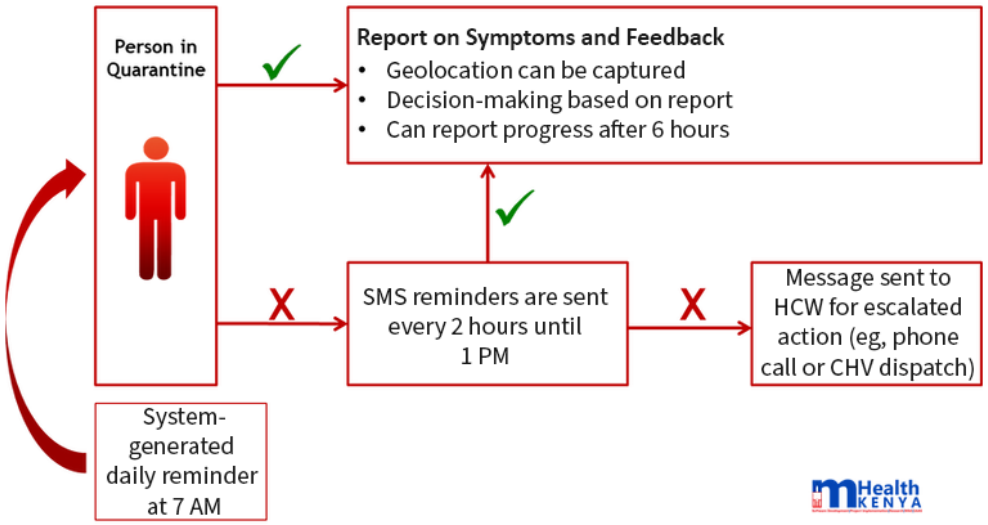
Kenya has a few different resources for virtual care using mHealth technology. mDaktari is providing virtual access to primary care via teleconsulting. Patients can access virtual care either through the app or web account, choose an available physician from a directory, or consult an expert using online

video or voice calling [27]. A recent Kenyan start-up, mTIBU, focuses on connecting patients to health care affordably without them having to leave home by using the mTIBU mobile app. They offer a variety of medical care including COVID-19 tests, medical consultation, and sample collection, all within the comfort of a patient's home [28]. In Samburu County, a rural area of northern Kenya, the cStock approach is being used by combining mobile technology, user-friendly dashboards, and quality improvement teams. It is customized for reporting and resupplying health products managed at the community level. Specifically, it is being implemented by the Samburu County Department of Health to supply COVID-19 primary protective gear with support by John Snow International. WelTel is also being used in Samburu County to assist HIV+ patient care and maternal, neonatal, and child health care. COVID-19 health care data are being reported using the DHIS2 tool. Safaricom is being used to send free text messages to educate the community on COVID-19. Health care workers are provided free courses on COVID-19, accessible via smartphones or computers.

There is a national mobile technology–focused group, called mHealth Kenya, which has developed a National Emergency System meant to capture, report, and view emerging epidemics. During COVID-19, the app Jitenge was developed, which allows registered users, either through self-registration or by Ministry of Health officials, to receive daily reminders and prompts to report their health status. The Jitenge system (Figure 1) manages and monitors home-based care management, self-quarantine for contacts, postisolation follow-up, and the monitoring of long-distance truck drivers [29].

With regards to contact tracing COVID-19 patients, an app is available called KoviTrace. It was developed at Mount Kenya University, and it uses a geo-sensing technology to track a patient's location over a 14-day period when they test positive for COVID-19. If an individual is in close proximity to a COVID-19–positive patient, SMS messages are sent with instructions, and it contacts a COVID-19 response team, depending on the contact's location [30]. Crowdsourcing movement in Red Zones using social media is also being implemented to aid in contact tracing.

Figure 1. Schematic of the basic functionality of Jitenge on how it works with individuals in quarantine. Image used with permission provided by mHealth Kenya.



Discussion

With an urgent, international emphasis on minimizing in-person communication, the development and implementation of mHealth technologies has undoubtedly increased in response to the current COVID-19 pandemic. It is evident that a wide variety of mHealth and virtual care tools with parallel functionality are being used in different sectors of health care throughout North America, Europe, and Africa. The key advantage that these technologies have is that they facilitate both the dissemination and collection of important health information while maintaining safe physical distancing. The most consistently used method of patient care involves teleconferencing and videoconferencing where patients are able to directly communicate with their health care practitioner. A few major players being used internationally include WelTel and Babylon Health, both emphasizing the importance of expedient and accurate delivery of health information and care.

Virtual care tools that facilitate direct communication with HCPs provide a secure way to disseminate or renew prescriptions and provide referrals, common patient needs that are easily addressed without the need for an in-person visit. There is also the added advantage of conversations between patients and HCPs being accessible by either party, especially in text message–based interactions where having access to prior conversations can help with care by verifying details. Traditionally, access to a patient's health information is held by their HCP; however, being able to refer to conversations is not only useful for the patients, but also for other HCPs to understand what was discussed with the patient. This feature is especially useful for patients communicating with a clinic where different staff are responsible for monitoring a patient's care.

Public health is also of primary concern, and all the countries discussed in this review are investing resources into national case contact tracing technologies to be able to track COVID-19 infections among their populations. A combination of locally

developed smartphone apps and government-implemented technology is being used in order to understand, and hopefully mitigate, the spread of COVID-19. Although a variety of different technologies is being used, there is an emphasis to ensure that those within the vicinity of potentially COVID-19–positive patients are notified and asked to self-isolate. Encouraging self-isolation is occurring at an international level, although regular contact with those in self-isolation varies among the countries we looked at. In general, contact tracing apps can provide more information than just COVID-19–proximity notifications. Data are generated whenever a COVID-19 notification is presented on an individual's phone, including how many individuals are in the proximity of a positive case as well as when the notification occurred. This information can be useful in figuring out trends of when higher instances of potential exposures are happening most often. Due to the nature of most apps, this information is available immediately and can be used to provide useful information for health authorities to inform decision-makers about how best to mitigate transmission. Moreover, depending on the location privacy of contact-tracing apps, it could be possible to identify zones where transmissions frequently occur. This information can also be stored for future reference.

The type of technology used to manage COVID-19 between the high-income and low-middle income countries was similar; however, Kenya and Rwanda had a more comprehensive approach when using novel technology. Studies evaluating the technical efficiency and overall performance of national COVID-19 management have consistently criticized the United Kingdom and Canada as underperforming despite having access to exceptional resources [31-33]. In fact, the Lowy Institute assessed the performance during the first wave of 98 countries by evaluating various parameters including confirmed cases and deaths per million people and confirmed the cases as a proportion of tests [32]. Their evaluation ranked Rwanda as 6th, Kenya as 48th, Canada as 61st, and the United Kingdom as 66th, indicating that many low-middle income countries were more effective in managing the initial impact of the COVID-19 pandemic than high-income countries. It is challenging to determine exactly how Rwanda and Kenya's mHealth and virtual care interventions played a role in providing efficient COVID-19 management, but their early development, adoption, and utilization of these tools likely contributed to their effective response to the first wave of the pandemic. In fact, a health care worker based in the United Kingdom responded to an international evaluation assessing COVID-19 management strategies stating that "There was a national plan but it was not effectively put into action" [33].

The implementation and utilization of mHealth and virtual care interventions have grown rapidly during the COVID-19 pandemic as a result of maintaining social distancing measures while providing much needed health care. Although it is not yet known which interventions are the most effective, it is evident that there is consistency with direct virtual health care provider and patient interactions as well as with case contact tracing to notify individuals in efforts to prevent the spread of COVID-19. Studies evaluating patients' satisfaction with virtual care and mHealth technologies, both before and during COVID-19, have been overwhelmingly positive, strengthening the likelihood that these interventions will become integrated as a regular component of patient care [34-38]. A significant observation of these studies was that patient care was comparable between in-person and virtual consultations; however, the increased convenience for both the HCP and patient and the decreased overall cost for virtual consultations were consistently noted as incentives for adopting mHealth interventions [34,36-38]. The staggering international increase in the adoption of mHealth and virtual care interventions facilitates researching and comparing their efficacy, paving the way to incorporating them into everyday health care globally.

Acknowledgments

We would like to thank Charles Kamau from mHealth Kenya and Martin Thuranira from Samburu County Director of Health. This project was funded by the Canadian Institute of Health Research, Grand Challenges Canada.

Conflicts of Interest

RL has a financial and professional interest in WelTel Incorporated, a company for which he is chief scientific officer. WelTel Incorporated is subcontracted to provide the SMS software platform and maintenance for the project. WelTel Incorporated is a company founded by RL's wife. The company develops software and related technologies to support mHealth through SMS-based, patient engagement platforms. WelTel Incorporated has been contracted by this research project to provide the software and the implementation service. As chief scientific officer of WelTel Incorporated, RL advises on the research and development aspects of WelTel and communicates with the scientific and customer communities concerning the support and utility of the software.

Multimedia Appendix 1
Summary of online resources available to the public to access virtual care and public health resources in Canada, the United Kingdom, Rwanda, and Kenya, during the initial wave of COVID-19.
[DOCX File , 17 KB - formative_v6i11e26041_app1.docx]

References

1. Digital contact tracing can slow or even stop coronavirus transmission and ease us out of lockdown. University of Oxford. URL: https://tinyurl.com/2tedufex [accessed 2021-12-30]

2. Willan J, King AJ, Jeffery K, Bienz N. Challenges for NHS hospitals during covid-19 epidemic. BMJ 2020 Mar 20;368:m1117. [doi: 10.1136/bmj.m1117] [Medline: 32198166]
3. Noel K, Messina C, Hou W, Schoenfeld E, Kelly G. Tele-transitions of care (TTOC): a 12-month, randomized controlled trial evaluating the use of Telehealth to achieve triple aim objectives. BMC Fam Pract 2020 Feb 07;21(1):27 [FREE Full text] [doi: 10.1186/s12875-020-1094-5] [Medline: 32033535]
4. Kvedar J, Coye MJ, Everett W. Connected Health: A Review Of Technologies And Strategies To Improve Patient Care With Telemedicine And Telehealth. Health Affairs 2014 Feb;33(2):194-199. [doi: 10.1377/hlthaff.2013.0992]
5. Vest J, Kern L, Silver M, Kaushal R, HITEC investigators. The potential for community-based health information exchange systems to reduce hospital readmissions. J Am Med Inform Assoc 2015 Mar;22(2):435-442. [doi: 10.1136/amiajnl-2014-002760] [Medline: 25100447]
6. Hamine S, Gerth-Guyette E, Faulx D, Green BB, Ginsburg AS. Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. J Med Internet Res 2015 Mar 24;17(2):e52 [FREE Full text] [doi: 10.2196/jmir.3951] [Medline: 25803266]
7. Virtual Care is real care: National poll shows Canadian are overwhelmingly satisfied with virtual health care. Canadian Medical Association. 2020. URL: https://tinyurl.com/4shcdvba [accessed 2021-12-30]
8. mHealth: New horizons for health through mobile technologies. World Health Organization. 2011. URL: https://www.who.int/goe/publications/goe_mhealth_web.pdf [accessed 2021-12-30]
9. World Bank Country and Lending Groups. The World Bank. URL: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups [accessed 2021-12-30]
10. Knottnerus JA, Tugwell P. Methodological challenges in studying the COVID-19 pandemic crisis. Journal of Clinical Epidemiology 2020 May;121:A5-A7. [doi: 10.1016/j.jclinepi.2020.04.001]
11. COVID Alert. Wikipedia. URL: https://en.wikipedia.org/wiki/COVID_Alert [accessed 2021-12-30]
12. Download COVID Alert today. Canadian Institutes of Health Research. URL: https://www.canada.ca/en/public-health/services/diseases/coronavirus-disease-covid-19/covid-alert.html [accessed 2021-12-30]
13. ABTraceTogether. Government of Alberta. URL: https://www.alberta.ca/ab-trace-together.aspx [accessed 2021-12-30]
14. A Smart and Ethical Contact-Tracing App to Fight COVID-19. Mila. URL: https://mila.quebec/en/a-smart-and-ethical-contact-tracing-app-to-fight-covid-19/ [accessed 2021-12-30]
15. COVID-19 Resources: Evidence-informed decision making. Canadian Institutes of Health Research. 2020. URL: https://cihr-irsc.gc.ca/e/52083.html [accessed 2021-12-30]
16. COVID-END. McMaster University. URL: https://www.mcmasterforum.org/networks/covid-end [accessed 2021-12-30]
17. COVIDmisinfo. Ryerson University: Ted Rogers School of Management - SocialMediaLab. URL: https://covid19misinfo.org/ [accessed 2021-12-30]
18. Strategy for Patient-Oriented Research (SPOR) Evidence Alliance. Canadian Institutes of Health Research. URL: https://sporevidencealliance.ca/ [accessed 2021-12-30]
19. medio.link is the #1 meeting platform for security and reliability in HEALTHCARE. Involve. URL: https://medio.link/nhs-video-conferencing/ [accessed 2021-12-30]
20. Fraser H, Keshavjee S. Video-observed therapy for tuberculosis: strengthening care. The Lancet 2019 Mar;393(10177):1180-1181. [doi: 10.1016/s0140-6736(19)30293-4]
21. Story A, Aldridge RW, Smith CM, Garber E, Hall J, Ferenando G, et al. Smartphone-enabled video-observed versus directly observed treatment for tuberculosis: a multicentre, analyst-blinded, randomised, controlled superiority trial. The Lancet 2019 Mar;393(10177):1216-1224. [doi: 10.1016/s0140-6736(18)32993-3]
22. Take-up of NHS contact-tracing app could only be 10%. The Guardian. 2020 Sep 24. URL: https://www.theguardian.com/world/2020/sep/24/take-up-of-nhs-contact-tracing-app-could-be-only-10 [accessed 2021-12-30]
23. UCL launches Connect to Protect reporting tool. UCL News. URL: https://www.ucl.ac.uk/news/2020/sep/ucl-launches-connect-protect-reporting-tool [accessed 2021-12-30]
24. Clarisse M, Muhamed S, Jean BM. Use of technologies in COVID-19 containment in Rwanda. Rw Public Health Bul 2020 2020;2(2):7-12.
25. Robots use in Rwanda to fight against COVID-19. World Health Organization: Africa. URL: https://www.afro.who.int/news/robots-use-rwanda-fight-against-covid-19 [accessed 2021-12-30]
26. Rwanda will be the world's most advanced country for digital health. Babylon. URL: https://www.babylonhealth.com/blog/business/rwanda-will-be-the-worlds-most-advanced-country-for-digital-health [accessed 2021-12-30]
27. mDaktari: Affordable convenient access to primary care: Virtual care combined with physical support. Connectmed. URL: https://connectmed.co.ke/ [accessed 2021-12-30]
28. Kirui D. Kenya launches KoviTrace app to curb COVID-19. IT Web Africa. IT Web. URL: https://itweb.africa/content/mQwkoq6PgWk73r9A [accessed 2021-12-30]
29. Tinu. Tibu. URL: https://tibu.africa/home [accessed 2021-12-30]
30. Technology for Better HEalth. mHealth Kenya. URL: https://www.mhealthkenya.org/ [accessed 2021-12-30]
31. Breltenbach M, Ngobeni V, Aye G. Efficiency of healthcare systems in the first wave of COVID-19 - a technical efficiency analysis. MPRA Paper No. 101440 2020 May 10:1-25.

32. Leng A, Lemahieu H. Covid Performance Index - Deconstructing Pandemic Responses. Lowy Institute. URL: https://interactives.lowyinstitute.org/features/covid-performance/ [accessed 2021-12-30]
33. Tartaglia R, La Regina M, Tanzini M, Pomare C, Urwin R, Ellis LA, et al. International survey of COVID-19 management strategies. Int J Qual Health Care 2021 Mar 20;33(1):1 [FREE Full text] [doi: 10.1093/intqhc/mzaa139] [Medline: 33219683]
34. Robb JF, Hyland MH, Goodman AD. Comparison of telemedicine versus in-person visits for persons with multiple sclerosis: A randomized crossover study of feasibility, cost, and satisfaction. Mult Scler Relat Disord 2019 Nov;36:101258. [doi: 10.1016/j.msard.2019.05.001] [Medline: 31472419]
35. Finkelstein SM, Speedie SM, Demiris G, Veen M, Lundgren JM, Potthoff S. Telehomecare: quality, perception, satisfaction. Telemed J E Health 2004;10(2):122-128. [doi: 10.1089/tmj.2004.10.122] [Medline: 15319041]
36. Agha Z, Schapira RM, Laud PW, McNutt G, Roter DL. Patient satisfaction with physician-patient communication during telemedicine. Telemed J E Health 2009 Nov;15(9):830-839. [doi: 10.1089/tmj.2009.0030] [Medline: 19919189]
37. Buchanan AH, Datta SK, Skinner CS, Hollowell GP, Beresford HF, Freeland T, et al. Randomized Trial of Telegenetics vs. In-Person Cancer Genetic Counseling: Cost, Patient Satisfaction and Attendance. J Genet Couns 2015 Dec;24(6):961-970 [FREE Full text] [doi: 10.1007/s10897-015-9836-6] [Medline: 25833335]
38. Portnoy JM, Waller M, De Lurgio S, Dinakar C. Telemedicine is as effective as in-person visits for patients with asthma. Ann Allergy Asthma Immunol 2016 Sep;117(3):241-245. [doi: 10.1016/j.anai.2016.07.012] [Medline: 27613456]

Abbreviations

DHIS: District Health Information Software
HCP: health care provider
mHealth: mobile health
NHS: national health service

Edited by G Eysenbach; submitted 25.11.20; peer-reviewed by G Pare, L Rabinovich, K Paik, A BenDor; comments to author 08.01.21; revised version received 26.02.21; accepted 08.10.21; published 30.11.22.

Please cite as:
Müller A, Cau A, Muhammed S, Abdullahi O, Hayward A, Nsanzimana S, Lester R
Digital mHealth and Virtual Care Use During COVID-19 in 4 Countries: Rapid Landscape Review
JMIR Form Res 2022;6(11):e26041
URL: https://formative.jmir.org/2022/11/e26041
doi:10.2196/26041
PMID:34932498

Original Paper

Using Digital Tools for Contact Tracing to Improve COVID-19 Safety in Schools: Qualitative Study Exploring Views and Experiences Among School Staff

Sofia Chantziara[1], PhD; Amberly Brigden L C[1], PhD; Claire H Mccallum[1], PhD; Ian J Craddock[1], PhD

Faculty of Engineering, University of Bristol, Bristol, United Kingdom

Corresponding Author:
Amberly Brigden L C, PhD
Faculty of Engineering
University of Bristol
Merchant Venturers Building
75 Woodland Rd
Bristol, BS81UB
United Kingdom
Phone: 44 0117 928
Email: amberly.brigden@bristol.ac.uk

Abstract

Background: Throughout the pandemic, governments worldwide have issued guidelines to manage the spread and impact of COVID-19 in schools, including measures around social distancing and contact tracing. Whether schools required support to implement these guidelines has not yet been explored in depth. Despite the development of a range of technologies to tackle COVID-19, such as contact-tracing apps and electronic vaccine certificates, research on their usefulness in school settings has been limited.

Objective: The aim of the study was to explore the needs of school staff in managing COVID-19 and their experiences and perspectives on technological support in relation to contact tracing. School staff are the ones likely to make key implementation decisions regarding new technologies, and they are also the ones responsible for using the new tools daily. Including both management staff and class teachers in the development of school-based technologies can lead to their successful adoption by schools.

Methods: Semistructured interviews were conducted with UK school staff, including primary and secondary school teachers and school managers. Thematic analysis, facilitated by NVivo, was used to analyze the data. Two of the authors independently coded 5 (28%) of the interviews and reached a consensus on a coding framework.

Results: Via purposive sampling, we recruited 18 participants from 5 schools. Findings showed that primary schools did not perform contact tracing, while in secondary schools, digital seating plans were used to identify close contacts in the classroom and manual investigations were also conducted identify social contacts. Participants reported that despite their efforts, high-risk interactions between students were not adequately monitored. There was a need to improve accuracy when identifying close contacts in common areas where students congregate. Proximity tracking, use of access cards, and closed-circuit television (CCTV) emerged as potential solutions, but there were concerns surrounding false alerts, burden, and security.

Conclusions: School staff have found it difficult to monitor and implement social distancing and contact-tracing provisions. There are opportunities for mobile digital technologies and CCTV to support school staff in keeping their students and colleagues safe; however, these must place minimal demands on staff and prioritize security measures. Study findings can help researchers and practitioners who work in different contexts and settings understand what particular challenges are faced by school staff, and inform further research on the design and application of digital solutions for contact tracing.

(JMIR Form Res 2022;6(11):e36412) doi:10.2196/36412

KEYWORDS

schools; contact tracing; COVID-19 mitigation; COVID-19; pandemic; disease prevention; health technology; COVID-19 management; technology support; digital tool; mobile health; mobile technology

Introduction

Background

Many governments worldwide used school closures as a way to temporarily reduce the spread of COVID-19. In the United Kingdom, schools closed the first time between March 18 and June 1, 2020, and a second time between January 4 and March 8, 2021 [1]. Since March 2021, schools have remained open; however, in the context of increased community transmission, high rates of COVID-19 have been observed among school-age children [2,3]. There are also strong indications that older children (10-16 years) can facilitate transmission and are more likely to introduce infections into a household compared to adults [3]. Previous research has suggested that there is a greater possibility for larger disease outbreaks in secondary schools compared to primary schools [4,5].

School closures should be avoided as they have a negative impact on children's social, physical, educational, and psychological development, with students from lower-income backgrounds impacted disproportionately [6]. They also affect parents' ability to work, (particularly women's), resulting in lower productivity and loss of income [6,7]. Supporting schools to stay open is crucial, yet it requires systems, such as testing, contact tracing, social distancing, and other support, to mitigate against infection transmission and to give students, staff, and parents confidence [8].

Over the course of the COVID-19 pandemic, countries have issued various sets of guidelines for schools [9-11]. These included measures for social distancing (eg, maintain distinct groups of pupils, minimize contact across the school site, stagger school start and finish times, maintain distance between teachers and pupils and between teachers) along with the use of regular testing, face coverings, contact tracing, and quarantine/testing of close contacts. Guidelines usually leave a certain degree of flexibility to schools to implement these measures in a way that meets the needs and age range of their students, the physical layout of each school, and the resources available [9]. School staff have nevertheless voiced concerns about how the guidelines can be followed in practice [12,13]. Since the implementation of these measures, little has been published on how school staff were managing and any challenges they have faced.

Over the past 2 years, technologies have been proposed to support COVID-19 management in several ways, including monitoring, surveillance, detection, and prevention [14,15]. One of the most high-profile uses of technology in tackling COVID-19 has been for contact tracing—a key measure in preventing the spread of infectious diseases [16]. Contact tracing involves identifying people who have been in contact with an infected individual and their subsequent isolation [16]. Digital technologies that support contact tracing, such as GPS chips capable of precise location tracking, Bluetooth radios that can sense the proximity between devices, and always-on connections to the internet, can increase efficiency over more labor-intensive manual methods [17,18]. Private companies have developed systems with the aim to assist schools and other businesses with contact tracing [19,20]; however, no studies so far have assessed their use for contact tracing within schools. A number of studies have aimed to investigate transmission models in schools and found that wireless sensors can more accurately identify close-proximity contacts of short durations compared to self-reported measures [21,22]. However, these studies were conducted before the COVID-19 pandemic and they did not look into the real-world adoption of these systems by schools.

Although contact-tracing tools have not been widely implemented in schools during the pandemic, surveillance technologies are already prevalent in UK, US, and Australian schools [23-25]. Biometric technologies, such as fingerprint scanners, have been used for library management and cashless catering [26,27]. Systems equipped with facial recognition technology, although less prevalent, have been installed in schools with the aim of detecting sex offenders or finding missing children [28] or to facilitate canteen payments [29]. The use of closed-circuit television (CCTV) is commonplace in many UK schools, and it is used primarily for the purpose of crime prevention and detection [24]. Previous studies on the acceptance of CCTV within schools have shown that perceived invasion of privacy relies upon a number of factors, such as the location of the cameras, their rationale, and whether individuals are being monitored continuously [26]. Studies have also suggested that a balance could be struck between the use of CCTV and the impact upon privacy [30].

Recent studies have highlighted that willingness to use contact-tracing technologies can be negatively affected by privacy concerns [31,32], even when concerns about COVID-19 remain equally high [31]. The use of these technologies raised concerns as individuals believed that contact tracing will involve increased surveillance by the government [30], were worried about third parties accessing their personal data, and had misunderstandings regarding what contact tracing will entail [30]. There are also examples of digital health programs that were abandoned as they failed to win public trust because of fears over privacy breaches and protection of anonymity [33,34].

It is important that any tools developed for schools be co-designed with potential users and stakeholders since this process ensures that they are more likely to be usable and engaging [35]. The first step in co-design is understanding the views and needs of potential users [36-38]. Exploring the context in which the tools would be deployed, and the views of potential users, can help ensure that any technology developed is likely to be useful, engaging, acceptable, and feasible to implement [39,40]. School staff, including both management staff and class teachers, are a key user group to involve in the development of school-based technologies [41,42]. The views of school staff can play a pivotal role in whether technologies in schools are successfully adopted [43,44], and they are also likely to make key implementation decisions and be responsible for using the tools daily.

Objective

In this study, we explored the challenges experienced by school staff in their efforts to limit the spread COVID-19 within the school, opportunities for technologies to support contact tracing, and considerations for the design of such technologies. The study was part of the wider COVID-19 Mapping and Mitigation in Schools (CoMMinS; R101587-103), a National Institute of

Health Research (NIHR), UK Research and Innovation (UKRI) project, which aimed to iterate and evaluate COVID-19 control and mitigation measures in schools through a program of active and responsive research conducted in partnership with schools.

Methods

Recruitment

Data for this study were collected from a subsample of primary and secondary schools in the wider area of Bristol (UK) that participated in the wider CoMMinS project. Special education needs schools were not included. We aimed to recruit a diverse range of schools based on the percentage of students from Black, Asian, and minority ethnic (BAME) groups; the percentage of students receiving funding to improve educational outcomes (student premium); the percentage of students eligible for free school meals (FSM); and indices of multiple deprivation (IMD). These data were ascertained from the UK government website [45]. Schools were also selected based on their capacity and willingness to engage in the study.

Participants were eligible if they were staff who held a teaching role, IT role, or school management role (eg, heads and deputy heads) or were otherwise tasked with managing COVID-19 within the school. Participants also needed to have access to video-call facility and be able to speak English. To recruit staff, a key contact was identified within each school, who liaised with the first author to advertise the study to the members of staff. This process included circulating an invitation letter to staff members, which included a link to the participant information sheet. In addition, a link to an online expression-of-interest form (via the electronic system REDCap [Vanderbilt University], a secure online data capture system designed exclusively for research [46]) was provided. On this expression-of-interest form, the participants were asked to provide contact details and were asked to provide consent for a researcher to contact them to provide full information about the study.

Ethical Considerations

Individuals who registered their interest in REDCap were contacted by the first author, who ensured that they were provided with and understood all the relevant information about the study. Informed consent in writing was obtained through REDCap from those individuals who agreed to participate. Participants were offered a £20 (US $22.18) voucher as a thank-you for taking part. All procedures were approved by the Faculty of Life Sciences Research Ethics Committee at the University of Bristol (reference no. 112284).

Data Collection

This qualitative study was grounded in the theory of phenomenology. Phenomenology aims to understand the meaning, structure, and essence of a lived experience of a particular phenomenon for individuals or a group of people [47]. In this context, the primary aim of qualitative research is to develop an understanding of how the world is constructed by the individuals involved in the research situation [48]. The aim of this study was to provide a deep understanding of participants' experience of the challenges they faced in their efforts to limit the spread COVID-19 within the school and their views on technologies that can support contact tracing.

Semistructured interviews were conducted between February and July 2021 by the first author via video calls. One of the benefits of interviews is the richness of data they can produce [49] compared to focus groups, which are more likely to give rise to attitudes, opinions, and third-person stories [50]. Semistructured interviews were chosen over fully structured ones as they promote a dialogue and allow the interviewer to explore in depth the thoughts, views, and experiences of participants, while additional questions can elicit more detailed narratives and stories [51]. As this study focused on an underinvestigated area, the aim was not to explore predetermined theories or themes and it was considered important to allow the participants during the interviews enough space to share their experiences and views.

A topic guide was developed by the research team to cover the main aims of the study. In the context of semistructured interviews, the topic guide was developed iteratively and refined and adjusted as interviews progressed (Textbox 1). The last section of the topic guide in particular (ie, digital solutions presented for discussion) was updated from one interview to the next as digital solutions were either suggested by participants (digital proforma, proximity tracking) or proposed by the researchers (CCTV, access cards, digital seating plans) based on participants' descriptions of the challenges they were facing when they were trying to identify close contacts of positive cases. During the first 2 interviews, no solutions were included in the topic guide. As the research team developed an understanding of how schools proceed with contact tracing and what were the blind spots in this process, they started formulating suggestions about digital tools that could be used to increase accuracy. Participants were also asked to provide their own ideas about digital tools. As interviews progressed, tools suggested by participants and researchers were added to the topic guide. Interviews lasted, on average, 30 minutes and were audio-recorded with participants' consent.

Textbox 1. Topic guide.

Participant/school characteristics

- What is your job role, and what does that involve in terms of managing COVID-19?
- Which areas, if any, can be characterized as hot spots for transmission (ie, areas where students/staff could interact)? How adequate are provisions for social distancing?

Current adherence to government guidance

- Can you tell us about the school's procedures for the management of COVID-19 in terms of social distancing and contact tracing?
- What happens if there is a positive case? What information is collected on the case?
- How do you identify contacts with the case? What information is used to identify contacts?
- Whom does the school notify (families, local authority), or who is alerted to the case?
- How do you record/evidence conversations?
- Does the school need to follow up/check-in on the case?
- Are there any steps to protect the anonymity of the case?

School staff needs and barriers to adherence

- What parts of the process are challenging?
- Do any parts/processes not work? If so, why not?
- What support, if any, do you/the school need to adhere to/implement guidelines?

Digital solutions

What technologies, if any, might help to support you/the school in managing COVID-19?

- Does your school have closed-circuit television (CCTV) and in which areas?
- Who has access to the footage? Do you need to get approval to review footage?
- What are your thoughts on using CCTV footage to identify close contacts?
- Do you think this could feasibly be reviewed to look at contact tracing outside the classroom (eg, lunch hall)?
- Do you foresee any challenges?

What are your thoughts on using proximity tracking for contact tracing?

Do you foresee any challenges? What are your thoughts on using access cards to enter common school areas? Do you foresee any challenges?

What are your thoughts on digital seating plans?

- Are pupils sticking to the seating plan? Are there are circumstances where they move around the classroom or change their desk?
- Primary schools: Do you have a seating plan, or are children mixing a lot within the class (may differ depending on the primary year group)?
- Are pupils allowed to have phones in schools, and are they allowed to use them?
- What are your thoughts on quick response (QR) codes on seats/tables to update seating plans?

What are your thoughts on a digital proforma that could be accessed and completed by the school and authorities?

- Would this make it easier/quicker for the school to collect and report information?
- Do you foresee any problems?

Analysis

Interviews were transcribed verbatim, and transcripts were pseudonymized. Thematic analysis was used to analyze the data, following the principles of a 6-stage process, as outlined by Braun and Clarke [52]. Analysis was facilitated by NVivo 12 (QSR International) [53]. All themes were produced inductively and were linked closely to participants' accounts. Two of the authors independently coded 5 (28%) of the interviews. The authors reached a consensus on a coding framework, and any disagreements were resolved through discussions before refining and finalizing themes and subthemes.

Results

Participants

The study included 5 schools: 1 primary (age range of students 5-11 years) and 4 secondary (age range of students 12-18 years). For 3 (60%) of the 5 schools, diversity measures were available. The sample included schools with very high to very low

deprivation scores, including 1 independent school. There was also diversity regarding the percentage of BAME students (38%-72%), students eligible for FSM (12%-36%), and students receiving a pupil premium (33%-56%). We initially aimed to purposively sample from the pool of interested participants, aiming for diversity on the role in school. However, because of the constraints in the number of participants expressing interest, recruitment was opportunistic, and we sampled all participants who completed an expression-of-interest form. All individuals (N=42) who provided their contact details via REDCap were contacted by the first author. Those who responded and agreed to participate were included in the study. Data collection ended when the research team concluded that any new information would have a minor or no influence on themes that already were emerging from participants' accounts and it was believed that saturation was reached [54]. Across the 6 schools, 18 (42.9%) participants were recruited. Of the 18 participants, 8 (44.4%) were female and 10 (55.6%) male; in addition, 4 (22.2%) were senior management, 12 (66.7%) were teachers, 1 (5.6%) a teaching assistant, and 1 (5.6%) a behavior support manager.

Themes

Three themes were identified that described the school staff's efforts and challenges in managing social distancing and contact tracing and their suggestions for digital solutions that could enhance existing provisions. These themes and subthemes along with quotes from interviews are outlined in Table 1 and described in detail later.

Table 1. Themes and subthemes.

Themes and subthemes	Interview quotes
Social distancing measures are in place to prevent and limit high-risk interactions; however, blind spots exist.	
Classrooms are a relatively controlled environment, with strict and easily observed social distancing measures in place in secondary schools, while more flexibility is allowed in primary schools.	• "The children, they were given seating plans, and the children had to stay in their seating plan, the teacher didn't mingle in with the pupils, so that...that was easy to manage inside the classroom." [Teacher S[a]] • "In early years and key stage 1, we do have group worktables, but they are a meter apart…the tables, so they're one either side. Or if they're next to each other, they're side by side rather than directly facing each other." [Teacher P[b]]
Mitigation measures are in place to prevent mixing between "bubbles"/year groups, and although social distancing is encouraged, it is not always possible.	• "We were able to stagger when different year groups were leaving school, entering school, and the different year groups had different areas in the school." [Teacher S] • "As soon as they all come out of those classrooms, you know, so there will always be excessive close contact in a corridor. I don't think logistically...that's very avoidable." [Teacher S]
Despite the existing measures, potentially risky interactions are still taking place, and mobile systems that monitor proximity between teachers and students could increase adherence.	• "Also, with teaching, it's effectively impossible to teach from the front of the class at all times, so there are certain times, um, that we find it a real struggle within our subjects." [Teacher S] • "A device that alerts you to the fact that you have been with this person or close to this person for more than 10 minutes now, it would be helpful." [Teacher S]
School staff use both manual and digital methods to support contact tracing; however, greater accuracy is needed.	
Contact tracing in secondary schools includes reviewing seating plans and consulting with the positive case and their circle of friends, staff, and family, and in this process, protecting the anonymity of the case is not the main priority.	• "Each case we had, we asked them were they happy with us sharing the information that they had tested positive so that we could trace any contacts, and every time they said yes, so with the pupils, there's not really any way around it." [Teacher S] • "We speak to the positive student, and if that means a phone call, that's a phone call home, and we just say, 'Who are your close contacts? Who are your friends? Who do you have lunch with?,' and then they tell us, and then, from there, we contact those students that have been named." [Teacher S]
In primary schools, contact tracing does not involve elaborate investigations, and in the presence of a positive case, whole bubbles are required to isolate.	• "[We] can't rely on them (the students) at all, and even, even with year 7 and 8, it's, it's difficult, because again I think the stigma around being isolated at the time meant they wouldn't want to tell you who they'd been in contact with." [Assistant head P and S] • "If they'd all been in their classes, no problem, you could just send one class home. If they'd been in mixed maths classes or a mixed PE[c] class, then we could have ended up in the situation where we would have to send potentially an entire year group home, depending on what they'd been doing, or at least a combination of bubbles might have to go." [Assistant head P and S]
Despite the manual investigations, students' cooperation, and the use of existing digital systems, contact tracing is not always comprehensive and accurate.	• "We had lots of students that wanted to be involved in the first case. You know, it was exciting they wanted to be a close contact. They thought going home would be great…and some of the students were right, and some of the students were just, you know, making it up, just to be involved in what was going on." [Assistant head S] • "You can ask the child, you can ask the parents, but they don't always necessarily know or they can't remember, so yeah, I guess that for me, that was the main thing is really like, I think I, I guess I just had to accept that it was never going to be fully accurate." [Teacher S]
The use of more flexible digital seating plans, CCTV[d], mobile proximity-tracking devices, and access cards could improve the accuracy of contact tracing, although concerns around privacy, acceptance, and technical limitations are prevalent among staff.	• "Okay (in relation to CCTV), my personal opinion is it wouldn't be ethical; I think it is the wrong way to use it…because you are literally spying on a person." [Teacher S] • "(in relation to proximity tracking) also I'm aware of the workforce needing to be in, and I don't want false alerts that mean that we can't run our school effectively because we've had to send people home, so some of that it needs to be utterly reliable." [Assistant head S]
No standardized procedures exist for storing and sharing information with authorities, and a digital system for interagency collaboration could assist schools' contact-tracing efforts.	• "Public Health England, actually the communication is dreadful. Our head teacher has had to sit on the phone spelling out her name, the school's name…it's all done in a really old-fashioned admin way." [Assistant head S] • "From first notification, we should just be able to upload all our information into a hub or a central record for our school." [Assistant head S]

Themes and subthemes	Interview quotes
Social distancing, contact-tracing provisions, and other COVID-19–related measures impact school functioning and place additional demands on staff and pupils.	
Measures have a negative impact on the teaching of practical subjects and other aspects of school life, although positive changes were also observed.	• "The more practical subjects have been definitely limited just because the teachers have been going to the students rather than the usual other way around, yes." [Assistant head S] • "We've seen things like cross-year bullying completely vanish by keeping those students separated." [Teacher S]
Staff have taken on additional responsibilities and roles, and their workload has increased.	• "We rewrote all of our schemes of learning to see if we could minimize the amount of physical contact the students had, while still being able to learn the skills we were trying to teach them." [Teacher S] • "Aa teacher, my role is [to] be sure that in the class, they're wearing masks, be sure that they sanitize their hands, and they're sitting where they're supposed to sit. Be sure that I have windows and door open." [Teacher S]
The behavior of students, along with factors inside and outside of school, could further complicate efforts to improve contact-tracing and social distancing provisions.	• "I think with technology, it's all dependent on students having something, and yes, without…with our all good intent and purposes, sometimes it's a battle trying to get students a pen [laugh] never even mind you know, something digital." [Teacher S] • "You know the…the breakages of equipment in school is…is…is unbelievable, I think, when you think about it…I don't know. I…I can't quite get my head round what that would look like." [Assistant head S]

[a]S: secondary.

[b]P: primary.

[c]PE: physical exercise.

[d]CCTV: closed-circuit television.

Theme 1: Social Distancing Measures Were Put in Place to Prevent and Limit High-Risk Interactions; However, Blind Spots Exist

School staff in both primary and secondary schools were attempting to follow national guidance; however, there appeared to be significant differences regarding the provisions they had in place to manage social distancing. In primary schools, where learning required group work and interaction (especially in early years), rules around social distancing and strict seating arrangements in the classroom were not considered appropriate. When possible, however, staff would choose seating arrangements that would minimize face-to-face contact, such as rows or a horseshoe. Each class would be considered a "bubble," whereby students would only mix with other students and staff in their bubble.

School staff in secondary schools used a number of additional systems and provisions that were in place. In the classrooms, strict seating arrangements were in place and participants reported using digital seating plans consistently; these had been generated by a digital system designed for classroom management. Bubbles were not used. During breaks, students would only mix with other students in their year group and there was not an expectation for students to socially distance within their year group. In the classroom, teachers would remain in a designated area 2 m apart from their students, and they continued to teach across different year groups.

To minimize interactions between different year groups in common areas, school staff in both primary and secondary schools implemented measures, such as staggered break times, keeping year groups separate, zoning, and 1-way systems. All staff were asked to keep at least 2 m distance from their colleagues, and changes were made to staff rooms to prevent close interactions.

Participants reported that high-risk interactions appeared to occur despite the measures that were put in place to prevent them. Staff members would not always keep their distance from their colleagues, and space limitations in some classrooms did not always allow them to stay 2 m apart from their students; and students in the same year groups were free to interact with each other in close proximity during break times in communal areas, such as halls and dining rooms. In secondary schools, science classes in particular presented challenges as there were instances when teachers where required to leave their designated area and approach the students to assist them with their experiments. One participant who was teaching science explained that it would be helpful to be reminded when they were in close proximity to their students. Receiving alerts from a digital system could potentially help prevent lengthy and risky interactions.

Theme 2: School Staff Use Both Manual and Digital Methods to Support Contact Tracing; However, Greater Accuracy Is Needed

In primary schools, where students and staff in each class formed a bubble, participants described contact tracing as a relatively simple process. When a positive case was identified, all students and staff in the same bubble were considered close contacts and were asked to isolate by the school. When students shared classes or activities with students from different bubbles, all those in the presence of a positive case would be sent into isolation. There was a perception among participants that primary school students could not be trusted to identify their close contacts. As a result, investigations often did not take

place and isolating everyone in the same class or activity with the positive case was considered a far more practical option.

In secondary schools, large numbers of students were mixing across their year group in common areas. Teachers also taught across multiple year groups. As a result, contact tracing involved elaborate investigations. Digital seating plans were used routinely to identify close contacts in the classroom. Manual investigations, such as interviews with teachers, students, and parents, were also conducted by the school's senior leadership team to identify social contacts. In 1 (16.7%) school, CCTV footage was reviewed routinely by staff to identify contacts in common school areas. Participants reported that maintaining the anonymity of the case was not a priority in contact-tracing investigations and stressed that no complaints had been raised by students, parents, or staff.

Participants, however, expressed strong concerns that manual investigations were not accurate. They admitted that seating plans were not always properly updated; participants would forget or struggle to update them if using alternative seating arrangements, and sometimes, students changed their allocated seat without first asking or informing their teacher. Participants explained that they would often keep separate files or instead take a quick photograph of the classroom. Recollection of close contacts by the positive cases was not always felt to be reliable, and participants would also come across contradictory accounts. Combined with the lack of any physical evidence of close interactions, participants had to manage these situations to the best of their ability and were relying heavily on their own memory and judgment.

Participants believed there was a need to improve the accuracy of contact tracing inside and outside the classroom and that digital solutions could help. Specifically, allowing students to update seating plans by scanning quick response (QR) codes on seats/tables was seen to have the potential to lead to more accurate plans and more flexible seating arrangements, along with more compliance from students. Participants felt CCTV cameras, which were already in place in all the schools, could help identify close contacts in areas such as dining halls. Mobile systems that could measure and track proximity between students and teachers (ie, proximity-tracking devices) were considered a valuable approach to improving accuracy. Less interactive technologies, including access cards, to monitor access in common school areas were also considered promising if they provided accurate information on the whereabouts of staff and students.

Despite their enthusiasm, participants did have reservations and concerns about these technological solutions. In relation to mobile proximity tracking, they were concerned about privacy violations as they felt that students would be monitored constantly within the school. They also highlighted the issue of consent and the need for everyone in the school to use this system. Furthermore, there were concerns about false proximity alerts that could send staff to self-isolation unnecessarily. Participants overwhelmingly expressed the view that manually reviewing CCTV footage would be time-consuming and therefore not feasible, and there were concerns regarding technical limitations of the CCTV system (eg, areas in the schools that are not covered by cameras, footage may not be clear and detailed enough to allow the identification of students). Furthermore, the majority of schools discussed by participants currently use CCTV footage only in exceptional circumstances to identify students who are involved in serious incidents, such as fights. Therefore, there were concerns among some participants about extending this to the infection control scenario as this would require frequent use of the system, which they considered invasive. Technical issues along the time required to review footage would make it difficult to make its use a common school practice.

Participants also highlighted challenges when contacting public health authorities (specifically the now disbanded Public Health England [PHE]) to report positive cases and to receive advice and support with contact tracing. The system required multiple phone calls and repeating the same information to different members of staff. Introducing a digital record that could be shared between schools and public health authorities was seen as a way to improve their communication and collaboration. However, this was not a challenge experienced by all participants, as the schools whose staff felt confident in managing contact tracing were not required to contact the authorities.

Theme 3: Social Distancing, Contact-Tracing Provisions, and Other COVID-19–Related Measures Impact School Functioning and Place Additional Demands on Staff and Pupils

Social distancing along with other provisions aiming to keep pupils and staff safe has had a profound impact on school life. Participants explained that the teaching of more practical subjects was especially difficult. In schools where zoning had been implemented, students in different year groups stayed in separate areas of the schools. As a result, access to rooms with equipment needed for drama, music, science classes, and outdoor facilities needed for physical exercise (PE) was restricted. Participants recounted trying to move equipment between classes but that this had significantly impacted the quality of teaching. PE sessions were also reduced in order to allow for cleaning of changing rooms.

To manage these changes, participants had taken on additional responsibilities, including moving equipment between classes, adjusting the curriculum, making sure to sterilize their own working areas, reminding students to wear masks and keep their distance, following up on students who tested positive, and managing the whole school response. Participants also highlighted that factors outside the school's control, such as mixed messages on what constitutes close contact in the school, not making testing obligatory for students, and change of rules around mask wearing and temperature testing, had further complicated their efforts to make the school a safe environment for their staff and students.

Despite the challenges, participants observed positive changes in the school environment. Cross-year bullying stopped in 1 (16.7%) school since year groups were kept separate, while enhanced cleaning and insistence on hygiene measures led to a perceived reduction in outbreaks of other illnesses. Some

participants also noted that their schools would permanently implement some of the changes.

Discussion

Principal Findings

Introducing CCTV, proximity tracking, and access cards into contact tracing could increase the accuracy and speed of contact tracing

This study explored the experiences of school staff in managing COVID-19, along with their requirements for digital solutions that can enhance their efforts to improve contact tracing and limit the spread of COVID-19 among students and staff. Although schools implemented government recommendations, school staff still found it difficult to limit and appropriately monitor high-risk contacts and strongly believed there was a need to improve accuracy in contact tracing. School staff in secondary schools faced more serious challenges as students interacted and congregated with other students in their year group, with year group sizes reaching 200, 300, or more students [55]. In primary schools, students were reported to have close contact with other students within their bubble, which includes an average of 30 students [56].

Findings suggest an opportunity for digital systems, particularly access cards, proximity tracking, and CCTV, to improve the accuracy and efficiency of contact tracing. Manual investigations (interviews with teachers, students, and parents) and reviews of seating plans were not considered accurate, as they relied on individuals' ability to recall interactions and to keep their records up to date. Introducing automation in contact tracing could increase the speed and accuracy of investigations and potentially have a more profound effect in secondary schools. The contacts of secondary school students can grow exponentially in the time it takes to conduct manual investigations or if some of the close contacts are missed in this process. Delays in quarantining or testing close contacts allows more space for the virus to spread within school.

Need to Understand and Address Privacy Expectations and Concerns Among Various School Stakeholders Before Introducing New Tools

Introducing mobile proximity-tracking devices and new uses for CCTV (primarily tracking of individuals for contact-tracing purposes) were felt to be promising but raised privacy concerns among the staff. Participants were specifically worried that the introduction of these solutions would result in the constant monitoring of students within the school. To address these concerns, more conceptual work is needed to unpick the notion of privacy and we should try to understand what the privacy expectations are across different settings within the schools. Although there is no universal definition of privacy, the concept of "reasonable expectations of privacy" determines in which places and activities a person can expect to have a right to privacy [57]. The concept highlights that expectations of privacy largely depend on the setting and circumstances. To this end, we should ask ourselves, What may be the reasonable expectations of privacy within schools? Schools are required to hold personal information about students and their families [58], and they are expected to monitor and supervise behavior to support the welfare and education of their students [59]. Individuals can also have different expectations, depending on where they are within the schools. There is probably no expectation of privacy in the classroom, where attendance and performance are strictly monitored, while school toilets are a place where individuals are not expected to be observed.

This study explored the experiences and perceptions of teachers, but there may be different expectations of privacy among the various stakeholders, including different members of the school community (teachers, parents, students). Parents many expect the school to inform them about their child's behavior, whereas students may not be keen for the school to observe or report on their behavior and keep their behavior private. Overall, digital solutions for contact tracing designed for schools should therefore prioritize security measures to address concerns around privacy. These measures, along with clear explanations regarding how they are going to be used and for what purposes, should be highlighted in the communication with the school community. Co-designing the solutions and exploring the concerns and expectations of privacy among different stakeholders will increase the likelihood that any solutions would be accepted with high uptake.

There were particular concerns about the technical requirements in using CCTV for contact-tracing purposes. Facial recognition technology, where captured images are compared against a database of pre-existing personally identifiable images in the system, should make it easier and faster to identify individuals [60]. However, this technology could be considered invasive, and its application has been met with resistance [25,29]. Exploring views regarding the application of facial recognition among members of the school community could provide more clarity on whether they would consider its use appropriate.

Contact-Tracing Tools Should Place Minimal Burden on the School, Staff, and Students

Participants were concerned about the demands of any contact-tracing systems on their schools' budget. Buying, fixing, and replacing equipment were suggested to be extremely difficult for schools, and there was also concern that mobile proximity tracking could generate false alerts and send the staff unnecessarily into isolation. This can lead to staff shortages and the need to bring in temporary staff at a great cost for the school. Any digital solution designed for schools should come with available technical support as schools are unlikely to cope with the ongoing demands of maintaining such a system. This study further suggests that accuracy (including avoiding false positives, which would unnecessarily require students and staff to isolate) should be 1 of the system's key features. Furthermore, reducing the need for involvement of staff and students in the application and use of any digital solution seems crucial. The staff have already taken on many responsibilities, and increasing their heavy workload would put additional pressure on them. Expecting students to take on an active role as users could create additional barriers, considering the different maturity levels among this population [61]. Therefore, digital solutions should be designed to fit in with schools' workflow and routine and require minimal interaction from individuals.

Limitations

This study concentrated on schools in the southwest of England. We chose this focused approach due to the lack of research in the area and the desire for an in-depth exploration of school staff's experiences. To build on our findings, a larger UK-wide quantitative study could enable broader generalizability. The topic guide was not pilot-tested, as the pace at which the COVID-19 pandemic has been unfolding required the research team to provide results and insight at a fast pace, thus creating additional time constraints.

Although the study achieved purposive sampling at the school level, that was not possible at the participant level. This study included a relatively small numbers of senior school staff (managers and teachers), mainly due to the high workload experienced by schools coping with the impact of the pandemic. Since senior members are tasked with managing the whole school response, they could have provided more insights regarding the impact and applicability of digital solutions. We did not manage to recruit staff members such as reception staff, other administration staff, and cleaners, and these staff members could have provided valuable insights into COVID-19 management in schools.

Conclusion

This qualitative study found that school staff reported a need for better COVID-19 mitigation measures, especially in secondary schools, and digital tools, such as CCTV and mobile proximity tracking and access cards, were described as potential solutions. It is important to ensure that any tools designed for schools prioritize privacy concerns and have minimal impact on staff, pupils, and day-to-day management. Further qualitative work would enable exploration of acceptability, feasibility, and engagement with the specific digital solutions that have emerged and explore the views of other key stakeholders, such as students, parents, and decision makers.

This study explored challenges faced by school staff in implementing COVID-19 measures and different provisions between secondary and primary to identify suitable digital tools. Overall, findings can help researchers and practitioners who work in different contexts and settings understand the particular challenges faced by school staff and inform further research on the design and application of digital solutions.

Acknowledgments

This work was supported by the National Institute for Health Research (NIHR) and UK Research and Innovation (UKRI) COVID-19 Rapid Response Initiative. The authors would like to thank all those involved in the COVID-19 Mapping and Mitigation in Schools (CoMMinS) study for their support and assistance with school recruitment.

Authors' Contributions

SC conducted data acquisition, analysis, and interpretation, and drafted and edited the manuscript. ABLC conceptualized and designed the study, contributed to the analysis and interpretation of data, and reviewed and edited the manuscript. CHM contributed to the analysis and interpretation of data, and reviewed and edited the manuscript. IJC conceptualized and designed the study, contributed to the analysis and interpretation of data, and reviewed and edited the manuscript.

Conflicts of Interest

None declared.

References

1. Institute for Government. Schools and Coronavirus: The Government's Handling of Education during the Pandemic. URL: https://www.instituteforgovernment.org.uk/sites/default/files/publications/schools-and-coronavirus.pdf [accessed 2022-01-10]
2. World Health Organization. Statement-Update on COVID-19: We Need Shift from Reaction to Stabilization Mode in This Crisis. URL: https://www.euro.who.int/en/about-us/regional-director/statements-and-speeches/2021/statement-update-on-covid-19-we-need-to-shift-from-reaction-to-stabilization-mode-in-this-crisis [accessed 2021-12-14]
3. UK Government. Covid-19 Children Transmission. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/935125/tfc-covid-19-children-transmission-s0860-041120.pdf [accessed 2021-12-14]
4. Aiano F, Mensah AA, McOwat K, Obi C, Vusirikala A, Powell AA, et al. COVID-19 outbreaks following full reopening of primary and secondary schools in England: cross-sectional national surveillance, November 2020. Lancet Reg Health Eur 2021 Jul;6:100120 [FREE Full text] [doi: 10.1016/j.lanepe.2021.100120] [Medline: 34278370]
5. Ismail SA, Saliba V, Lopez Bernal J, Ramsay ME, Ladhani SN. SARS-CoV-2 infection and transmission in educational settings: a prospective, cross-sectional analysis of infection clusters and outbreaks in England. Lancet Infect Dis 2021 Mar;21(3):344-353. [doi: 10.1016/s1473-3099(20)30882-3]
6. UNESCO. Covid-19 Education Response. URL: https://en.unesco.org/covid19/educationresponse/consequences [accessed 2021-08-20]
7. Royal Society. DELVE Initiative Balancing the Risks of Pupils Returning to Schools. URL: https://rs-delve.github.io/reports/2020/07/24/balancing-the-risk-of-pupils-returning-to-schools.html#schools-and-parental-employment-and-earnings [accessed 2021-08-20]

8. Viner RM, Bonell C, Drake L, Jourdan D, Davies N, Baltag V, et al. Reopening schools during the COVID-19 pandemic: governments must balance the uncertainty and risks of reopening schools against the clear harms associated with prolonged closure. Arch Dis Child 2021 Feb 03;106(2):111-113 [FREE Full text] [doi: 10.1136/archdischild-2020-319963] [Medline: 32747375]
9. UK Government. Education, University and childcare during Coronavirus. Action for Schools during Coronavirus Outbreak. URL: https://www.gov.uk/government/publications/actions-for-schools-during-the-coronavirus-outbreak [accessed 2021-08-21]
10. Johansen T, Astrup E, Jore S, Nilssen H, Dahlberg B, Klingenberg C. Infection prevention guidelines and considerations for paediatric risk groups when reopening primary schools during COVID-19 pandemic, Norway, April 2020. Eurosurveillance 2020 Jun 4;25(22):2000921. [doi: 10.2807/1560-7917.es.2020.25.22.2000921]
11. Centers for Disease Control and Prevention (CDC). Operational Strategy for K-12 Schools through Phased Prevention. URL: https://tinyurl.com/46sdnrth [accessed 2021-08-20]
12. Lorenc A, Kesten JM, Kidger J, Langford R, Horwood J. Reducing COVID-19 risk in schools: a qualitative examination of secondary school staff and family views and concerns in the South West of England. BMJ Paediatr Open 2021 Mar 10;5(1):e000987 [FREE Full text] [doi: 10.1136/bmjpo-2020-000987] [Medline: 34192194]
13. Kim L, Leary R, Asbury K. 'We need clear directions, if we're going to move forward. It's as simple as that': Teachers' narratives during partial school reopenings in the COVID-19 pandemic. psyarxiv Preprint posted online September 29, 2020. [doi: 10.31234/osf.io/m8scj]
14. Ting DSW, Carin L, Dzau V, Wong TY. Digital technology and COVID-19. Nat Med 2020 Apr 27;26(4):459-461 [FREE Full text] [doi: 10.1038/s41591-020-0824-5] [Medline: 32284618]
15. Mbunge E. Integrating emerging technologies into COVID-19 contact tracing: opportunities, challenges and pitfalls. Diabetes Metab Syndr 2020 Nov;14(6):1631-1636 [FREE Full text] [doi: 10.1016/j.dsx.2020.08.029] [Medline: 32892060]
16. World Health Organization. Infection Prevention and Control: Contact Tracing. URL: https://www.who.int/news-room/q-a-detail/contact-tracing [accessed 2021-08-20]
17. Fahey RA, Hino A. COVID-19, digital privacy, and the social limits on data-focused public health responses. Int J Inf Manage 2020 Dec;55:102181 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2020.102181] [Medline: 32836638]
18. Rodríguez P, Graña S, Alvarez-León EE, Battaglini M, Darias FJ, Hernán MA, RadarCovidPilot Group, et al. A population-based controlled experiment assessing the epidemiological impact of digital contact tracing. Nat Commun 2021 Jan 26;12(1):587 [FREE Full text] [doi: 10.1038/s41467-020-20817-6] [Medline: 33500407]
19. Estimote. Workplace Safety with Wearables. URL: https://estimote.com/ [accessed 2021-08-20]
20. Volan Technology. Solutions for Industry 4.0: Reimagine Manufacturing, Warehousing and Construction with New Labor-Utilization Insights. URL: https://volantechnology.com/solutions [accessed 2021-08-20]
21. Leecaster M, Toth DJA, Pettey WBP, Rainey JJ, Gao H, Uzicanin A, et al. Estimates of social contact in a middle school based on self-report and wireless sensor data. PLoS One 2016 Apr 21;11(4):e0153690 [FREE Full text] [doi: 10.1371/journal.pone.0153690] [Medline: 27100090]
22. Mastrandrea R, Fournet J, Barrat A. Contact patterns in a high school: a comparison between data collected using wearable sensors, contact diaries and friendship surveys. PLoS One 2015 Sep 1;10(9):e0136497 [FREE Full text] [doi: 10.1371/journal.pone.0136497] [Medline: 26325289]
23. Nemorin S. Post-panoptic pedagogies: the changing nature of school surveillance in the digital age. Surveill Soc 2017 May 08;15(2):239-253. [doi: 10.24908/ss.v15i2.6033]
24. Taylor E. UK schools, CCTV and the Data Protection Act 1998. J Educ Policy 2011 Jan;26(1):1-15. [doi: 10.1080/02680939.2010.493226]
25. Andrejevic M, Selwyn N. Facial recognition technology in schools: critical questions and concerns. Learn Media Technol 2019 Nov 05;45(2):115-128. [doi: 10.1080/17439884.2020.1686014]
26. Bryce TGK, Nellis M, Corrigan A, Gallagher HG, Lee P, Sercombe H. Biometric surveillance in schools : cause for concern or case for curriculum? Scottish Educ Rev 2010;42(1):3-22.
27. Taylor E. I spy with my little eye: the use of CCTV in schools and the impact on privacy. Sociol Rev 2010 Aug 01;58(3):381-405. [doi: 10.1111/j.1467-954x.2010.01930.x]
28. Kupchik A, Monahan T. The New American School: preparation for post - industrial discipline. Br J Sociol Educ 2006 Nov;27(5):617-631. [doi: 10.1080/01425690600958816]
29. Guardian. ICO to Step in after Schools Use Facial Recognition to Speed up Lunch Queue. URL: https://tinyurl.com/c29xxchy [accessed 2021-10-10]
30. Taylor E. Surveillance Schools: Security, Discipline and Control in Contemporary Education. Berlin: Springer; 2013.
31. Chan EY, Saqib NU. Privacy concerns can explain unwillingness to download and use contact tracing apps when COVID-19 concerns are high. Comput Human Behav 2021 Jun;119:106718 [FREE Full text] [doi: 10.1016/j.chb.2021.106718] [Medline: 33526957]
32. Williams SN, Armitage CJ, Tampe T, Dienes K. Public attitudes towards COVID-19 contact tracing apps: a UK-based focus group study. Health Expect 2021 Apr 12;24(2):377-385 [FREE Full text] [doi: 10.1111/hex.13179] [Medline: 33434404]

33. Godlee F. What can we salvage from care.data? BMJ 2016 Jul 14;354:i3907. [doi: 10.1136/bmj.i3907]
34. van Staa T, Goldacre B, Buchan I, Smeeth L. Big health data: the need to earn public trust. BMJ 2016 Jul 14;354:i3636. [doi: 10.1136/bmj.i3636] [Medline: 27418128]
35. 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 30;17(1):e30 [FREE Full text] [doi: 10.2196/jmir.4055] [Medline: 25639757]
36. Bowen S, Sustar H, Wolstenholme D, Dearden A. Engaging teenagers productively in service design. Int J Child Comput Interact 2013 Sep;1(3-4):71-81 [FREE Full text] [doi: 10.1016/j.ijcci.2014.02.001] [Medline: 26516621]
37. Boyd H, McKernon S, Mullin B, Old A. Improving healthcare through the use of co-design. N Z Med J 2012 Jun 29;125(1357):76-87. [Medline: 22854362]
38. Rogers Y, Marshall P. Research in the Wild. Synth Lect Hum-Centered Inform 2017 Apr 04;10(3):i-97. [doi: 10.2200/s00764ed1v01y201703hci037]
39. Bevan Jones R, Stallard P, Agha SS, Rice S, Werner-Seidler A, Stasiak K, et al. Practitioner review: co-design of digital mental health technologies with children and young people. J Child Psychol Psychiatry 2020 Aug 22;61(8):928-940 [FREE Full text] [doi: 10.1111/jcpp.13258] [Medline: 32572961]
40. Thabrew H, Fleming T, Hetrick S, Merry S. Co-design of eHealth interventions with children and young people. Front Psychiatry 2018 Oct 18;9:481 [FREE Full text] [doi: 10.3389/fpsyt.2018.00481] [Medline: 30405450]
41. Polman J, Kohnen A, Whitacre M. In: Bonsignore E, DiSalvo B, DiSalvo C, Yip J, editors. Introduction to Participatory Design in the Learning Sciences. New York, NY: Routledge; 2017:59-70.
42. Kucirkova N. iRPD—a framework for guiding design‐based research for iPad apps. Br J Educ Technol 2016 Jan 27;48(2):598-610. [doi: 10.1111/bjet.12389]
43. Bhati N, Mercer S, Rankin K, Thomas B. Barriers and facilitators to the adoption of tools for online pedagogy. Int J Pedagog Learn 2014 Dec 17;5(3):5-19. [doi: 10.5172/ijpl.5.3.5]
44. Molebatsi M. Socio-technical factors that influence learning management systems' adoption in developing countries. In: Brown T, van der Merwe H, editors. The Mobile Learning Voyage - From Small Ripples to Massive Open Waters. mLearn 2015. Communications in Computer and Information Science, vol 560. Cham: Springer; 2015:391-404.
45. UK Government. Explore Our Statistics and Data. URL: https://explore-education-statistics.service.gov.uk/ [accessed 2021-03-03]
46. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
47. Lyncoln E, Cuba C. Conceptualizing issues in qualitative inquiry. In: Patton MQ, editor. Qualitative Research and Evaluation Methods. London, UK: Sage; 2002.
48. McLeod J. Qualitative Research in Counselling and Psychotherapy. London, UK: Sage; 2011.
49. Harvey-Jordan S, Long S. The process and the pitfalls of semi-structured interviews. Commun Pract 2001 Jun 1;74(6):219.
50. Smith J, Fieldsend M. Interpretative phenomenological analysis. In: Camic PM, editor. Qualitative Research in Psychology: Expanding Perspectives in Methodology and Design. Washington, DC: American Psychological Association; 2021:147-166.
51. Dicicco-Bloom B, Crabtree BF. The qualitative research interview. Med Educ 2006 Apr;40(4):314-321. [doi: 10.1111/j.1365-2929.2006.02418.x] [Medline: 16573666]
52. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
53. Edhlund B, McDougall A. NVivo 12 Essentials. Morrisville, NC: Lulu Press; 2019:11.
54. Guest G, Bunce A, Johnson L. How many interviews are enough? An experiment with data saturation and variability. Field Methods 2016 Jul 21;18(1):59-82. [doi: 10.1177/1525822x05279903]
55. UK Government. Number of Pupils on Roll in Secondary Schools by Year Group. URL: https://data.gov.uk/dataset/27d4c0ea-9148-44a9-b6f5-9a70401a339e/number-of-pupils-on-roll-in-secondary-schools-by-year-group [accessed 2021-12-10]
56. Department for Education. Class Size and Education in England Evidence Report. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/183364/DFE-RR169.pdf [accessed 2021-12-10]
57. Selbst AD. Contextual expectations of privacy. Cardozo Law Rev 2013;35:643. [doi: 10.2139/ssrn.2093594]
58. Department for Education. Keeping and Maintaining Records. URL: https://dera.ioe.ac.uk/22630/8/Keeping%20and%20maintaining%20records%20-%20GOV_UK.pdf [accessed 2022-01-10]
59. Department for Education. Behaviour and Discipline in Schools. URL: https://tinyurl.com/2p9483fk [accessed 2022-01-10]
60. Celine J, Sheeja AA. Face recognition in CCTV systems. 2019 Presented at: ICSSIT 2019: 2nd International Conference on Smart Systems and Inventive Technology; November 27-29, 2019; Tirunelveli, India p. 111-116. [doi: 10.1109/ICSSIT46314.2019.8987961]
61. Galambos NL, Barker ET, Tilton-Weaver LC. Who gets caught at maturity gap? A study of pseudomature, immature, and mature adolescents. Int J Behav Dev 2016 Jun 30;27(3):253-263. [doi: 10.1080/01650250244000326]

Abbreviations

BAME: Black, Asian, and minority ethnic
CCTV: closed-circuit television
CoMMinS: COVID-19 Mapping and Mitigation in Schools
FSM: free school meals
PE: physical exercise
PHE: Public Health England
QR: quick response

Edited by A Mavragani; submitted 14.01.22; peer-reviewed by E Loh, S Pit; comments to author 05.08.22; revised version received 27.09.22; accepted 03.10.22; published 01.11.22.

Please cite as:
Chantziara S, Brigden L C A, Mccallum CH, Craddock IJ
Using Digital Tools for Contact Tracing to Improve COVID-19 Safety in Schools: Qualitative Study Exploring Views and Experiences Among School Staff
JMIR Form Res 2022;6(11):e36412
URL: https://formative.jmir.org/2022/11/e36412
doi:10.2196/36412
PMID:36191172

Original Paper

The Experience of Health Professionals With Misinformation and Its Impact on Their Job Practice: Qualitative Interview Study

Nashwa Ismail[1], PhD; Dhouha Kbaier[2], PhD; Tracie Farrell[3], PhD; Annemarie Kane[4], PhD

[1]School of Education, Durham University, Durham, United Kingdom

[2]School of Computing and Communications, The Open University, Milton Keynes, United Kingdom

[3]Knowledge Media Institute, The Open University, Milton Keynes, United Kingdom

[4]Faculty of Arts and Social Sciences, The Open University, Milton Keynes, United Kingdom

Corresponding Author:
Dhouha Kbaier, PhD
School of Computing and Communications
The Open University
Walton Hall, Kents Hill
Milton Keynes, MK7 6AA
United Kingdom
Phone: 44 07778649257
Email: dhouha.kbaier@open.ac.uk

Abstract

Background: Misinformation is often disseminated through social media, where information is spread rapidly and easily. Misinformation affects many patients' decisions to follow a treatment prescribed by health professionals (HPs). For example, chronic patients (eg, those with diabetes) may not follow their prescribed treatment plans. During the recent pandemic, misinformed people rejected COVID-19 vaccines and public health measures, such as masking and physical distancing, and used unproven treatments.

Objective: This study investigated the impact of health-threatening misinformation on the practices of health care professionals in the United Kingdom, especially during the outbreaks of diseases where a great amount of health-threatening misinformation is produced and released. The study examined the misinformation surrounding the COVID-19 outbreak to determine how it may have impacted practitioners' perceptions of misinformation and how that may have influenced their practice. In particular, this study explored the answers to the following questions: How do HPs react when they learn that a patient has been misinformed? What misinformation do they believe has the greatest impact on medical practice? What aspects of change and intervention in HPs' practice are in response to misinformation?

Methods: This research followed a qualitative approach to collect rich data from a smaller subset of health care practitioners working in the United Kingdom. Data were collected through 1-to-1 online interviews with 13 health practitioners, including junior and senior physicians and nurses in the United Kingdom.

Results: Research findings indicated that HPs view misinformation in different ways according to the scenario in which it occurs. Some HPs consider it to be an acute incident exacerbated by the pandemic, while others see it as an ongoing phenomenon (always present) and address it as part of their daily work. HPs are developing pathways for dealing with misinformation. Two main pathways were identified: first, to educate the patient through coaching, advising, or patronizing and, second, to devote resources, such as time and effort, to facilitate 2-way communication between the patient and the health care provider through listening and talking to them.

Conclusions: HPs do not receive the confidence they deserve from patients. The lack of trust in health care practitioners has been attributed to several factors, including (1) trusting alternative sources of information (eg, social media) (2) patients' doubts about HPs' experience (eg, a junior doctor with limited experience), and (3) limited time and availability for patients, especially during the pandemic. There are 2 dimensions of trust: patient-HP trust and patient-information trust. There are 2 necessary actions to address the issue of lack of trust in these dimensions: (1) building trust and (2) maintaining trust. The main recommendations of the HPs are to listen to patients, give them more time, and seek evidence-based resources.

(JMIR Form Res 2022;6(11):e38794) doi:10.2196/38794

KEYWORDS

health misinformation; social media; health professional; patients; trust; communication, COVID-19; intervention; qualitative research; interpretive phenomenological analysis; thematic analysis; misinformation; health practitioner; infodemiology

Introduction

Background

Health misinformation is currently recognized as a significant issue [1-3]. Misinformation has been defined as information that has no scientific evidence to support it and is contradictory to the most recent, most reliable evidence [4,5]. Wang et al [3] distinguish further between *misinformation* and *disinformation* posted online and on social media platforms in particular. According to a study published by the Council of Europe recently [6], the 2 terms are defined in terms of intent to harm. *Misinformation* occurs when inaccurate information is disseminated, and it was not *intended* to cause harm. Misinformation that is *intended* to harm is called *disinformation*. False propaganda containing harassment, hate speech, and an intent to harm is considered *malinformation* [6]. Obviously, the spread of misinformation is not new but dates back to the early days of printing [3], and concerns about fake news and misinformation in traditional media have been prevalent since the early decades of the 20th century [7]. With the advent of digital technology and the internet, misinformation and disinformation have significantly changed how they are communicated worldwide and amplified rapidly. For health professionals (HPs), digital technology and the internet can play a critical role in combatting medical misinformation but have not got an opportunity to fully address medical misinformation.

During the COVID-19 pandemic, misinformation has been increasingly presented, based on anecdotal evidence, false information, or misleading information due to a lack of existing scientific knowledge. This information is false but not created with the intention of causing harm [8]. However, the onslaught of misinformation can lead to risky behaviors or reduced trust in authorities. In addition to investigating ways to identify and counter misinformation, researchers have focused on the consumers of misinformation [9], not only as recipients, but also as potential amplifiers. Who is being misinformed, and what does it mean to be misinformed?

Certain personal characteristics or demographic features have been implicated in the spread of misinformation, for example, the characteristics of extroversion and cooperativeness [10,11], dogmatism and religious beliefs [12], and overconfidence in one's knowledge and critical analytic skills [13]. With regard to demographic features, Guess et al [14] found that conservatives are more likely to share news from fake news sources, such as Facebook, and are 7 times more likely to share news from fake domains. Guess et al [14] differentiated Facebook as an example of an echo chamber where beliefs are amplified or reinforced by communication and repetition inside a closed system. Cinelli et al [15] argued that echo chambers limit exposure to diverse perspectives and favor and reinforce presupposed narratives and ideologies. There also exist fake domains (eg, articles and websites with unknown sources). The term "fake domain" refers to a situation in which an adversary creates a fake website or social media profile for a variety of reasons, including creating confusion among a targeted community. More recently, Guess et al [16] addressed the opposite, as sharing information via Facebook was a relatively rare activity. It was found that conservatives are more likely to share articles from fake domains. Consequently, such studies can be difficult to interpret. The same study, for example, determined that older users are also more likely to share facts. Studies on the role of partisan thinking and misinformation have also had mixed results. Some studies show that conservatives share more misinformation [17,18], while other studies have argued that this correlation may be related to other potentially confounding factors, such as perceived bias in the media [19] or shared information processing tendencies of conservative versus liberal individuals [20]. Harper and Baguley [20] demonstrated that liberals and conservatives are equally vulnerable to believing misinformation but for different reasons. The authors found that the greater the partisan attachment (on either side), the more willing individuals appear to be in engaging in "cognitive distortion" to protect their views. Even whole communities of individuals can be misinformed due to their exclusion from mainstream society, such as migrant networks [21] or niche online communities [22].

This study investigates disease outbreaks where a large amount of health misinformation is produced and released. We focused on the misinformation that is related to HPs' specialist areas using COVID-19 as a case study of a disease outbreak (1) to investigate how misinformation may have impacted health practitioners' job practice, how they witnessed it occurring, and how they interpreted it, intervened, and responded to it and (2) to examine whether any reshaping of practices occurred or was expected to occur in response to misinformed patients. Due to the uncertainty regarding the sources of misinformation in this study, we considered all sources of misinformation, including echo chambers both online (eg, Facebook) and offline (eg, committees in religious worship places, such as mosques and churches).

Health Professionals' Terminology in This Study

Various terms are used in the literature to describe staff members who work in the health sector and provide health services to patients. For this research, we combined definitions from Medscape [23] and the National Health Service (NHS) [24] to determine the job description of individuals in the health sector who can serve the purpose of the study. According to Medscape [23], an HP is a provider of health care treatment and advice based on formal training and experience. The field includes those who work as nurses and physicians of all specialties and those who perform services in allied health professions. Public health and community health experts are also health professionals (HPs). According to the NHS, allied health professionals (AHPs) also include 14 categories (eg, osteopaths) that provide solution-focused, goal-centered care to support patients' independence and help them with day-to-day living. Therefore, in this study, HPs include these 2 groups under the

condition of having 1-to-1 communication and discussions with the patients. More details about the inclusion and exclusion of participants can be found in the Methods section.

Research Questions

This qualitative study specifically explored responses to the following research questions (RQs):

- RQ1: How do HPs react when they learn that a patient has been misinformed?
- RQ2: What misinformation do they believe has the greatest impact on medical practice?
- RQ3: What aspects of change and intervention in HPs' practice are in response to misinformation?

Due to the many potential variables influencing health practitioners' experience and recommendations around dealing with misinformation, this research was determined to be best served by a qualitative approach, allowing us to collect rich data from a smaller subset of health care practitioners working in the United Kingdom.

Methods

Methodology

This study used a qualitative approach that shows the data findings and results from semistructured interviews that were conducted with HPs and that were audio-recorded. The interviews lasted between 30 and 45 minutes, with an average duration of 40 minutes. HPs included doctors and nurses from different areas in the United Kingdom, as well as 1 HP from the United States. Participant interviews were conducted either online (Teams or Zoom) or via phone call at the participant's convenience.

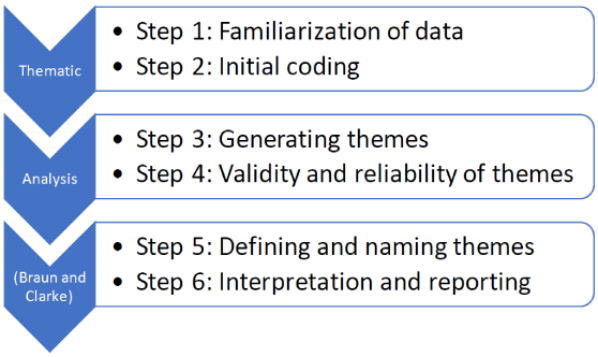
Ethical Considerations

The study was approved by the Human Research Ethics Committee at the Open University (reference no. HREC/3960/Kbaier, approval date April 14, 2021). The conducted interviews followed a predefined semistructured interview protocol informed by the study's aims and objectives.

Data Collection

Data were collected through 1-to-1 online interviews (Zoom/Teams/online) with 13 health practitioners, including junior and senior physicians as well as nurses. These interviews were transcribed and analyzed using the thematic analysis (TA) approach to understand and interpret the perceived experiences of HPs during COVID-19. Participants in the study were recruited via a combination of convenience and snowball sampling methods. The HPs' job titles were doctors and nurses. The doctors were a mix of men and women. Their professional experience ranged from junior to senior positions in different specialties, such as psychiatry and hematology. Geographically, the majority of HPs (n=12, 92%) were from the United Kingdom, except for 1 (8%) from the United States. It was useful to include a non-UK perspective to enhance understanding of the contextual factors. The HPs from the United Kingdom worked in various regional locations, including the Southeast, the Southwest, and the Midlands. Figure 1 illustrates the applied framework to analyze the qualitative data thematically, starting with reading and familiarization with the narrative across interviews, progressing to reporting the study results [25].

Figure 1. Framework for conducting a TA. TA: thematic analysis.



Recruitment of Participants

The research team contacted (by email or telephone) the communities where the target population work in the health sector (eg, doctors and nurses in general practice in their local areas). These individuals were then asked to suggest other prospective colleagues for the study. This method is referred to as snowball sampling [26] because (in theory) once the ball starts rolling, it picks up more "snow" along the way (recruits more individuals) and the sample size grows progressively larger. In this research, the snowball sampling process consisted of 2 steps:

- Identifying potential participants: HPs they may work with or know in their workplace
- Asking those HPs to recruit other people (and then asking those people to recruit others and so on)

Inclusion Criteria

Participants were selected based on 3 main criteria:

- Participants were currently working in the health care sector, with a job description that requires direct, day-to-day interaction, including discussions with patients. For example, staff members who were working in data entry without having any contact with patients were excluded.

- Participants have witnessed misinformation regardless of how they define what misinformation is.
- Participants' experience extended to the period before and during the pandemic.

It is important to emphasize that misinformation that participants narrated their experience with is not about COVID-19 but about their own specialist area, as the research investigated the impact of misinformation more generally using the pandemic as a practical example of a disease outbreak. For example, an endocrine doctor narrates their experience about patients being misinformed by a family member to eat a specific fruit, believing that it would lower the level of blood sugar. In this example, the HP explained that in response to the lockdown during the pandemic, patients did not get an opportunity to have frequent face-to-face meetings with the allocated doctors and that justifies the reason for misinformation from the HP's view.

Participants who expressed their interest in taking part in the research were then sent a participant information sheet (PIS) to explain the purpose of the research, what participants will be required to do, and how they will be involved. In addition, they were asked to sign a written consent form and send it back to the researchers. Within this consent, they were informed that the interview would be audio-recorded, and they were given the choice of whether to continue their participation in the research interviews and how to withdraw, if they wished. Finally, they were sent an email with an invitation to the interview.

Interrater Reliability

Figure 2 illustrates the different steps involved in the interrater reliability (IRR) process. The 2 researchers coded and analyzed the 3 agreed interviews independently, followed by comparing their coding results (superthemes and subthemes) to highlight any agreements and disagreements on the sufficiency and adequacy of the baseline data [27]. An important component of qualitative research is the identification and negotiation of disagreements throughout the iterative process of developing a codebook [28].

Reliability is calculated as the number of agreements divided by the total number of agreements and disagreements. Both node structures between the 2 coders were compared as follows:

- Score the same existing theme in both node structures as "1 agreement."
- Score nonexisting themes in both node structures as "1 disagreement."
- Add up all agreed themes versus adding up disagreed themes and calculate the percentage difference to determine the percentage of agreement between the 2 researchers.

We (researchers/coders) obtained a result of 78.3% agreement in our approach. Miles and Huberman [27] recommend that when coding 50 statements of transcripts, 80% of agreement between coders is used as a percentage difference target. However, we argue for this result of 78.3% because the 2 coders agreed to share the coding of the 2 full transcripts (about 600 statements), which resulted in minimizing the expected percentage of agreement between the 2 coders.

To ensure that the data collected within qualitative and quantitative research were correctly interpreted by the research team and can be used to build new insights, it is imperative that data analysis be conducted using best practices [29]. These best practices should include methods to safeguard the trustworthiness and quality of the research. According to McAlister et al [30], when one is using qualitative coding techniques, establishing the IRR is a recognized method of ensuring the trustworthiness of the study when multiple researchers are involved with coding. In the course of our study, we used the IRR to test the strength of our categories for our codebook development. Being under the threshold of 80% agreement may indicate that there are categories that need to be further negotiated in subsequent studies. Consequently, in this research, trustworthiness through IRR gauged how well the evidence presented supported the value of the results, while quality analysis measured how likely systematic error and bias were prevented through the design of the study.

Figure 2. Applied steps in the IRR exercise. IRR: interrater reliability.

Codebook Development

Our RQ was about HPs' experience with misinformation and its impact on their professional practice. Two researchers, with shared experiences in collecting and analyzing qualitative data, collaboratively developed the codebook informed by many meetings before, during, and after the data collection to develop the RQs, following up on the flow of the interviews and discussing the emerging themes after the interviews.

A total of 13 interviews were conducted and transcribed. The initial plan was to select randomly 1 (8%) interview of the 13 to share coding and run the IRR exercise. However, running the "Word Frequency" query over the 13 interviews using NVivo (QSR International) identified possible themes, particularly in the early stages of the project [31]. It showed that there were repeated uses of words and their synonyms in the interviews that reflected perspectives of the interviews that underpinned the interviewees' opinions about misinformation. For example, in the political perspective interview, words such as "political affiliation," "bias," "league," "party," and "political" were frequently repeated. Therefore, after identifying the 3 main perspectives (professional, social and cultural, and political) both coders agreed to select and code 1 interview from each perspective.

The 13 interviews were then grouped according to these perspectives into 5 professional, 7 social and cultural, and 1 political interview. Next, the researchers selected 1 interview sample that related to each of the 3 themes. In total, 3 interviews were selected to practice the coding together and to conduct the IRR exercise. After that, the 2 coders started to code the 3 selected interviews separately. Following O'Connor et al [32], the researchers identified initial themes through group discussions of overlaps and divergences. Before coding, researchers started to cobuild the baseline of understanding the research aims, RQs, and interviewees' professional backgrounds, how HPs witnessed misinformation in the job practice, and how to communicate with their patients accordingly. Moreover, in-depth details about the HPs' job practice and approach to dealing with misinformed patients and the applied approach, training, and resources that the HPs follow or recommend were determined. This phase ended with developing the codebook (see Table 1), followed by the IRR practice (as shown in Figure 2). The structure of the combined deductive and inductive codes is provided in Multimedia Appendix 1. Table 1 shows an example of the developed codebook.

Table 1. Example of subthemes in the codebook, definitions, and quote examples.

Code	Definition	Example
Authority	Subcode to the supercode "blind trust" or "status of information and perception	• "Sometimes, it could be their religious leaders; sometimes it could be their elders."
Logic	Statements about patient epistemology relying on logic as "evidence" of misinformation or information	• "I think it's easier when you're treating somebody from the leafy suburbs of, let's say, Southampton, where we live, people who are well-off, high up on this economical, know how to search for up-to-date health information, look at the guidelines, and they come armed with a lot of knowledge which then, and you can easily win them over in an argument."
Availability	Absence of or lack of exposure to verified information leading patients to trust easily accessible information (eg, Google or social media)	• "When you don't have information, you tend to start to pick up sources which are not necessarily valid, not recommended by a health practitioner, or could be related to other parts of the world." • "Sometimes, it's the patient; they resist the treatment because they, someone from the family, told them that, something."

Results

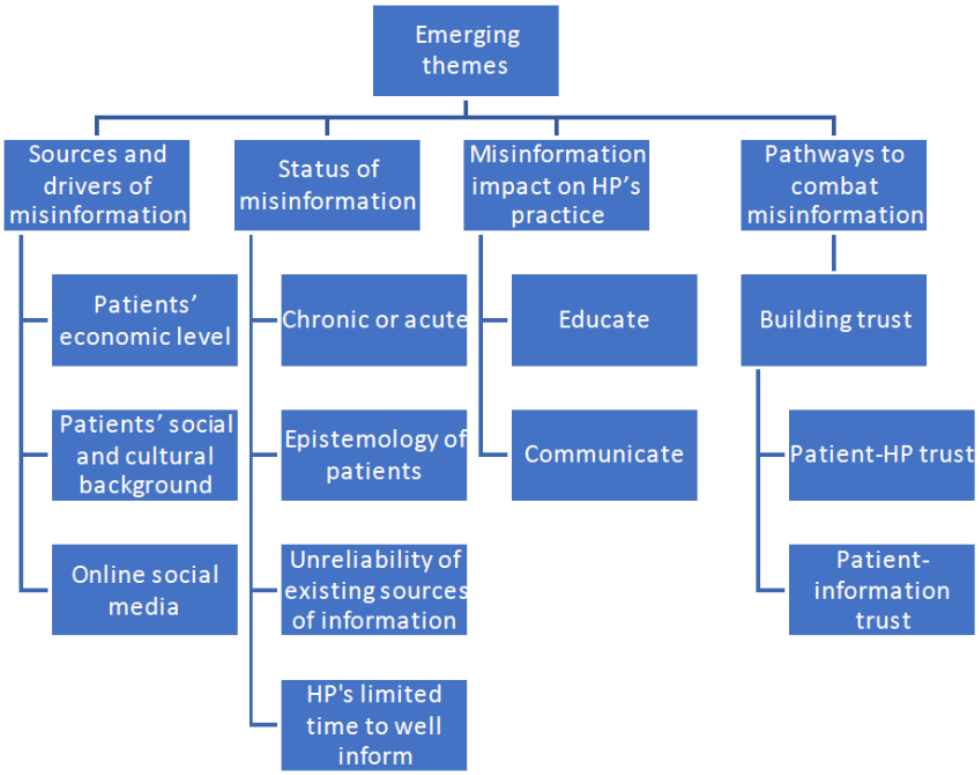
Data Analysis and Findings

Qualitative data were thematically analyzed to focus on the key aspects related to the study RQs. TA is the search for and extraction of general patterns found in the data through careful reading and re-reading across the qualitative data set [25]. TA flexibility involves making several decisions regarding data collection and analysis before they are undertaken. In this research, the initial analysis started with ongoing discussions between the researchers; these discussions informed the development of the codebook. Within the analysis and flexibility of TA, there were opportunities for newly developed themes (other than those in the codebook) to emerge within the data analysis. TA began while the interviews were ongoing, and transcripts were analyzed one by one using NVivo 12. As the analysis progressed, a figure of emergent codes was developed and refined (see Figure 3). Each new transcript led to codes being further expanded or adjusted. Once all the transcripts were analyzed, researchers refined each code in order to identify any duplicate coding or emerging patterns. Upon completing this process, the themes were developed, refined, and named.

Findings from the analysis of qualitative interviews navigated us on a journey starting from the origins of misinformation (patient's education, online sources) to the status of misinformation in HPs' job practice regarding its occurrence—whether it is a novelty or an ongoing phenomenon. The status of misinformation includes how patients think about, accept, and trust this knowledge, highlighting 2 main challenges addressed by HPs that are considered catalysts for the spreading of misinformation. First, existing information is somehow unreliable (eg, out of date). Second, HPs have limited time to meet, listen to, and talk with their patients. For HPs, misinformation impacted their job practice through applying 2 different approaches: (1) educate the patients through multiple methods (eg, HP's advisory role) and (2) communicate with the patient. However, the pathways to implementing these approaches focus on building patient-HP relationships and building trust between them. "Trust" is an emerging theme that includes patient-HP trust and patient-information trust (see Figure 3).

The following section details the identified themes with supporting quotes.

Figure 3. Summary of emerged themes. HP: health professional.



Emergent Themes

Sources and Drivers of Misinformation

HPs narrated 2 main reasons that are behind misinformation: (1) patients' backgrounds and (2) the unreliability of the information available. With regard to the patients themselves, reasons such as education, culture, religion, and political affiliation were listed as sources of information.

Patients' Educational Background and Misinformation

Interestingly, according to HP4, the educational level (high or middle level of education) may drive the patients' information toward 2 alternative routes, either follow or unfollow.

> *You'll be surprised to know that…I'm not going to say that the very poor people, but I'm just going to say the working category among the people. They were very keen to follow the right information. And as you go up in the society, when you go to, a little bit higher-educated people; unfortunately, they was the people who was not getting the right information because of their political backgrounds or whatever kind of media they listen to.* [HP4]

Patients' Social and Cultural Background and Misinformation

Furthermore, the interviews with the HPs revealed the role of the community and word of mouth in spreading misinformation. As emphasized in the following 3 quotes, online social networks create amplification, an illusion of significance, and false credibility for these resources.

> *We have people from different backgrounds who will look up information in their home countries, which not necessarily could be medical information, but still use them in this country to request for a treatment or trying to sort of push for a particular way of management.* [HP9]

> *Well, it is because it has the echo chamber phenomena, which means that you only need 1 or 2 people in the community to relay a message across certain networks, like WhatsApp or whatever. And then suddenly, those messages will amplify themselves like an echo chamber and become so popular and track a lot of...and you see some people resending the messages so many times that the message become significantly more present in those networks from multiple sources, although they might have started with 1 or 2 people. But that then somehow infer [sic] or confirm legitimacy or authenticity on that message and becomes difficult to tackle. Now, because [the] social network has no limits, is like the physical encounter between 2 people, you need to see somebody and let's say gossip, or misinformation…you have somebody who will pass a message to you, and then you need to wait for another person who to meet to say that message again to you and a third person.* [HP5]

> *I think it comes from all sorts of sources. Yeah, so I don't think it's always just the internet…think it comes from TV. I think it comes from friends. It comes from family, it comes from generations, it comes from old books. I think it comes from other health care professionals. I think health care professionals within our own spheres are not updated…I think it's all over the place.* [HP12]

Online Social Media and Misinformation

> *The social platform, the groups, especially the groups. They are believing them, especially the groups. Some of them, they are dark web, I can see, dark web. And last is Google…because Google gives them the good information.* [HP11]

> *We have people who search the internet and come up with all sorts of information, which has not been substantiated in any way and come up requesting a treatment.* [HP2]

Status of Misinformation

Misinformation (Chronic or Acute?)

HPs look at misinformation as something more "chronic" rather than "acute." It reminds us that what makes misinformation during COVID-19 has less to do with how much misinformation is shared and more to do with existing inequalities, cultural exclusion, and lacking communication structures that make the consumers of misinformation more likely to believe and, crucially, act on misinformation. HPs deal with all kinds of health-related misinformation, and it appears to be related to many of the same factors. As stated, "Misinformation is always there, not only because of the pandemic" (HP10).

Unreliable Sources of Misinformation

Asked whether patients would have access to the right information, HPs provided varying responses regarding information resources. Although some health care professionals described the resources of information as being available, they also highlighted the lack of trust patients have in them. This contrasts with the high level of trust placed in social media. However, some HPs have frequently complained that the misinformation is based on outdated sources.

> *Some of them [is] outdated information, where people have been not aware of new developments on treatments. So that's…we can call it misinformation or lack of information. And then you have people who come up with some ideas about different kinds of treatments or lack of treatment.* [HP2]

It is important to mention that a few HPs addressed the difficulty of facilitating this reliable information.

> *I think it's very difficult to give people key health messages that are current, up to date, and researched and consistent.* [HP3]

> *Yes. So, when you don't have information, you tend to start to pick up sources which are not necessarily valid, not recommended by a health practitioner, or could be related to other parts of the world.* [HP9]

Epistemology of Patients

According to many HPs, patients blindly trust some of the sources of misinformation. This blind trust is based on 3 factors: authority of the sources, availability, and logical reasoning (plausibility).

Authority

Using the following example, HP2 explains that when the information is provided by a "gatekeeper" (a native speaker of the patient's own language), it is considered more authoritative than information provided by an "outsider" (a nonnative speaker).

> *So, the guy was translating my chat in Turkish because he knew English and he was very respectful and very appreciative. And he said, "My apologies for having to second, remention your message." And he knew that a lot of them knew English. And they could know what I'm saying. But they wanted to hear it in Turkish from somebody who's in Istanbul, while they are in England.* [HP2]

Availability

The availability of information (potentially in the absence of, or lack of exposure to, verified information) is 1 of the reasons why patients may choose to trust information to which they have easy access. There are many people who share information online (eg, from Google or social media), and the accessibility of this information is attractive, especially when the HP is difficult to reach.

> *I think they just follow whatever suggestion is given by the search engine, and later on, just after the consultation and coming across to the specialist, they start to be aware of the actual accredited body of the society of the disease that they are dealing with.* [HP3]

Logic

It is clear from the narrative that patients blindly trust the information when it occurs logically in the context of the community surrounding them and when it aligns with their experiences and perspectives. The patient, for example, anticipates that, since their friend had an illness, they will also have a similar experience.

> *She has explained to me that she has had these symptoms for a while, and I was like, "So what makes you think now this is the problem?" and she was like, "Well, my friend just recently got diagnosed and she has had this for a while as well," and she said she's got some of the symptoms her friend has, so it's not even like she completely has what her friend has…she's like, "Oh, yeah, it's just the painful periods and the heavy periods."* [HP7]

> *Again, same problem, irregular periods, but this time her friend had just been diagnosed…or they had found cysts on her ovaries, so she has been diagnosed with polycystic ovaries, and she was like, "Yeah, she's the same age as me. I have really painful periods as well and this must be it," even though this has been the situation, it's not like it has gotten worse over time, it's just…yeah, they will have a conversation with someone, and they are like, "Oh, this is what's happening, this makes sense," and it is kind of confirmed when they search Google, as well.* [HP2]

HP's Limited Time to Well Inform

> *You [HP] will take a lot of time, time they visit or make an appointment, meet their GP if they do not have this time they may become more inclined to rely on other sources of information.* [HP10]

Misinformation Impact on HPs' Practice

Educate the Patient

HPs adopted 2 approaches to dealing with misinformation. A majority of HPs agreed on the first point, which is to educate the patients. HPs use different strategies to educate patients, such as instructing, patronizing, presenting evidence, coaching, or researching the patient to determine the appropriate channel to use. Nevertheless, it has been observed that some doctors patronize their patients and treat them with an air of superiority. As part of a treatment path, other strategies may also be used, such as setting aside time to listen, talk, and communicate and building trust.

> *"You go online and look it up yourself," which obviously sometimes, because I wouldn't give them a specific website. I'd just say, "Look it up yourself for head injury advice."* [HP8]

> *Actually, just you explain that you understand the culture, but this is totally different of this, and this is classic mental illness. And we just do like psychoeducation about the signs and symptoms of mental illness and how we treat this. And we try to convince them and, like, reassure them about the medication. It's not causing addiction. It's not severely affecting the patient. Just, we are trying to treat them and make them feel better and make them more functioning.* [HP10]

> *If you are in a very busy inpatient ward, it's a very heavy workload, but still if you have 1 patient who is refusing treatment, it is worth to go and talk to him face to face. Sometimes, the patients refuse to do, like, a blood test, or sometimes, they are refusing to go to the general hospital if they are having high blood sugar with ketones and they are going to diabetic ketoacidosis, so if I talk to him, I will save his life, literally. So, it's worth doing it even if you have something else.* [HP10]

> *I would probably take more time with them, because I think it would take more…you maybe would have to spend more time. And sometimes, there would be resistance, I think, to maybe, you're telling them something that is the opposite to what they've been told, possibly by somebody that they would trust, or somebody...We do get doctors telling people the opposite of what we say sometimes. And then having to kind of change that advice is quite difficult. So, a doctor may say to somebody, "You must change your catheter leg back(?) every day or something, for example. And actually, you only change them once a week. And so, if a doctor has told somebody something, misinformed them, and then a nurse who is seen as a, you know, of a lesser.* [HP12]

> *But I think sometimes, certain treatment, from my experience from going on insulin, you can tell that people from, for example, [an] Asian background, they just have lots of concern about going on insulin, because they have heard so many stories, or they think that going on insulin doesn't mean they are really badly with diabetes, they don't want to accept that. So, this is 1 example, for example, changing the treatment from tablet to insulin, you just see lots of resistance from some patients. For example, like 1 patient told me, "Oh, my cousin lost her leg after she went on insulin." But actually, just when I explained, "No, it's not when she went on insulin. It's because the diabetes wasn't treated well."* [HP1]

Communicate With the Patient

HPs highlighted the importance of communication with the patient as an important pathway in their job practice to spot and witness misinformation, know its drivers and sources, and find pathways to confront it.

> *There's something called Sugar Buddies that if somebody was type 1 diabetes can be paired with somebody who had type 1 diabetes for 20 years, and really very well insulin management, so we pair them with somebody who we know that he's been really very well educated, managing. So, that has been helpful as well. And then the patient[s] like it, and they actually meet other patient[s] with the same condition.* [HP2]

For interviewees, communication has been addressed differently. For some, it is to talk and provide information; for others, it is to listen to the patient and discuss further health-related topics that have been spotted misinformed. As stated by HP3 and HP8:

> *It is how you communicate with people who are coming up with these kinds of challenges.* [HP3]

> *It is down to communication skills. Adapting maybe my tone, my voice, my speech, my physical gestures to a patient.* [HP8]

Interestingly, the words "listen" and "listening" to patients as part of the role of HPs were mentioned in the narrative 60 times across the 13 interviews. For the majority of interviewees, time is a common requirement to communicate with their patients.

Time Is a Requirement to Communicate With Patients

> *I would have to spend some more time building some trust and showing that you know what you talk about. Because I think sometimes, people, when they see that you know what you're talking about, they will see that actually after some time, then they will build some trust in you. But sometimes, yeah, so you probably would spend more time with that person. And feel you needed to build some trust and get them to, you know, invest in your point of thinking.* [HP10]

> *You will take a lot of time. And not only 1 session…also you should…this is really my practice is like that, it is not from the first take, you will never convince him from the first take, you should give him*

the basic information first, let him think about, then another talk with him at another, you know, in another situation, and in another meeting. They are partly convinced when you speak, you know, more frequently about it—in many sessions, not only 1 session. And don't give the final decision from 1 session or 2 sessions; you should give him more time to digest this information. [HP11]

Pathways to Combat Misinformation

HP-Patient Relationship

Misinformation is often influenced by the HP-patient relationship, whether it is short term or long term. For instance, if patients speak to a different health care provider each time they visit or make an appointment, if they do not have a family doctor, they may become more inclined to rely on other sources of information.

I think so, I think so, just because when you are in the [general practice], it's quite early on, and the GP [general practitioner] is kind of a filter as well, that's how you know who needs to be passed on to secondary care, and who has just got very severe health anxiety, and they need someone to sit and talk to them. You can even harshly sometimes just tell them, "You're okay, go away." Even with the GPs, sometimes, they have to do that, over the phone. So, they've had like diarrhea for a day, or they had diarrhea this morning, it's like, "Why are you calling, we all get diarrhea" kind of thing…do you know what I mean? They're like…I think they even quote things from Google, like, "This is malabsorption." The GP is, "Do you know what malabsorption is?" Do you know what I mean? So, I think GP, you get a lot of that stuff, but you kind of have to sieve through it, and sometimes you have to be like, "Stop, you're overreacting" kind of thing. [HP3]

Patients need to see a face; they need to see somebody on their territory or in areas like social networks they have. If they have a Teams or a Zoom or a Skype meeting or a function or whatever, you need to show some presence there so that you then...because as a clinician, we don't treat people virtually. We have to, at some point, meet them and give them the treatment. So, you need to be a physical presence in their life at some point. [HP2]

There is a long-term investment; [a] short-term one will be to make sure that your message is clear and short and concise. And the medium term will be to get access to networks which those communities are using through your liaisons…are your champions in that community who can then give you some access into them so that you can pass on the messages into those communities, either directly or through proxies, through them to counteract any kind of misinformation. [HP1]

It is worth to go and talk to him face to face. Sometimes the patients refuse to do, like, a blood test, or sometimes, they are refusing to go to the general hospital if they are having high blood sugar with ketones and they are going to diabetic ketoacidosis, so if I talk to him, I will save his life, literally. So, it's worth doing it even if you have something else. [HP10]

Building Trust

Trust was 1 of the strong themes in the study findings; it was cross-referenced across several interviews. It included different parties—the trust between the HP and the patient and the trust of the health system itself, including the existing information (eg, websites and leaflets).

HP-Patient Trust

If the patient can trust me, he would be more convinced when I talk to him about the misinformation that he had and about the right information if we can see(?) or the valid information about his disease and the treatment options. So, if I manage to get him trusting me, it would make a big step in our relationship. But the problem is a lot of our patients have no insight at that time. A lot of our patients are being paranoid. [HP10]

Trusting the Health System Information

For combating the misinformation, but increasing the credibility and the trust of your local health system here, more representation from those communities within the health system, within the hierarchy as well. So, it's not only a global level, even at managerial, and where they can see that our people are in high positions, and they are endorsing a message of health or health awareness or a treatment or a campaign on increasing information on this and that. So, this will then allow them to drop those misinformation. [HP2]

Discussion

Principal Findings

How Do Health Professionals Respond When Experiencing That a Patient Has Been Misinformed?

According to the study findings, during the conversation between HPs and their patients, HPs explore the sources of information. For them, identifying the origin of this information helps them identify the impact and consequences of the information and find pathways to confront the misinformation. For example, if the patient is a social media follower, that informs the HP to provide online accredited sources instead of the misinformative followed source [33]. Therefore, identifying sources and drivers of misinformation is considered a start to the pathway to confront misinformation [33].

The second manifested response to misinformation is how HPs see the misinformation: as chronic and as ongoing.

What Misinformation Has the Greatest Impact on Medical Practice?

For the second RQ in this study, which questions what HPs feel the biggest impact of misinformation is on health care practice.

The study does not specify certain misinformation that is considered to be leading to the biggest impact. HPs deal with all kinds of health-related misinformation, and it appears to be related to many of the same factors. For HPs, it all affects health and it is all big. This may be because according to the collected data, patients are misinformed because of the implications of different and diverse sources, such as online (eg, Google), family members, and offline (eg, social communities). It could also be that misinformation is an ongoing issue and is not linked with the pandemic solely. For HPs, all types of impacts of misinformation are major and need almost the same pathways to combat. For example, misinformation negatively impacts chronic patients (eg, those with bowel cancer). Misinformed patients resist following up on the chemotherapy protocol recommended by the HP.

What Aspects of Change and Intervention in Health Professionals' Practice Are in Response to Misinformation?

Educate the Patient

Lilley [33] confirmed that there is a need to educate patients to prevent errors and improve the quality of health care. This education influences patient behavior and produces the changes in knowledge, attitudes, and skills necessary to maintain and improve health. However, educating patients is not an easy and direct job. According to Ward et al [34], it is a myriad of interventions to support patients' education and adherence to doctor or HP recommendations for diet, exercise, medications, and advice. This conclusion is well aligned with what HPs narrated in this study, as they use different methods to educate their learners, including being advisors, providing resources, and acting as tutors and counsellors.

Communicate With the Patient

According to Palmieri and Stern [35], effective bidirectional communications (between the patient and the HP) are pathways to making an accurate diagnosis. The impact of communicating with the patient is not limited to the issue of

misinformation, but it would lead to better health service provision to the patient. According to Davis [36], a patient who listens and has ongoing communication with the HP (eg, doctor) first and then decides has done a better job of deliberating than a patient who first consults web pages or friends and acquaintances and makes a decision before talking to the doctor. Such patients are simply not as good at deliberating about medical matters as patients who do not engage in premature consent. Knowing which sources to heed when making medical decisions is part of being competent at making such decisions. It is part of the skill of deliberation. The challenge to address in relation to encouraging communication is the limited time that HPs can allocate to meet, talk, and listen to their patients. This is confirmed by Palmieri and Stern [35] as managing care and time constraints, adding further pressure on HPs that need to be facilitated. Communication is a 2-way effort to maintain, and there is a role for each party [37]. HPs are encouraged to rehearse different communication strategies and to seek supervision and consultation around matters that are challenging. Patients have a role in fostering honest communication with their providers, while physicians can best promote such interactions by being thoughtful, deliberate, and self-aware.

What Is the Intervention in Health Professionals' Practice in Response to Misinformation?

For this third RQ, according to the study findings, trust is the dimension that needs to be tackled to combat misinformation. In the issue of trust, the study encompasses 2 aspects: patient-HP trust and patient-source of information trust that is issued by a health body, such as the World Health Organization (WHO).

Building Trust

Trust involves relationships and not just facts. Trust is most likely in situations in which people directly encounter a health care professional in person (at least virtually) rather than in situations in which people are presented with information in other ways [38]. HPs who spoke about gatekeepers from the different communities they serve, or building long-term relationships with patients, understand that many patients will trust people they know—people who share the same language or cultural experiences. Making inroads with different communities and representing them in health care are strategies that can mitigate the impacts of misinformation. This is a time-consuming activity, but it has a long view, and this is important. If HPs view misinformation as chronic, then they understand that the treatment pathway must be more in depth and contextualized.

Patient-HP Trust

Regarding the personal characteristics or demographic features that may play a role, HPs named education, culture, and political affiliation as playing a role. This is consistent with the existing literature. However, similar to Harper and Baguley [20], the correlations are not always in 1 specific direction (more or less related to misinformation-sharing behavior). What appears to be important is how they impact trust. If patients have a trust relationship with their health care provider, and trust in their relevant authorities to provide accurate and timely information, it may not matter how much misinformation a person sees. They will be able to make good decisions for their health and take the necessary actions to protect themselves and their families.

Patient-Source of Information Trust

HPs in the study findings linked misinformation to trust, both between patients and their health care providers as well as between patients and the information provided by experts (eg, health organizations and accredited websites). Some of these official resources do not provide up-to-date information. Although this issue does not directly relate to HPs, it affects their practice in terms of patient-information trust. This mistrust is confirmed by Davis [36], as the rise of premature consent cases is that trust in the health care system has been undermined by instantaneously disseminated information about medically related errors without any details. Making an accurate diagnosis relies on the provision of reliable information [35]. Nonetheless,

this information leads the tech-savvy patient to be skeptical about the physician-employee's management. Therefore, evaluating sources of medical information and advice are pathways to support patients to decide whether to believe and trust the provider.

Limitations

The study had 2 main limitations: difficulty in recruiting participants and the lack of patients' voices.

Recruitment of Participants

This study was conducted during the pandemic when the majority of HPs were extremely busy and overloaded with work duties and pressure. Consequently, recruitment of a satisfactory number of participants to take part was 1 of the main limitations of this study. The further difficulty, after finding HPs willing to be interviewed, was to find a time slot in their agenda to schedule the interview in the middle of a global pandemic.

Patient's Voice

Although HPs in this study could give an in-depth view about their patients' misinformation, this is considered to be only 1 side of the coin. The other important aspect that needs to be explored in depth are the patients' views about misinformation, including its drivers, perception, and how to confront it from their angle. With the limited time and funding for this pilot, we thought it would be more effective to speak to HPs about many different diverse experiences with patients, rather than collecting a small sample of patient views, at this time.

Future Work

Many HPs recommended that this study be complemented with another empirical study that incorporates patients' voices and explores their views. According to several interviewees, patients' views, understandings, beliefs, and attitudes can shed light on different angles from those narrated by HPs. Including those views will aid in understanding misinformation more thoroughly and in depth. However, recruitment of patients who have experienced misinformation might be a challenge in this study, because it may be difficult for some individuals to admit that they have been misinformed.

Conclusion

Misinformation affects patients' decisions to follow a treatment or guidance prescribed by their HPs. According to the study findings, patients follow misinformation resources for 3 reasons: (1) available resources (eg, Google); (2) meaningful resources, as they reflect their personal or cultural beliefs; and (3) authorized resources, as they have been disseminated by a source of power for the patient (eg, a political party). The qualitative research presented in this paper revealed that patients do not always trust their HPs or the authorities about health-related information. As a result, they may choose not to follow HP advice on matters that impact their health, including COVID-19. The lack of trust in HPs was identified as a prominent theme in this study, and it was attributed to several factors, including trusting other sources of information (eg, social media), patient's doubts about HP experience (eg, a junior doctor with only a few years' experience), and patients' doubts about the available sources of information that are provided by the HP (eg, out-of-date resources). There are 2 dimensions of trust: patient-HP trust and patient-information trust. There are 2 necessary actions to address the issue of lack of trust in these dimensions: (1) building trust and (2) maintaining trust. The main recommendations of the HPs are to listen to the patients, to give them more time, and to seek evidence-based resources. Finally, misinformation is an ongoing phenomenon; it is not solely manifested during the pandemic and the spread of fake news where some patients resisted COVID-19 vaccination. Misinformation has been shown for patients with other chronic diseases (eg, bowel cancer). These patients, because of misinformation, resisted following up the chemotherapy protocol recommended by the HPs. Consequently, for HPs, finding out the sources and drivers of misinformation is a pathway to identify, track, and confront misinformation.

Acknowledgments

This research was funded by the School of Computing and Communications at the Open University and the Knowledge Media Institute (KMi), more specifically the HERoS project (European Union's Horizon 2020 research and innovation program under grant no. 101003606). It allowed researchers across different faculties to collaborate and build a research team that focused on the experience of health practitioners with misinformation and its impact on their job practice.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Structure of combined deductive and inductive codes.
[DOCX File , 15 KB - formative_v6i11e38794_app1.docx]

References

1. Broniatowski DA, Jamison AM, Qi S, AlKulaib L, Chen T, Benton A, et al. Weaponized health communication: Twitter bots and Russian trolls amplify the vaccine debate. Am J Public Health 2018 Oct;108(10):1378-1384. [doi: 10.2105/AJPH.2018.304567] [Medline: 30138075]

2. Mokhtari H, Mirzaei A. The tsunami of misinformation on COVID-19 challenged the health information literacy of the general public and the readability of educational material: a commentary. Public Health 2020 Oct;187:109-110 [FREE Full text] [doi: 10.1016/j.puhe.2020.08.011] [Medline: 32942170]
3. Wang Y, McKee M, Torbica A, Stuckler D. Systematic literature review on the spread of health-related misinformation on social media. Soc Sci Med 2019 Nov;240:112552 [FREE Full text] [doi: 10.1016/j.socscimed.2019.112552] [Medline: 31561111]
4. Tully M, Bode L, Vraga EK. Mobilizing users: does exposure to misinformation and its correction affect users' responses to a health misinformation post? Soc Media Soc 2020 Dec 10;6(4):205630512097837 [FREE Full text] [doi: 10.1177/2056305120978377]
5. Zhao Y, Da J, Yan J. Detecting health misinformation in online health communities: incorporating behavioral features into machine learning based approaches. Inf Process Manag 2021 Jan;58(1):102390 [FREE Full text] [doi: 10.1016/j.ipm.2020.102390]
6. Wardle C, Derakhshan H. Information Disorder: Toward an Interdisciplinary Framework for Research and Policy Making. 2017. URL: https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-researc/168076277c [accessed 2022-10-19]
7. McKernon E. Fake News and the Public. 1925. URL: https://harpers.org/archive/1925/10/fake-news-and-the-public/ [accessed 2022-10-15]
8. Chou WS, Oh A, Klein WMP. Addressing health-related misinformation on social media. JAMA 2018 Dec 18;320(23):2417. [doi: 10.1001/jama.2018.16865]
9. Karlova NA, Fisher KE. A Social Diffusion Model of Misinformation and Disinformation for Understanding Human Information Behaviour. 2013. URL: http://informationr.net/ir/18-1/paper573.html#.Y0-PCnZBxD8 [accessed 2022-10-19]
10. Chen X. Remedial Coursetaking at US Public 2-and 4-Year Institutions: Scope, Experiences, and Outcomes; Statistical Analysis Report. 2016. URL: https://nces.ed.gov/pubs2016/2016405.pdf [accessed 2022-10-19]
11. Chen X, Sin SJ. Misinformation? What of it? Motivations and individual differences in misinformation sharing on social media. Proc Am Soc Inf Sci Technol 2014 May 08;50(1):1-4. [doi: 10.1002/meet.14505001102]
12. Bronstein MV, Pennycook G, Bear A, Rand DG, Cannon TD. Belief in fake news is associated with delusionality, dogmatism, religious fundamentalism, and reduced analytic thinking. J Appl Res Memory Cogn 2019 Mar;8(1):108-117. [doi: 10.1016/j.jarmac.2018.09.005]
13. Pennycook G, Cannon TD, Rand DG. Prior exposure increases perceived accuracy of fake news. J Exp Psychol Gen 2018 Dec;147(12):1865-1880 [FREE Full text] [doi: 10.1037/xge0000465] [Medline: 30247057]
14. Guess A, Nyhan B, Reifler J. Selective exposure to misinformation: evidence from the consumption of fake news during the 2016 US presidential campaign. Eur Res Council 2018;9(3):4.
15. Cinelli M, De Francisci Morales G, Galeazzi A, Quattrociocchi W, Starnini M. The echo chamber effect on social media. Proc Natl Acad Sci U S A 2021 Mar 02;118(9):e2023301118 [FREE Full text] [doi: 10.1073/pnas.2023301118] [Medline: 33622786]
16. Guess A, Nagler J, Tucker J. Less than you think: prevalence and predictors of fake news dissemination on Facebook. Sci Adv 2019 Jan 04;5(1):eaau4586 [FREE Full text] [doi: 10.1126/sciadv.aau4586] [Medline: 30662946]
17. Ecker UKH, Ang LC. Political attitudes and the processing of misinformation corrections. Political Psychol 2018 Oct;40(2):241-260. [doi: 10.1111/pops.12494]
18. Grinberg N, Joseph K, Friedland L, Swire-Thompson B, Lazer D. Fake news on Twitter during the 2016 U.S. presidential election. Science 2019 Jan 25;363(6425):374-378. [doi: 10.1126/science.aau2706] [Medline: 30679368]
19. Allcott H, Gentzkow M. Social media and fake news in the 2016 election. J Econ Perspect 2017 May 01;31(2):211-236. [doi: 10.1257/jep.31.2.211]
20. Harper CA, Baguley T. "You are fake news": ideological (a)symmetries in perceptions of media legitimacy. PsyArXiv Preprint posted online January 23, 2019. [doi: 10.31234/osf.io/ym6t5]
21. Elsner B, Narciso G, Thijssen J. Migrant networks and the spread of information. Oxf Bull Econ Stat 2017 Oct 26;80(3):659-688. [doi: 10.1111/obes.12216]
22. Cadet P, Feldman MD. Pretense of a paradox: factitious intersex conditions on the internet. Int J Sex Health 2012 Apr;24(2):91-96. [doi: 10.1080/19317611.2011.629287]
23. Medscape. HCWs with Long COVID Report Doubt, Disbelief from Colleagues. 2021. URL: https://www.medscape.com/viewarticle/963766 [accessed 2022-10-19]
24. Moriarty R, Rastrick S. Allied Health Professionals Personal Health Budgets. 2022. URL: https://www.england.nhs.uk/ahp/ [accessed 2022-10-19]
25. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
26. Everitt B, Skrondal A. Snowball sampling. In: The Cambridge Dictionary of Statistics. Cambridge, UK: Cambridge University Press; 2002:2.
27. Miles MB, Huberman AM. Qualitative Data Analysis: An Expanded Sourcebook. New York, NY: Sage; Dec 1994.

28. Barbour RS. Checklists for improving rigour in qualitative research: a case of the tail wagging the dog? BMJ 2001 May 05;322(7294):1115-1117 [FREE Full text] [doi: 10.1136/bmj.322.7294.1115] [Medline: 11337448]
29. Lincoln YS, Guba EG. Naturalistic Inquiry. Newbury Park, CA: Sage; 1985.
30. McAlister AM, Lee DM, Ehlert KM, Kajfez RL, Faber CJ, Kennedy MS. Qualitative coding: an approach to assess inter-rater reliability. 2017 Presented at: 2017 ASEE Annual Conference & Exposition; June 24-28, 2017; Columbus, OH. [doi: 10.18260/1-2--28777]
31. NVIvo. Word Frequency Query. URL: https://help-nv.qsrinternational.com/12/win/v12.1.112-d3ea61/Content/queries/word-frequency-query.htm [accessed 2022-10-19]
32. O'Connor C, Joffe H. Intercoder reliability in qualitative research: debates and practical guidelines. Int J Qual Methods 2020 Jan 22;19:160940691989922. [doi: 10.1177/1609406919899220]
33. Lilley LL, Guanci R. Educate the patient. Am J Nurs 1996;96(4):14. [doi: 10.1097/00000446-199604000-00009]
34. Ward M, Ní Shé É, De Brún A, Korpos C, Hamza M, Burke E, et al. The co-design, implementation and evaluation of a serious board game 'PlayDecide patient safety' to educate junior doctors about patient safety and the importance of reporting safety concerns. BMC Med Educ 2019 Jun 25;19(1):232 [FREE Full text] [doi: 10.1186/s12909-019-1655-2] [Medline: 31238936]
35. Palmieri JJ, Stern TA. Lies in the doctor-patient relationship. Prim Care Companion J Clin Psychiatry 2009 Aug 15;11(4):163-168. [doi: 10.4088/pcc.09r00780]
36. Davis J. Dr. Google and premature consent: patients who trust the internet more than they trust their provider. HEC Forum 2018 Sep;30(3):253-265 [FREE Full text] [doi: 10.1007/s10730-017-9338-z] [Medline: 29063996]
37. Barclay JS, Blackhall LJ, Tulsky JA. Communication strategies and cultural issues in the delivery of bad news. J Palliat Med 2007 Aug;10(4):958-977. [doi: 10.1089/jpm.2007.9929] [Medline: 17803420]
38. Southwell BG, Wood JL, Navar AM. Roles for health care professionals in addressing patient-held misinformation beyond fact correction. Am J Public Health 2020 Oct;110(S3):S288-S289. [doi: 10.2105/ajph.2020.305729]

Abbreviations

HP: health professional
IRR: interrater reliability
NHS: National Health Service
RQ: research question
TA: thematic analysis

Edited by A Mavragani; submitted 15.04.22; peer-reviewed by I Wilson, W Evans; comments to author 04.07.22; revised version received 14.10.22; accepted 17.10.22; published 02.11.22.

Please cite as:
Ismail N, Kbaier D, Farrell T, Kane A
The Experience of Health Professionals With Misinformation and Its Impact on Their Job Practice: Qualitative Interview Study
JMIR Form Res 2022;6(11):e38794
URL: https://formative.jmir.org/2022/11/e38794
doi:10.2196/38794
PMID:36252133

Original Paper

The Dutch COVID-19 Notification App: Lessons Learned From a Mixed Methods Evaluation Among End Users and Contact-Tracing Employees

Joris Elmar van Gend[1], MSc; Jan Willem Jaap Roderick van 't Klooster[1], IR, PhD; Catherine Adriana Wilhelmina Bolman[2], PhD; Julia Elisabeth Wilhelmina Cornelia van Gemert-Pijnen[3], PhD

[1]The BMS Lab, Faculty of Behavioural, Management and Social Sciences, University of Twente, Enschede, Netherlands

[2]Department of Psychology, Open University of the Netherlands, Heerlen, Netherlands

[3]Section of Psychology Health & Technology, Department of Technology, Human and Institutional Behavior, Faculty of Behavioural, Management and Social Sciences, University of Twente, Enschede, Netherlands

Corresponding Author:
Joris Elmar van Gend, MSc
The BMS Lab
Faculty of Behavioural, Management and Social Sciences
University of Twente
De Zul 10
Enschede, 7522 NB
Netherlands
Phone: 31 0534892525
Email: bmslab@utwente.nl

Related Article:

This is a corrected version. See correction statement: https://formative.jmir.org/2022/12/e44497

Abstract

Background: The Dutch CoronaMelder (CM) app is the official Dutch contact-tracing app (CTA). It has been used to contain the spread of the SARS-CoV-2 in the Netherlands. It allows its users and those of connected apps to anonymously exchange warnings about potentially high-risk contacts with individuals infected with the SARS-CoV-2.

Objective: The goal of this mixed methods study is to understand the use of CTA in the pandemic and its integration into the Municipal Health Services (MHS) efforts of containment through contact tracing. Moreover, the study aims to investigate both the motivations and user experience–related factors concerning adherence to quarantine and isolation measures.

Methods: A topic analysis of 56 emails and a web-based survey of 1937 adults from the Netherlands, combined with a series of 48 in-depth interviews with end users of the app and 14 employees of the Dutch MHS involved in contact tracing, were conducted. Mirroring sessions were held (n=2) with representatives from the development (n=2) and communication teams (n=2) responsible for the creation and implementation of the CM app.

Results: Topic analysis and interviews identified procedural and technical issues in the use of the CTA. Procedural issues included the lack of training of MHS employees in the use of CTAs. Technical issues identified for the end users included the inability to send notifications without phone contact with the MHS, unwarranted notifications, and nightly notifications. Together, these issues undermined confidence in and satisfaction with the app's use. The interviews offered a deeper understanding of the various factors at play and their effects on users; for example, the mixed experiences of the app's users, the end user's own fears, and uncertainties concerning the SARS-CoV-2; problematic infrastructure at the time of the app's implementation on the side of the health services; the effects of the society-wide efforts in containment of the SARS-CoV-2 on the CM app's perception, resulting in further doubts concerning the app's effectiveness among MHS workers and citizens; and problems with adherence to behavioral measures propagated by the app because of the lack of confidence in the app and uncertainty concerning the execution of the behavioral measures. All findings were evaluated with the app's creators and have since contributed to improvements.

Conclusions: Although most participants perceived the app positively, procedural and technical issues identified in this study limited satisfaction and confidence in the CM app and affected its adoption and long-term use. Moreover, these same issues

negatively affected the CM app's effectiveness in improving compliance with behavioral measures aimed at reducing the spread of the SARS-CoV-2. This study offers lessons learned for future eHealth interventions in pandemics. Lessons that can aid in more effective design, implementation, and communication for more effective and readily adoptable eHealth applications.

(JMIR Form Res 2022;6(11):e38904) doi:10.2196/38904

KEYWORDS

eHealth; contact tracing; digital contact tracing; contact-tracing apps; COVID-19; adherence; public health; mobile health; topic analysis; health service; user experience; eHealth intervention; mobile phone

Introduction

Background

One of the measures that numerous countries have implemented to mitigate the spread of the SARS-CoV-2 is the use of contact-tracing apps (CTAs) [1], defined as "software that can be installed on a user's personal device, such as a smartphone to notify the user when they come into contact with a person or persons infected with SARS-CoV-2" [2]. Studies have revealed epidemiological impact; large numbers of cases of COVID-19 infection would be averted by this digital contact tracing [3] mainly through the potential increase in speed of the contact-tracing process. However, estimates show that more than half of the population must simultaneously use the app for it to be effective [4].

The Dutch CoronaMelder App

The Dutch government has developed a CTA called CoronaMelder (CM) based on the Google Apple Exposure Notification system jointly developed by Apple and Google [5]. The app was implemented and made available to the public in October 2020. The Dutch CM app aimed to (1) relieve the Municipal Health Service (MHS) in contact tracing, (2) identify and trace high-risk contacts earlier and faster than manual contact tracing of the MHS, and (3) inform the users who have been in the proximity of an infected app user about the measures that were strongly advised to be taken (eg, undergo a polymerase chain reaction test, quarantine, not to have visitors at home, and social distancing) [6] to avoid (further) potential spread of the virus. The basic mechanism of the app is to classify the prolonged (for at least 15 minutes) proximity of one app user to a second app user who has tested positive for COVID-19 infection as a high-risk contact. The app will then anonymously send the first user a notification, alerting them about their high-risk contact in the app. This notification recommends that the user undergo a test at the MHS and quarantine until the test result is received. If the user receives a positive test result, the app offers the user the ability to notify other app users with whom they in turn have been in contact. In the Dutch situation, the MHS is a linking pin in most cases, that is, in the case of a positive test, the user is flagged as infected in the MHS' systems and contacted to start the (manual) contact-tracing process. The MHS official (health official) subsequently asks the end user to read a code generated by the app, which the health official then adds to their system to be validated as a code for an infected individual. Next, the end user can choose to send a notification to warn others through the app [7]. Alternatively, since October 2021, the user can validate the code themselves through a 2-step verification–protected MHS website that provides them with positive test results [8].

State of CTAs and Research

This study is the second qualitative study on the Dutch CM app and part of a larger effort to evaluate the effectiveness of the app as a tool during the COVID-19 pandemic. Earlier research [9] focused on the usability and user-centeredness of the app during its initial release. Changes were made to the app because of this first study. CTAs have also been introduced in other countries. Although algorithmic improvements were made to increase the accuracy of CTAs in some countries (including the Netherlands), most countries relied on the Google Apple Exposure Notification reference framework. The implemented possibilities to contact the authorities after a positive test differ by country, and adoption rates vary by country as well. Before the spread of the omicron variant of the SARS-CoV-2, epidemiological studies in the United Kingdom have found that CTAs are effective when adopted by the public [3]. A German study [10] found that a higher level of education led to improved adoption of pandemic apps. In Belgium, a survey found a large majority of nonusers This refusal of use was largely due to privacy concerns and ambivalence about the app's utility [11]. In Greece, an in-depth additional contact-tracing training study found that contact-tracing efforts work best in a systematic and coordinated manner, and systematic and organized training of contact-tracing workers can greatly increase its effectiveness [12]. Thus, it is important to consider not only the app but also the privacy concerns and embedding in the public health infrastructure.

The Dutch CM App and Contact-Tracing Process Over Time

The Dutch app is continuously being monitored and evaluated by the Dutch Ministry of Health, Welfare and Sport for its adoption and effectiveness. Studies have revealed that approximately one-third of the Dutch population downloaded the app approximately 1 year after its launch [13]. The lowest adoption has been found among people with lower education levels, lower monthly incomes, and those aged >80 years [6]. The CM app has proven to trace and alert a substantially higher number of contacts than MHSs; 77% of those who were notified by the app had not been contacted by the MHS [14] in their manual contact-tracing process. The results are less positive regarding behavioral outcomes, such as adherence to measures designed to prevent the spread of the virus that are provided in the app. Only 45% of the users stayed at home after notification, and only 41% of the users applied for a polymerase chain reaction test. Little is known about the causes of behavioral

outcomes. For example, no insights have been gained at that point into the user friendliness of the app and its comprehensibility [9]. This changed after the initial usability study of the CM app by Bente et al [8]. The same goes for the extent to which the behavioral measures (presented as recommendations to users, such as self-isolation) the app provides are perceived as both clear and unambiguous and to what degree the app's notifications might trigger potentially unintended and undesired effects (eg, anger or panic). Finally, it was not known why users might have chosen not to follow these measures and how the app's integration with the work processes of the MHS might have affected the app's effectiveness. However, these insights are essential to determine whether and why the app is used correctly and does or does not enforce the proposed behavioral measures. Eventually, such information is vital for improving the CTA and its integration into the contact-tracing process of the MHS and ultimately, to support its sustainability and scalability. The latter will remain of paramount importance as the COVID-19 pandemic has been shown to be persistent with periodically fluctuating levels of infections and social control measures. Digital contact tracing remains an important part of the strategy to suppress the spread of the SARS-CoV-2.

Challenges From Privacy by Design

Because the app was designed in accordance with the *privacy by design* principle, it is complicated to gather the aforementioned information quantitatively via the app. Moreover, options for data collection that do exist are limited to numbers, such as the number of notifications sent in a period, which do not shed light on underlying motives or causes for behavior, such as adherence to behavioral measures or a lack thereof. This study builds upon and follows up on the results from the evaluation of the CM app and is part of the continuous evaluation of the CM CTA [8,13]. The integration of the CM app in contact tracing of the MHS will also be explored. As such, this study takes 2 viewpoints into account: those of end users and employees of the Dutch MHS involved in the contact-tracing process. Moreover, the study uses a mirroring approach on the development and communication teams involved in the creation and implementation of the app to gain insights into the context of the findings and the feasibility of implementing improvements that are advised. The research questions are as follows: (1) What hinders, deters, or motivates end users (citizens) in adopting the app and their adherence to the app's instructions and advice (behavioral measures)? (2) How is the app implemented in the work processes of the Dutch MHS' contact-tracing teams? (3) How does CTA use affect adherence to isolation and quarantine measures?

Methods

Overview

To gain the necessary insights into the first and third research questions (use of the app, adherence of CM app users, and user friendliness), 3 methods were applied. First, a topic analysis (n=56) to get a first understanding of the pain points encountered when using the app was conducted using emails sent by users of the app. Second, a short web-based survey (n=1802) was conducted to gain insights into the adoption of the app and the degree to which app notifications were received and sent and to gather the contact information of positively tested individuals who used the app and who would be invited to participate in in-depth interviews. Third, semistructured interviews (n=48) were conducted to gain deep insight into adherence to behavioral measures and perceptions of the app. The second research question was studied through semistructured interviews with contact-tracing employees of the MHS. The findings were then discussed with the teams of the Dutch Ministry of Health, Welfare and Sport and the MHS responsible for building and maintaining the app and communicating about the app. The design and procedures of this study were evaluated by the responsible ethics committee (Multimedia Appendix 1) of the University of Twente's BMS faculty.

Topic Analysis on Emails Sent by Users of the App

The topic analysis was based on 56 participants who reacted by email on a call in a regional newspaper (November 11, 2020, Twentsche Courant Tubantia) [14], which requested that they share their experiences with the CM app as part of scientific research. Out of 63 mails, 56 (89%) met the inclusion criteria. Participants were included if they had installed and activated the CM app, received a notification, and possibly shared their app's (MHS) key with MHS employees to warn other CM app users. The 56 participants selected formed a varied sample from across the Netherlands. The sentiment in their emails was analyzed using manual coding. This means that 2 coders would independently read each email and determine whether the sentiment of the email was generally positive, negative, or neutral. Similarly, a list of possible topics was identified in the first round of analysis and then coded by 2 independent coders in the final analysis. Conflicting assessments of coders were resolved through a third reviewer's evaluation or the email in question.

Web-Based Survey

A web-based survey was sent between December 1, 2020, and December 21, 2020, via a Dutch panel service called "Panelclix." The main goal of this survey was to recruit potential participants for the interviews. The selection criteria were age (>18 years) and education; the aim was to have a sample group in which 70% of the participants had a lower and middle level of education. Participants were asked 9 questions concerning the use of the CM app, whether they had received a notification or, in case of a corona infection, whether they had shared their CM app MHS authentication key with the MHS. Panel members who received notification were asked to participate in the interviews (on the web or by phone). The full set of questions from the survey is included in Multimedia Appendix 2 and is in the Dutch language.

Interviews

Interviews With CM App Users

A total of 48 web-based interviews were conducted between December 1, 2020, and January 21, 2021, with participants who used the CM app, received a notification (38/48, 79%) or had tested positive for COVID-19 infection, after which they had or had not shared the MHS key with the MHS (8/48, 16%).

Some participants (3/48, 6%) fulfilled both the criteria. The targeted sample consisted of individuals from the general population, with a focus on hard-to-reach groups, such as individuals with low education levels, limited reading or digital skills, a migration background, and adults over the age of 65 years. Multiple recruitment channels were used, such as the media (outreach through the newspapers), the CM app helpdesk, schools, companies, a library, a football club, a health care organization, and organizations that support individuals with a low level of literacy such as "Pharos" and "Stichting ABC." Participants were asked questions about factors that influenced their adherence behavior regarding COVID-19–related behavioral measures, their underlying motivations for their adherence (or lack thereof), and the limiting or facilitating factors they experienced within the CM app during their use. Moreover, questions were asked to explore their experience with the work processes of the MHS if the interviewed individual was contacted by an MHS contact-tracing employee. Important factors (eg, technical problems) identified during the topic analysis and the survey were used in the design of the questions for the interview with contact-tracing employees. The full list of interview questions and subjects can be found in Multimedia Appendices 3 and 4 and are in the Dutch language.

Interviews With Contact-Tracing Employees

To explore the embedment of the CM app in the contact-tracing process at the MHS, 11 semistructured web-based interviews were conducted with contact-tracing employees of the Dutch MHS. The interviews were conducted between January 28, 2021, and February 10, 2021. The participants were recruited through a network of researchers, and an open invitation was posted at 2 MHS locations. Participants were asked about the process of contact tracing, the manner in which they communicated with the index (person tested positive for COVID-19), the extent to which the CM app was embedded in this process, the limitations and difficulties they experienced in their work, and how these could be reduced. Moreover, the results of both the topic analysis and interviews with CM app users served as a basis for the questions in the interviews with MHS employees. Issues identified by users in the MHS processes and uncertainties about the MHS practices were integrated into the interview questions.

All interview participants, both app users and MHS employees, received a gift card after the interview in exchange for their time investment. The overall results were communicated and mirrored by the designers and creators of the CM app and the communication team of the Dutch Ministry of Health, Welfare and Sport responsible for the campaigns concerning the CM app to gain additional insights and aid in the app's further development and implementation campaign. The app and the workflow and training of the MHS employees were adapted because of this mirroring (see Discussion).

Analysis of Both Interview Series

Semistructured interviews were conducted in accordance with the list of subjects and questions. The interviews were recorded, transcribed verbatim, and coded by 2 independent coders. The subjects mentioned and statements made by the participant related to the questions, and subjects on the list were coded and then added to the count of the appropriate code on the coding scheme or added as a new code. Because of this method, the sum of the frequencies of individual codes under a code group (subject) does not always correspond to the total number of participants. There are 2 reasons for this finding. First, frequencies represent the counts of codes in statements made by the participant. Not all participants made statements that qualified for a code for each code group, resulting in a lower sum of frequencies than the total number of participants. Second, some participants may have made statements to which multiple codes of the same code group applied, resulting in a higher sum of frequencies. The coding scheme consisted of a limited set of codes created beforehand based on the characteristics of the app (thematic coding), procedures from the contact-tracing teams, and information and adherence recommendations that were provided in the app and input from both topic analyses, which were supplemented with codes identified while analyzing the interviews. Coders returned to previously analyzed transcripts to verify that statements to which newly created codes would apply were not missed. Conflicts between coders were resolved through discussion and a third coder that could weigh in.

Feedback by Stakeholders

The findings of the 3 substudies were presented to and discussed with the teams responsible for the realization and implementation of the app (MHS, design team, and communication team). In total, 2 sessions were held for approximately an hour with 2 representatives from the development team and 2 from the communication team of the Dutch Ministry of Health, Welfare and Sport. This allowed them to apply the findings to the redesign of the CM and implementation of the CM in the work processes of the MHS and to tailor the communication campaign. Moreover, it provided this study with the opportunity to incorporate the challenges faced from a policy, technical and communication perspective in the realization, implementation, and use of the CM app and its processes. Thus, this study incorporated insights from various stakeholders.

Ethical Considerations

Participants of all 3 substudies were informed about their participation beforehand and could withdraw their participation at any point of time. In the case of topic analysis, this was done by explicitly mentioning that the emails sent to researchers would be analyzed and experiences distilled and anonymously used for scientific research. In the case of both the survey and interviews, an informed consent procedure was followed, as outlined in the application to the University of Twente's Ethical Review Board (approval number 201323) for the Behavioral Sciences (Multimedia Appendix 1). The purpose of this evaluation was to independently assess and address (potential) ethical concerns regarding the study or its compliance with applicable legislation.

Results

The results of the topic analysis are presented first, followed by the interviews with the CM app users and the results of the interviews with the contact-tracing employees. Selected quotes from participants were translated literally and used to illustrate

the results. The original untranslated quotes are provided in Multimedia Appendix 5.

Topic Analysis

Overview

In total, 56 people sent usable emails after the publication of a call (outreach) to share their experiences with the CM app. Of the 56 participants, 22 (39%) were male, 30 (54%) were female, and of the remaining 4 (7%) participants, sex could not be determined with certainty from the contents of their email. Participants were aged between 20 and 80 (mean 55, SD 15.0388) years. A relatively large group (10/56, 18%) was ≥65 years. The contents of the emails were unprompted and thus reflected the experiences that were most focal to the participants.

All but one of the participants (55/56, 98%) indicated that they had received a notification from the CM app warning them of high-risk contact with an individual who tested positive for COVID-19 infection. Initial responses to the notification were mostly emotional and strongly negative and manifested in the form of anger, fear, and disbelief. A common theme in the emails was the uncertainty participants felt regarding the contact that triggered the notification and the meaning and implications of the notification (eg, How high is the risk? When was the contact exactly? What to do now?). Moreover, during the early phases of the app's use, the phone's operating system would periodically send notifications that confused the participants. The notifications would inform the user that they had spent a week without having any high-risk contact. These notifications were often mistaken for notifications intended to warn about past high-risk contacts and caused dissatisfaction and anger:

> *On Monday 31 Augustus, I read on my CoronaMelder that I had been close for more than 15 minutes to someone who reportedly was infected. It was a big shock! That, according to me, wasn't possible! But, anyway, I called the phone number that was provided. I got the advice to go in quarantine and to call my general practitioner or the MHS if symptoms developed. I called my friends and family! The Monday afterwards I received another notification from the CM app. I had not been in contact with anyone infected according to the app. I didn't understand this at all!* [MAIL0116]

Notifications regarding high-risk contacts were often perceived as arriving too late. Most commonly (17/56, 30%), participants indicated that they had received a notification within 5 days after their high-risk contact. However, a significant group was indicated to have received the notification between 5 and 10 days (11/56, 20%) or more than 10 days after the high-risk contact:

> *After reading your article in the newspaper, I would like you to know that I have removed the app [CM] from my phone and this is why: On 2 November I received a message that I had been in contact for more than 15 minutes with someone that has corona and that I had to stay inside until 30 October.* [MAIL0072]

Response to Notifications

Participants (7/56, 13%%) expressed severe doubts about the validity of the notification they received. For example, participants knew that they had not left the house on the day of the supposedly high-risk contact. Others could identify the source of the notification as their neighbor they had not seen in person. They then deduced that the Bluetooth signal must have traveled through the walls of their homes. A different subsection consisting of 7 (13%) participants indicated that their first course of action after receiving a notification was to verify its validity because of doubts. Participants indicated that these occurrences undermined their confidence in the validity of the notifications and the CM app, which resulted in a more complicated decision-making process regarding whether they should go into quarantine or isolation:

> *My experience is that the phones of the neighbours and myself had been in contact. ... It can happen that people get a notification [from CM] that is incorrect, but better this than the other way around.* [MAIL0074]

To relieve anxiety and doubts created by the receipt of notifications, participants, for example, chose to contact the MHS or their general practitioner. Their aim was to remove some of the uncertainty and gain insights into the meaning and implications of the notification and actions to be taken:

> *After yet another sleepless night I concluded that I still did not experience any symptoms. To get rid of all the brooding, I made an appointment with the branch of [a commercial test provider] in [the city of] Hengelo.* [MAIL0209]

When participants were able to secure a COVID-19 test and received a positive test result, the MHS had to contact them to start the contact-tracing process. However, participants indicated that this contact moment had not occurred, and others indicated that when this did happen, the MHS did not ask about their use and thus the possibility of sending a notification through the CM app.

Overall Sentiment

Of the emails analyzed, 34 were coded as containing a primarily negative sentiment regarding the CM app. By contrast, 12 emails were coded as having neutral and 10 as positive sentiments.

Web-Based Survey

An invitation to fill in the web-based survey was sent to 7489 individuals in the Netherlands who were selected based on their age (>18 years) and education level (mainly those with medium or lower education level were targeted). In total, 1802 (24.06%) people participated in the survey (Table 1). A total of 54.61% (984/1802) of the participants had never installed the CM app at that point in time (autumn 2020). A large group of 692 (38.40%) participants had the app installed at that moment, and a smaller group of 126 (6.99%) participants indicated that they had installed the app at some time but had removed it since then. Participants that had never installed the app were from here on out excluded from the analysis. Out of the participants that had the app installed at one point (n=818), most participants

(728/818, 89%) had not received a notification about high-risk contact from the app, while 55 (6.7%) participants received 1 notification and 26 (3.2%) received multiple notifications. A small group of participants (37/818, 4.5%) tested positive at some point before completing the survey. At the time of the survey, most participants (755/818, 92.2%) had not tested positive for COVID-19. Some participants (5/818, 0.6%) did not wish to divulge this information or provided no answer (9/818, 1.1%). Of the 37 participants who tested positive, 22 (59%) decided to share the app's key with the MHS to anonymously notify their contacts, while 15 (41%) did not. Of the 22 participants who decided to share the key with the MHS, 3 (14%) decided not to warn their contacts through the CM app. Participants who had installed the app at some point, received a notification or had tested positive (86/818, 11%) were asked to provide demographic data and participate in the follow-up interviews. Of the 86 participants, most were male (45/86, 52%). The participants had a mean age of 44.8 (SD 17.2) years. A total of 40.7% (35/86) of the participants had a low education level, 38.3% (33/86) had a medium education level, and 20.9% (18/86) had a high education level.

Table 1. Overview of results from the web-based survey (N=1802).

Subject and code	Count, n (%)
Installation of app (n=1802)	
Never	984 (54.61)
Currently installed	692 (38.4)
Installed at some point	126 (6.99)
Received a notification (n=818)	
None	728 (88.9)
One notification	55 (6.7)
Multiple notifications	26 (3.1)
No answer	9 (1.1)
Tested positive (n=818)	
Yes	37 (4.5)
No	755 (92.2)
Other or do not want to share	5 (0.6)
No answer	21 (2.6)
Shared key with Municipal Health Service (n=37)	
Yes	22 (59)
No	15 (41)

The CM App User Interview Results

The results of the app user interviews were grouped and presented for each subject. Each subject is briefly introduced and a table with the results can be found at the end of each section.

Sample Characteristics

In total, 48 CM app users were interviewed. They were recruited using various methods. A large group was recruited through outreach efforts by various organizations, such as local newspapers (33/48, 69%), low-literacy support organizations (2/48, 4%), and local professional football club supporters (4/48, 8%). Furthermore, a web-based panel (2/48, 4%, derived from the aforementioned survey), the Dutch CM app helpdesk (2/48, 4%), a local trade school (1/48, 2%), and other means (5/48, 10%) provided the remaining participants. Of the participants who were interviewed, a little more than half were identified as female (27/48, 56%) and the rest as male (21/48, 44%). Participants were aged between 21 and 80 years with a mean age of 51 (SD 16.2) years. The level of education among the participants was not representative of the Dutch population and skewed toward a high level of education (34/48, 71%) as opposed to middle (7/48, 15%) or low (7/48, 15%). The recordings of 3 (6%) interviews were not usable because of technical problems. Thus, these participants were excluded from the analysis, leaving 45 interviews for this analysis.

Installation of the App

Table 2 displays the division of the installation date across the sample. among the CM app users that were interviewed, there were 3 major points in time at which the app could be installed. First, when the app was in a local pilot-testing stage (22/49, 46%). Second, when it initially became available to the public in September 2020 (22/45, 49%). Third, after December 1, 2020 (1/45, 2%), at which point in time a major update was released for the app and the associated MHS processes [13]. The major update included the ability for CM to exchange codes with CTAs from other European countries, changes to texts in the app, and the update to MHS procedures, which allowed people to schedule a COVID-19 test if they indicated not to experience any COVID-19–related symptoms. A small group (5/45, 11%)

decided to uninstall the app between its initial release and the interview period, and a slightly larger group (7/45, 16%) predicted that they would uninstall the app in the near future:

> *My daughter in law owns a hair salon in [city], which is now closed as well. She had 5 people working for her and had the CoronaMelder. She removed it. She indicated that the app drove her crazy.* [CME008, male, 71 years, high level of education]

Table 2. Overview of results on installation and deletion of CoronaMelder app (N=45).

Subject and code	Value, n (%)
Date of installation	
During pilot	22 (46)
After pilot and before December 1, 2020	22 (46)
After December 1, 2020	1 (2)
Removed app	
Before December 1, 2020	5 (11)
After December 1, 2020	0 (0)
Will remove app in future	7 (16)

Receiving a Notification

Users of the Dutch CM app received a notification if the app determined that they were at an increased or high risk of being infected with the SARS-CoV-2. This is the case when the app registered that the user was near (closer than 1.5 meters) another app user for a prolonged (>15 minutes) period who had tested positive for COVID-19 infection through an official Dutch test center. The notification becomes visible in the notification bar of the user's phone and in the main menu of the CM app (Figure 1). The results regarding this subject are found in Table 3. Most participants (30/45, 67%) received 1 notification, some (8/45, 18%) received multiple notifications, and a smaller group (7/45, 16%) did not receive a notification.

Figure 1. Screenshot from the main menu of the app when the user was warned about a high-risk contact containing 4 main elements: (1) Title reading "You have been in the vicinity of someone with corona". (2) Description reading "You were at an increased risk of infection on <date> (x days ago).". (3) Button reading "What can I do now?". (4) Button reading "Delete notification".

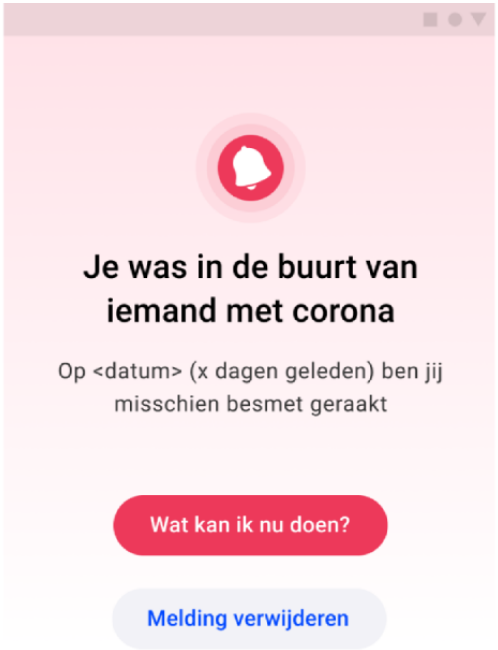


Table 3. Overview of results on receipt of notifications (N=45).

Subject and code	Value, n (%)
Number of notifications received	
None	7 (16)
One notification	30 (67)
Multiple notification	8 (18)
Prevalence of symptoms at the time of notification (n=38)	
None	30 (79)
Some	8 (21)
Prevalence of symptoms after receiving the notification (n=38)	
None developed	26 (68)
Some did develop	4 (11)
Already had symptoms	8 (21)
Actions with notification (n=38)	
Removed notification	10 (26)
Saved notification	12 (32)
First response about notification (n=38)	
Negative (eg, angry or upset)	4 (11)
Fear or anxiety	23 (61)
Disbelief or amazement	14 (37)
Beliefs about validity of notification (n=38)	
Doubtful or invalid	26 (68)
Valid	12 (32)
First action after receiving notification (n=38)	
Contacted friends or family	18 (47)
Contacted work	4 (11)
Investigated the validity and origin of notification	29 (76)
Removed the CoronaMelder app	3 (8)
Contacted a medical professional (eg, Municipal Health Service or general physician)	18 (47)
Requested test	9 (24)
Remained watchful of symptoms developing	4 (11)
Response to advice given in notification (n=38)	
Followed fully or partially	23 (61)
Not followed	9 (24)
Advice was read	6 (16)

Among those that had received a notification (n=38), the most common initial responses to receiving a notification were shock (23/38, 61%) and anger and outrage (4/38, 11%). The latter response was largely due to notification being perceived as arriving very late by the participants (9/38, 24%). This was days after the participants had already received news of high-risk contact through other channels or thought that they had already endangered others:

> *I had already received a WhatsApp message in the morning from someone whom I had exercised with that they were positive. Therefore, I already knew that I had to take action. I had a sore throat ache, but I doubted a bit. Then the CM notification came and that felt more serious.* [CME0146, female, 58 years, high level of education]

> *I was scared to death. You receive a notification that you were in contact with someone who was infected with corona on the 17th of October. That made me think: Where was I on that Saturday? What was I doing? I don't go to the market or store. I started thinking and I really thought hard about it. It gave me sleepless nights and I never figured out where I*

> *had been and who I had been in contact with.* [CME0017, male, 65 years, middle level of education]

Participants reported problems in the process of receiving notifications, such as receiving notifications late at night that woke them up or scared them. Others attempted to contact the MHS and were unable to do so, owing to busy lines or contact information that was misunderstood. Most participants (26/38, 68%) doubted whether the notification was accurate or justified, while only 12 (32%) participants were convinced of its correctness:

> *I know that I was supposed to have had a contact on 29 October and that I received the notification on the Sunday (8 November) afterwards. This, on the moment when my quarantine period had elapsed, I received the notification.* [CME0109, female, 45 years, middle level of education]

> *Because that same day I also called the MHS, but it was not possible to get through. So, then I called again the next day.* [CME0104, male, 59 years, high level of education]

Participants undertook various (and often multiple) actions after receiving a notification and often mentioned multiple actions. These actions varied from trying to verify the notification's correctness (29/38, 76%), contacting a branch of the health services or a general practitioner (19/38, 50%), contacting friends and family (18/38, 47%), contacting work (4/38, 11%), being watchful of symptoms (4/38, 11%), and uninstalling the app (3/38, 8%). Overall, participants indicated that they were uncertain or completely unaware of what to do:

> *But I did call all those people in the meantime. And those people were not afraid either, they were happy that I had called them and that was it.* [CME0008, male, 71 years, high level of education]

> *What do you do: You are going to think about where I had been and what I had done that day. And you can't figure it out properly and you thus become irritated.* [CME0125, female, 67 years, low level of education]

Perception of and Complying With Isolation and Quarantine Advice

The Dutch CM app provides information on whether one should go into quarantine after receiving a notification or should isolate themselves when receiving a positive COVID-19 test result. The advice was shown to all participants (n=45) in the study. Most of the advice given was understood, and the participants indicated that they had complied with them in general (23/45, 51%). Table 4 provides the values per advice. Participants, however, indicated that they did not understand the advice to keep distance from coinhabitants of the same house; in-depth interview questions revealed that they misunderstood the advice (7/45, 16%). The participants did not follow up on this advice (15/45, 33%):

> *Yes, but my husband and I live with just the two of us, so that's not applicable. ... My husband and I lie next to each other in 1 bed, so that is not going to work out.* [CME0006, female, 62 years, high level of education]

> *It was my daughter's birthday when my son was infected. My daughter then sat downstairs with some friends and my son was upstairs. We didn't say, at that point in time: No one is allowed into the house.* [CME0031, female, 50 years, high level of education]

Table 4. Overview of results on response, understanding, and adherence to advice provided in the CoronaMelder app (N=45).

Subject and code	Value, n (%)
Advice to stay at home	
Was recognized	35 (78)
Was executable	24 (53)
Did follow	28 (62)
Did not follow	8 (18)
Advice to have others do grocery shopping	
Was recognized	28 (62)
Was executable	24 (53)
Did follow	24 (53)
Did not follow	3 (7)
Advice to maintain distance from room- and housemates	
Was recognized	27 (60)
Not recognized or interpreted wrongly	7 (16)
Was executable	17 (38)
Not executable	5 (11)
Did follow	13 (29)
Did not follow	15 (33)
Advice to have not visitors	
Was recognized	32 (71)
Was executable	20 (44)
Did follow	23 (51)
Did not follow	2 (4)
Advice to seek medical help when symptoms worsen	
Was recognized	20 (44)
Was executable	13 (29)
Did follow	19 (42)
Advice to get tested for COVID-19 infection when symptoms present themselves	
Was recognized	25 (56)
Was executable	12 (27)
Did follow	16 (36)
Did not follow	2 (4)

Acquiring a COVID-19 Test and the Test Result

A key mechanism of the process behind the Dutch CM app is the acquisition of an official COVID-19 test after receiving a notification. A total of 24 people requested a test (Table 5), of which 13 (54%) did so at a branch of the Dutch MHS, of which 4 (31%) participants did so at a location run by general practitioners. The other 9 (69%) participants went for a COVID-19 test at another type of location, run either commercially or by their employers. Most participants (8/11, 73%) who did not go to a testing location run by the Dutch MHSs indicated that the reason for this was the perceived unavailability and slowness of the Dutch health services in scheduling tests and communicating results. Most participants received their test results between 24 and 48 hours (7/24, 29%) or within 24 hours (6/24, 25%) after the test. Results regarding the receival of test results are found in Table 6. Some participants (5/24, 21%) received their test results after 48 hours. Most participants (13/24, 54%) acquired their results through a protected government website, indicating that this was the easiest way. Participants who tested positive (13/24, 54%) were all called by the health services, which served as the start of the contact-tracing process. A total of 14 participants reported having gone in isolation sometime after a positive test, 11 (79%) of whom reported having received a notification from the app before going to the test:

At the MHS, the problem at that point in time was that their capacity was too limited for the number of requests they received. It was known at that time to be the case. That was quite a big hassle and irritator. However, it was easily arranged at a commercial test centre, so that was a good experience. [CME0227, male, 35 years, high level of education]

So, then I was tested, but my symptoms weren't severe enough. I also couldn't get a test immediately. Well, that was what the girl on the phone said, "You don't have symptoms." However, I then told her that I had been in contact with someone that had been infected and that I'd like to be tested to be sure. I said that I might have experienced mild symptoms in the throat, but I don't think that I would have had myself tested with those symptoms under normal circumstances. I did it to make the situation more severe so I could get tested. [SME 0042, female, 62 years, high level of education]

He was tested on Tuesday evening and was called on Thursday afternoon that his results were negative. In the meantime, we had drawn the conclusion that his test result would be negative, but it caused him a lot of stress in the meantime. Particularly because he knew the test result was available. But that they then waited for 1.5 days to call him, was pretty frustrating. [CME0028, female, 36 years, middle level of education]

Table 5. Overview of results on requesting COVID-19 tests (n=45).

Subject and code	Value, n (%)
COVID-19 test requested	
Test requested	24 (53)
Method or channel of request (n=24)	
Employer	1 (4)
Coronatest website (MHS[a]-run website)	6 (25)
Phone	7 (29)
General practitioner	1 (4)
Unknown or uncertain	9 (38)
Reason for use of method or channel of request (n=24)	
Speed or ease of web-based channel	6 (25)
Ease of starting a call from the CoronaMelder app	2 (8)
Higher availability	3 (13)
Priority or employer arranged	4 (17)
Unknown or other	7 (29)
Type of testing facility used (n=24)	
MHS	13 (54)
Facility run by general practitioners	4 (17)
Commercial	3 (13)
Unknown	4 (17)
Reason for using a non-MHS test (n=11)	
Speed or capacity of the MHS test insufficient	8 (72)
Time between receipt of notification and execution of test (n=24)	
<24 hours	6 (25)
<48 hours	2 (8)
>48 hours	1 (4)
Not applicable (no notification)	2 (8)
Unclear or uncertain	11 (46)
Isolation (n=24)	
Went into isolation after test	14 (58)
In isolation after notification and test	11 (46)

[a]MHS: Municipal Health Service.

Table 6. Overview of the results on receival of test results (n=24).

Subject and code	Value, n (%)
Channel through which test result was received	
Phone	4 (17)
SMS text messaging or email	2 (8)
On the web	14 (58)
Test result	
Negative	10 (42)
Positive	13 (54)
Unknown	1 (4)
Time between COVID-19 test and test result	
<24 hours	5 (21)
24-48 hours	7 (29)
>48 hours	5 (21)
Unknown	4 (17)
Actions after test result	
Removed app	4 (17)
Informed acquaintances or work	1 (4)
Searched for support (eg, Municipal Health Service)	3 (13)

Sending a Notification After a Positive Test Result

The Dutch CM app offered the option of warning others to whom a user had been in close proximity after they had received a positive test result for COVID-19 infection themselves. See Table 7 for the actions taken by participants regarding sharing their result and sending a notification. In total, 13 participants received positive test results. Participants were divided with regard to the action to be taken after receiving a positive test result. Some participants removed the app (4/13, 31%); sought help from branches of health services or general practitioners (3/13, 23%); or immediately started contacting relatives, friends, and colleagues (1/13, 7%). The participants could send notifications to their contacts. They first had to share their app's key ("the MHS key") with the health services while on the phone, which 11 (85%) participants did. They then had to send the notification through the app (Figure 2). In total, 5 out of 11 (46%) participants succeeded in doing so and thus completed the notification-sending process. In several cases (5/11, 45%), the option to send a notification was either not offered or unavailable. Overall, 2 out of 13 (15%) participants reported a positive experience in sending notifications:

> *Because of this, you also start doubting the effectiveness of other things, like the CM app. I had assumed that the hospital would have shared the positive test result. ... But in my case, the hospital hadn't communicated the positive test result with the MHS.* [CME0426, male, 69 years, education unknown]

> *And that was the next day, so on Monday the 19th I had a person on the phone about contact tracing. They asked whether I had been in contact with people. I told them that I hadn't been in contact with many people, I had been in contact with the physical therapist. They asked me if I could inform them myself and wanted to hang up. I then said that I had the CM app and asked if I could do something with it. They ten told me to provide the code, so they could report it. That I did. However, if I hadn't told them ... Then they would have just asked me to inform others and nothing else.* [CME0179, male, 70 years, high level of education]

> *No, because I assumed that the MHS would do that. ... Because the MHS asked me for the key and it wasn't clear to me that I had to finish the rest of the procedure.* [CME0179, male, 70 years, high level of education]

The screenshot in Figure 2 has been translated as follows: (1) title: "sending a notification." (2) description: "Have you been tested, and do you have COVID-19 infection? Then an MHS employee will call you. The employee will help you warn others who have been in your vicinity. You will need the MHS key for this." (3) blue hyperlinked text: "How does this work?" The 3-step plan needed for notification sending and reading, "1. Pass this MHS key through to the employee: A56-34F. 2. Wait on the MHS employee for the next step. 3. Warn others by sending an anonymous notification." (4) A button saying "Continue," which allows users to continue the process and afterward confirm the sharing of their anonymously gathered codes through a pop-up from the operating system.

Table 7. Overview of results on sharing of the Municipal Health Service (MHS) key with the MHS (n=13).

Subject and code	Value, n (%)
Shared MHS key with MHS	
Did share	11 (85)
Did not share	0 (0)
Did not mention or uncertain	2 (15)
Experience with sharing the key	
Positive	9 (69)
Negative	4 (31)
Send notification	
Did send	5 (38)
Did not send	5 (38)
Did not mention or uncertain	3 (23)
Experience with sending notification (n=11)	
Positive	2 (18)
Negative	2 (18)
Reason for not sending notification	
MHS will take care of this	5 (45)

Figure 2. Screenshot showing the screen where CM app users would start the process of sharing their anonymously gathered keys. The screen shows the three steps needed before a notification is send and other users are warned. The translation of the steps is as follows: "1 Give this MHS key to the MHS employee", "2 Wait on the MHS employee for the next step" and "3 Warn others by sending an anonymous notification.".

Overall Attitude Toward the App and the Improvements in and Strengths of the App

Most participants (37/45, 82%), as shown in Table 8, perceived the app positively overall:

> *But if you get a notification and you stay inside, then you cannot infect anyone else and then it won't spread as much. So yes, absolutely.* [CME0006, female, 62 years, high level of education]

However, a significant group (13/45, 29%) had doubts (also among those that perceived the app positively) about its effectiveness or was decidedly negatively inclined (3/45, 7%) toward the CM app. Participants experienced much uncertainty in case they received a notification and indicated (5/45, 11%) that they wanted clarification on when and why notifications were sent and the level of certainty that the high-risk contact could be determined:

> *My brother-in-law, for example, received a notification as well, but afterwards it turned out that he wasn't infected. You have to be close to someone for 15 minutes, but he says he wasn't. So that makes you doubt a bit, whether that was good. Is it [CM app] functioning 100%?* [CME0006, female, 62 years, high level of education]

Others (4/45, 9%) noted that the time between the contact for which they received a notification and the actual moment of receiving the notification was too long (eg, sometimes more than 4 days later). Respondents indicated that this undermined their confidence in the CM app's efficacy. The notification-sending process itself was unclear to some extent (4/45, 9%). Some wanted more functionality or capability (5/45, 11%) of the app and better and more graphically oriented content (5/45, 11%). Finally, some expressed that CM app use (5/45, 11%) or adherence to the recommendations provided by the app after a notification (2/45, 4%) needed to be communicated and stimulated among the public. Participants classified the app's ease of use (18/45, 40%) and user friendliness and the recommendations provided by the app as its strengths.

Table 8. Overview of results on overall attitude, strengths, and points for improvements of the CoronaMelder (CM) app (n=45).

Subject and code	Value, n (%)
Overall attitude toward the CM app	
Positive	37 (82)
Had doubts on effectiveness	13 (29)
Negative	3 (7)
Potential improvements to the CM app	
Clarify sending of notification	2 (4)
Reason and process of receiving notification is unclear	5 (11)
Time between high-risk contact and notification is too long	4 (9)
Increase functionality	5 (11)
Stimulate app use	5 (11)
Stimulate adherence to advice	2 (4)
Add more graphics in the CM app	3 (7)
Strengths of the CM app	
Clear advice	7 (16)
Easy to use or clear	18 (40)
User-friendly and good layout	9 (20)
Anonymity or privacy	1 (2)

Results of the Interviews With Contact-Tracing Employees

In total, 14 interviews were conducted with the employees of the Dutch MHS. Of these participants, 13 (93%) worked as contact-tracing employees who were responsible for contacting individuals with a positive COVID-19 test result and finding their source of infection and those that they might have infected. A participant worked as a medical adviser for the policy branch of Dutch MHS and was involved in the writing and creation of protocols and procedures for contact-tracing employees. The results are presented in text without tables, as the relatively low number of participants did not warrant the use of tables. Most participants that were active in contact tracing had worked in their roles for between 3 and 6 months (9/13, 69%). Others had worked as contact-tracing employees for <3 months (2/13, 15%). Most participants (8/13, 61%) indicated to have worked for the health services of a single region, while a single participant indicated that they had worked for more than one region. All participants that were active in contact tracing indicated that they had received some form of training in contact tracing. Most participants (10/13, 77%) indicated that their training involved some form of training on using the Dutch CM app. Of those, 7 (70%) indicated that they had been sufficiently trained in both

their primary contact-tracing process and the use of the Dutch CM app. The most often cited reason for those that reported to have received insufficient training (n=6) is that they lacked practical examples and ways to handle them (4/6, 67%) and that they would have liked more training in conversations (2/6, 33%) or parts of the CM app (4/6, 67%):

> *And the training is very theoretically good, and you think that this is how it will go. And then you start working, and you notice that it doesn't go like they said in the instructions. So you do learn useful skills, but in the end those conversations are different. Mainly because they are less scary. The instruction makes it seem like everyone is constantly angry and is going to threaten your life, but I never experienced this.* [CME 1008, male, 26 years, high level of education]

> *What we are missing, and we pointed that out yesterday, is that we need more depth. It all remains a bit superficial.* [CME1007, female, 28 years, high level of education]

Overall, 9 out of 14 (64%) interviewed participants had a generally favorable attitude toward the Dutch CM app, and 4 (28%) had a neutral or negative attitude. Only half of the participants (7/14, 50%) were app users.

Most participants (11/14, 79%) indicated that the app was only helpful in some cases, was never helpful, or that they were unsure about the app's effectiveness. The most often cited reason for this (4/11, 36%) was that participants experienced that only a small portion of their contact-tracing indexes using the app:

> *I don't use the [CM] app myself because I have the idea that it gets used too little to be truly useful. I think, that if I were to be infected, that I wouldn't be warned through the app.* [CME1008, male, 26 years, high level of education]

CM App in the Contact-Tracing Employee's Work

Procedures and Workload

Once on the phone with their contact-tracing subject ("index"), the participants that worked in contact tracing (n=13) had to walk through extensive checklists and procedures that took >2 hours on average per subject, according to most contact-tracing employees (7/13, 54%). Only 3 (23%) contact-tracing employees reported spending <2 hours on average per index:

> *Lately, I have conversations of an hour and the time needed for administration is at least an hour as well. Sometimes even longer, depending on where someone has been. Thus, it is very hard for me to make an estimation of it [time spent per index]. Often, it takes longer than expected.* [CME1009, female, 25 years, high level of education]

Participants indicated that they were often (6/13, 46%) able to contact their indexes between 24 and 48 hours after the positive test result. Moreover, 2 (15%) participants reported that they often had to contact indexes well after 48 hours had passed since the positive test result was known. For most indexes, this was their first opportunity to learn about test results.

A majority of participants (5/13, 38%) reported working according to the procedures provided by the National Office of the Health Service or a local variation in these procedures (2/13, 15%). A total of 3 (13%) participants reported that they found the procedures and their status to be unclear. Only 2 (15%) participants indicated that there was a clear set of instructions to use the CM app. A majority participants (10/13, 77%) reported that they had inquired about whether the CM app was installed by the contact-tracing indexes who they contacted. However, inquiries were of limited scope, as only 4 (30%) participants reported asking whether the contact-tracing indexes had received a notification from the app themselves. Moreover, the registration of these data about the app was also reportedly limited. Only 2 (15%) participants indicated that they had asked for the first day of the onset of the symptoms, and only 8 (62%) inquired whether the app was currently in use by the contact-tracing subject:

> *The way in which I do my contact tracing, I do not explicitly ask for it [use of CM app]. Maybe a link could be added that if you make an appointment and it is because of the CM app, that it automatically adds this in the system. When we get the test results, it then reads "This person is warned by the app." That would be a way to make that clear. At this moment there are several ways into the test centres. You can call because you experience symptoms, you can also call because you were warned by the CM app or because you are in quarantine for 5 days. We currently do not see information about this in the system.* [CME1000, male, 29 years, highly educated]

Participants working in contact tracing reported that they had to register data about the CM app in a myriad of different systems, but most commonly in one of the following systems called the "HP zone" (10/13, 77%), "CoronIT" (7/13, 54%), or special checklists (4/13, 31%). They reported that they often had to register the same data in multiple systems, and hence, the overlap in reported counts. When asked to estimate the percentage of their contact-tracing indexes using the CM app, most reported <15% (5/13, 38%) or between 15% and 30% (4/13, 31%). Moreover, according to 4 (31%) participants, a data breach at the health services at a certain moment in the CM implementation period lowered the willingness of people to share the required code to send a notification.

Only 4 (31%) MHS contact tracing employees who were interviewed reported that they had no difficulties with the CM app. Participants reported providing information about the CM app to contact-tracing indexes only to a limited degree. A total of 8 (62%) participants reported that they emphasized the importance of the subject sharing the MHS key, of which 2 (25%) participants reported explaining the process behind it as well. Only 5 (38%) participants indicated that they had separately mentioned the importance of the subject pressing the button in their CM app to send the actual notification. Participants (3/13, 23%) reported that the indexes would change their willingness to share their CM app key after having the process explained to them. A total of 8 participants reported having encountered troubles with the key sharing and notification-sending process. Most often (3/8, 38%), it was

owing to a (temporary) disruption of service; other times (2/8, 25%), the subject fell into a demographic (eg, high-school students) who at that time were not asked to share their key, and 2 (25%) participants reported having trouble completing the process itself:

> *The reason why they don't share the keys is because they didn't know that they had to provide information. Why they don't know that, is something I don't know.* [CME1006, male, 44 years, high level of education]

Improvements to the Contact-Tracing Process and CM App

The participants (n=14) were asked to describe the weaknesses of this system and some improvements. A total of 6 (43%) participants reported that the process was too slow. Some of these participants (3/6, 50%) thought that part of the process or workflow concerning the CM app could be shortened. They suggested achieving this by allowing users of the app to share their keys themselves and notify their contacts without needing contact with the MHS. Moreover, 5 (36%) participants would have liked to pay more attention to the CM app in the protocol for contact tracing. They indicated that they saw potential leads for their contact-tracing work by inquiring more deeply about CM app use; for example, on whether the index had decided to have themselves tested because of a notification. Moreover, 5 (36%) participants liked to see the app use promoted among the public.

Results Mirroring the Approach

The results of the combined research methods were discussed during 2 meetings with both the team responsible for developing and implementing the app and the team responsible for the (mass media) communication about the CM app. The results were discussed chronologically and resulted in the following changes to the CM app.

First, both end users and contact-tracing employees pointed out that the need for the MHS to be on the phone with an index to send a notification through the CM app was a limitation of the app's mechanisms. Delays were introduced into the MHS chain during periods when the number of infections was very high. Consequently, the delay between the initial infection and the moment of notification sending would increase to such a degree that it would hamper the CM app's ability to send out the warning in a timely manner. An update in December 2021 made it possible for infected individuals to send their MHS code to the MHS through a website on which they could see their test results to speed up notification.

Second, early in the interviews, it became apparent that hospitals that administer their own COVID-19 tests did not share their results in such a way, with the MHS, that a CM app notification could be sent to the patient's contacts. It was made part of the hospital protocols to participate in the CM app notification-sending process before January 2021.

Third, the mirroring sessions with the build and communication teams revealed that there was a lack of awareness concerning the importance of the total time between the first and second generations of potentially infected individuals going into quarantine. In other words, the time between individual A being infected and being notified about this and individual B, who had been in contact with A, receiving a notification from individual A about their own potential infection is a crucial time window. Stakeholders, such as build and communication teams, often used a narrower definition of this window, which excluded the part where individual B would be informed and quarantined. This resulted in an underestimation of the time window and a delay or prioritizing of changes that could shrink the time window. Both the communication and build teams embraced this broader definition, and measures were taken to reduce delays at each step of the process as a result.

Fourth, notifications sent during the night were perceived as annoying and scary. Participants would wake up from it and would not be able to go back to sleep. The app now takes the time of day into account and does not send notifications during the night.

Fifth, fears, misconceptions, and concerns identified to be prevalent among the CM app users were addressed on various information channels. The CM app's in-app information was updated, processes clarified, and the information provided by the official Dutch (government) channels was amended.

Discussion

Overview

This study investigated the implementation, adoption, and use of the Dutch CM app using a mixed methods approach. The app provides a set of features that, in theory, can greatly enhance the capability to perform contact tracing to control the spread of an infectious disease as, for example, studies of this app [8,13] and international equivalents [15] have shown. Adoption of the app and adherence to the advice (behavioral measures) are, however, key elements in the app's effectiveness This study identified issues within the app that undermine the adherence (use of the app) and implementation (adoption). This chapter answers the main research questions, makes comparisons to earlier work, discusses the strengths and weaknesses of the study, and provides a conclusion.

Principal Findings

Findings are discussed using the 3 research questions posed in the Introduction.

Factors Affecting End-User Adherence and Adoption of the App

First, the elements of the app, such as the way notifications were structured, the time of day at which notifications would appear, and the seemingly high error rate of notifications as reported by a large majority of participants, caused distress and dissatisfaction and undermined user trust.

Second, because of a lack of understanding of the mechanisms behind the app on the side of the users and MHS personnel, users reported not starting or aborting the key sharing and notification-sending processes. Moreover, MHS personnel were not adequately trained and motivated to consistently explain and encourage the use of key sharing and notification-sending functionalities.

Third, the interval between an individual getting tested, receiving a positive result, sending out a notification, and their contact receiving the notification and quarantining was key to creating and maintaining confidence in the app's effectiveness. Participants reported intervals longer than desired. Various technical and administrative actions have been suggested to reduce this interval and have since been implemented. This illustrates the importance of paying attention to the larger system and context in which the app exists. Participants indicated that they had lost (parts of) adherence to behavioral measures as a result.

Finally, the scope of the CM app was limited to an app that could warn potentially infected contacts. Users expected and indicated that they desired more of the app. A study by Blasimme et al [16] shows that adding new features may "be seen as one way to deliver more personal utility to app users, thus incentivizing participation." Instead of adding features, the reopening of society was, for example, facilitated by a separate app "CoronaCheck," and the needs around the psychological aspects of isolation were not met at all [17]. Thus, the opportunity to enhance the CM app was missed, which the German "CoronaWarn," for example, used.

Implementation in the MHS Work Processes

The CM app is intended as a tool within the broader contact-tracing process in which its purpose is to speed up parts of this process, allow for earlier warning, and reduce the load on the MHS. However, this study found that it was underused. This was especially relevant in times when infections were high, and the manual contact-tracing process was less effective. Hence, the integration with the MHS systems and processes is important for its effective functioning. However, the CM app received too little importance within the MHS; training in using and understanding the value of the app was perceived as inadequate, the data provided by the app were scarcely used, and motivation among MHS employees in using the CM app and trust in the app's effectiveness were lacking. Moreover, the overall MHS processes were perceived as fragmented and cumbersome by the employees, which further negatively affected the MHS employees' motivation to include and use the CM app. As a result, the CM was not consistently part of the MHS process, and its potential strengths in allowing for faster and earlier contact tracing did not materialize in full. Here, an app such as the CM offers a unique value that was not used.

CTA Use and Adherence to Quarantine and Isolation

The COVID-19 pandemic presents the first case in the Netherlands, in which technologies such as the CM app were used as a society-wide intervention to increase adherence to behavioral measures. As a result, this study brought to light lessons on the mistakes and importance of integration between different interventions and communication campaigns from different organizations. First, the CM app itself provided coherent and consistent messaging on behavioral measures, and its users were motivated to follow this. Second, the effectiveness of the CM app was partly dependent on its inclusion in a wider narrative set through communication campaigns and health policies. Conflicting or changing narratives and changing and nontransparent policies contributed to a lowering of trust in the CM app as a tool and adherence to its advice. Moreover, the reverse was observed during the study, and the performance of the CM app affected the overall perception of those narratives and policies. Third, the mirroring and action research approach used in this study served as an effective method to achieve rapid change and helped stakeholders in the app improve their understanding of both the daily use and reality of the app's users and the broader context in which they were the stakeholders themselves. Such changes can help increase adherence in situations such as during the COVID-19 pandemic, where uncertainty is high and change rapid. The effectiveness of the CM app in encouraging adherence behavior was limited, but its potential was great.

Comparison With Prior Work and Strengths

This mixed methods study provided a broad and in-depth insight into the functioning of the CM app, adherence to the actions, and contact-tracing process of MHS. To our knowledge, this study is the first to provide such a broad and simultaneously deep view. Other studies have focused on singular aspects, such as usability [9], its epidemiological impact [14], and factors contributing to the adoption of similar apps [9,16]. This study provides both a deep and broad view of the app and the environment in which it exists.

A strength of this study is the "mirroring approach." The results have been translated into concrete recommendations for optimization of the CM app for designers as well as for the national system for contact tracing under the MHS care. Recommendations were discussed within and between these teams and resulted in feasible and actionable changes that affected users' adherence. For example, sharing the key has changed, which means that users can upload the key without contacting the MHS. The findings were also used to set up a novel communication campaign to improve the adoption of the CM. Thus, this study has proven to be a valuable contribution to better use of the CM and optimization of the contact-tracing process. To our knowledge, this is the first study to have attempted such an approach in the context of a global health crisis.

Limitations

This study had several limitations. The first limitation is that a major part of the interviews with the CM users and the entire topic analysis relate to the period before December 1, 2020; the influence of changes in the test policy and the app after that date could not be properly determined. The second limitation was the representativeness of this study. People with a migration background, middle and lower education levels, and those aged >70 years were underrepresented. This is pertinent because these groups are less likely to work from home [18] and run a higher risk of developing a more severe illness because of lower vaccination coverage or age [19]. Another issue related to generalizability is that critical CM users may have participated in the study. The latter is related to our enrollment procedure. A third limitation related to the sample was that all interviewees had reasonable to good digital skills. It is expected that problems identified with the app will be greater or might differ for people with more limited digital skills and people with a low educational level, mild intellectual disability, or migration

background. Although the issues mentioned here negatively affect generalization, the results from the first CM evaluation study [9] and quantitative studies [6] show a picture similar to this second qualitative study. Finally, the contact-tracing study was conducted among only 3 of 25 geographically spread MHSs, which may also limit generalizability.

Conclusions

The evaluation of the CM app with end users, designers, and the MHS provided useful recommendations for the CM CTA. The lessons learned can be used to position CTAs as digital solutions for the next global health crisis and provide insights for the development and implementation of digital solutions in general.

The CM app is a CTA that is easy to use and supports intuitive use. However, the adherence to the prescribed actions (eg, sharing a "key" in case of being infected to alert other users) is low, owing to misunderstanding of the working mechanism of the app, a design that is not based on the mental logic of the end user, problems with the accuracy of notifications in 2020, the testing policy, and the uncertainty associated with receiving notifications in general. Issues that could have, at least in part, been foreseen and proactively tackled by an intervention in the technical or communication domain. Moreover, Blasimme et al [16] showed that (national) governments and health departments are responsible for the infrastructure and education on CTAs, which strongly affects the end user's adoption and use. For example, choices made during the build process of the app allow for the minimization of collected data and a fluent user experience through a sufficient level of infrastructure, thereby lowering barriers to use. Moreover, the choices made in the area of communication result in a level of education on CTAs that affects the user's familiarity and trust in such techniques [20]. In these areas, this study shows shortcomings in the approach of the (national) government and where future eHealth initiatives can improve.

The added value of the CM app is tracing risks at an earlier time than traditional contact tracing and through this, increasing the chances of breaking the chain of infections. To realize this, CM should complement or relieve traditional contact tracing. However, this potential was not realized because the interaction between the infected individual and the MHS was problematic and slow, and embedment of the CM app in MHS processes was too severely limited.

The adoption of the CM app was low. In particular, the use of the app among older adults and younger people is low [6]. The CM app was prototyped by those with lower digital literacy skills. However, the communication campaigns did not motivate younger people to use the app [21]. The incentives to download the app were based on solidarity and vulnerability to protect others, such as the older adults, which now seem to be less effective among the young.

Lack of leadership and fragmentation in governance (between the Dutch Ministry of Health, Welfare and Sport, MHS, and local government) caused distrust in the COVID-19 measurements. Besides, there was no attention by the Dutch Ministry of Health, Welfare and Sport in press conferences to the added value of the CM app, for example, to use the app as an instrument that facilitates reopening society when people increasingly come into contact with others (van't Klooster et al [17]).

Overall, this study is a testament to how a coherent strategy and process in the design, implementation, and embedment of eHealth apps, especially digital CTAs, can contribute to pandemic preparedness. To that end, Multimedia Appendix 6 [14,16,19,22,23] contains a proposal about how the lessons of this study can contribute to a strategy of that increases pandemic preparedness. Moreover, the mirroring and action approach used in this study and the focus on distilling actionable improvements to both the procedure and app provides a template for a mechanism through which future eHealth apps can rapidly be evaluated and improved even during a crisis.

Acknowledgments

The authors would like to thank the following groups and individuals for their contribution to the research project. First, Prof Dr Wolfgang Ebbers in his role as head of the research team CoronaMelder. Second, the researchers would like to acknowledge all the anonymous participants from the end users and MHS ("GGD") employee groups. Lastly, the team would like to thank those employees of the Dutch MHS communication and CM app development teams that helped reflect on and refine the findings and the local newspaper that helped with the recruitment of participants.

Conflicts of Interest

None declared.

Multimedia Appendix 1
I. Application to Ethical commission.
[DOCX File , 52 KB - formative_v6i11e38904_app1.docx]

Multimedia Appendix 2
II. Survey questions (Dutch).
[DOCX File , 16 KB - formative_v6i11e38904_app2.docx]

Multimedia Appendix 3
III. Interview questions for the interviews with CM app users.
[DOCX File , 239 KB - formative_v6i11e38904_app3.docx]

Multimedia Appendix 4
IV. Interview questions for the interviews with MHS employees.
[DOCX File , 18 KB - formative_v6i11e38904_app4.docx]

Multimedia Appendix 5
V. Original quotes and their translations.
[DOCX File , 23 KB - formative_v6i11e38904_app5.docx]

Multimedia Appendix 6
VI. Measures for preparedness of future pandemics.
[DOCX File , 15 KB - formative_v6i11e38904_app6.docx]

References

1. Thorneloe R, Fynn W, Daly M, Stanulewicz N, Kassianos A, Shorter GW, et al. Scoping review of mobile phone app uptake and engagement to inform digital contact tracing tools for COVID-19. PsyArXiv 2020 Apr 30. [doi: 10.31234/osf.io/qe9b6]
2. Wymant C, Ferretti L, Tsallis D, Charalambides M, Abeler-Dörner L, Bonsall D, et al. The epidemiological impact of the NHS COVID-19 app. Nature 2021 Jun 12;594(7863):408-412. [doi: 10.1038/s41586-021-03606-z] [Medline: 33979832]
3. Trang S, Trenz M, Weiger WH, Tarafdar M, Cheung CM. One app to trace them all? Examining app specifications for mass acceptance of contact-tracing apps. Eur J Inform Syst 2020 Jul 27;29(4):415-428. [doi: 10.1080/0960085x.2020.1784046]
4. Apple and Google partner on COVID-19 contact tracing technology. Google. 2020 Apr 10. URL: https://blog.google/inside-google/company-announcements/apple-and-google-partner-covid-19-contact-tracing-technology/ [accessed 2022-03-18]
5. Evaluation Corona Reporter: an overview Results of the study into the effect of CoronaMelder. Central Government. URL: https://www.rijksoverheid.nl/documenten/publicaties/2021/05/28/rapporten-evaluatie-coronamelder-9-maanden [accessed 2022-10-25]
6. Clarification report CoronaMelder: information security and privacy protection. Central Government. 2020 Aug 28. URL: https://www.rijksoverheid.nl/documenten/rapporten/2020/08/28/duidingsrapportage-coronamelder-informatiebeveiliging-en-privacybescherming-stand-van-zaken-lanceringsadvies [accessed 2022-10-25]
7. Op Coronatest.nl meteen besmetting doorgeven via CoronaMelder. Central Government. URL: https://www.rijksoverheid.nl/documenten/rapporten/2020/08/28/duidingsrapportage-coronamelder-informatiebeveiliging-en-privacybescherming-stand-van-zaken-lanceringsadvies [accessed 2022-03-18]
8. Bente BE, van 't Klooster JW, Schreijer MA, Berkemeier L, van Gend JE, Slijkhuis PJ, et al. The Dutch covid-19 contact tracing app (the CoronaMelder): usability study. JMIR Form Res 2021 Mar 26;5(3):e27882 [FREE Full text] [doi: 10.2196/27882] [Medline: 33724198]
9. Buhr L, Schicktanz S, Nordmeyer E. Attitudes toward mobile apps for pandemic research among smartphone users in Germany: national survey. JMIR Mhealth Uhealth 2022 Jan 24;10(1):e31857 [FREE Full text] [doi: 10.2196/31857] [Medline: 35072646]
10. Walrave M, Waeterloos C, Ponnet K. Reasons for nonuse, discontinuation of use, and acceptance of additional functionalities of a COVID-19 contact tracing app: cross-sectional survey study. JMIR Public Health Surveill 2022 Jan 14;8(1):e22113 [FREE Full text] [doi: 10.2196/22113] [Medline: 34794117]
11. Riza E, Kakalou E, Nitsa E, Hodges-Mameletzis I, Goggolidou P, Terzidis A, et al. Appraisal of a contact tracing training program for COVID-19 in Greece focusing on vulnerable populations. Int J Environ Res Public Health 2021 Sep 02;18(17):9257 [FREE Full text] [doi: 10.3390/ijerph18179257] [Medline: 34501844]
12. CoronaMelder werkt nu samen met andere Europese corona-apps. Rijksoverheid. URL: https://www.rijksoverheid.nl/actueel/nieuws/2020/12/02/coronamelder-werkt-nu-samen-met-andere-europese-corona-apps [accessed 2022-03-18]
13. Eindrapportage Evaluatie CoronaMelder-meldingen van kans op besmetting. Rijksoverheid. 2021 Mar 23. URL: https://www.rijksoverheid.nl/documenten/publicaties/2021/03/23/eindrapportage-evaluatie-coronamelder-eur-23-maart-2021 [accessed 2022-10-25]
14. Timmers F. University of Twente investigates effect of CoronaMelder: 'How did you react to a report?'. Tubantia. 2020 Nov 11. URL: https://www.tubantia.nl/enschede/universiteit-twente-onderzoekt-effect-coronamelder-hoe-reageerde-jij-op-een-melding~a2cc486d/?referrer=https%3A%2F%2Fwww.google.com%2F [accessed 2022-10-25]

15. Jenniskens K, Bootsma MC, Damen JA, Oerbekke MS, Vernooij RW, Spijker R, et al. Effectiveness of contact tracing apps for SARS-CoV-2: a rapid systematic review. BMJ Open 2021 Jul 12;11(7):e050519 [FREE Full text] [doi: 10.1136/bmjopen-2021-050519] [Medline: 34253676]
16. Blasimme A, Ferretti A, Vayena E. Digital contact tracing against COVID-19 in Europe: current features and ongoing developments. Front Digit Health 2021 Jun 17;3:660823 [FREE Full text] [doi: 10.3389/fdgth.2021.660823] [Medline: 34713135]
17. van 't Klooster JW, van Gend JE, Schreijer MA, de Witte ED, van Gemert-Pijnen L. The value of eCoaching in the COVID-19 pandemic to promote adherence to self-isolation and quarantine. In: Intelligent Human Computer Interaction. Cham: Springer; 2022.
18. De impact van de COVID-19 pandemie op werknemers. TNO. 2020. URL: https://www.monitorarbeid.tno.nl/nl-nl/publicaties/de-impact-van-de-covid-19-pandemie-op-werknemers/ [accessed 2021-09-01]
19. Grotere kans op een ernstig verloop van COVID-19. Rijksinstituut voor Volkgsgezondheid en Milieu. 2022 Jun 8. URL: https://www.rivm.nl/coronavirus-covid-19/risicogroepen [accessed 2020-09-01]
20. Fahey RA, Hino A. COVID-19, digital privacy, and the social limits on data-focused public health responses. Int J Inf Manage 2020 Dec;55:102181 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2020.102181] [Medline: 32836638]
21. Grip op de CoronaMelder. Rijksoverheid. URL: https://www.rijksoverheid.nl/documenten/publicaties/2021/04/14/eindrapport-kwalitatief-onderzoek-grip-op-coronamelder-universiteit-twente-open-universiteit [accessed 2020-09-01]
22. Jacob S, Lawarée J. The adoption of contact tracing applications of COVID-19 by European governments. Policy Design Pract 2020 Nov 28:1-15. [doi: 10.1080/25741292.2020.1850404]
23. Huisman C. Megalabs, sneltests en XL-teststraten: zo moet het testbeleid eindelijk op orde komen. de Volkskrant. URL: https://www.volkskrant.nl/nieuws-achtergrond/megalabs-sneltests-en-xl-teststraten-zo-moet-het-testbeleid-eindelijk-op-orde-komen~bb9851a7/ [accessed 2021-09-01]

Abbreviations

CM: CoronaMelder
CTA: contact-tracing app
MHS: Municipal Health Service

Edited by A Mavragani; submitted 20.04.22; peer-reviewed by S Jansen-Kosterink, A Rodrigues; comments to author 24.05.22; revised version received 10.08.22; accepted 17.08.22; published 04.11.22.

Please cite as:
van Gend JE, van 't Klooster JWJR, Bolman CAW, van Gemert-Pijnen JEWC
The Dutch COVID-19 Notification App: Lessons Learned From a Mixed Methods Evaluation Among End Users and Contact-Tracing Employees
JMIR Form Res 2022;6(11):e38904
URL: https://formative.jmir.org/2022/11/e38904
doi:10.2196/38904
PMID:36074930

Original Paper

Early Learning From a Low-Resource COVID-Response Virtual Mental Health Crisis Ward: Mixed Methods Study

Katherine Lee[1], BSc; Shay-Lee Bolton[2], PhD; Ravit Shterenberg[1], BSc; James M Bolton[2], MD; Jennifer M Hensel[2], MD, MSc

[1]Max Rady School of Medicine, University of Manitoba, Winnipeg, MB, Canada

[2]Department of Psychiatry, University of Manitoba, Winnipeg, MB, Canada

Corresponding Author:
Jennifer M Hensel, MD, MSc
Department of Psychiatry
University of Manitoba
PsycHealth Centre
771 Bannatyne Ave
Winnipeg, MB, R3E 3N4
Canada
Phone: 1 204 787 3621
Email: jhensel@hsc.mb.ca

Abstract

Background: The COVID-19 pandemic was accompanied by the accelerated uptake of virtual care, leading to a proliferation of virtual ward models as alternatives to facility-based care. Early in the pandemic, our program implemented a virtual mental health crisis ward (vWard) to provide options for individuals requiring intense psychiatric and/or crisis support but who preferred to remain in the community and were deemed safe to do so.

Objective: The aim of this study was to identify early learnings from the vWard, which was implemented rapidly in a resource-constrained environment, to inform the future state should it be sustained beyond the pandemic.

Methods: Mixed methods of data collection were used to evaluate provider perspectives on the vWard, develop archetypes for individuals who are a good fit for the vWard model, and create a driver diagram. Data sources included an anonymous survey of clinical and managerial staff involved in the vWard, a service planning workshop, and program discharge forms for all individuals admitted between March 2020 and April 2021. Survey responses were coded for themes under categories of “benefits” and “challenges.” Discharge forms where the team indicated that the vWard was a good fit for an individual were examined for characteristics common to these admissions. These findings were reviewed in the service planning workshop and refined with input from the participants into patient archetypes. A driver diagram was created for the future state.

Results: Survey respondents (N=60) represented diverse roles in crisis services and the vWard team. Ten providers took part in the service planning workshop. A total of 467 discharge forms were reviewed. The vWard was felt to be a model that worked by 39 survey respondents, one respondent felt it did not work, and the remaining participants had no response. Several benefits for the individual and the system were identified alongside challenges, including certain processes and materials related to the nature of rapid implementation during the pandemic, and others due to lack of fit for certain individuals. The model was felt to be a good fit for 67.5% of admissions. Four patient archetypes representing a good fit with the model were developed. The driver diagram connected the program aim with primary drivers of (1) reduce barriers to care; (2) improve outcomes; and (3) provide collaborative, patient- and family-centered care to secondary drivers and interventions that leveraged virtual technology among other crisis care interventions.

Conclusions: Despite some challenges, the vWard demonstrated high levels of provider acceptance and a range of mechanisms by which the model works for a variety of patient archetypes. These early learnings provide a foundation for growth, sustainability, and spread of this model going forward beyond the pandemic.

(JMIR Form Res 2022;6(11):e39861) doi:10.2196/39861

KEYWORDS

virtual ward; mental health; COVID-19; implementation; driver diagram; virtual care; virtual health care; acceptance; psychiatric support; crisis support; provider perspectives

Introduction

The COVID-19 pandemic drastically accelerated the uptake of virtual care [1-3], as many health systems sought ways to continue to provide care in accordance with new public health guidelines and to prevent the spread of SARS-CoV-2. The mental health field was uniquely positioned as a leader in telemedicine, as telemental health had a strong real-world evidence base for its effective use prior to the pandemic [2]. This was in part facilitated by the fact that physical exams are generally not required for mental health visits, whereas this is a frequently cited concern among practitioners in other fields when faced with the transition to telemedicine [3-5]. Therefore, as physical spaces in medical facilities closed or reduced capacity to accommodate physical distancing during the pandemic, mental health care programs around the world, aided by relaxed regulatory constraints, were able to rapidly transition to or expand virtual-based care that supported patients in their homes [2,6].

The pandemic also saw the proliferation of virtual ward models that leveraged technology to reduce the need for hospitalization. Examples of virtual wards specifically designed to assess and manage individuals with COVID-19 infection emerged across the globe, with studies reporting on these models in the Americas, Europe, Australia, and Asia [7]. Virtual ward models are expansions of hospital-at-home models, where a health care team provides treatment to a patient with an acute condition in the patient's home [8]. A recent systematic review of hospital-at-home programs showed evidence of benefit on patient/family- and system-level outcomes for chronic respiratory and cardiac disease, with greater cost-savings potential in the admission avoidance models compared to the models aimed at accelerating discharge from hospital [8]. Mental health crisis outreach models represent another form of hospital-at-home models in that they target individuals at high risk of hospitalization to offer crisis stabilization through in-person care in the community [9]. Prior to COVID-19, technology-enhanced virtual wards were building momentum but were not widely adopted and were largely serving individuals with chronic medical diseases [7]. While virtual mental health care delivery to individuals in their homes was increasing prior to the pandemic with the use of personal videoconferencing [10-12], there were no examples of virtual wards specifically designed for the management of a mental health crisis that leveraged technology for remote assessment and intervention.

Within weeks of the COVID-19 pandemic reaching Winnipeg, Manitoba, Canada, our center rapidly opened two virtual mental health crisis units aimed at avoiding admission to the corresponding facilities: one focusing on high-acuity psychiatric presentations to avoid admission to hospital, and the other being a lower-acuity unit to replace several beds on the Crisis Stabilization Unit (CSU) that were closed to adhere to public health guidelines for social distancing. The CSU is a community-based voluntary short-stay facility, staffed by a multidisciplinary team that acts as an alternative to hospitalization for individuals experiencing a mental health crisis [13]. The virtual units (collectively referred to as the virtual ward [vWard]) were designed to deliver all care remotely by email, text messaging, telephone, and/or videoconferencing to patients who were deemed suitable to remain in their homes.

We here report on the early learnings from the rapid implementation and delivery of these collective models in a low-resource environment. We examined the strengths and challenges of the model from the perspectives of the health care teams and the profiles of patients who were felt to be the best fit for the model. Based on these findings, we further developed a driver diagram for the future vision of a comprehensive virtual mental health crisis ward that can act as an alternative to hospital- and facility-based care in the post-COVID era, where the objective is no longer limited to keeping individuals in their homes for public health reasons. Going forward, understanding the mechanisms by which the virtual crisis ward model can be effective and for whom will inform the future of home-based virtual mental health crisis care.

Methods

Study Design

This study employed a mixed methods approach to data collection and an integrated analysis to achieve the objectives. Data sources included a voluntary online survey of providers, a service planning workshop, and discharge forms completed on all patients who accessed the virtual ward programs between March 2020 and April 2021.

Ethics Approval

Research ethics approval for this study was obtained from the University of Manitoba (HS23878 [H2020:196]).

Setting

The study was based in Winnipeg, the capital city of Manitoba, Canada. Winnipeg has a population of ~780,000, which is served by 3 emergency departments (EDs) and 3 urgent care centers offering 24/7 service. Adults can access mental health crisis care in Winnipeg via these centers or the centrally located Crisis Response Centre (CRC). The CRC is a stand-alone 24/7 walk-in mental health center, which also offers a telephone crisis line, postcrisis follow-up services, and is linked with the CSU. The virtual wards were housed at the CRC and CSU. Eligible individuals were assessed at the CRC, EDs, or urgent care centers in the city, and were deemed safe to remain in the community. This required that the individual was not at imminent risk of harm to themselves or others, participated actively in safety planning, and, when possible, had a secondary contact who would be involved in their care. The higher-acuity vWard required that individuals were assessed by the psychiatry team at the referring sites prior to referral. There were no restrictions on diagnosis or concomitant care.

Referrals were submitted to the vWard teams and patients were contacted the next day for an initial consultation. Initial contact was made by phone or email to arrange a detailed assessment, which was conducted via videoconferencing or phone when video was not available. At this assessment, a plan was made for the individual's stay, including treatment goals, safety planning, and program offerings. Families and other supports were often involved in these meetings. The higher-acuity beds were managed by a rotation of physician assistants and psychiatrists working at the CRC who offered assessment up to multiple times a day as needed, including diagnostic assessment, risk monitoring, medication management, and supportive care. This team was usually also responsible for providing care at the CRC for individuals presenting on site in crisis. The lower-acuity vWard was managed by a dedicated crisis clinician 7 days a week, usually having training in nursing or social work. The crisis clinician provided 1:1 daily virtual crisis assessment and support, along with optional group classes teaching skills derived from cognitive behavioral therapy and dialectical behavioral therapy. Psychiatric support was available to the lower-acuity beds as required. All patients had access to the 24/7 crisis phone line at the CRC for after-hours support. Medication management, when applicable, was coordinated with the individual's community pharmacy. vWard admissions were documented in the same electronic patient record used by the CRC, facilitating common access to clinical care details. The target length of stay for the virtual beds was 3-5 days. On average, there were 2 higher-acuity beds and 6 lower-acuity beds available at a given time.

Although the virtual wards operated independently of one another, we have examined them collectively with the natural progression being the merging of the units with shared infrastructure and management to provide a more collaborative, full-spectrum model of care with a dedicated team.

Data Collection

Provider Survey

A voluntary, anonymous open online survey was created using SurveyMonkey and distributed to all clinical and managerial staff who were involved with the vWard anywhere from the point of referral through discharge. An invitation to complete the survey was sent out through the email listservs for the involved services by the principal investigator of the study and forwarded by service leads who encouraged participation. The survey was open for a period of approximately 2 weeks (mid-June to early July 2021). There was no monetary incentive to participate. The survey items were developed by the principal investigator with input from the research team. Respondents were asked to identify their roles, indicate whether they felt that the vWard worked as a model of care (yes/no), and answer a series of free-response questions that included: (1) Who does the virtual unit serve? Describe the patient population. (2) What does the virtual unit do? (3) Why does it work? (4) Why doesn't it work? (5) What impact of the virtual unit do you perceive for the health care providers?

Patient Discharge Forms

Each individual admitted to the vWard had a discharge form completed by a clinical team member at the time of discharge. The discharge forms captured the key elements of the individual's condition and care delivered during the virtual admission. At the end of the form, the team member was asked to rate if the vWard was a good fit for the individual on a 5-point Likert scale ("strongly disagree" to "strongly agree"), and comment on specific successes and challenges in free text.

Service Planning Workshop

An invitation to participate in a virtual service planning workshop was emailed to the same recipients of the provider survey. Survey completion was not required to participate. Workshop recruitment favored a diversity of roles within crisis services to gain a breadth of perspectives. The workshop aimed to expand on findings from the survey and discharge forms by presenting some results for discussion. Participants were engaged in several rounds of feedback as well as a series of planned exercises drawn from quality improvement toolkits (eg, driver diagram, generating change ideas, impact vs feasibility matrix). The main objectives of the workshop were (1) to refine profiles of patients who were felt to be best suited to this model of care and (2) to create a driver diagram for a future virtual crisis ward building on the identified benefits and mechanisms from the survey. In addition to participating in the 2-hour workshop, participants were required to do some preparatory work and participate in some postworkshop follow-up totaling approximately another 2 hours. Compensation was provided in the form of a Can $200 (approximately US $140) electronic gift card to a retailer of their choosing. The 2-hour workshop was held over Zoom, facilitated by a psychiatrist and a medical student who presented the survey results, engaged deeper discussion of those results to develop the impacts and mechanisms of the vWard model for specific patient profiles, and led the participants through the planned exercises. The facilitators were also active participants in the exercises. The workshop was recorded, and all whiteboards, PowerPoint slides, and chat content were saved.

Data Analysis

The roles and involvement of survey and workshop participants were summarized descriptively. All survey respondents were included regardless of whether every question was answered. Qualitative survey responses were reviewed by 2 study team members and coded for themes pertaining to "benefits" and "challenges" of the vWard. We reviewed responses to all questions and extracted the responses that specifically addressed a benefit or challenge of the intervention. We then used a qualitative content analysis approach [14] to code individual responses and group them into thematic categories. All discharge forms where the fit rating was "agree" or "strongly agree" were selected and reviewed for demographics, clinical presentation, and the free-text comments from the team to cluster into profiles that shared common features. The preliminary profiles were presented to the workshop participants for validation, and feedback was incorporated to further develop them. In addition, a list of individual and clinical features that were generally felt to fit well with the vWard were developed. A driver diagram

was drafted during the workshop, which the investigators then further developed with reference to survey responses and workshop participant input. A complete draft was circulated back to the workshop participants for additional feedback prior to finalizing.

Results

Survey and Workshop Respondents

In total, there were 60 survey responses, including those in a role of decision-maker/manager/leadership, psychiatrist/physician assistant, crisis unit clinicians, CRC clinical staff, and other/unspecified. There were 10 participants in the provider workshop (Table 1).

Table 1. Roles of survey and workshop participants.

Virtual ward role	Survey participants, n (%)	Workshop participants, n (%)
Decision-maker/manager/leadership	10 (16)	1 (10)
Psychiatrist/physician assistant	8 (13)	2 (20)
CSU[a] clinical staff	16 (26)	3 (30)
CRC[b] clinical staff	17 (28)	2 (20)
Peer support[c]	0 (0)	2 (20)
Other/unspecified	9 (15)	0 (0)
Total	60 (100)	10 (100)

[a]CSU: Crisis Stabilization Unit.

[b]CRC: Crisis Response Centre.

[c]Peer support did not receive the survey invitation; this role was not formally involved in the virtual units at the time of the study.

Discharge Forms

Discharge forms were reviewed for 335 admissions to the lower-acuity unit and for 132 admissions to the higher-acuity psychiatric unit. Responses to the statement "Virtual care was a good fit for this patient" were missing for 7 low-acuity and 1 high-acuity admissions. In the remaining admissions, staff agreed or strongly agreed that the lower-acuity unit was a good fit in 214/328 cases (65.2%) and the higher-acuity unit was a good fit in 96/131 cases (73.3%).

Provider Perspectives: Benefits and Challenges

Survey respondents mostly stated that the vWard worked as a model of care (39/60 responded yes, 1 responded no, and the remaining 20 did not provide a response). Provider perspectives on the benefits of the vWard fell into five thematic categories: (1) provides support to stabilize acute crisis, (2) allows patients to stay in their homes, (3) increases options for patients and care providers, (4) acts as an entry point into the mental health system, and (5) can have better outcomes compared to usual care (Table 2). Providers identified the vWard's immediate and daily check-ins (for support, monitoring, and early detection of deterioration) and ease of medication support as key factors that facilitated stabilization of acute mental health crises. Providers stated that allowing patients to remain in their homes acted to reduce some of the common patient barriers to typical care, with stated examples including caregiving/work responsibilities, physical distance from the care site, disabilities, and stigma. Furthermore, allowing patients to remain in their homes had the added benefits of reducing overall hospitalizations; maximizing inpatient beds for other users; respecting patient choice; and avoiding patient-oriented risks of inpatient care, such as communicable diseases, violence, and trauma. Providers identified that having vWards available increased the number of options and flexibility of care that they could recommend to patients by offering increased hours in which appointments could occur, different types of care options, and various communication modalities (phone, videoconferencing, virtual resources). Providers also liked having the option to work remotely within the vWard. Regardless of whether the structure of the vWard met the needs of a given patient, providers also identified the benefit of having communication with patients to facilitate referrals to community supports, follow-up care, and/or inpatient care as needed. Lastly, providers identified several potential ways that vWards could provide better outcomes for patients compared to usual care. These included allowing provider assessment of the patient's function in their usual environment, encouraging practice of coping strategies in real-life situations, facilitating family involvement in care, and creating a smoother transition to the community following discharge.

Despite survey respondents agreeing with the vWard as a model of care, participants outlined several challenges with the existing program. These fell into 4 categories: (1) staff and resource limitations, (2) need for process refinement, (3) limitations of the virtual model compared to usual care, and (4) potential lack of fit for certain individuals (Table 2). The staff and resource limitations outlined by providers included concerns about redeployed and insufficient staff, inadequate equipment/resources due to a limited budget, as well as frustrating technical issues and a steep learning curve associated with the rapid pivot to virtual-based care. The category of "limitations of virtual model compared to usual care" included responses where the virtual model was contrasted with usual inpatient care in terms of assessment/diagnostic accuracy,

patient-provider rapport, ability to observe patient behavior, and access to interprofessional supports (eg, social work, nursing).

Providers identified several changes they felt were needed in the administration of the vWard process, including having scheduled appointments to reduce the administrative burden on the provider, and having a standardized protocol for admission, care delivery, and discharge. Providers also felt that the length of stay should be extended for patients needing longer periods to stabilize their crises. Provider responses captured several types of individuals that the vWard model may not be a good fit for. Referred individuals must be self-motivated and actively engaged to see improvements. The virtual aspect of care may be "too convenient" in some cases and lead to disengagement. Patients may also choose virtual care despite a clinical recommendation for inpatient care due to other factors such as social anxiety. Finally, providers highlighted clinician presentations where treatment is more suited to a supervised/closed environment, such as addictions, active suicidal intent/other safety concerns, mania and psychosis, and crises that are a result of the individual's environment (eg, intimate partner violence, problematic relationships). Additionally, the issue of access to virtual-enabled devices (ie, telephone, internet) was identified.

Table 2. Provider perspectives on virtual crisis wards with representative quotes.

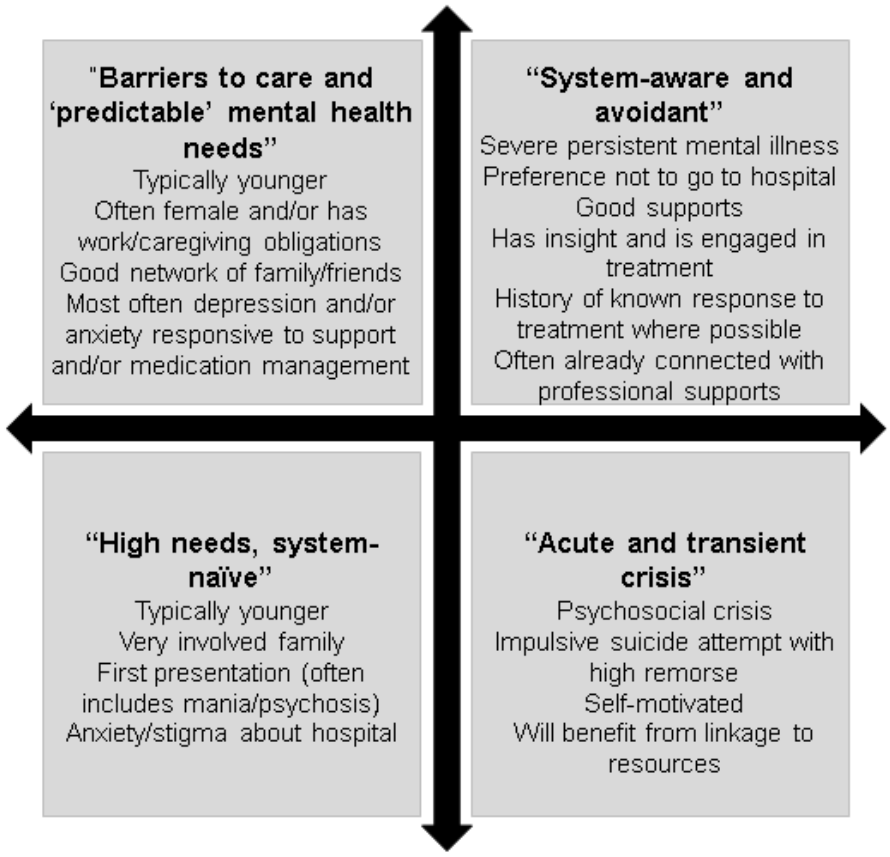
Perspectives	Description	Supporting quotes
Benefits		
Provides support to stabilize acute crisis	• Immediate and daily check-ins for support and monitoring • Daily assessment promotes early detection of deterioration • Regular medication support (reminders, adjustments)	• "With daily monitoring it serves as an access point to clients who may require further support if their current mental health further deteriorates"
Allows patients to stay in their homes	• Respects patient preference • Reduces barriers to typical care (eg, caregiving or work-related responsibilities, physical distance, disabilities, stigma) • Avoids patient-oriented risks of inpatient care (communicable diseases, violence, trauma) • Reduces hospitalizations and frees up inpatient beds	• "This provides individuals with the opportunity to continue with their daily activities and/or remain in their personal environment and still attain support"
Potential for better outcomes compared to usual care	• More flexible care (hours, types of care) • Integrates technology (phone, videoconferencing, virtual resources) • Allows providers to work remotely	• "It allows providers to see patients in their home environment and make sustainable treatment plans"
Entry point into the mental health system	• Facilitates referral to community supports and follow-up care • Seamless transition to inpatient care if needed	• "The opportunity to bring a client into the in-person unit if they aren't doing well is a further advantage…"
Challenges and limitations		
Staff and resource limitations	• Additional responsibilities contribute to staff burnout • Learning curve for providers to pivot to virtual care • Technical issues can be frustrating • Lack of adequate staffing leads to limited capacity and increased wait times • Lack of adequate equipment due to limited budget	• "Adding it on to an already very busy service can overwhelm health care providers and contribute to burnout/resentment of the work. These services would likely benefit from their own dedicated team" • "Having a 'waitlist' defeats the purpose of access to virtual care to those in community requiring supports"
Processes need refinement	• Standardized protocols for admission, care delivery, and discharge • Optimization of strategies required for scheduling appointments • Maximum length of stay should be extended • Risk of shifting individuals who require inpatient care being shifted to virtual due to bed shortages	• "A concise procedure/process in writing regarding what to do if there is no contact with a client; how long/how many attempts [to] make" • "Reevaluating the length of stay—considering longer"
Limitations of virtual model compared to usual care	• Challenges with virtual assessment accuracy and rapport • Lacks observation level of inpatient care • Patient must be self-motivated and engaged with care (ie, can be difficult to connect) • Lack of typical inpatient interprofessional supports (eg, social work, nursing)	• "It's easy for someone to 'tune-in' via Zoom, but also 'tune out.' Virtual lacks accountability that one would have with in-person stay" • "Sometimes you are stuck trying to manage a very complex case without any of the actual supports you would've gotten in the in-person setting"

Lack of fit for certain individuals	• Certain mental health presentations may require a supervised environment and/or closer observation
	  • Addictions (sober environments)
	  • Active suicidal intent/other safety concerns
	  • Mania and psychosis
	  • Crises that are a result of the environment (eg, relationship issues)
	• Patients who clinically require inpatient care may opt for virtual (eg, due to social anxiety)
	• Lack of access to resources needed for virtual care
	  • Those without access to phone or internet
		• "[Virtual ward] does not work for patients whose acute crisis presentation had to do with their environment–you can't always send the patient who [overdose]'d after an argument with tumultuous partner right back to that environment and call them the next day"
		• "Unfortunately, the use of virtual services and the requirement of technology excludes a significant portion of our client population, including those who have lower [socioeconomic status] or experience homelessness"

Patient Profiles and General "Good" Fit Factors

Individuals who were felt to be a good fit for the vWard fell into four profiles (Figure 1): (1) barriers to care and "predictable" mental health needs; (2) acute and transient crisis; (3) system-aware and avoidant; and (4) high needs, system-naïve. The first profile (barriers to care and "predictable" mental health needs) represents an individual who typically has a common mental disorder likely to be responsive to treatment with medication and/or supportive intervention (ie, depression or anxiety), who has a strong support network and/or is highly reliable to engage and follow through, and who has barriers to seeking traditional hospital- or clinic-based care. A high proportion of this group were females in the postpartum year or with young children. Other barriers to typical care pathways included work and physical disabilities.

Figure 1. Profiles of patients who were a good fit for the virtual ward (vWard) model.



Good fit factors that span all profiles:

- Self/family-motivated
- Involvement of family/personal supports
- Preference to be at home and receive care virtually
- Barriers to usual hospital/clinic-based care: school/work/other obligations/physical disabilities, etc.
- Complicated follow-up plan or needing transitional support: benefits from touchpoints to engage follow through on plan
- Client and/or family need psychoeducation/navigation support
- Able to access and use videoconferencing (particularly if diagnostic clarity needed and or concern about risk assessment)

The "acute and transient crisis" group is one with a specific event or interpersonal difficulty precipitating the crisis admission that was amenable to brief intervention. This ranged from events such as a relationship breakup to an impulsive suicide attempt with high remorse. This group was motivated to move forward and was open to support and assistance with navigating resources for the longer term.

The "system-aware and avoidant" and "high needs, system-naïve" groups represent individuals who usually have more severe mental health presentations, including severe depression, personality disorders, bipolar, and psychotic disorders. The system-aware and avoidant group was very familiar with the system, often with prior hospitalizations, but preferred not to go to hospital. They were often connected with community resources, including psychiatrists and other mental health professionals, and if experiencing a recurrence of a preexisting problem, had some insight into what the trigger was and/or what would work to improve the symptoms based on prior experience. Conversely, the system-naïve group had little to no prior contact with mental health services; were experiencing a new-onset, usually severe problem; and were

very opposed to the idea of hospitalization. These latter two profiles often necessitated the involvement of community-based supports such as family. In a significant number of cases, these presentations did lead to hospitalization, but these transitions were smoother and more acceptable to the individuals and the families who benefitted from the time to attempt recovery at home, receive more education, and collaboratively make the decision that hospitalization was needed.

A list of general "good" fit factors that crossed all patient profiles regarding vWard fit was also created (Figure 1). These included patient factors such as preference for virtual and home-based care, ability to use videoconferencing, self-motivated, involvement of family/personal supports, and the presence of barriers to seeking usual hospital-based care (eg, school or work obligations, mobility challenges). Additionally, the vWards were beneficial to patients and/or families requiring increased support, whether due to a complicated follow-up plan that required points of engagement to ensure follow-through or an increased need for psychoeducation/navigation.

Driver Diagram

A driver diagram (Multimedia Appendix 1) for the vWard program was developed that included an overall aim (aspiration and scope of the effort), primary drivers (key system components that contribute directly to accomplishing the system aim), secondary drivers (components that contribute to achieving the primary drivers), and interventions (action ideas that can be implemented with the purpose of achieving the aim, with relationship to the drivers). Provider workshop participants identified the overall aim of the vWard:

> *To deliver an alternative, home-based, collaborative model for acute mental health crisis care that works flexibly with patients and families to reduce barriers and improve outcomes.*

The primary drivers that contributed to the aim were identified as (1) reduce barriers to care; (2) improve outcomes; and (3) provide collaborative, patient- and family-centered care. The primary driver of "reduce barriers to care" included both the secondary drivers of system factors such as total system capacity and wait times, and patient "convenience and privacy" factors such as home-based, flexible, personalized, and low-stigma care options. These factors leveraged virtual technology for interventions in communication, direct care delivery, and other program components such as psychotherapy and recreational programming. The primary driver of "improve outcomes" linked to the secondary drivers of promotion of self-management techniques/in vivo stabilization (ie, stabilization in the home environment), stabilization of mental health crises, and seamless transition to other services such as follow-up and in-patient care. The transitions are further aided by providing informational continuity to patients, family, and other care providers. Finally, the primary driver of "collaborative, patient- and family-centered care" links to the secondary drivers of involving a multidisciplinary team and family members/social supports in patient care, as well as performing a patient- and family-centered needs assessment.

Discussion

Principal Results

In this paper, we report on our learning from the rapid implementation of vWards for individuals in mental health crises as alternatives for admission to in-person facilities. These models were low-resource, created in response to the public health restrictions imposed by the pandemic, and can provide the foundation to plan for more comprehensive models based on early success and evidence of feasibility. Providers from diverse roles who were involved with care delivery overwhelmingly perceived these to be models that worked and provided benefit to the patient and system. These benefits included patient choice, reduction of barriers to care, improved transitions to other services, and avoidance of risks of in-patient care such as violence or communicable disease. On a system level, the benefits included increased system capacity, improved access, and potential cost-savings. Limitations noted reflected the rapid nature of the implementation with limited resources and a sudden transition to virtual care delivery, in addition to limitations inherent to virtual acute care delivery. Distinct and diverse patient profiles that could particularly benefit from the virtual model were developed. Together, these findings allowed the development of a driver diagram for a comprehensive model that could be delivered in a higher-resource setting beyond the pandemic driven by patient choice, outcomes, and optimization of resources.

Comparison With Prior Work

Not surprisingly, many of the benefits and mechanisms of impact that were identified mirror those reported in the implementation of community-based crisis resolution teams [9]. Some additional advantages of the virtual model included the ability to reach patients in settings beyond the home to include workplaces. However, the availability of community outreach options was noted as a desired intervention for a future model, recognizing the need to connect more directly with individuals in person whose condition is deteriorating or who are unreachable. To address this, virtual models could partner with police-involved crisis outreach teams [15] as a way to increase the spectrum of care that could be offered and to provide additional comfort for the team when managing more acute presentations. Furthermore, many of the positive effects of the vWard align with those reported in the home-based telemental health literature. The general benefits included decreased barriers to treatment (eg, stigma, social anxiety, physical disabilities), patient convenience, safer environment for providers and patients at risk of violence/behavioral issues, reduced disease transmission, and the option of remote work for providers [4,5,12,16]. It has been noted that the patient-centered approach of telemental health (ie, focusing on removal of barriers and patient convenience) can lead to improved treatment adherence as patients are more satisfied with their care [10,12]. Improved treatment adherence is one of the possible mechanisms by which the vWard could result in improved mental health outcomes compared to usual care.

While providing low-barrier, flexible options for care is beneficial, there were also situations where the vWard was not

felt to be appropriate or introduced limitations to proper assessment and management of certain individuals. The model relies on patients being self-motivated and engaged with their care, as there are more distractions in the home environment that could impede focus on self-improvement. If a patient missed virtual appointments, there were few options to contact them for follow-up. Home disturbances and interfering factors have been previously described in the virtual care literature [17] and for mental health care specifically [11], and require attention if care outcomes are to be optimized. Patients who require contained environments, such as those with highly agitated and disorganized presentations, or those needing sober environments will also not be good candidates for vWard admission. Zimmerman et al [18] reported on the virtual transformation of a partial hospitalization program during the pandemic, demonstrating feasibility and retention compared to a historical comparison cohort; however, the users of the virtual model did have lower levels of psychosis, and presence of a primary substance use disorder was an exclusion criterion. Access to virtual care is also a limitation as there is a risk to expose inequities [19]. While the majority of Canadians do have access to virtual technology, a portion do not, and this is often correlated with other measures of marginalization and poor health [19]. This is an area that requires more attention at a population level to ensure equitable access for all.

Many of these factors were also identified in our patient profile based on good fit. The generally "good" factors capture the presence of barriers and preference, ability, and motivation for the virtual care model compared with in-person care. The specific profiles that were elucidated from the data exemplify the diversity of individuals who can be managed in this model, recognizing that each profile likely needs a unique approach necessitating a team with a wide skill set. Some studies have discussed individual patient characteristics that are suitable to outpatient virtual care, including transportation issues (eg, living far from the location of service or lack of vehicle), busy work or family schedules that make seeking in-person care difficult, and those whose conditions impair treatment-seeking (eg, anxiety, agoraphobia) [20]. The patient profiles identified from our data hone in on the archetypes that benefit most so that services can be designed and delivered with the needs of these groups in mind. Depending on priority areas, gaps in other services, or availability of resources, the model could pivot to focus more or less on certain groups [21].

In addition to limitations of the model itself, providers also identified challenges with resources and processes. Many of these were a result of the model being rapidly deployed due to the threat of COVID-19 and relatively low availability of resources, not necessarily a limitation of the vWard model itself. Staff were encountering rapid change alongside uncertainty about infection risk, leading to increased stress and potential for burnout [22,23]. Prior to COVID-19, the learning curve associated with the pivot to delivering care virtually had been documented as a common limitation to the general adoption of virtual care [3,21]. This model was resourced with minimal levels of equipment and staffing, as well as a lack of the interprofessional teams typical of hospital care. Although these resource limitations were significant, the model was sustained through dedicated staff, a shared vision, and adaptive leadership styles. As discussed by Laur et al [21], these facets are critical to managing rapid change in the health care system. Going forward, these limitations could be rectified with additional investment. There is evidence of cost-effectiveness for home-based virtual care delivery compared to in-person care [12,24]. Beyond savings to the health care system, patients also report direct savings when able to receive care at home [25]. This makes a case for greater investment in the growth of these models beyond the pandemic, alongside ongoing evaluation of impact.

Through the creation of the driver diagram, we propose a blueprint for the future vWard as an alternative model of care that leverages technology. The COVID-19 pandemic has had a negative impact on the mental health of the population due to increased stress, isolation, and reduced treatment access [26]. Additionally, in our region, we have a higher burden of mental illness in the population and a higher rate of mental health presentations to EDs compared to the rest of Canada [27]. New strategies are needed. Furthermore, our province experiences significant regional variation in access to mental health care, with disparities increasing when moving north toward rural areas [27]. Although this has not been a focus to date, virtual models of care could address geographical barriers to access with the vWard having the potential to fill a major service gap. Urgent telemental health programs have been developed in rural settings to provide assessment, with some programs offering follow-up care [28]. The vWard expands on this with the goal of reducing hospitalization that often takes individuals away from their communities and families, aiming to provide intensive care to support crisis resolution. The ability to refer to a follow-up service provides additional options for emergency mental health teams; for example, access to a telephone-based peer-led navigation service reduced the rate of admission following emergency telepsychiatry assessment to urban and suburban areas in North Carolina, United States. Although not significant, this low-resource intervention provided a signal of possible impact [29]. Drawing on the evaluation of home-based crisis resolution teams in the UK National Health Service [9], a follow-up service that includes a prescriber and well-trained multidisciplinary team members to support a range of health and social needs may be most successful at reducing rates of admission and repeat acute care use.

Limitations

One limitation of this study is the unique circumstance under which this model was developed and launched. The catalyst for virtual care delivery and innovation provided by the pandemic was indeed an opportunity but also created conditions that do not normally exist in health care service design [30]. The survey was voluntary and thus subject to response bias; however, we achieved a very good representation of roles and range of perspectives as evidenced by the variety of staff respondent groups, which are proportional in size to the total number of individuals working in these roles. We did not collect demographic or other respondent characteristics to assess representativeness across the workforce profile. A significant gap in this work is the lack of patient perspectives, which we were unable to comprehensively collect due to resource

limitations during the study period. With a plan to sustain the model locally, we are now building in patient satisfaction and further evaluation of patient experience. Other studies of outpatient virtual care have assessed patient satisfaction, generally finding high ratings, with many of the themes overlapping with those identified by our providers [4,25]. This unique blend of crisis support and virtual care requires further exploration from the patient perspective. The findings of this study must also be taken in context of its implementation: rapid, low-resource, and limited budget. Many of the "problems" of this home-based virtual model identified by the providers were instead areas that could be improved with increased investment of staff and equipment and are not necessarily intrinsic issues to the model. This is also why we focused on developing a driver diagram for a future vision based on the learnings of this rapid pivot in care delivery. These drivers and interventions will need to be validated and tested. The next step to completing the driver diagram is to map on an implementation plan along with process and outcome measures that can be collected over time, as illustrated in the Action Effect Diagram described by Reed et al [31]. Collection of these measures is crucial to evaluate the impact of the interventions.

Conclusions

The COVID-19 pandemic advanced an opportunity to develop a novel model by leveraging technology to provide care virtually to a high-acuity population. Despite some challenges with resources in a rapidly changing health care context, we have demonstrated high levels of provider acceptance and a range of mechanisms by which the model works for a variety of patient archetypes. These findings highlight barriers to be anticipated and overcome in the design of similar models and identify the patients who may benefit most from virtual crisis intervention as an alternative to staying in a facility, be it a hospital or CSU-type environment. There is still room to improve and optimize this model. These early learnings provide a foundation for growth, sustainability, and spread going forward beyond the pandemic to increase access to quality care using novel means that are highly patient-centered. Other jurisdictions interested in developing similar initiatives may use these learnings as a starting point in the design and implementation of local programs.

Acknowledgments

The study team acknowledges the participants for their valued contributions, and appreciates all of the health care providers and support staff involved in the implementation and day-to-day operations of these models of care during a period of high stress and rapid change. The study team also acknowledges Jocelyne Lemoine, MA, Research Associate with the Department of Psychiatry at the University of Manitoba, for assisting with manuscript preparation. JMH was awarded a grant from the Canadian Medical Association for this research.

Conflicts of Interest

The authors have no potential conflicts of interest to declare.

Multimedia Appendix 1
Driver diagram for the future state of the virtual mental health crisis ward (vWard) implementation.
[DOCX File , 167 KB - formative_v6i11e39861_app1.docx]

References

1. Glazier RH, Green ME, Wu FC, Frymire E, Kopp A, Kiran T. Shifts in office and virtual primary care during the early COVID-19 pandemic in Ontario, Canada. CMAJ 2021 Feb 08;193(6):E200-E210 [FREE Full text] [doi: 10.1503/cmaj.202303] [Medline: 33558406]
2. Shore JH, Schneck CD, Mishkind MC. Telepsychiatry and the coronavirus disease 2019 pandemic-Current and future outcomes of the rapid virtualization of psychiatric care. JAMA Psychiatry 2020 Dec 01;77(12):1211-1212. [doi: 10.1001/jamapsychiatry.2020.1643] [Medline: 32391861]
3. Woolliscroft JO. Innovation in response to the COVID-19 pandemic crisis. Acad Med 2020 Aug;95(8):1140-1142 [FREE Full text] [doi: 10.1097/ACM.0000000000003402] [Medline: 32282372]
4. Chen JA, Chung W, Young SK, Tuttle MC, Collins MB, Darghouth SL, et al. COVID-19 and telepsychiatry: early outpatient experiences and implications for the future. Gen Hosp Psychiatry 2020 Sep;66:89-95 [FREE Full text] [doi: 10.1016/j.genhosppsych.2020.07.002] [Medline: 32750604]
5. Connolly SL, Gifford AL, Miller CJ, Bauer MS, Lehmann LS, Charness ME. Provider perceptions of virtual care during the coronavirus disease 2019 pandemic: a multispecialty survey study. Med Care 2021 Jul 01;59(7):646-652 [FREE Full text] [doi: 10.1097/MLR.0000000000001562] [Medline: 34009880]
6. Zangani C, Ostinelli EG, Smith KA, Hong JSW, Macdonald O, Reen G, et al. Impact of the COVID-19 pandemic on the global delivery of mental health services and telemental health: systematic review. JMIR Ment Health 2022 Aug 22;9(8):e38600 [FREE Full text] [doi: 10.2196/38600] [Medline: 35994310]

7. Vindrola-Padros C, Singh KE, Sidhu MS, Georghiou T, Sherlaw-Johnson C, Tomini SM, et al. Remote home monitoring (virtual wards) for confirmed or suspected COVID-19 patients: a rapid systematic review. EClinicalMedicine 2021 Jul;37:100965 [FREE Full text] [doi: 10.1016/j.eclinm.2021.100965] [Medline: 34179736]
8. Arsenault-Lapierre G, Henein M, Gaid D, Le Berre M, Gore G, Vedel I. Hospital-at-home interventions vs in-hospital stay for patients with chronic disease who present to the emergency department: a systematic review and meta-analysis. JAMA Netw Open 2021 Jun 01;4(6):e2111568 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.11568] [Medline: 34100939]
9. Wheeler C, Lloyd-Evans B, Churchard A, Fitzgerald C, Fullarton K, Mosse L, et al. Implementation of the crisis resolution team model in adult mental health settings: a systematic review. BMC Psychiatry 2015 Apr 08;15(1):74 [FREE Full text] [doi: 10.1186/s12888-015-0441-x] [Medline: 25879674]
10. Pruitt LD, Luxton DD, Shore P. Additional clinical benefits of home-based telemental health treatments. Prof Psychol Res Pr 2014 Oct;45(5):340-346. [doi: 10.1037/a0035461]
11. Hensel JM, Yang R, Vigod SN, Desveaux L. Videoconferencing at home for psychotherapy in the postpartum period: identifying drivers of successful engagement and important therapeutic conditions for meaningful use. Couns Psychother Res 2020 Dec 14;21(3):535-544. [doi: 10.1002/capr.12372]
12. Fletcher TL, Hogan JB, Keegan F, Davis ML, Wassef M, Day S, et al. Recent advances in delivering mental health treatment via video to home. Curr Psychiatry Rep 2018 Jul 21;20(8):56. [doi: 10.1007/s11920-018-0922-y] [Medline: 30032337]
13. Adams CL, El-Mallakh RS. Patient outcome after treatment in a community-based crisis stabilization unit. J Behav Health Serv Res 2009 Jul 3;36(3):396-399. [doi: 10.1007/s11414-008-9141-3] [Medline: 18766444]
14. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
15. Lamanna D, Shapiro GK, Kirst M, Matheson FI, Nakhost A, Stergiopoulos V. Co-responding police-mental health programmes: service user experiences and outcomes in a large urban centre. Int J Ment Health Nurs 2018 Apr 17;27(2):891-900. [doi: 10.1111/inm.12384] [Medline: 29044920]
16. Barnett P, Goulding L, Casetta C, Jordan H, Sheridan-Rains L, Steare T, et al. Implementation of telemental health services before COVID-19: rapid umbrella review of systematic reviews. J Med Internet Res 2021 Jul 20;23(7):e26492 [FREE Full text] [doi: 10.2196/26492] [Medline: 34061758]
17. Stuesson L, Groth K. Effects of the digital transformation: qualitative study on the disturbances and limitations of using video visits in outpatient care. J Med Internet Res 2018 Jun 27;20(6):e221 [FREE Full text] [doi: 10.2196/jmir.9866] [Medline: 29950290]
18. Zimmerman M, Terrill D, D'Avanzato C, Tirpak JW. Telehealth treatment of patients in an intensive acute care psychiatric setting during the COVID-19 pandemic: comparative safety and effectiveness to in-person treatment. J Clin Psychiatry 2021 Mar 16;82(2):20m13815. [doi: 10.4088/JCP.20m13815] [Medline: 33989463]
19. Crawford A, Serhal E. Digital health equity and COVID-19: the innovation curve cannot reinforce the social gradient of health. J Med Internet Res 2020 Jun 02;22(6):e19361 [FREE Full text] [doi: 10.2196/19361] [Medline: 32452816]
20. Morland LA, Poizner JM, Williams KE, Masino TT, Thorp SR. Home-based clinical video teleconferencing care: clinical considerations and future directions. Int Rev Psychiatry 2015 Nov 30;27(6):504-512. [doi: 10.3109/09540261.2015.1082986] [Medline: 26619273]
21. Laur CV, Agarwal P, Mukerji G, Goulbourne E, Baranek H, Pus L, et al. Building health services in a rapidly changing landscape: lessons in adaptive leadership and pivots in a COVID-19 remote monitoring program. J Med Internet Res 2021 Jan 13;23(1):e25507 [FREE Full text] [doi: 10.2196/25507] [Medline: 33417588]
22. Hannemann J, Abdalrahman A, Erim Y, Morawa E, Jerg-Bretzke L, Beschoner P, et al. The impact of the COVID-19 pandemic on the mental health of medical staff considering the interplay of pandemic burden and psychosocial resources-A rapid systematic review. PLoS One 2022;17(2):e0264290 [FREE Full text] [doi: 10.1371/journal.pone.0264290] [Medline: 35192662]
23. Liu Q, Luo D, Haase JE, Guo Q, Wang XQ, Liu S, et al. The experiences of health-care providers during the COVID-19 crisis in China: a qualitative study. Lancet Glob Health 2020 Jun;8(6):e790-e798 [FREE Full text] [doi: 10.1016/S2214-109X(20)30204-7] [Medline: 32573443]
24. Bounthavong M, Pruitt LD, Smolenski DJ, Gahm GA, Bansal A, Hansen RN. Economic evaluation of home-based telebehavioural health care compared to in-person treatment delivery for depression. J Telemed Telecare 2016 Nov 12;24(2):84-92. [doi: 10.1177/1357633x16678147]
25. Yang R, Vigod SN, Hensel JM. Optional web-based videoconferencing added to office-based care for women receiving psychotherapy during the postpartum period: pilot randomized controlled trial. J Med Internet Res 2019 May 20;21(6):e13172 [FREE Full text] [doi: 10.2196/13172] [Medline: 31199291]
26. 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]
27. Virgo Planning and Evaluation Consultants Inc. Improving access and coordination of mental health and addiction services: a provincial strategy for all Manitobans. Government of Manitoba. 2018 Mar 31. URL: https://www.gov.mb.ca/mh/mh/docs/mha_strategic_plan.pdf [accessed 2019-04-26]

28. Saurman E, Johnston J, Hindman J, Kirby S, Lyle D. A transferable telepsychiatry model for improving access to emergency mental health care. J Telemed Telecare 2014 Nov 16;20(7):391-399. [doi: 10.1177/1357633x14552372]
29. Roberge J, McWilliams A, Zhao J, Anderson WE, Hetherington T, Zazzaro C, et al. Effect of a virtual patient navigation program on behavioral health admissions in the emergency department: a randomized clinical trial. JAMA Netw Open 2020 Jan 03;3(1):e1919954 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.19954] [Medline: 31995214]
30. Hayden L, Hensel J, Bhattacharyya O. Health services and designing for uncertainty. Stanf Soc Innov Rev 2015:1-5 [FREE Full text] [doi: 10.48558/KW2Z-6A08]
31. Reed JE, McNicholas C, Woodcock T, Issen L, Bell D. Designing quality improvement initiatives: the action effect method, a structured approach to identifying and articulating programme theory. BMJ Qual Saf 2014 Dec 15;23(12):1040-1048 [FREE Full text] [doi: 10.1136/bmjqs-2014-003103] [Medline: 25319412]

Abbreviations

CRC: Crisis Response Centre
CSU: Crisis Stabilization Unit
ED: emergency department
vWard: virtual mental health crisis ward

Edited by G Eysenbach; submitted 25.05.22; peer-reviewed by MM Patwary, M Brown; comments to author 07.09.22; revised version received 26.09.22; accepted 07.10.22; published 04.11.22.

Please cite as:
Lee K, Bolton SL, Shterenberg R, Bolton JM, Hensel JM
Early Learning From a Low-Resource COVID-Response Virtual Mental Health Crisis Ward: Mixed Methods Study
JMIR Form Res 2022;6(11):e39861
URL: https://formative.jmir.org/2022/11/e39861
doi:10.2196/39861
PMID:36252139

Original Paper

The Incidence and Effect of Adverse Events Due to COVID-19 Vaccines on Breakthrough Infections: Decentralized Observational Study With Underrepresented Groups

Irene S Gabashvili[1], PhD

Aurametrix, MEBO Research, Miami, FL, United States

Corresponding Author:
Irene S Gabashvili, PhD
Aurametrix
MEBO Research
13220 SW 35th Terrace
Miami, FL, 33175
United States
Phone: 1 4083419355
Email: irene@aurametrix.com

Abstract

Background: Despite continuing efforts to improve the inclusion of underserved groups in clinical research, gaps in diversity remain. Participation of special populations is especially important when facing problems of unprecedented complexity such as the COVID-19 pandemic. A better understanding of factors associated with the immune response in diverse populations would advance future preventive and curative approaches.

Objective: The objective of this study was to investigate the factors potentially responsible for adverse events following COVID-19 immunization. The study population included adults from rural areas, transitional countries, and those with medically understudied conditions, across a broad age range.

Methods: The study evolved from peer support networks developed during the COVID-19 pandemic. Participants were recruited digitally through online neighborhood and health communities. Some of the participants volunteered as study investigators assisting with offline recruitment and safety monitoring. Individuals who consented to participate were asked to share their vaccination experiences either using constantly evolving web-based surveys or via one-on-one communication. Inferential statistical analysis to estimate differences between study groups was performed using parametric and nonparametric tests.

Results: Of 1430 participants who shared their vaccination experiences, 648 had outcome measures at their 1.5-year follow-up. Significant differences were found between age groups, types of vaccine adverse events (VAEs), incidences of breakthrough infections, and health conditions linked to the microbiome. Pairwise comparisons showed that VAEs interfering with daily activities were significantly higher in both younger (18-59 years) and older age groups (80-100 years, $P<.001$) than in the 60-79–year age group. Short-term VAEs were associated with lower incidence of breakthrough COVID-19 infections relative to those who reported either minimal or long-term adverse events ($P<.001$). A genetic origin was suggested for some adverse reactions.

Conclusions: The findings of this study demonstrate that vaccine adverse reactions in older individuals are being overlooked, and the incidence of VAEs impairing immunity may be higher than previously perceived. Better preventive measures are needed for all those at risk for life-threatening and long-term adverse events due to vaccination. Supportive community-based studies focusing on these populations could add important data to the current body of knowledge. Further and more comprehensive studies should follow.

Trial Registration: ClinicalTrials.gov NCT04832932; https://clinicaltrials.gov/ct2/show/NCT04832932

International Registered Report Identifier (IRRID): RR2-10.1101/2021.06.28.21256779

(JMIR Form Res 2022;6(11):e41914) doi:10.2196/41914

KEYWORDS

COVID-19; COVID-19 vaccines; vaccine adverse events; breakthrough infections; decentralized participatory study; elderly; older individuals; medically underserved populations; aging; elderly population; vaccination; genetic disparity; microbiome disparity; impaired immunity

Introduction

Background

While COVID-19 vaccines are highly effective, they can lead to a range of vaccine adverse events (VAEs). The mechanisms underlying VAEs and their association with vaccination efficacy are not completely clear.

Traditional clinical trials are essential for understanding the safety profiles of new interventions, but many people would not enroll themselves. The COVID-19 Citizen Science study [1] and social media analytics [2,3] address some knowledge gaps, but not social- and age-related barriers to participation. Passive surveillance data [4] are subject to the same limitations and insufficient follow-up. Alternative research strategies are needed to complement the existing evidence base.

Research Objectives

The idea for this study evolved as minority populations were struggling to make vaccination appointments owing to digital inequality, or seeking answers to questions about their pre-existing conditions, which were not addressed by funded research. A previously reported protocol for an ambispective study [5] hypothesized that the safety profile and immune response to COVID-19 vaccines depend on pre-existing health conditions, metabolism, and microbiomes. The objective of this study was to investigate the factors influencing adverse events following COVID-19 immunization in communities that include underrepresented groups.

Methods

Study Population

Participants were recruited from private e-neighborhood networks and health support groups via direct emails and social media posts. Particular efforts were made to recruit populations underrepresented in existing research data sets. Individuals with socially debilitating metabolic body odor (MEBO) [6], including extraoral halitosis, “People Allergic to me,” and trimethylaminuria were invited to contribute to the goals of this study along with their relatives. Support groups for autoimmune disorders, neuropathy, tinnitus, and irritable bowel syndrome were also contacted. To ensure the inclusion of digitally disadvantaged individuals, several volunteers served as study investigators providing face-to-face support. To prevent double-counting, investigators had access to lists of subjects maintained by other investigators in their community but not to personally identifiable information. Cookies were not used to guarantee the anonymity of those sharing information without enrollment. To boost engagement, research insights were periodically communicated via the internet. Data collection was automated using rule-based parsing of emails, alerts from social networks, survey spreadsheets, and group application programming interfaces. An inquiry into medical history was made at initial and follow-up data collections.

Inclusion Criteria

Inclusion criteria were age≥18 years, intention to get vaccinated, and intended availability throughout the study period. No one was excluded for reasons other than age.

Outcomes

The study’s primary outcome was the incidence of adverse reactions within 14 days of immunization. The secondary outcome was long-term health conditions and the incidence of breakthrough COVID-19 infections that occurred post vaccination with either a single dose of a COVID-19 vaccine or with 2 or more doses.

Ethical Considerations

Details of the study procedures (the optional nature of all questions, how the information will be used, the ability to withdraw from research, risks, and benefits) were explained to the participants, and informed e-consent was obtained. Ethics approval for primary and secondary analysis was granted by the institutional review board of MEBO Research (IRB00010169, protocol 20210103MEBO). To ensure participant confidentiality, all their data were coded and stored in a decentralized manner with no individual having complete access to sensitive information. All identifiable data were removed from survey and interview responses. Neither participants nor investigators were compensated.

Data Analysis

Demographic and clinical characteristics of study groups were compared using the Fisher exact test and relative risk calculations for categorical variables and the Mann-Whitney *U* test for continuous variables. All statistical tests were 2-sided, for which a *P* value of ≤.05 and a 95% CI were used to indicate statistical significance. All analyses were conducted using Python (version 3.10).

Results

Participant Characteristics

Participant characteristics are described in Table 1. Of the 1430 vaccinated adults and the 648 participants who were followed up, 51% (n=732 and n=329, respectively) were female. Prevalence of chronic disease was age- and sex-matched between the study cohort and the general population. The age at vaccination ranged from 18 to 119 years. The age group of ≥100 years includes 20 vaccinated semisuper- and supercentenarians (10 men and 10 women) with official social media and Gerontology Wiki accounts [7]. These individuals were added to the study database and followed up from early 2021. Since no information is available about their

postimmunization symptoms, this group is not included in the VAE analysis.

Table 1. Basic descriptive and inferential statistics of the study population.

Characteristics	Vaccinated (n=1430)				1-year follow up (n=648)				All-cause mortality (n=30)	
	Total (n=1430)	No or minimum VAEs[a] (n=1113)	VAEs (n=317)	*P* value	Total (n=648)	No COVID-19 infection post vaccination (n=389)	Breakthrough COVID-19 infection (n=259)	*P* value	Mortality (n=30)	*P* value
Age at receipt of the first dose (years), median (IQR)										
All	62 (40-70)	65 (49-72)	42 (30-61)	<.001	58 (38-70)	61 (47-62)	42 (30-66)	<.001	95 (75-110)	<.001
Female	63 (41-70)	65 (49-72)	45 (31-64)	Ref.[b]	59 (40-70)	63 (50-72)	43 (32-65)	Ref.	104 (84-114)	Ref.
Male	62 (38-71)	65 (46-72)	40 (27-55)	<.001	56 (35-71)	60 (43-72)	42 (29-67)	.90	91 (75-108)	.07
Sex, n (%)										
Female	732 (51)	558 (50)	174 (55)	Ref.	329 (51)	201 (52)	128 (49)	Ref.	12 (40)	Ref.
Male	698 (49)	555 (50)	143 (45)	.20	319 (49)	188 (48)	131 (51)	.60	18 (60)	.30
Adverse events, n (%)										
Short-term	174 (12)	0 (0)	174 (56)	N/A[c]	93 (15)	77 (20)	16 (6)	<.001	1 (3.5)	.20
Long-term	143 (10)	5 (0)	138 (44)	N/A	103 (16)	52 (13)	51 (20)	.20	1 (3.5)	.40
No or minimal	1113 (78)	1113 (100)	0 (0)	N/A	450 (69)	260 (67)	187 (74)	Ref.	26 (93)	Ref.
Age groups (years), n (%)										
18-29	169 (7)	97 (9)	72 (23)	<.001	85 (13)	29 (7)	56 (22)	<.001	0 (0)	N/A
30-39	173 (9)	102 (9)	71 (22)	<.001	85 (13)	34 (9)	51 (21)	<.001	0 (0)	N/A
40-49	139 (12)	93 (9)	46 (15)	<.001	81 (13)	47 (12)	34 (13)	.10	0 (0)	N/A
50-59	164 (17)	127 (11)	37 (12)	<.001	90 (14)	63 (17)	27 (11)	>.99	0 (0)	N/A
60-69	401 (24)	356 (32)	45 (14)	Ref.	135 (21)	93 (24)	42 (16)	Ref.	2 (7)	Ref.
70-79	301 (20)	272 (24)	29 (9)	.50	116 (17)	78 (20)	39 (14)	.70	8 (29)	.02
80-100	62 (7)	45 (4)	17 (5)	<.001	35 (6)	27 (7)	8 (3)	.90	7 (23)	<.001
>100	21 (4)	N/A	N/A	N/A	21 (3)	20 (4)	1 (0)	.06	13 (41)	<.001

[a]VAE: vaccine adverse event.

[b]Ref.: Reference.

[c]N/A: not applicable.

Evaluation Outcomes

The CONSORT (Consolidated Standards of Reporting Trials) flow diagram in Figure 1 shows the progression of the study, including that of the 1430 subjects who received their first vaccine dose between December 2020 and August 2022 and the 648 individuals whose most recent update was between May and October 2022.

The bar charts in Figure 1 illustrate a balanced representation of both sexes in all age groups, which is also evident in Table 1. Out of 1430 study participants who self-reported outcomes after vaccination, 317 (22%) experienced side effects that prevented them from performing daily activities after receiving at least one of the doses (Table 1). The overall Kruskal-Wallis comparison of VAEs in all age groups was significant (P<.001). Pairwise comparisons showed that while the rate of adverse events was similar in the 60-69–year and 70s-79–year age group, the incidence of VAEs in both younger (18-59 years) and older age groups (80-100 years) was significantly higher (P<.001).

Figure 2 displays the distribution of 648 participants by adverse events, or the absence thereof, for each age group.

Incidence of breakthrough infections was significantly higher in participants younger than 40 years (Figure 3), as observed previously in many different healthy cohorts [8-10].

Another significant association was found between adverse reactions and the ability to avoid infection, but only if symptoms did not last longer than a week (Table 1). We observed higher incidence of pulmonary VAEs in the MEBO [6] subgroup of individuals with halitosis when compared to the group with the

second highest prevalence (P=.03). We also identified 5 pairs of first-degree relatives with nearly identical sets of VAEs including cardiological events or severe nausea.

Overall, 5%-10% of individuals in chronic disease groups experienced exacerbation of their respective conditions following COVID‐19 immunization. The difference in the incidence of VAEs between healthy individuals and those with chronic conditions was not significant for all age groups (P>.05).

There were no significant gender differences in the incidence of VAEs, but age distributions revealed significant differences: a greater risk of VAEs among younger males and higher all-cause mortality in older individuals (Table 1).

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) flow diagram. Participant flow through the study. Asterisks denote data for the centenarians with publicly available profiles, followed up from early 2021. The population pyramid chart (blue: males; pink: females; transparent shades: no follow-up) shows age and sex distribution of vaccinated individuals (n=1430) and those who were recently followed up (n=648). VAE: vaccine adverse event.

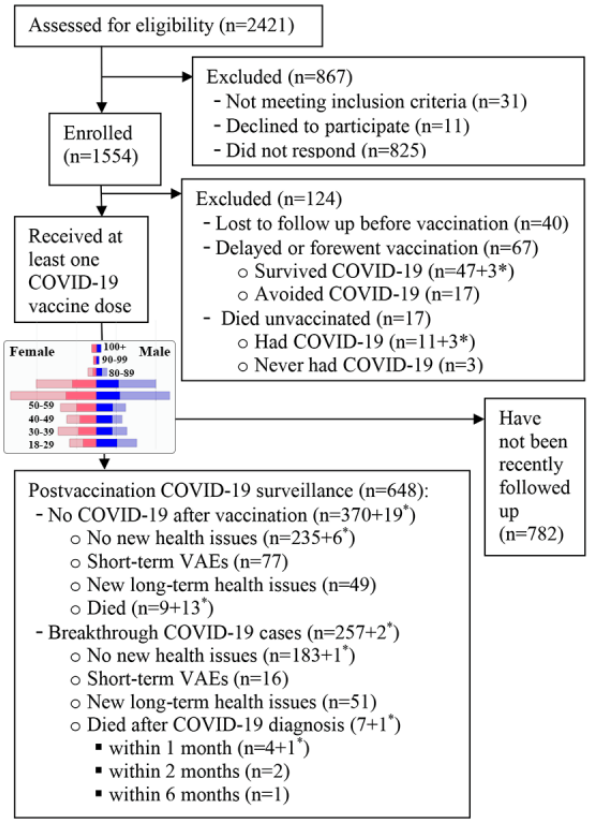


Figure 2. Distribution of participants by adverse event. A white background indicates no or minimal VAEs, while an orange background depicts participants who reported VAEs that prevented them from performing daily activities, diagonal stripes denote breakthrough COVID-19 infection, and solid black represents fatal outcomes. VAE: vaccine adverse event.

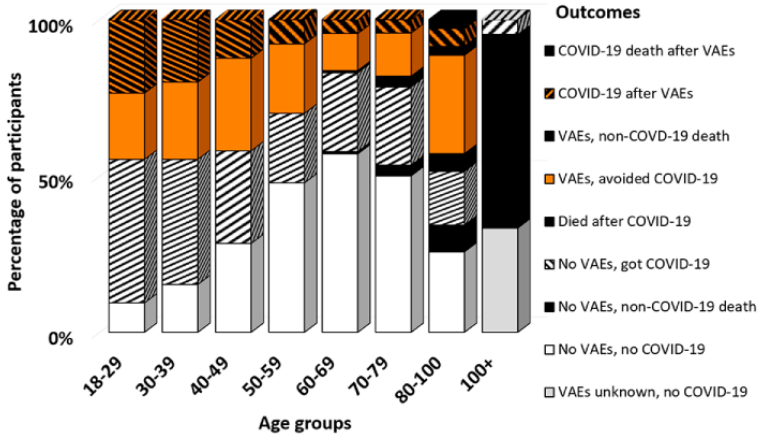
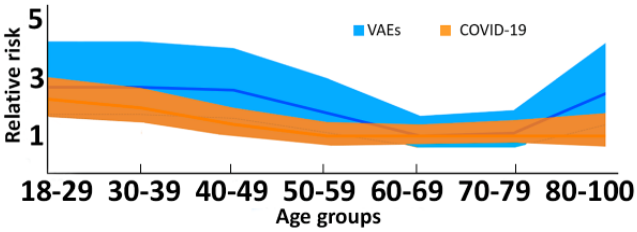


Figure 3. Area charts of CIs for relative risk of vaccine adverse events (VAEs; light blue) vs breakthrough infections (orange) shown for the 648 individuals with recently updated information. Relative risk is shown as a solid line (dark blue: VAEs; orange: COVID-19).



Discussion

Principal Findings

In summary, we found that adverse events following the COVID-19 immunization are likely to be influenced by a combination of demographic, genetic, and environmental factors. The principal findings of this study were (1) higher incidence of VAEs in both younger and the "oldest old" groups than in "younger old" populations, (2) association of VAEs with immunogenicity observed for short-term but not long-term adverse reactions, and (3) indications that disparities in host genetics and microbiomes in VAEs may exist.

Comparison With Prior Work

Age is a known factor contributing to reactogenicity. The heterogeneity is commonly addressed by splitting the sample in 2 groups with the cutoff age between 50 and 65 years [11]. Our study suggests that this simple assumption is not sufficient.

Reporting of VAEs is higher in younger and more educated individuals [12], but it does not appear to translate to higher rates of hospitalization or life-threatening events [13]. We observed a higher incidence of VAEs in both younger and older age groups than in the "young old" populations. We speculate that more adverse events are identified when younger and more educated individuals monitor the oldest participants. The oldest-old participants were more likely to either ignore the side effects or attribute them to aging; for example, duodenal bleeding, markedly decreased appetite [14], and transient amnesia [15] following immunization.

Systemic adverse reactions have been found to be associated with a higher antibody response in mostly healthy younger [16,17] and diseased populations [18,19]. We observed this association across all age groups, health conditions, and genders, but only with respect to short-term effects of vaccination. An impaired humoral immune response was observed after longer-lasting neurological side effects [20,21]. Lower antibody titers were also associated with depressive symptoms after the first and before the second dose of mRNA vaccines [22]. We suggest that more studies are needed on all types of serious and longer-lasting side effects of vaccination.

Host genetic factors are known to contribute to the severity of COVID-19 [23] and stronger short-term reactions to COVID-19 vaccines [24]. Genetic contributions are also being considered for several serious adverse reactions [25,26]. Our preliminary data support the contributions of genetic and microbiome to cardiological and respiratory VAEs. More comprehensive multiomic analyses are needed to draw definite conclusions.

Limitations

The primary limitation of this study was that the data were obtained from self-reports.

Conclusions

Our results demonstrate that vaccine adverse reactions in older populations can be easily overlooked. Long-term effects of vaccination in all age groups could outweigh the benefits of this preventive measure in some populations. More research is needed for genetic, epigenetic, metabolome-, and microbiome-associated risk factors of serious VAEs. The prohibitive cost of comprehensive studies [16-24] disproportionally affects underserved populations. Observational trials such as this study therefore represent an effective alternative prescreening strategy for multiomics research.

Acknowledgments

I thank all participants for helping others in their communities and contributing data for the study.

Data Availability

Raw deidentified data and case reports are available upon request to the author, subject to ethics approval.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1153 KB - formative_v6i11e41914_app1.pdf]

References

1. Beatty AL, Peyser ND, Butcher XE, Carton TW, Olgin JE, Pletcher MJ, et al. The COVID-19 Citizen Science study: protocol for a longitudinal digital health cohort study. JMIR Res Protoc 2021 Aug 30;10(8):e28169 [FREE Full text] [doi: 10.2196/28169] [Medline: 34310336]
2. Jarynowski A, Semenov A, Kamiński M, Belik V. Mild adverse events of Sputnik V vaccine in Russia: social media content analysis of Telegram via deep learning. J Med Internet Res 2021 Nov 29;23(11):e30529 [FREE Full text] [doi: 10.2196/30529] [Medline: 34662291]
3. Khademi Habibabadi S, Delir Haghighi P, Burstein F, Buttery J. Vaccine adverse event mining of Twitter conversations: 2-phase classification study. JMIR Med Inform 2022 Jul 16;10(6):e34305 [FREE Full text] [doi: 10.2196/34305] [Medline: 35708760]
4. Hause AM, Baggs J, Marquez P, Abara WE, Baumblatt J, Blanc PG, et al. Safety monitoring of COVID-19 mRNA vaccine second booster doses among adults aged ≥50 years - United States, March 29, 2022-July 10, 2022. MMWR Morb Mortal Wkly Rep 2022 Jul 29;71(30):971-976 [FREE Full text] [doi: 10.15585/mmwr.mm7130a4] [Medline: 35900925]
5. Gabashvili I. Community-based phenotypic study of safety, tolerability, reactogenicity and Immunogenicity of emergency-use-authorized vaccines against COVID-19 and viral shedding potential of post-vaccination infections: protocol for an ambispective study. medRxiv. Preprint posted online July 21, 2021. [FREE Full text] [doi: 10.1101/2021.06.28.21256779]
6. Gabashvili IS. Cutaneous bacteria in the gut microbiome as biomarkers of systemic malodor and People Are Allergic to Me (PATM) conditions: insights from a virtually conducted clinical trial. JMIR Dermatol 2020 Nov 4;3(1):e10508 [FREE Full text] [doi: 10.2196/10508]
7. List of oldest people vaccinated against COVID-19. Gerontology Wiki. URL: https://gerontology.fandom.com/wiki/List_of_oldest_people_vaccinated_against_COVID-19 [accessed 2021-10-01]
8. Stouten V, Hubin P, Haarhuis F, van Loenhout J, Billuart M, Brondeel R, et al. Incidence and risk factors of COVID-19 vaccine breakthrough infections: a prospective cohort study in Belgium. Viruses 2022 Apr 13;14(4):802 [FREE Full text] [doi: 10.3390/v14040802] [Medline: 35458532]
9. Antonelli M, Penfold RS, Merino J, Sudre CH, Molteni E, Berry S, et al. Risk factors and disease profile of post-vaccination SARS-CoV-2 infection in UK users of the COVID Symptom Study app: a prospective, community-based, nested, case-control study. Lancet Infect Dis 2022 Jan;22(1):43-55 [FREE Full text] [doi: 10.1016/S1473-3099(21)00460-6] [Medline: 34480857]
10. Nishimi K, Neylan TC, Bertenthal D, Seal KH, O'Donovan A. Association of psychiatric disorders with incidence of SARS-CoV-2 breakthrough infection among vaccinated adults. JAMA Netw Open 2022 Apr 01;5(4):e227287 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.7287] [Medline: 35420660]
11. Sutton N, San Francisco Ramos A, Beales E, Smith D, Ikram S, Galiza E, et al. Comparing reactogenicity of COVID-19 vaccines: a systematic review and meta-analysis. Expert Rev Vaccines 2022 Sep 15;21(9):1301-1318 [FREE Full text] [doi: 10.1080/14760584.2022.2098719] [Medline: 35796029]
12. Mangat HS, Musah A, Luedtke S, Syed AA, Maramattom BV, Maruthanal J, et al. Analyses of reported severe adverse events after immunization with SARS-CoV-2 vaccines in the United States: One year on. Front. Public Health 2022 Oct 13;10:972464 [FREE Full text] [doi: 10.3389/fpubh.2022.972464] [Medline: 36311588]
13. Hamilton E, Oversby S, Kitchener S, Ratsch A. Post COVID-19 vaccination: AusVaxSafety survey participation and adverse events - a community-based regional Queensland study. Aust N Z J Public Health 2022 Oct 03 [FREE Full text] [doi: 10.1111/1753-6405.13300] [Medline: 36190203]
14. Osozawa S. Case report: anorexia as a new type of adverse reaction caused by the COVID-19 vaccination: a case report applying detailed personal care records. F1000Res 2022 Jan 4;11:4 [FREE Full text] [doi: 10.12688/f1000research.75277.1]
15. Katsuki M, Higo Y, Komagata S, Kashiwagi K, Koh A. Transient global amnesia related to the third coronavirus disease-19 (COVID-19) vaccination. Cureus 2022 Jul;14(7):e27121 [FREE Full text] [doi: 10.7759/cureus.27121] [Medline: 36004026]
16. Uwamino Y, Kurafuji T, Sato Y, Tomita Y, Shibata A, Tanabe A, Keio Donner Project Team. Young age, female sex, and presence of systemic adverse reactions are associated with high post-vaccination antibody titer after two doses of BNT162b2 mRNA SARS-CoV-2 vaccination: An observational study of 646 Japanese healthcare workers and university staff. Vaccine 2022 Feb 11;40(7):1019-1025 [FREE Full text] [doi: 10.1016/j.vaccine.2022.01.002] [Medline: 35033389]
17. Kanizsai A, Molnar T, Varnai R, Zavori L, Tőkés-Füzesi M, Szalai Z, et al. Fever after vaccination against SARS-CoV-2 with mRNA-based vaccine associated with higher antibody levels during 6 months follow-up. Vaccines (Basel) 2022 Mar 14;10(3):447 [FREE Full text] [doi: 10.3390/vaccines10030447] [Medline: 35335080]
18. Braun E, Horowitz NA, Leiba R, Weissman A, Mekel M, Shachor-Meyouhas Y, et al. Association between IgG antibody levels and adverse events after first and second Bnt162b2 mRNA vaccine doses. Clin Microbiol Infect 2022 Jul 15:S1198-743X(22)00366-4 [FREE Full text] [doi: 10.1016/j.cmi.2022.07.002] [Medline: 35843565]
19. Lin T, Hung N, Hung S. Association of reactogenicity with immunogenicity of the ChAdOx1 nCoV-19 vaccine in patients undergoing hemodialysis. Vaccines (Basel) 2022 Aug 21;10(8):1366 [FREE Full text] [doi: 10.3390/vaccines10081366] [Medline: 36016253]

20. Kaulen LD, Doubrovinskaia S, Mooshage C, Jordan B, Purrucker J, Haubner C, et al. Neurological autoimmune diseases following vaccinations against SARS-CoV-2: a case series. Eur J Neurol 2022 Feb 31;29(2):555-563 [FREE Full text] [doi: 10.1111/ene.15147] [Medline: 34668274]
21. Cicalese MP, Ferrua F, Barzaghi F, Cerri F, Moro M, Aiuti A, et al. Third cranial nerve palsy in an 88-year-old man after SARS-CoV-2 mRNA vaccination: change of injection site and type of vaccine resulted in an uneventful second dose with humoral immune response. BMJ Case Rep 2022 Mar 08;15(2):e246485 [FREE Full text] [doi: 10.1136/bcr-2021-246485] [Medline: 35135792]
22. Kaneko H, Tsuboi H. Depressive symptoms predict antibody titers after a second dose of the SARS-CoV-2 BNT162b2 vaccine among hospital workers in Japan. Brain Behav Immun 2022 Oct 15:S0889-1591(22)00379-8 [FREE Full text] [doi: 10.1016/j.bbi.2022.09.004] [Medline: 36116693]
23. Niemi MEK, Daly MJ, Ganna A. The human genetic epidemiology of COVID-19. Nat Rev Genet 2022 Sep 02;23(9):533-546 [FREE Full text] [doi: 10.1038/s41576-022-00478-5] [Medline: 35501396]
24. Bolze A, Neveux I, Schiabor Barrett KM, White S, Isaksson M, Dabe S, et al. HLA-A∗03:01 is associated with increased risk of fever, chills, and stronger side effects from Pfizer-BioNTech COVID-19 vaccination. HGG Adv 2022 May 14;3(2):100084 [FREE Full text] [doi: 10.1016/j.xhgg.2021.100084] [Medline: 35005651]
25. Salter B, Jessome M, Tarnopolsky M, Yousuf H. Possible association between rhabdomyolysis and mRNA SARS-CoV-2 vaccination in a patient with gene mutation. CMAJ 2022 Feb 22;194(7):E252-E256 [FREE Full text] [doi: 10.1503/cmaj.211856] [Medline: 35193861]
26. Nishibayashi H, Kishimoto S, Sekiguchi K, Okada S, Ihara K. Myocarditis in 13-year-old monochorionic diamniotic twins after COVID-19 vaccination. J Clin Immunol 2022 Aug 26:1-3 [FREE Full text] [doi: 10.1007/s10875-022-01360-z] [Medline: 36008643]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
MEBO: metabolic body odor
VAE: vaccine adverse event

Edited by A Mavragani; submitted 22.08.22; peer-reviewed by H Zhai, O Navarro; comments to author 05.10.22; revised version received 13.10.22; accepted 29.10.22; published 04.11.22.

Please cite as:
Gabashvili IS
The Incidence and Effect of Adverse Events Due to COVID-19 Vaccines on Breakthrough Infections: Decentralized Observational Study With Underrepresented Groups
JMIR Form Res 2022;6(11):e41914
URL: https://formative.jmir.org/2022/11/e41914
doi:10.2196/41914
PMID:36309347

Original Paper

Risk Assessment of COVID-19 Cases in Emergency Departments and Clinics With the Use of Real-World Data and Artificial Intelligence: Observational Study

Evangelos Logaras[1], MSc; Antonis Billis[1], PhD; Ilias Kyparissidis Kokkinidis[1], MSc; Smaranda Nafsika Ketseridou[1], MSc; Alexios Fourlis[2], MSc; Aristotelis Tzotzis[3], MD; Konstantinos Imprialos[3], MD; Michael Doumas[3], MD, PhD; Panagiotis Bamidis[1], PhD

[1]Laboratory of Medical Physics and Digital Innovation, School of Medicine, Aristotle University of Thessaloniki, Thessaloniki, Greece

[2]Department of Research and Development, Vidavo S.A., Thessaloniki, Greece

[3]2nd Propaedeutic Department of Internal Medicine, School of Medicine, Aristotle University of Thessaloniki, Thessaloniki, Greece

Corresponding Author:
Antonis Billis, PhD
Laboratory of Medical Physics and Digital Innovation
School of Medicine
Aristotle University of Thessaloniki
University Campus, Building D, Gate 8
Thessaloniki, 54124
Greece
Phone: 30 2310999237
Email: ampillis@med.auth.gr

Abstract

Background: The recent COVID-19 pandemic has highlighted the weaknesses of health care systems around the world. In the effort to improve the monitoring of cases admitted to emergency departments, it has become increasingly necessary to adopt new innovative technological solutions in clinical practice. Currently, the continuous monitoring of vital signs is only performed in patients admitted to the intensive care unit.

Objective: The study aimed to develop a smart system that will dynamically prioritize patients through the continuous monitoring of vital signs using a wearable biosensor device and recording of meaningful clinical records and estimate the likelihood of deterioration of each case using artificial intelligence models.

Methods: The data for the study were collected from the emergency department and COVID-19 inpatient unit of the Hippokration General Hospital of Thessaloniki. The study was carried out in the framework of the COVID-X H2020 project, which was funded by the European Union. For the training of the neural network, data collection was performed from COVID-19 cases hospitalized in the respective unit. A wearable biosensor device was placed on the wrist of each patient, which recorded the primary characteristics of the visual signal related to breathing assessment.

Results: A total of 157 adult patients diagnosed with COVID-19 were recruited. Lasso penalty function was used for selecting 18 out of 48 predictors and 2 random forest–based models were implemented for comparison. The high overall performance was maintained, if not improved, by feature selection, with random forest achieving accuracies of 80.9% and 82.1% when trained using all predictors and a subset of them, respectively. Preliminary results, although affected by pandemic limitations and restrictions, were promising regarding breathing pattern recognition.

Conclusions: This study represents a novel approach that involves the use of machine learning methods and Edge artificial intelligence to assist the prioritization and continuous monitoring procedures of patients with COVID-19 in health departments. Although initial results appear to be promising, further studies are required to examine its actual effectiveness.

(JMIR Form Res 2022;6(11):e36933) doi:10.2196/36933

KEYWORDS

COVID-19 pandemic; risk assessment; wearable device; respiration evaluation; emergency department; artificial intelligence; real-world data

Introduction

Background

The recent COVID-19 pandemic has highlighted the weaknesses of health care systems around the world and has put and continues to put great pressure on public health organizations and, in particular, reference hospitals [1]. As the number of clinical cases increases, the COVID-19–related knowledge base becomes increasingly detailed. The vast majority of patients has a good outcome. However, since there are few effective drugs for COVID-19 treatment, some patients with COVID-19 get worse as a result of progressive pneumonia, severe dyspnea, or multiple organ failure and some of them die [2-4]; therefore, the strain on hospitals continues to be a major concern. With a view to improving the monitoring of cases admitted to emergency departments (EDs) or treated in the corresponding COVID-19 wards, it has become increasingly necessary to adopt new innovative technological solutions in clinical practice. This necessity is greatly reinforced in view of the chronic understaffing of health care wards in Greece.

The application of machine learning algorithms has been studied in a variety of fields related to COVID-19 including, but not limited to, the detection of outbreaks, the identification and classification of COVID-19 cases based on medical images, rapid diagnosis, mortality risk prediction, intensive care transfers prediction, and others [5-9]. Available machine learning techniques have promising prognostic added value. Fever and breathlessness (difficulty in breathing) are considered 2 of the main symptoms of patients infected with SARS-CoV-2 [2]; both symptoms can be objectively detected by monitoring vital signs, such as body temperature and respiratory rate or blood oxygen saturation levels, respectively. Currently, the continuous monitoring of vital signs is only performed in patients admitted to the intensive care unit (ICU). For COVID-19 cases in EDs or inpatient units, only sporadic checks are performed, depending on the availability of health care staff, which is usually limited and under enormous pressure.

Using the demographics and medical history of individuals who have tested positive for SARS-CoV-2 in combination with the continuous monitoring of inpatient cases via a wearable biosensor device [10], it is possible to overcome these issues. By considering the more detailed information provided at an individual level, as opposed to macroscopically variable information at a population level, it may allow for higher accuracy. We hypothesize that by using the continuous monitoring of important physiological variables to predict the progress and severity of each case, we can get more precise forecasting results, which are critical for optimized resource management in health care facilities.

Objectives

The aim of this study was to develop a clinical decision support system that will (1) dynamically prioritize patients through real-world data and the continuous monitoring of vital signs using a wearable biosensor device [10] and (2) estimate the likelihood of deterioration of each case using artificial intelligence (AI) models. More specifically, 2 discrete tools based on AI techniques will be developed: (1) a predictive model to estimate the probability of admission of a COVID-19 case from the ED to an inpatient unit and (2) a neural network (NN) for the respiration evaluation of inpatients in COVID-19 units.

Methods

Study Environment

The data for the study was collected from the ED and COVID-19 inpatient unit of the 2nd Propaedeutic Department of Internal Medicine, General Hippokration Hospital, Thessaloniki, Greece.

Timeline

The study took place over a period of 4 months, from September 2021 to January 2022.

Framework

The study was carried out in the framework of the COVID-X [11] H2020 project (101016065), which was funded by the European Union.

Participants

In this study, participants were confirmed COVID-19 cases from the ED and COVID-19 inpatient unit, aged ≥18 years, and referred by a study partner physician, provided that they have signed the consent form authorizing the collection and processing of the data collected for the purpose of the study. Regarding the exclusion criteria, cases with a known chronic disease and a very short survival expectancy, which may adversely affect the purpose and aim of the study, were avoided.

Data Collection Plan

Overview

Data collection (clinical data, screening data, and vital sign recordings) was performed in the context of an observational cohort study, during which data were observed and collected randomly for selected COVID-19 cases in the ED or those hospitalized in a corresponding unit.

Data Collection From the COVID-19 ED

Upon entry of a confirmed COVID-19 incident into the COVID-19 ED, information relating to medical and social history, demographics, and patient screening data were recorded using a smartphone app. At the same time, reference values of vital signs were obtained, as well as test results used to assess the status of COVID-19 cases, based on the applicable protocols and guidelines.

Patients with COVID-19 that were indicated by a collaborating physician received a wearable biosensor device on their wrist, which continuously recorded heart rate, blood oxygen saturation, and skin temperature. The use of the wearable devices took place until the outcome of each incident—that is, discharge and home monitoring or admission to a COVID-19 inpatient unit.

Data Collection From Inpatient Units

For the training of the NN, data collection was performed from COVID-19 cases hospitalized in a respective COVID-19 inpatient unit. Cases were selected by a collaborating physician. A wearable biosensor device was placed on the wrist of each patient, which recorded the primary characteristics of the visual

signal related to breathing. At the same time, a certified medical device was also placed on the patient's other hand to record the characteristics of breathing, which during the analysis process were used as a ground truth measurement to train the NN. As the wearable biosensor device had a certain battery life, in cases of extended monitoring time, more than 1 wearable device were used, which were sequentially replaced.

Provisions for the Ethical Conduct of the Study

Study participants were informed during the consent process regarding their rights over the collected data according to General Data Protection Regulation and Greek law (4624/2019). Each participant provided written consent with full knowledge of the procedures involved. Consent forms and procedures were fully explained by the investigator or a member of the study staff, including the study aims, methods, benefits, and risks, and signed by the subject before enrollment into the study. Prospective participants were informed that study participation is voluntary; they may withdraw at any time; choosing against participation will not affect the care received for treatment; and they have sufficient time to read the study information and consent form and ask any questions. Once the informed consent was signed, the participant received a copy of the document.

Ethics Approval

This study was submitted and approved by the Research Ethics and Conduct Committee of the Aristotle University of Thessaloniki (219970/2021).

Data Management

All research data collected during this study are stored and encrypted, in pseudonymized form, in dedicated, password-protected computers of the Laboratory of Medical Physics and Digital Innovation of Aristotle University of Thessaloniki, with restricted access. Study hard-copy documents, including signed informed consent documents, are kept in locked cabinets with restricted access. All data will be kept for 5 years and deleted afterward.

Access to Data

Only investigators from the 2nd Propaedeutic Department of Internal Medicine of the Hippokration Hospital of Thessaloniki and the Laboratory of Medical Physics and Digital Innovation of Aristotle University of Thessaloniki who are authorized by the principal investigator have access to the acquired, deidentified data sets to ensure participant confidentiality. Data transfer over a computer network will take place through safe processes, and data will be stored in secure digital structures with limited access.

Study participants have access to the data they have contributed to the study, in compliance with the respective mandate of the General Data Protection Regulation. Participants are also able to request for their data to be deleted. Instructions for requesting, receiving, or deleting data were included in the study information sheet.

Data Analysis

Prediction of Admission to an Inpatient Unit

As a first step, a time series analysis of vital signs collected for COVID-19 cases via the wearable biosensor device (heart rate, blood oxygen saturation, and skin temperature) was performed. Through the analysis, an attempt was made to extract features concerning trends and statistically significant changes in the time series.

Next, machine learning algorithms, such as support vector machine (SVM), random forest (RF), and logistic regression (LR), etc, were studied, which was used for the training of the data set collected in the study. The outcome of each incident—0=discharge and home follow-up and 1=admission to a COVID-19 inpatient unit—was used as the predictor variable. The handling of missing values included k-nearest neighbor imputation. Regarding the training process, standard procedure was followed. Model evaluation was performed based on 2-fold cross validation. Initially, feature importance was computed based on RF. The *DataSynthesizer* [12] tool was exploited for increasing the initial training sample. Finally, the algorithm with the highest prediction accuracy was selected for use in the predictive model and was further studied along with dimensionality reduction based on Lasso penalty function.

Neural Network for Respiration Evaluation

An important innovation of the proposed solution is the use of AI techniques for the respiration evaluation of patients hospitalized in COVID-19 units. More specifically, a machine learning model using NNs was designed and developed to be executed locally on the wearable electronic biosensor device (wearable). The developed model is able to classify the breathing of patients with COVID-19, with minimal energy consumption, as no connection to the internet or third-party systems is required at any stage of data processing. The NN was trained to recognize 3 basic patterns regarding breathing: (1) normal breathing (12-18 breaths per minute), (2) tachypnea (more than 18 breaths per minute), and (3) bradypnea (fewer than 10 breaths per minute).

For the whole toolchain of creating the machine learning model optimized for embedded devices, the Edge Impulse platform [13] and, more specifically, the EON Tuner tool was used. The tool analyzed the inputted data, the signal processing blocks, and the available NN architectures and outputted an overview of possible model architectures that were able to fit the chosen device's latency and memory requirements. For the raw photoplethysmography data acquisition and ingestion, a firmware application that exploits the open-source Edge Impulse API [14] was developed.

Results

The aforementioned evaluation process regarding the predictive algorithms involved a total of 157 registered patients that met the inclusion criteria (see Participants). Overall, 70 (44.6%) out of the 157 COVID-19 cases were admitted to an inpatient unit. The data collection questionnaire included the 48 variables presented in Table 1. The RF algorithm outperformed SVM and LR in predicting hospitalization outcome according to all the

metrics listed in Table 2; thus, it was selected for further evaluation using the feature selection method. The subset of predictors that was selected based on Lasso penalty function included 18 variables, namely age; smoking history; systolic blood pressure; oxygen saturation; heart rate; arrival category (ambulance, ambulatory, and other means); radiographic and computed tomographic findings; estimated glomerular filtration rate; and some chief complaints (pharyngalgia, general weakness, headache, fever, vomiting, cough, shortness of breath, obesity, and hypertension).

The results listed in Table 2 indicate that the RF algorithm might be able to satisfactorily predict which COVID-19 cases require hospitalization. Comparing the results before and after feature selection, there was a slight improvement in most of the metrics, namely accuracy, recall, F_1-score, receiver operating characteristic, and precision-recall area under the curve. Even though these differences are not quite significant, feature selection method might potentially benefit in terms of the required computational cost in case of an increase in the set of training data.

Regarding the respiration evaluation model development, due to COVID-19 pandemic limitations and restrictions (eg, the prohibition of entry to nonmedical personnel in COVID-19 inpatient clinics and ED, burdened clinical staff, and emerged shortages due to suspensions, etc), for the training process of the NN alongside with the collected data, an open-source photoplethysmography data set [15] was used additionally to enrich the training data set, which consisted of 53 recordings, each of 8-minute duration.

The confusion matrix of the selected trained model is presented in Table 3. The preliminary results, although based on small data set, seem promising. The F_1-score on the tachypnea label revealed that the accuracy on real-world conditions was lower for that label. The main issue is the small size of tachypnea data samples at the current point, since tachypnea is rarer than bradypnea in patients with COVID-19.

Table 1. Patient cohort characteristics.

Variable	Value (N=157)
Sex, male, n (%)	88 (56.1)
Age (years), mean (SD)	52.8 (19.6)
Systolic blood pressure, mean (SD)	140.4 (24.1)
Diastolic blood pressure, mean (SD)	82.1 (15)
Temperature (°C), mean (SD)	36.4 (3)
Blood oxygen saturation (SpO2; %), mean (SD)	94.6 (5.8)
Heart rate, mean (SD)	87.5 (18)
Estimated glomerular filtration rate, mean (SD)	80.8 (26.9)
C-reactive protein, mean (SD)	53.4 (69.7)
D-dimers, mean (SD)	1196.9 (3124.8)
History of alcohol use, n (%)	49 (31.2)
History of nicotine dependence, n (%)	64 (40.8)
Polypharmacy, n (%)	50 (31.8)
Radiographic findings, mean (SD)	1.8 (2)
Computed tomographic findings, mean (SD)	2.2 (1)
Generalized abdominal pain, n (%)	24 (15.3)
Shiver, n (%)	12 (7.6)
Fever, n (%)	93 (59.2)
Generalized fatigue, n (%)	41 (26.1)
Dyspnea, n (%)	52 (33.1)
Cough, n (%)	88 (56.1)
Anosmia, n (%)	14 (8.9)
Ageusia, n (%)	16 (10.2)
Cephalalgia, n (%)	25 (15.9)
Catarrh, n (%)	13 (8.3)
Nausea, n (%)	15 (9.6)
Emesis, n (%)	13 (8.3)
Diarrhea, n (%)	20 (12.7)
Dysphagia, n (%)	7 (4.5)
Pharyngalgia, n (%)	23 (14.6)
Thoracic pain, n (%)	26 (16.6)
Abdominal pain, n (%)	13 (8.3)
Type 1 diabetes, n (%)	7 (4.5)
Type 2 diabetes, n (%)	12 (7.6)
Hypertension, n (%)	45 (28.7)
Heart failure, n (%)	7 (4.5)
Obesity, n (%)	11 (7)
Chronic obstructive pulmonary disease, n (%)	2 (1.3)
Bronchial asthma, n (%)	10 (6.4)
Coronary heart disease, n (%)	3 (1.9)
Ischemic stroke, n (%)	4 (2.5)
Arrhythmia, n (%)	8 (5.1)

Variable	Value (N=157)
Dyslipidemia, n (%)	13 (8.3)
Chronic kidney disease, n (%)	2 (1.3)
Cancer, n (%)	6 (3.8)
Acute liver failure, n (%)	1 (0.6)
Immunodeficiency, n (%)	3 (1.9)
Rheumatic diseases, n (%)	4 (2.5)

Table 2. Random forest evaluation metrics.

Metric (%)	All predictors			Selected predictors
	Random forest	Support vector machine	Logistic regression	Random forest
Accuracy	81	66	69	82
Recall	76	43	57	81
Specificity	85	89	81	83
Precision	80	79	74	79
F_1-score	87	66	69	88
ROC-AUC[a]	78	75	76	80
Precision-recall AUC[b]	81	66	69	82

[a]ROC-AUC: area under the receiver operating characteristic curve.

[b]AUC: area under the curve.

Table 3. Confusion matrix of the trained model.

	Actual case	
	Bradypnea	Tachypnea
Predicted case (%)		
Bradypnea	99.8	0.2
Tachypnea	27	73
F_1-score	0.98	0.84

Discussion

Principal Findings

In this study, we examined the effectiveness of a smart system that is able to (1) dynamically prioritize patients diagnosed with COVID-19 through the continuous monitoring of vital signs using a wearable biosensor device and recording of meaningful clinical records and (2) estimate the likelihood of deterioration of each case using AI models. The results showed that the developed predictive model appears to be able to function as an auxiliary tool, in conjunction with the statistical analysis of vital signs, offering an initial indication of the patient's risk. By combining the predictive model with the continuous monitoring of the vital signs and augmenting the wearable with a NN classifier for breathing assessment to serve as a relapse detector, this integrated system can potentially reduce the time to manage patients at greatest risk.

Comparison to Prior Work

Several attempts to develop solutions that accurately predict deterioration among patients with COVID-19 at various levels have been made since the beginning of the pandemic. An observational cohort study developed in Brazil [16] attempted to evaluate the risk of severe forms of COVID-19, based on clinical, laboratory, and imaging markers in patients initially admitted to the ward. The data they used were acquired from the electronic medical records of inpatients, with laboratory confirmation of COVID-19. They concluded that the presence of some laboratory markers, clinical criteria, and findings in imaging exams as elucidated in the study may have a significant relationship with the patient's evolution to the ICU. In Singh et al [5], a proprietary prediction model was used as an assessment of the Epic Deterioration Index (EDI), which is used in over 100 US hospitals to aid medical decision-making. A composite outcome of intensive care unit–level treatment, mechanical ventilation, or in-hospital mortality was predicted using EDI. They demonstrated that the EDI identifies small subgroups of individuals with high-risk and low-risk COVID-19

with strong discrimination, but its clinical application as an early warning system is restricted by its poor sensitivity.

A previously disclosed algorithm, Predicting Intensive Care Transfers and Other Unforeseen Events [17], which was originally designed to predict patient deterioration in general wards, was retrained to target the outcomes thought to be the most relevant to COVID-19, such as ICU level of care, mechanical ventilation, and death. The method was then tested on a group of patients with COVID-19. The algorithm was capable of accurately anticipating adverse events up to 24 hours ahead of time. They also compared the model to the aforementioned EDI and were able to show that using a head-to-head comparison, their model outperformed the EDI by a statistically significant margin. Another approach [6] looked into the possibility that machine learning could be used to provide reliable COVID-19 patient outcome predictions by combining data points from several sources, such as the electronic health record. The incorporation of electronic health record repositories in a combined model could improve risk prediction and disease driver identification at specific times of the year. In this scenario, they included diagnosed individuals outside of hospitals, including the complete SARS-CoV-2 spectrum. Their findings showed that by concentrating on a small number of demographic characteristics, such as age, gender, and BMI, it is feasible to predict the likelihood of hospital and ICU admission, mechanical ventilation usage, and mortality as early as the time of diagnosis.

Finally, on a similar note, in Gao et al [18], the authors developed a Mortality Risk Prediction Model for COVID-19 based on more complicated clinical data points from the admission of 2520 consecutive patients with COVID-19 with known outcomes to stratify patients by mortality risk. Their model was created using 4 machine learning methods: LR, SVM, gradient boosted decision tree, and NN. They demonstrate that the ensemble model can forecast physiological decline and mortality up to 20 days in advance. These findings provide support for the use of machine learning algorithms to create a real-time, predictive clinical support system that could assist in the monitoring and prediction of potential adverse outcomes based on previously identified risk factors and patient demographics.

In this study, the key differentiator—and possibly, the improvement—in patient screening, monitoring, and management processes is the use of a low-cost and energy-efficient wearable biosensor device. As previously mentioned, fever and dyspnea are considered 2 of the main symptoms of COVID-19 patients, and these 2 symptoms can be easily and objectively detected using a wearable biosensor device.

By aiming to improve the monitoring of COVID-19 cases that enter the EDs or that are hospitalized in the corresponding units, it becomes necessary to adopt new innovative technological solutions in clinical practice, such as the use of wearable devices. This need is substantially strengthened considering the chronic understaffing of health units in many countries around the world.

Limitations

The limitations of this study, as mentioned before, are mainly related to the difficulty in collecting data due to pandemic restrictions (eg, the prohibition of entry to nonmedical personnel in health units, burdened clinical staff, and shortages due to suspensions, etc).

Conclusions

This study offered the required data to guarantee that existing wearable initiatives could have great impact and efficiency in meeting pressing demands among COVID-19 response units. Moreover, the study will guide the next steps in our research to customize and test NN algorithms and predictive machine learning algorithms for the best accuracy and reach among caregivers in EDs and COVID-19 units. More precise and accurate estimates of each case's chance of deterioration, as well as remote monitoring and the evaluation of respiration, will allow caregivers to timely respond. This improvement would also help minimize the danger to frontline health care staff working in EDs and COVID-19 units by potentially decreasing unnecessary traffic in these facilities. Furthermore, it will serve as additional evidence of how such wearable devices may be used to improve the efficiency of patient monitoring in general.

Acknowledgments

COVID-X received funding from the European Union's Horizon 2020 research and innovation program (grant 101016065).

Data Availability

The data sets generated during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

References

1. Phelan AL, Katz R, Gostin LO. The novel coronavirus originating in Wuhan, China: challenges for global health governance. JAMA 2020 Feb 25;323(8):709-710 [FREE Full text] [doi: 10.1001/jama.2020.1097] [Medline: 31999307]

2. Islam MA, Kundu S, Alam SS, Hossan T, Kamal MA, Hassan R. Prevalence and characteristics of fever in adult and paediatric patients with coronavirus disease 2019 (COVID-19): a systematic review and meta-analysis of 17515 patients. PLoS One 2021 Apr 06;16(4):e0249788 [FREE Full text] [doi: 10.1371/journal.pone.0249788] [Medline: 33822812]
3. Lai CC, Shih TP, Ko WC, Tang HJ, Hsueh PR. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): the epidemic and the challenges. Int J Antimicrob Agents 2020 Mar;55(3):105924 [FREE Full text] [doi: 10.1016/j.ijantimicag.2020.105924] [Medline: 32081636]
4. Xie J, Tong Z, Guan X, Du B, Qiu H. Clinical characteristics of patients who died of coronavirus disease 2019 in China. JAMA Netw Open 2020 Apr 01;3(4):e205619 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.5619] [Medline: 32275319]
5. Singh K, Valley TS, Tang S, Li BY, Kamran F, Sjoding MW, et al. Evaluating a widely implemented proprietary deterioration index model among hospitalized patients with COVID-19. Ann Am Thorac Soc 2021 Jul;18(7):1129-1137 [FREE Full text] [doi: 10.1513/AnnalsATS.202006-698OC] [Medline: 33357088]
6. Jimenez-Solem E, Petersen TS, Hansen C, Hansen C, Lioma C, Igel C, et al. Developing and validating COVID-19 adverse outcome risk prediction models from a bi-national European cohort of 5594 patients. Sci Rep 2021 Feb 05;11(1):3246 [FREE Full text] [doi: 10.1038/s41598-021-81844-x] [Medline: 33547335]
7. Cummings BC, Ansari S, Motyka JR, Wang G, Medlin RP, Kronick SL, et al. Predicting intensive care transfers and other unforeseen events: analytic model validation study and comparison to existing methods. JMIR Med Inform 2021 Apr 21;9(4):e25066 [FREE Full text] [doi: 10.2196/25066] [Medline: 33818393]
8. Cho A. AI systems aim to sniff out coronavirus outbreaks. Science 2020 May 22;368(6493):810-811. [doi: 10.1126/science.368.6493.810] [Medline: 32439771]
9. Mei X, Lee HC, Diao KY, Huang M, Lin B, Liu C, et al. Artificial intelligence-enabled rapid diagnosis of patients with COVID-19. Nat Med 2020 Aug 19;26(8):1224-1228 [FREE Full text] [doi: 10.1038/s41591-020-0931-3] [Medline: 32427924]
10. Kamišalić A, Fister I, Turkanović M, Karakatič S. Sensors and functionalities of non-Invasive wrist-wearable devices: a review. Sensors (Basel) 2018 May 25;18(6):1714 [FREE Full text] [doi: 10.3390/s18061714] [Medline: 29799504]
11. Covid X. URL: https://www.covid-x.eu/ [accessed 2022-10-14]
12. Ping H, Stoyanovich J, Howe B. DataSynthesizer: Privacy-Preserving Synthetic Datasets. In: ACM Digital Library. 2017 Jun 27 Presented at: SSDBM '17: Proceedings of the 29th International Conference on Scientific and Statistical Database Management. [doi: 10.1145/3085504.3091117]
13. We put ML into real products: Edge Impulse is the leading development platform for machine learning on edge devices, free for developers and trusted by enterprises. Edge Impulse. URL: https://www.edgeimpulse.com/ [accessed 2022-08-31]
14. alexisicte / aviateNNsprint2. GitHub. URL: https://github.com/alexisicte/aviateNNsprint2/tree/master/respiration_evaluation_NN [accessed 2022-08-31]
15. Pimentel MAF, Johnson AEW, Charlton PH, Birrenkott D, Watkinson PJ, Tarassenko L, et al. Toward a robust estimation of respiratory rate from pulse oximeters. IEEE Trans Biomed Eng 2017 Aug;64(8):1914-1923. [doi: 10.1109/tbme.2016.2613124]
16. Costa VO, Nicolini EM, da Costa BMA, Teixeira FM, Ferreira JP, Moura MA, et al. Evaluation of the risk of clinical deterioration among inpatients with COVID-19. Adv Virol 2021 Jun 25;2021:6689669 [FREE Full text] [doi: 10.1155/2021/6689669] [Medline: 34257657]
17. Gillies CE, Taylor DF, Cummings BC, Ansari S, Islim F, Kronick SL, et al. Demonstrating the consequences of learning missingness patterns in early warning systems for preventative health care: a novel simulation and solution. J Biomed Inform 2020 Oct;110:103528 [FREE Full text] [doi: 10.1016/j.jbi.2020.103528] [Medline: 32795506]
18. Gao Y, Cai G, Fang W, Li H, Wang S, Chen L, et al. Machine learning based early warning system enables accurate mortality risk prediction for COVID-19. Nat Commun 2020 Oct 06;11(1):5033 [FREE Full text] [doi: 10.1038/s41467-020-18684-2] [Medline: 33024092]

Abbreviations

AI: Artificial Intelligence
ED: Emergency Department
EDI: Epic Deterioration Index
ICU: Intensive Care Unit
LR: Logistic Regression
NN: Neural Network
RF: Random Forest
SVM: Support Vector Machine

Edited by T Leung; submitted 31.01.22; peer-reviewed by Z Butt, B Mesko; comments to author 20.07.22; revised version received 31.08.22; accepted 12.09.22; published 08.11.22.

Please cite as:
Logaras E, Billis A, Kyparissidis Kokkinidis I, Ketseridou SN, Fourlis A, Tzotzis A, Imprialos K, Doumas M, Bamidis P
Risk Assessment of COVID-19 Cases in Emergency Departments and Clinics With the Use of Real-World Data and Artificial Intelligence: Observational Study
JMIR Form Res 2022;6(11):e36933
URL: https://formative.jmir.org/2022/11/e36933
doi:10.2196/36933
PMID:36197836

Original Paper

Stressors and Destressors in Working From Home Based on Context and Physiology From Self-Reports and Smartwatch Measurements: International Observational Study Trial

Danielle Tump[1], BSc, MSc; Nitin Narayan[1], BE, MSc; Vera Verbiest[1], BSc, MSc; Sander Hermsen[1], BSc, MSc, PhD; Annelies Goris[1], BSc, MSc, PhD; Chui-De Chiu[2], BSc, MSc, PhD; Ruud Van Stiphout[1], BSc, MSc, PhD

[1]OnePlanet Research Center, Stichting imec Nederland, Wageningen, Netherlands

[2]Department of Psychology, The Chinese University of Hong Kong, Hong Kong, Hong Kong

Corresponding Author:
Danielle Tump, BSc, MSc
OnePlanet Research Center
Stichting imec Nederland
Plus Ultra 2
Bronland 10
Wageningen, 6708 WE
Netherlands
Phone: 31 611231567
Email: danielle.tump@imec.nl

Abstract

Background: The COVID-19 pandemic has greatly boosted working from home as a way of working, which is likely to continue for most companies in the future, either in fully remote or in hybrid form. To manage stress levels in employees working from home, insights into the stressors and destressors in a home office first need to be studied.

Objective: We present an international remote study with employees working from home by making use of state-of-the-art technology (ie, smartwatches and questionnaires through smartphones) first to determine stressors and destressors in people working from home and second to identify smartwatch measurements that could represent these stressors and destressors.

Methods: Employees working from home from 3 regions of the world (the United States, the United Kingdom, and Hong Kong) were asked to wear a smartwatch continuously for 7 days and fill in 5 questionnaires each day and 2 additional questionnaires before and after the measurement week. The entire study was conducted remotely. Univariate statistical analyses comparing variable distributions between low and high stress levels were followed by multivariate analysis using logistic regression, considering multicollinearity by using variance inflation factor (VIF) filtering.

Results: A total of 202 people participated, with 198 (98%) participants finishing the experiment. Stressors found were other people and daily life getting in the way of work (P=.05), job intensity (P=.01), a history of burnout (P=.03), anxiety toward the pandemic (P=.04), and environmental noise (P=.01). Destressors found were access to sunlight (P=.02) and fresh air (P<.001) during the workday and going outdoors (P<.001), taking breaks (P<.001), exercising (P<.001), and having social interactions (P<.001). The smartwatch measurements positively related to stress were the number of active intensity periods (P<.001), the number of highly active intensity periods (P=.04), steps (P<.001), and the SD in the heart rate (HR; P<.001). In a multivariate setting, only a history of burnout (P<.001) and family and daily life getting in the way of work (P<.001) were positively associated with stress, while self-reports of social activities (P<.001) and going outdoors (P=.03) were negatively associated with stress. Stress prediction models based on questionnaire data had a similar performance (F_1=0.51) compared to models based on automatic measurable data alone (F_1=0.47).

Conclusions: The results show that there are stressors and destressors when working from home that should be considered when managing stress in employees. Some of these stressors and destressors are (in)directly measurable with unobtrusive sensors, and prediction models based on these data show promising results for the future of automatic stress detection and management.

Trial Registration: Netherlands Trial Register NL9378; https://trialsearch.who.int/Trial2.aspx?TrialID=NL9378

(JMIR Form Res 2022;6(11):e38562) doi:10.2196/38562

KEYWORDS

stress; telework; wearables; COVID-19; pandemic; remote working; employees; stressors; destressors; remote work; mental health; psychological health; smartphone; digital questionnaire; stress management; occupational health; stress detection; prediction model

Introduction

In March 2020, the World Health Organization declared the outbreak of COVID-19 a global pandemic, which lasted until the beginning of 2022 [1]. It forced people to work from home full-time to minimize the spread of the COVID-19 virus. Employees had to quickly adapt their way of working and their personal lives, as not only offices but also schools, sport facilities, and day cares were closed. This new way of working, however, was associated with social deprivation, lack of exercise, home confinement, and additional stress from managing work while taking care of children [2]. Even though some companies have returned to the office full-time, many companies continue to offer hybrid or completely remote work after the pandemic [3-5].

Prior to the pandemic, research into the link between stress and remote employment had yielded conflicting results, with small effect sizes [6-9]. Known mediators of remote work–induced stress were job autonomy, work-life conflict, and work-life balance [7,8]. Working remotely during the pandemic, however, differs radically from prior remote work arrangements, in that it is involuntary; independent of employees' wishes, personal characteristics, and circumstances; independent of organizational culture; and coinciding with the staying at home of other members of the household. All this means that we can expect previously reported small effects of stressors and destressors (ie, factors reducing stress) on stress levels when working from home to be enhanced. Indeed, such results are currently starting to emerge in published literature, with findings that sometimes indicate increased stress responses [10-12] and sometimes reduced ones [13]; similarly, productivity is sometimes reduced [13,14] and sometimes increased [10]. This suggests a greater impact of stressors and destressors on stress levels when working from home during the pandemic.

When stress is prolonged, it can lead to mental and physical health issues, such as depression and burnout [15]. Burnout and even short-term stress lead to a decrease in overall performance and increased amounts of sick leave, costing companies billions of dollars every year [16]. The new hybrid working environment asks for accessible and dedicated solutions to help employees cope with their experienced stress at home. An effective solution would not only increase the mental well-being of employees but also significantly reduce costs for their employers. To create such solutions for stress management, insights into stressors or destressors when working from home are needed. A broad range of potential stressors and destressors have been included in recent research, most notably gender [10,12,17], age [17], the work-from-home experience [14,17], the presence of children [17], the fear of COVID-19 [11,17], work-life balance conflicts [17], social isolation [12,17], a distractive work environment [14,17], job characteristics [13], and sleep quality [12].

However, yet, by far, the greatest part of research on the relationship between stress and working from home has been panel studies with retrospective questionnaires [10-14,17]. No study has used unobtrusive measuring of stress levels, or in-the-moment polling of stress and other emotions through ecological momentary assessments (EMAs), a method of delivering short, recurring questionnaires at multiple times each day to tap into participant experiences without overburdening (see Ref. [18], page 98, for an overview of EMA use in a broad range of populations and research questions). In stress management, it is key to create solutions that require minimum effort for the employee to ensure extensive and effective use. The use of digital technologies that can sense and assess stress and its risk factors continuously, and EMAs for experience sampling, can be beneficial for both the growth of knowledge and the development of solution spaces.

The main aim of this study, therefore, was to find stressors and destressors when working from home using digital technologies (ie, questionnaires on a smartphone and physiological and activity state from a smartwatch). Both types of devices are increasingly prevalent in society and used to assess health. As a secondary objective, this study evaluated whether stressors or derivatives thereof can be assessed continuously and nonobtrusively by means of a smartwatch only. Continuous monitoring of stress and its risk factors on these devices would also allow for future intervention recommendations to avoid stressful moments, which could reduce the risk of developing long-term stress-related conditions [19].

The working-from-home conditions during the pandemic and the advancement and availability of smartwatches that measure physiology have created the possibility to conduct this study fully remote. Data of 202 participants from 3 different regions around the world (the United States, the United Kingdom, and Hong Kong) were collected within 12 weeks, without any physical meetings with the researchers and from the comfort of the participants' own home.

Methods

Study Design

Recruitment

Participants were included if they were active employees of Cigna International (a global insurance company) at the time of the study and were working from home for at least 2 days a week. Only participants who met all inclusion criteria and none of the exclusion criteria (which can be found in Multimedia Appendix 1) could take part in the study. In total, 202 participants were included in the study: 50 (24.8%) in the United States, 70 (34.7%) in Hong Kong, and 82 (40.5%) in the United Kingdom. Participants were allowed to keep the Garmin smartwatch that they used in the study regardless of whether

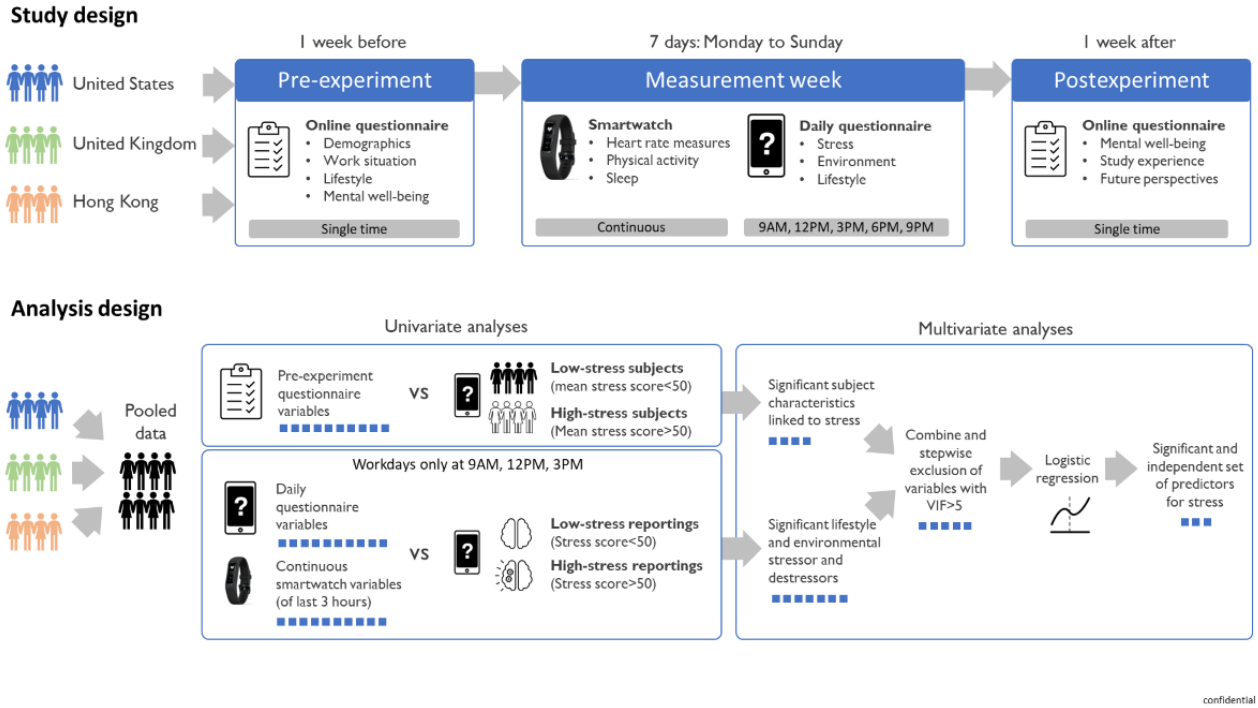
they had finished the study or stopped their participation before the end.

Procedure

The experiment was conducted in 12 weeks between April and September 2021. Measurements for the US participants were taken between April and June, UK participants between May and July, and Hong Kong participants in August. During these periods, there were no major COVID-19 infection peaks in any of the study regions; we therefore expected similar anxiety levels toward the virus during the entire measurement period [20]. Due to the COVID-19 pandemic, employees in all 3 regions were asked by their employers to work from home for at least a few days per week if work activities permitted this. The measurements lasted 7 days per participant, hereafter called "measurement week," starting on Monday morning and ending on Sunday evening. A pre-experiment questionnaire was filled in the week before the measurement week, and a postexperiment questionnaire was filled in the week after the measurement week. See Figure 1 for a schematic overview of the study design.

During the measurement week, participants had to fill in the 5 daily questionnaires (EMAs) at set times. These EMAs were kept as short as possible to keep the participant burden as low as possible (see Ref. [21] for criteria). The EMAs were prompted in a questionnaire app on their smartphone. Participants wore the smartwatch continuously during the week (day and night) and got notifications in the questionnaire app to remind them to synchronize their watch to the smartwatch app on their phone.

Figure 1. Study design (top) and analysis design (bottom). Blue squares in the analysis design represent a fictional number of variables remaining in each step. VIF: variance inflation factor.



Consumer Smartwatch

The Garmin Vivosmart 4 was used, which measures the user's heart rate (HR), steps, classification of activity, sleep duration, sleep stages, and stress measurements, as derived by Garmin's algorithms. Except for sleep characteristics [22,23] and stress measurements, the Garmin sensors have proven to be fairly accurate [24-27]. Participants had to download the Garmin Connect app to their own smartphone to set up and synchronize the data of the smartwatch during the study. The used smartwatch has a thin band and was therefore expected to not have any influence on the daily life of a participant.

Questionnaires

There were 3 sets of questionnaires: a pre-experiment questionnaire, a postexperiment questionnaire, and daily questionnaires (all questionnaires are available as Multimedia Appendix 2). The pre-experiment questionnaire contained questions on stable determinants and traits known from the literature to affect work-from-home stress: age; gender; job type; children present; work-from-home facilities (ie, a dedicated office), work-family interference, and social support from the boss, colleagues, and family members, respectively; current mental well-being; impact of the pandemic; and fear of (the debilitating consequences of) COVID-19. Moreover, the questionnaire contained questions on stress-mitigating activities, such as mindfulness and yoga, which are known to reduce work-related stress, both before [28,29] and during [17,30] the pandemic. The postexperiment questionnaire focused on participants' experiences with the protocol and wearable device during the measurement week, their mental state of the past week, outlook on working from home, and stress detection by means of a wearable. The pre- and postexperiment questionnaires were filled in through Castor electronic data capture (EDC), a digital clinical trial platform to which participants got a link through their email.

Daily questionnaires were delivered 5 times a day using EMAs. To keep the participant burden acceptable, the questionnaires contained single questions on participants' mental state (perceived stress, focus, motivation, productivity), the occurrence of stressors (work, children, house chores and other daily life distractions, noise, distractions from family, technical issues, others), and the occurrence of destressors (breaks, going outside, access to sunlight and fresh air, physical activity, social contacts). Morning questionnaires (9:00 a.m.) also contained questions on sleep quality and outlook on the upcoming day, whereas the midday questionnaires (noon, 3:00 p.m., and 6:00 p.m.) contained questions on current stress levels and activities and the mental state over the past 3 hours. The evening questionnaire (9:00 p.m.) had additional questions regarding the overall outlook on the past day, work times, and activities. Where possible, items were delivered as visual analog scales (VASs), a "gliding" scale ranging from "not at all" to "extremely," which participants could seamlessly set to their preferences. VASs are at least as discriminating as questionnaires when it comes to highlighting differences in experience (see Ref. [31]), and known strengths of VASs are their relatively low burden for participants [32], their responsiveness to change [33], and their reliability when used repeatedly [34]. All daily questionnaires were delivered through the imec Q app, which also notified the user when a new questionnaire was ready. This questionnaire was only available for 90 minutes after it was prompted. The imec Q app was downloaded from the app store (iOS or Android) on the participants' own smartphone, and registration to the study was done by a participant-specific quick response (QR) code.

Ethical Considerations

This observational study was approved by the Western Institutional Review Board, USA (study number: 1303528; IRB tracking number: 20210874), and the Survey and Behavioral Research Ethics Committee at the Chinese University of Hong Kong, Hong Kong (study number: SBRE-20-798). No further approval was necessary for the United Kingdom. This study complies with the guidelines of the Declaration of Helsinki and is registered (registration ID NL9378) in the Dutch Trial Register (NTR).

To ensure anonymous participation of the subjects, recruitment emails were sent through the intranet of the company, asking employees to sign up by emailing the researchers from a personal email account of their choice. To make sure they understood the study design, participants had to correctly answer questions about the procedure of the study before being officially included in the study. All participants signed the informed consent form before participating in the study.

Statistical Analysis

Sample Size

The sample size calculation was based on the aim of finding at least 1 known stressor. Research shows that 82.5% of people with high stress levels experience high noise in their environment [35]. Only 55.3% of people with low stress levels had high noise in their environment. Additionally, we used an outcome of a public opinion poll of 2019 where 38% of employees were estimated to experience stress in the workplace [36]. A 2-sample, 2-sided equality sample size calculation (p_A=0.553, p_B=0.825, k=0.62/0.38=1.632, 1 – β=.95, α=.05) resulted in a minimal sample size of 162. A power of 95% was chosen to account for the potential differences in the participant group per location. Accounting for dropout (10%) and loss of compliance during the week (10%), similarly to Smets et al [37], the actual required sample size of 194 was obtained.

Data

The data from the 3 different regions were combined to obtain the necessary sample size. Due to limited data of people working in the office and low compliance during the evenings and weekends, only daily data when working from home during work hours were used for analysis, including the prequestionnaire data. All data were divided into 3-hour blocks: 9:00 a.m.-noon, noon-3:00 p.m., and 3:00 p.m.-6:00 p.m. Each 3-hour block was associated with smartwatch data collected during the corresponding period and a questionnaire filled at the end of the period (at noon, 3:00 p.m., and 6:00 p.m.). Smartwatch data were averaged over the 3-hour period. Sleep quality from the 9:00 a.m. questionnaire was added to all other data points for that workday.

In the analysis, the stress level, as indicated by the participant, in the daily questionnaires on a VAS of 0 (no stress) to 100 (extremely stressed) was taken as the dependent variable, while the other daily questionnaire answers and smartwatch data were taken as independent variables. The independent variables were further divided into participant characteristics (ie, general demographics and environment) and environmental stressors (access to sunlight, access to fresh air, noise, distractions of daily life, and distractions of people), which were reported on a VAS of 0 (none) to 100 (a lot), and lifestyle stressors (taking a break, social interactions, going outside, and exercising), which were reported as a "yes" or "no" answer. Sleep quality was also part of the lifestyle stressors but was reported on a 0 (bad) to 100 (good) VAS. The smartwatch measurements of interest in the study were the HR, total sleep time, and activity data (step count and activity intensity). Activity intensity was scored by Garmin as "sedentary," "active," or "highly active" periods of maximally 15 minutes. No sleep stages were analyzed, due to the low or unknown performance in previous research characteristics [22,23]. Measurements from a 3-hour block were excluded in the analysis if either the smartwatch or the questionnaire data were missing. The percentage of excluded blocks is reported in the Results section.

Analysis and Modeling

In data exploration, a bimodal distribution of stress scores was observed (Multimedia Appendix 3). Furthermore, we found a relatively high frequency for stress scores of 50. Since the VAS was set to 50 by default in the questionnaires, there was a possibility of an unintentional increase of the scores corresponding to 50 in the case there was no answer from the participants or if the participants rushed though the questionnaire without moving the slider. The score distribution of various VAS variables also showed a relatively high frequency for the 50 values. We therefore excluded the exact value of 50 from the analysis, both for the stress scores and for the independent

variables from the daily questionnaires. This exclusion emphasized the bimodal stress score distribution. Based on these findings and the increased interpretability of a binary outcome, we divided self-reported stress into 2 categories: low stress (stress scores lower than 50) and high stress (stress scores higher than 50). Furthermore, others have also reported robust findings after dichotomization of continuous questionnaire-based outcome variables [38,39] or binary stress classifications in wearable studies [40].

The variables from the pre-experiment questionnaire, daily questionnaire, and smartwatch were compared between the low- and high-stress groups by means of a univariate analysis and a multivariate analysis, as depicted in Figure 1.

Given the different data types for the low- and high-stress comparisons, different statistical tests were performed: a Mann-Whitney U (MWU) test for continuous/ordinal data, the Fisher exact test for 2 proportions, or a test of proportions for more than 2 proportions. The threshold for significance in all tests was set to $P<.05$.

Significant variables from the univariate analysis were used as input for the multivariate logistic regression. To minimize the effects of multicollinearity among the variables identified as participant characteristics, stressors, destressors, and smartwatch measurements in the univariate analysis, the variance inflation factor (VIF) was calculated. Stepwise reduction of variables was performed based on the VIF calculated on the initial set of significant variables; variables were dropped 1 by 1 (starting with the highest VIF) until a variable set with all VIFs less than 5 was obtained. A logistic regression model was trained on the resulting variables from the VIF analysis to classify instances as low stress (0) or high stress (1). The variables were scaled between 0 and 1 to make the coefficients comparable. Using the *P* values and coefficients obtained from the logistic regression model, the most important characteristics, stressors, destressors, and smartwatch measurements were identified ($P<.05$). All data analyses were performed using Python (version 3.8), Scikit-learn (version 1.0), and Scipy (version 1.8).

Results

Data Set Description

This study included 202 participants from 3 regions. This number of participants was divided over the 3 regions relative to the number of employees at those locations; see Table 1 for the exact number of participants per location.

Table 1 shows the characteristics of the sampling population. This study had a low rate of dropouts during the study (n=4, 2%). Dropouts were mainly caused by irritations from the watch or no specified reason. The low data availability in Hong Kong was caused by a bug in the Garmin data retrieval, which caused a loss of smartwatch data for 9 (13%) of 70 participants. Overall, population characteristics were similar across regions. Trends in differences of more female participants in United States and more anxiety toward the pandemic and more stress in Hong Kong were found.

Table 1. Population and data set characteristics, also split per region.

Characteristics	US participants (n=50)	UK participants (n=82)	Hong Kong participants (n=70)	Total participants (N=202)
Participants who completed the experiment, n (%)	47 (94)	82 (100)	69 (99)	198 (98)
Questionnaires completed during the workweek per participant (%), mean (SD)	92 (25)	80 (23)	73 (28)	81 (25)
Age (years), mean (SD)	43 (11.1)	41 (8.5)	38 (8.7)	40 (9.4)
Gender, n (%)				
Male	12 (23)	36 (44)	30 (43)	79 (39)
Female	38 (77)	46 (56)	40 (57)	123 (61)
Role, n (%)[a]				
Manager	23 (46)	33 (40)	32 (46)	89 (44)
Analyst	12 (23)	25 (31)	15 (21)	51 (25)
Customer service	5 (10)	9 (11)	1 (1)	16 (8)
Other	11 (21)	15 (18)	22 (32)	46 (23)
Participants with children aged <10 years in the house during the workday, n (%)	11 (21)	25 (30)	16 (23)	51 (25)
Job very to extremely demanding/intense, n (%)	17 (33)	36 (44)	27 (38)	81 (40)
Moderate-to-extreme anxiety toward COVID-19, n (%)	18 (35)	23 (28)	36 (52)	79 (39)
Stress level throughout the workweek (0-100), mean (SD)	34.5 (26.2)	36.2 (27.7)	43.8 (22.8)	37.5 (26.4)

[a]The numbers could be more than the total because of rounding.

Participant Characteristics Related to Stress

Table 2 shows the statistics to indicate an enrichment of pre-experiment questionnaire variables for either the low- or the high-stress group based on the entire sampling population. Burnout experience, job intensity, whether family and everyday life events get in the way of work, and anxiety due to the pandemic were all significantly different between low- and high-stress groups. There was no significant difference in region between low and high stress levels.

Table 2. Differences in stress levels for participant profile items based on the entire population[a].

Profile items	Low stress	High stress	Ratio	*P* value	Test
Age (years)	Mean 40.9 (SD 9.0)	Mean 38.4 (SD 9.4)	0.94	.10	*U*[b]
Gender, male/female	0.39	0.31	0.81	.39	*F*[c]
Children, yes/no	0.35	0.26	0.75	.35	*F*
Dedicated office space, yes/no	0.81	0.74	0.90	.34	*F*
Experienced burnout in the past, yes/no	0.45	0.76	1.69	.03[d]	*F*
Perform stress reduction activities, yes/no	0.48	0.64	1.32	.09	*F*
Family/life events get in the way of work (0-4)	Mean 1.1 (SD 0.9)	Mean 1.3 (SD 0.8)	1.25	.05[d]	*U*
Support from colleagues (0-4)	Mean 2.7 (SD 1.0)	Mean 2.4 (SD 1.0)	0.91	.08	*U*
Job intensity (0-4), mean (SD)	Mean 2.2 (SD 0.8)	Mean 2.6 (SD 0.8)	1.18	.01[d]	*U*
Relationship with colleagues (0-4)	Mean 3.0 (SD 0.7)	Mean 2.8 (SD 0.5)	0.94	.08	*U*
Life affected by pandemic (0-4)	Mean 2.2 (SD 0.9)	Mean 2.4 (SD 0.7)	1.08	.17	*U*
Living a full life during the pandemic (0-4)	Mean 1.6 (SD 1.0)	Mean 1.7 (SD 1.0)	1.04	.67	*U*
Help from others in dealing with life (0-4)	Mean 1.9 (SD 1.1)	Mean 1.9 (SD 1.2)	0.98	.95	*U*
Anxiety due to pandemic (0-4)	Mean 1.2 (SD 0.8)	Mean 1.4 (SD 0.9)	1.24	.04[d]	*U*
Region					
Hong Kong (n=59)	38 (64%)	21 (36%)	0.55	.10	Pr[e]
United Kingdom (n=82)	65 (79%)	17 (21%)	0.26	.10	Pr
United States (n=43)	34 (79%)	9 (21%)	0.26	.10	Pr

[a]The "Low stress" and "High stress" columns represent the mean (SD) values for continuous/ordinal variables, such as age, and a ratio of the participant count for binary variables, such as gender. The "Ratio" column represents the ratio of these values between high and low stress levels. Responses ranged from 0 for very low/little to 4 for very high/much. The *P* value is for the performed test that is depicted in the "Test" column.

[b]*U*: Mann-Whitney *U* (MWU) test.

[c]*F*: Fisher exact test.

[d]$P<.05$

[e]Pr: test of proportions.

Identification of Stressors and Destressors

Environmental and lifestyle factors that were subjectively scored 3 times per workday were compared between high- and low-stress groups. The results of these comparisons are depicted in Table 3. Significant stressors and destressors are also visualized in Figure 2. The significant environmental stressors when working from home were distractions of other people, distractions of daily life, and noise of the surroundings. The significant environmental lifestyle destressors were access to fresh air and sunlight, and the lifestyle destressors were taking a break, having social interactions outside of work, going outside, and exercising. There was no significant relationship found between low and high stress levels during work hours and the quality of sleep during the night before.

Table 3. Ratio of values of variables between low- and high-stress groups and the coefficients of stress prediction modeling[a].

Characteristics	Univariate analysis					Multivariate analysis			
	Low stress	High stress	Ratio	*P* value	Test	VIF[b] initial	VIF final	B (LR)[c]	*P* value
Participant characteristics									
Family/life events get in the way of work (0-4)	Mean 1.1 (SD 0.9)	Mean 1.3 (SD 0.8)	1.25	.05[d]	*U*[e]	3.46	2.24	1.18	<.001[f]
Job intensity (0-4)	Mean 2.2 (SD 0.8)	Mean 2.6 (SD 0.8)	1.18	.007[d]	*U*	9.97	N/A[g]	N/A	N/A
Anxiety due to pandemic (0-4)	Mean 1.2 (SD 0.8)	Mean 1.4 (SD 0.9)	1.24	.04[d]	*U*	7.69	N/A	N/A	N/A
Experienced burnout in the past, yes/no	0.45	0.76	1.69	.03[d]	*F*[h]	4.98	3.37	1.08	<.001[f]
Environment									
Sunlight (0-100)	Mean 47.2 (SD 32.3)	Mean 42.9 (SD 30.4)	0.91	.02[d]	*U*	5.18	4.99	0.18	.49
Fresh air (0-100)	Mean 40.7 (SD 33.6)	Mean 34.5 (SD 31.7)	0.84	<.001[f]	*U*	4.99	4.93	–0.08	.78
Noise (0-100)	Mean 23.0 (SD 23.9)	Mean 26.2 (SD 25.0)	1.13	.008[d]	*U*	2.70	2.65	0.08	.79
Distraction of daily life (0-100)	Mean 22.2 (SD 23.4)	Mean 27.7 (SD 26.7)	1.24	<.001[f]	*U*	3.50	3.49	0.23	.51
Distraction of people (0-100)	Mean 18.4 (SD 22.2)	Mean 23.9 (SD 26.0)	1.30	<.001[f]	*U*	3.22	3.21	0.52	.17
Lifestyle									
Social, yes/no	0.59	0.39	0.66	<.001[f]	*F*	2. 27	2.23	–0.89	<.001[f]
Outdoor, yes/no	0.41	0.27	0.66	<.001[f]	*F*	2.86	2.85	–0.44	.03[d]
Break, yes/no	0.57	0.44	0.77	<.001[f]	*F*	2.47	2.40	–0.03	.83
Exercise, yes/no	0.29	0.16	0.55	<.001[f]	*F*	2.37	2.35	–0.28	.20
Sleep quality of night before (0-100)	Mean 61.9 (SD 27.7)	Mean 61.5 (SD 27.1)	0.99	.64	*U*	N/A	N/A	N/A	N/A
Smartwatch									
Active intensity (0-12)	Mean 7.92 (SD 3.0)	Mean 5.99 (SD 2.9)	0.85	<.001[f]	*U*	5.28	4.57	0.18	.83
Highly active intensity (0-12)	Mean 1.87 (SD 2.2)	Mean 1.29 (SD 1.8)	0.83	.04[d]	*U*	1.84	1.68	0.37	.37
Steps (count)	Mean 1858 (SD 1863)	Mean 875 (SD 1203)	0.66	<.001[f]	*U*	4.65	4.07	–0.27	.62
Mean HR[i,j] (beats per minute [bpm])	Mean 78.78 (SD 9.5)	Mean 78.35 (SD 9.4)	0.99	.57	*U*	N/A	N/A	N/A	N/A
SD HR[k] (bpm)	Mean 10.9 (SD 6.5)	Mean 9.0 (SD 3.9)	0.82	<.001[f]	*U*	9.01	N/A	N/A	N/A
Sleep duration (hours)	Mean 7.7 (SD 1.0)	Mean 7.8 (SD 1.2)	1.01	.66	*U*	N/A	N/A	N/A	N/A

[a]The "Low stress" and "High stress" columns represent the mean (SD) values. The "Ratio" column represents the ratio between high and low stress levels. Yes/no questions are presented as the ratio of "yes" answers. Responses ranged from 0 for very low/little to 4 for very high/much.

[b]VIF: variance inflation factor.

[c]B(LR): binary (logistic regression) weight.

[d]*P*<.05.

[e]*U*: Mann-Whitney *U* (MWU) test.

[f]*P*<.001.

[g]N/A: not applicable.

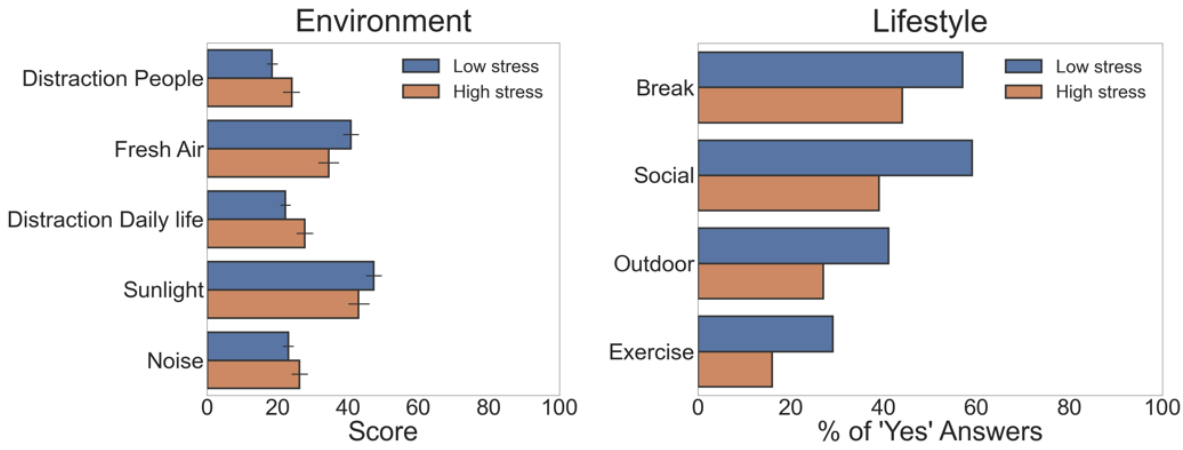
[h]*F*: Fisher exact test.

[i]HR: heart rate.

[j]The "Low stress" and "High stress" values represent the mean and SD of the mean HR feature within the specified time window.

[k]The "Low stress" and "High stress" values represent the mean and SD of the SD HR feature within the specified time window.

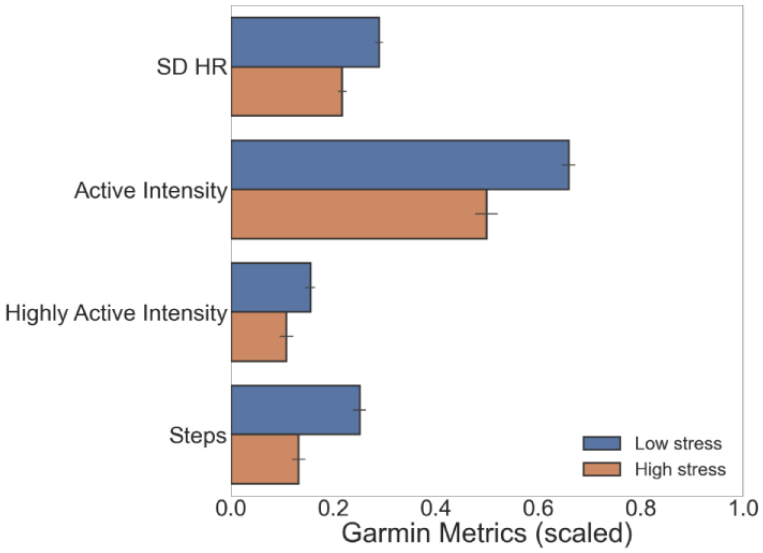
Figure 2. Significant environmental and lifestyle stressors and destressors subjectively scored in the daily questionnaires (P<.05). The bars represent the mean value of the group, and the error bars on the environmental stressors represent confidence.



Measurements With the Smartwatch

The smartwatch measured and calculated variables for physical activity, HR, and sleep duration. We excluded 14% of 3-hour data blocks of questionnaire data due to missing smartwatch data. Table 3 shows the results of the comparisons made between low- and high-stress groups, and significant predictors are visualized in Figure 3. The SD of the HR, active and highly active intensity periods, and the number of steps of the past 3 hours were associated with stress. No significant relationships were found between the stress levels and the average HR and sleep duration.

Figure 3. Significant activity and physiological predictors of stress from the smartwatch with P<.05. The bars represent the mean value of the stress group, and the error bars represent CIs of the mean. The metrics are scaled to values between 0 and 1 for visualization purposes. HR: heart rate.



Combining Predictors for Stress

All identified variables from participant characteristics, lifestyle, environment, and smartwatch that were significantly associated with stress were combined in a multivariate analysis to identify a set of independent predictors of stress. By calculating the VIF over this set of variables, 5 (29%) of 17 predictors (with VIF>5) were identified as predictors contributing significantly to multicollinearity (Table 3). Stepwise reduction of variables with VIF>5 resulted in a set of 14 variables that was used as input for binary logistic regression. Of these input variables, 4 (29%) were found significantly associated with stress in the regression model and they were all self-reported in the pre-experiment questionnaire or daily lifestyle questions. Higher stress was significantly associated with more disruptive family or life events while working (P<.001), previous experience with

burnout (P<.001), fewer social events (P<.001), and less outdoor time (P=.03).

Since all multivariate predictors for stress result from questionnaire data, the question arises whether continuous and nonobtrusive sensing of environmental conditions and human activity/physiology can replace repetitive questionnaires. To answer that question, we ran 2 logistic regression models with a 3-fold cross-validation scheme: one reference model with the significant questionnaire predictors from the original logistic regression (whether they had disruptive family or life events while working, experience with burnout, social events, and outdoor events) and another model with predictors that can be (potentially) continuously and passively sensed (fresh air, noise, sunlight, steps, SD of the HR, active intensity, highly active intensity, and the 2 participant characteristics of whether they had disruptive family or life events and experience with burnout). The questionnaire model resulted in F_1 scores of 0.66, 0.4, and 0.48 in the 3-fold cross-validation (mean F_1=0.51, SD F_1=0.11). The continuous sensing model resulted in F_1 scores of 0.59, 0.38, and 0.44, which were not much lower than those for the questionnaire model (mean F_1=0.47, SD F_1=0.09). In addition, both models were better than guessing randomly using known class distributions (mean F_1=0.33) or by majority class guessing (mean F_1=0). This indicates that continuous and nonobtrusive sensing has the potential to replace questionnaires when monitoring momentary stress.

Bidirectional Relationship Between Stress and Sleep

Results in Table 3 show that the sleep quality of the night before, as answered in the 9:00 a.m. daily questionnaires, was not associated with stress levels across an entire workday, the same counts for the smartwatch measurements of sleep duration. We additionally checked whether sleep quality affected experienced stress levels reported in the 9:00 a.m. questionnaire on all measured days (workdays and non-workdays). The average sleep quality was significantly higher (P<.001, MWU test) in the low stress samples (Mean: 66.8, *SD*: 26.3) than in the high stress samples (Mean: 55.7, *SD*: 28.5). Similarly, we checked whether stress reported at the end of the day (9:00 P.M.) was associated with sleep quality in the following night across all measured days. Low stress moments resulted in significantly (P=.004, MWU test) higher sleep quality scores (mean 65.1, SD 26.6) compared to the high-stress moments (mean 59.7, SD 26.8). These results indicate that high evening stress is related to poor sleep quality during the night after and that low sleep quality is correlated with higher-stress moments in the morning after.

Discussion

Principal Findings

This study had 2 main objectives: first, using EMAs to sample the effect of stressors on work-from-home–related mental well-being. The study showed that participants working from home had stressors and destressors related to their characteristics, environment, and lifestyle. Family and daily life getting in the way of work, job intensity, a history of burnout, and anxiety toward the COVID-19 pandemic were all related to higher stress levels. Environmental stressors were noise and distractions from both daily lives as well as other people. The destressors found to influence the stress levels of the participants were having access to sunlight and fresh air during workhours. Further destressors related to lifestyle were going outdoors, taking breaks, exercising, and having social interactions outside of work. These stressors and destressors are comparable to those found in previous research [17,41], reporting that decreased physical activity, social isolation, family-work conflict, and more distractions when working from home are significant predictors of decreased mental well-being. However, not all potential determinants in previous research showed to affect mental well-being were significant predictors in this study. Most notably, age, gender, the presence of children, a dedicated office space, social support from colleagues, impact of the pandemic, and sleep quality did not significantly impact stress levels. Further research, especially meta-analyses, should shed light upon the causes and value of these results. A major difference between this study and the greater part of previous research is that the latter made use of single, retrospective questionnaires, whereas this study applied recurrent, brief measurements to sample current experiences. Further research using experience sampling methodology, such as EMAs, could further elucidate the impact of measurement timing on determinant strength.

The second objective was to compare EMA measurements to more unobtrusive physiological measurements with a smartwatch. Our results indeed suggest that some of the activity-related stressors and destressors found in this study could be indirectly measured with a smartwatch. The significant activity measurements of the smartwatch, such as steps and activity classifications, were negatively related to stress and may represent destressors, such as exercise, breaks taken, and going outside, as these could all be represented by movement. The same reasoning can be applied for the significant negative relationship between the SD of the HR and stress. A potential hypothesis is that the HR will fluctuate more during activity compared to sitting still behind a desk, which is associated with lower stress levels, with the important notion that these associations were not tested for causality.

When all significant participant characteristics, stressors, destressors, and smartwatch measurements were combined into a model, only some participant characteristics and destressors remained significant in stress prediction. Two participant characteristics significantly contributed to the stress prediction model, namely the disruption of family and everyday life events and the previous experience of burnout. The former could be a direct effect of the working-from-home regime, where families are asked to stay and work from home to limit the spread of the COVID-19 virus, causing personal and work lives to no longer be separate. The latter could be a residual effect of low stress tolerance during the recovery of stress-related exhaustion [42]. These residual effects are shown to still be prevalent after at least 7 years since seeking help. Future research should examine whether these participant characteristics persist being significant or whether their influence is reduced due to the opening of schools, offices, and other locations. Two destressors significantly contributed to the model: having social interactions and being outdoors. These findings may indicate that

encouraging employees to go outside during breaks and talking to people outside of work meetings could already have a beneficial effect on their mental state.

Wearables and Stress

This study showed a minimal influence of smartwatch data in the most optimal model when added to the questionnaire data. This redundancy could be the result of the questionnaires already covering the amount of time spent outdoors, for example, which would also be represented in the steps, active periods, and highly active periods measured by the smartwatch. Although performances for the prediction model based on questionnaire data alone and the model that uses smartwatch and contextual data were relatively low (F_1 score of 0.66 and 0.59, respectively), their similar performance could indicate that in the future, in potential, questionnaire measurements can be replaced by wearable or other unobtrusive sensor measurements.

Ideally, someone's stressors and destressors could be automatically measured with a smartwatch such that no effort is needed from the employee to gather data of their mental state and surroundings. Even though the current state of smartwatches might not be able to inform accurate stress prediction models by themselves, it is expected that performance will improve in the future. As companies are adding better and more sensors to their smartwatches in nearly every new version, it is likely that more stressors and destressors can be measured automatically. Additional physiological sensors, such as skin conductance, skin temperature, and HR variability, could improve stress detection [37], while additional environmental sensors, such as sound levels and ambient light sensors, could measure noise, distractions, the environment, and outdoor activities and therefore replace questions to the user. Furthermore, additional sensing technology, such as smart chairs [43], smart glasses, office equipment use [44], and data collected with a smartphone [45] could improve stress detection.

However, future studies concerning the accuracy of devices in measuring physiological signals need to be conducted as previous studies have shown a wide variability in the accuracy of different consumer devices [22,46-48]. Furthermore, more data must be collected to build reliable models for stress detection. So far, stress prediction models by means of wearables have shown promising, although not 100% accurate, results in a controlled environment with momentary stress tasks [49]. However, stress levels in daily life can be different from those in a controlled laboratory setting. Consequently, model performance of studies in daily life settings is lower [37,40,50]. Next steps in this field of research would need to focus on detecting stress by means of more and improved unobtrusive measurements in a daily dynamic setting. If stress detection models would eventually be reliable enough, it would provide the possibility to monitor stress levels and intervene when high stress levels are detected in order to prevent stress-related diseases in the long term.

Limitations

This study consisted of 202 participants from 3 different regions of the world: the United States, the United Kingdom, and Hong Kong. As Table 1 has shown, there are some differences between Hong Kong, the United States, and the United Kingdom, specifically in average stress levels reported and anxiety toward the COVID-19 pandemic in Hong Kong participants. Although these 2 can be related to each other, the reason could also be cultural differences. Cultural differences in norms and values might cause a difference in the definition or the view of mental health [51], which makes reported mental health scores hard to compare between regions. The higher stress levels in Hong Kong might also be due to the specific company location, which can be caused by a difference in organizational structures or the intensity of the assignments to be performed. The same counts for the anxiety toward the pandemic, as COVID-19 measures differed per region. Even though location (with their cultural and societal differences) might affect stress levels, we were still able to find common stressors and destressors, possibly due to the robust split in low- and high-stress groups. Regional differences in stress scores also did not differ significantly in any of the statistical tests. Furthermore, a difference between regions was accounted for in the sample size calculation by using a high power of 0.95.

Another limitation of this study is that the participant group consisted of employees of the same company, as certain stressors and destressors might be specific to this company. However, we expect the influence to be minimal. These employees were chosen due to the international character of Cigna International, representing multiple cultural backgrounds, and the wide variety of roles within the company. In addition, most employees went from working full-time in the office to working from home overnight similarly to a large part of the population. In addition, similar research among other companies has shown similar stressors and destressors when working from home [17,41], showing similar influences on stress in different working environments. Furthermore, the significant stressors and destressors found in this study are largely unrelated to the job itself, except job intensity. However, job intensity has been proven to cause stress in a variety of job sectors worldwide [52-54]. Therefore, we expect similar results to be obtained performing this study with employees of different companies.

For many of the single-item measures used in the questionnaires, a broad range of longer, validated versions exist [55], such as the Perceived Stress Scale (PSS) [56]. However, we opted not to use those in order to reduce participant burden; especially in recurrent questionnaires, longer answering times and higher burden are known to significantly reduce adherence [18,21]. This can be seen as a limitation of this study because unvalidated questionnaires could potentially be less valid and reliable. However, the literature shows that single-item questions, especially the VASs used in this study, are well suited to measure a multitude of constructs, often as effectively as longer, multi-item questionnaires [57,58]. We therefore argue that the current questionnaires, with shorter, repetitive measurements, often in the form of VASs, have sufficient reliability and validity to aptly sample participants' experiences. Future replications and further studies can support (or, of course, reject) this assumption.

Finally, the dichotomous split in stress scores can be argued. Our first argument is that the stress score distributions were bimodal in nature. Second, a clear and interpretable outcome

of both univariate and multivariate analyses was preferred, also to standardize the analysis design. These arguments were considered having more weight than a potential slight loss of statistical power that dichotomization may cause. Third, previous studies have followed a similar dichotomous approach to continuous questionnaire-based outcomes and have made robust findings [38,39].

Remote Study

This study showed that valuable results can be obtained in research that is conducted fully remote in different regions of the world. Participants lived their normal daily lives and participated in the study from the comfort of their own homes, without any effort to go to a specific research location, therefore minimizing moderator and recall bias. They always had the possibility of asking questions and requesting videocalls, and data were monitored by the researchers to make sure everything was still working. Even though contact was limited, the dropout rate was low and the average percentage of filled-in questionnaires per participant was high (81%). Furthermore, many participants could be measured at the same time, significantly reducing time investments of the researchers and also reducing the influence of the measurement period. In this case, all participants had to be measured within the same COVID-19 time frame to have comparable situations for all participants. However, even future studies will benefit from simultaneous measurements. It is shown that mental states are affected by outside temperature or the duration of sunlight during the day [59,60], and measuring all participants within a short time frame could limit this effect.

These results show that the fast advancement of digital technology, including wearables or other unobtrusive sensors, secure communication platforms, and smartphone applications, extend our possibilities of conducting daily-life research with large sample sizes. A small note must be made in that relying on technologies also brings dependencies on the reliability of that technology as access to backups or support is often limited. This was demonstrated by the smartwatch issue we had in Hong Kong, causing us to lose a small part of data in that region.

Conclusion

This study shows that people have different stressors and destressors when working from home; these should be considered when managing stress in employees. Some of these stressors and destressors are (in)directly measurable with unobtrusive sensors, and prediction models based on these data show promising results for the future of automatic stress detection and management.

These findings may contribute to people optimizing the home office environment (enough sunlight, fresh air, and low noise levels) and activity suggestions during work hours (having social interactions, going outside, and exercising during breaks) to reduce stress levels.

Acknowledgments

This work was supported by Cigna International. The company did not have any influence on the results, nor did it intervene during the data collection. Anonymized data and results were shared according to General Data Protection Regulation (GDPR) guidelines. The OnePlanet Research Center was supported by the Province of Gelderland.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Inclusion and exclusion criteria.
[DOCX File , 32 KB - formative_v6i11e38562_app1.docx]

Multimedia Appendix 2
Questionnaires: the general inclusion questionnaire, the daily questionnaire, and the pre- and postexperiment questionnaires.
[DOCX File , 34 KB - formative_v6i11e38562_app2.docx]

Multimedia Appendix 3
Distributions of the answers to 6 of the daily questionnaires.
[DOCX File , 120 KB - formative_v6i11e38562_app3.docx]

References

1. World Health Organization. Listings of WHO's Response to COVID-19. 2021. URL: https://www.who.int/news/item/29-06-2020-covidtimeline [accessed 2022-10-21]
2. Salari N, Hosseinian-Far A, Jalali R, Vaisi-Raygani A, Rasoulpoor S, Mohammadi M, et al. Prevalence of stress, anxiety, depression among the general population during the COVID-19 pandemic: a systematic review and meta-analysis. Global Health 2020 Jul 06;16(1):57 [FREE Full text] [doi: 10.1186/s12992-020-00589-w] [Medline: 32631403]
3. Meta Careers. Meta. URL: https://www.metacareers.com/life/what-remote-and-flexible-work-will-look-like-at-facebook [accessed 2022-10-20]

4. Google. A Hybrid Approach to Work. URL: https://blog.google/inside-google/life-at-google/hybrid-approach-work [accessed 2022-10-21]
5. PwC. The Benefits and Challenges of Hybrid Working. URL: https://www.pwc.nl/en/topics/blogs/the-benefits-and-challenges-of-hybrid-working.html [accessed 2022-10-21]
6. Mann S, Holdsworth L. The psychological impact of teleworking: stress, emotions and health. New Technol Work Employ 2003 Oct 02;18(3):196-211. [doi: 10.1111/1468-005X.00121]
7. Song Y, Gao J. Does telework stress employees out? A study on working at home and subjective well-being for wage/salary workers. J Happiness Stud 2019 Nov 04;21(7):2649-2668. [doi: 10.1007/s10902-019-00196-6]
8. Russell HO. The impact of flexible working arrangements on work-life conflict and work pressure in Ireland. Gender Work Organ 2009;16:73-97. [doi: 10.1111/j.1468-0432.2008.00431.x]
9. Sullivan C. Remote working and work-life balance. In: Work and Quality of Life. Dordrecht, the Netherlands: Springer; 2012.
10. Sandoval-Reyes J, Idrovo-Carlier S, Duque-Oliva EJ. Remote work, work stress, and work-life during pandemic times: a Latin America situation. Int J Environ Res Public Health 2021 Jul 02;18(13):7069 [FREE Full text] [doi: 10.3390/ijerph18137069] [Medline: 34281007]
11. Shao Y, Fang Y, Wang M, Chang C, Wang L. Making daily decisions to work from home or to work in the office: the impacts of daily work- and COVID-related stressors on next-day work location. J Appl Psychol 2021 Jun;106(6):825-838. [doi: 10.1037/apl0000929] [Medline: 34138589]
12. Şentürk E, Sağaltıcı E, Geniş B, Günday Toker Ö. Predictors of depression, anxiety and stress among remote workers during the COVID-19 pandemic. WOR 2021 Sep 28;70(1):41-51. [doi: 10.3233/wor-210082]
13. Shimura A, Yokoi K, Ishibashi Y, Akatsuka Y, Inoue T. Remote work decreases psychological and physical stress responses, but full-remote work increases presenteeism. Front Psychol 2021 Sep 30;12:730969 [FREE Full text] [doi: 10.3389/fpsyg.2021.730969] [Medline: 34659039]
14. Kumar P, Kumar N, Aggarwal P, Yeap JA. Working in lockdown: the relationship between COVID-19 induced work stressors, job performance, distress, and life satisfaction. Curr Psychol 2021 Mar 14;40(12):6308-6323 [FREE Full text] [doi: 10.1007/s12144-021-01567-0] [Medline: 33746462]
15. Lindblom KM, Linton SJ, Fedeli C, Bryngelsson I. Burnout in the working population: relations to psychosocial work factors. Int J Behav Med 2006 Mar;13(1):51-59. [doi: 10.1207/s15327558ijbm1301_7] [Medline: 16503841]
16. Hassard J, Teoh KRH, Visockaite G, Dewe P, Cox T. The cost of work-related stress to society: a systematic review. J Occup Health Psychol 2018 Jan;23(1):1-17. [doi: 10.1037/ocp0000069] [Medline: 28358567]
17. Galanti T, Guidetti G, Mazzei E, Zappala S, Toscano F. Work from home during the covid-19 outbreak: the impact on employees' remote work productivity, engagement, and stress. J Occup Environ Med 2021:e426-e432. [doi: 10.1097/jom.0000000000002236]
18. Myin-Germeys I, Kuppens P. The Open Handbook of Experience Sampling Methodology: A Step-by-Step Guide to Designing, Conducting, and Analyzing ESM Studies. 2nd ed. Leuven, Belgium: Center for Research on Experience Sampling and Ambulatory Methods; 2022.
19. Heber E, Lehr D, Ebert DD, Berking M, Riper H. Web-based and mobile stress management intervention for employees: a randomized controlled trial. J Med Internet Res 2016 Jan 27;18(1):e21 [FREE Full text] [doi: 10.2196/jmir.5112] [Medline: 26818683]
20. World Health Organization. WHO Coronavirus (COVID-19) Dashboard. URL: https://covid19.who.int/ [accessed 2022-10-21]
21. Eisele G, Vachon H, Lafit G, Kuppens P, Houben M, Myin-Germeys I, et al. The effects of sampling frequency and questionnaire length on perceived burden, compliance, and careless responding in experience sampling data in a student population. Assessment 2022 Mar 10;29(2):136-151. [doi: 10.1177/1073191120957102] [Medline: 32909448]
22. Stone JD, Rentz LE, Forsey J, Ramadan J, Markwald RR, Finomore VS, et al. Evaluations of commercial sleep technologies for objective monitoring during routine sleeping conditions. Nat Sci Sleep 2020 Oct;12:821-842. [doi: 10.2147/nss.s270705]
23. Mouritzen NJ, Larsen LH, Lauritzen MH, Kjær TW. Assessing the performance of a commercial multisensory sleep tracker. PLoS One 2020 Dec 11;15(12):e0243214 [FREE Full text] [doi: 10.1371/journal.pone.0243214] [Medline: 33306678]
24. Claes J, Buys R, Avila A, Finlay D, Kennedy A, Guldenring D, et al. Validity of heart rate measurements by the Garmin Forerunner 225 at different walking intensities. J Med Eng Technol 2017 Aug 04;41(6):480-485. [doi: 10.1080/03091902.2017.1333166] [Medline: 28675070]
25. Lamont RM, Daniel HL, Payne CL, Brauer SG. Accuracy of wearable physical activity trackers in people with Parkinson's disease. Gait Posture 2018 Jun;63:104-108. [doi: 10.1016/j.gaitpost.2018.04.034] [Medline: 29729611]
26. Reddy RK, Pooni R, Zaharieva DP, Senf B, El Youssef J, Dassau E, et al. Accuracy of wrist-worn activity monitors during common daily physical activities and types of structured exercise: evaluation study. JMIR Mhealth Uhealth 2018 Dec 10;6(12):e10338 [FREE Full text] [doi: 10.2196/10338] [Medline: 30530451]
27. Lee J, Byun W, Keill A, Dinkel D, Seo Y. Comparison of wearable trackers' ability to estimate sleep. Int J Environ Res Public Health 2018 Jun 15;15(6):1265 [FREE Full text] [doi: 10.3390/ijerph15061265] [Medline: 29914050]
28. Murphy LR. Stress management in work settings: a critical review of the health effects. Am J Health Promot 1996 Aug 25;11(2):112-135. [doi: 10.4278/0890-1171-11.2.112] [Medline: 10163598]

29. Janssen M, Heerkens Y, Kuijer W, van der Heijden B, Engels J. Effects of mindfulness-based stress reduction on employees' mental health: a systematic review. PLoS One 2018;13(1):e0191332 [FREE Full text] [doi: 10.1371/journal.pone.0191332] [Medline: 29364935]
30. Zadok-Gurman T, Jakobovich R, Dvash E, Zafrani K, Rolnik B, Ganz AB, et al. Effect of inquiry-based stress reduction (IBSR) intervention on well-being, resilience and burnout of teachers during the covid-19 pandemic. Int J Environ Res Public Health 2021 Apr 01;18(7):3689 [FREE Full text] [doi: 10.3390/ijerph18073689] [Medline: 33916258]
31. Lesage F, Berjot S, Deschamps F. Clinical stress assessment using a visual analogue scale. Occup Med (Lond) 2012 Dec;62(8):600-605 [FREE Full text] [doi: 10.1093/occmed/kqs140] [Medline: 22965867]
32. McCormack HM, Horne DJ, Sheather S. Clinical applications of visual analogue scales: a critical review. Psychol Med 1988 Nov 09;18(4):1007-1019. [doi: 10.1017/s0033291700009934] [Medline: 3078045]
33. Huskisson EC, Jones J, Scott PJ. Application of visual-analogue scales to the measurement of functional capacity. Rheumatol Rehabil 1976 Aug;15(3):185-187. [doi: 10.1093/rheumatology/15.3.185] [Medline: 968347]
34. Bowsher D. Acute and chronic pain and assessment. In: Wells P, Frampton V, Bowsher D, editors. Pain Management by Physiotherapy. 2nd ed. Oxford: Butterworth-Heinemann; 1994.
35. Mursali A, Endang B, Dharmono S. Relationship between noise and job stressat a private thread spinning company. Universa Medicina 2009:8-16. [doi: 10.18051/UnivMed.2009.v28.8-16]
36. Gallup. State of the Global Workplace: 2022 Report. URL: https://www.gallup.com/workplace/349484/state-of-the-global-workplace-2022-report.aspx [accessed 2022-10-21]
37. Smets E, Rios Velazquez E, Schiavone G, Chakroun I, D'Hondt E, De Raedt W, et al. Large-scale wearable data reveal digital phenotypes for daily-life stress detection. NPJ Digit Med 2018 Dec 12;1(1):67 [FREE Full text] [doi: 10.1038/s41746-018-0074-9] [Medline: 31304344]
38. Cho Y, Julier SJ, Bianchi-Berthouze N. Instant stress: detection of perceived mental stress through smartphone photoplethysmography and thermal imaging. JMIR Ment Health 2019 Apr 09;6(4):e10140 [FREE Full text] [doi: 10.2196/10140] [Medline: 30964440]
39. Van Der Feltz-Cornelis CM, Varley D, Allgar VL, de Beurs E. Workplace stress, presenteeism, absenteeism, and resilience amongst university staff and students in the covid-19 lockdown. Front Psychiatry 2020 Nov 27;11:588803 [FREE Full text] [doi: 10.3389/fpsyt.2020.588803] [Medline: 33329135]
40. Han HJ, Labbaf S, Borelli JL, Dutt N, Rahmani AM. Objective stress monitoring based on wearable sensors in everyday settings. J Med Eng Technol 2020 May 26;44(4):177-189. [doi: 10.1080/03091902.2020.1759707] [Medline: 32589065]
41. Xiao Y, Becerik-Gerber B, Lucas G, Roll SC. Impacts of working from home during COVID-19 pandemic on physical and mental well-being of office workstation users. J Occup Environ Med 2021 Mar 01;63(3):181-190 [FREE Full text] [doi: 10.1097/JOM.0000000000002097] [Medline: 33234875]
42. Glise K, Wiegner L, Jonsdottir IH. Long-term follow-up of residual symptoms in patients treated for stress-related exhaustion. BMC Psychol 2020 Mar 19;8(1):26 [FREE Full text] [doi: 10.1186/s40359-020-0395-8] [Medline: 32188513]
43. Hyun Jae Baek, Gih Sung Chung, Ko Keun Kim, Kwang Suk Park. A smart health monitoring chair for nonintrusive measurement of biological signals. IEEE Trans Inform Technol Biomed 2012 Jan;16(1):150-158. [doi: 10.1109/titb.2011.2175742]
44. Pepa L, Sabatelli A, Ciabattoni L, Monteriu A, Lamberti F, Morra L. Stress detection in computer users from keyboard and mouse dynamics. IEEE Trans Consumer Electron 2021 Feb;67(1):12-19. [doi: 10.1109/tce.2020.3045228]
45. Stutz T, Blechert J, Tiefengrabner M, Wilhelm F. Smartphone based stress prediction. 2015 Presented at: UMAP 2015: 2015 International Conference on User Modeling, Adaptation, and Personalization; June/July 2015; Dublin, Ireland. [doi: 10.1007/978-3-319-20267-9_20]
46. Guo F, Li Y, Kankanhalli M, Brown M. An evaluation of wearable activity monitoring devices. 2013 Presented at: PDM '13: 1st ACM International Workshop on Personal Data Meets Distributed Multimedia; October 22, 2013; Barcelona, Spain p. 31-34. [doi: 10.1145/2509352.2512882]
47. Case MA, Burwick HA, Volpp KG, Patel MS. Accuracy of smartphone applications and wearable devices for tracking physical activity data. JAMA 2015 Feb 10;313(6):625-626. [doi: 10.1001/jama.2014.17841] [Medline: 25668268]
48. Gillinov S, Etiwy M, Wang R, Blackburn G, Phelan D, Gillinov AM, et al. Variable accuracy of wearable heart rate monitors during aerobic exercise. Med Sci Sports Exerc 2017 Aug;49(8):1697-1703. [doi: 10.1249/MSS.0000000000001284] [Medline: 28709155]
49. Han L, Zhang Q, Chen X, Zhan Q, Yang T, Zhao Z. Detecting work-related stress with a wearable device. Comput Ind 2017 Sep;90:42-49. [doi: 10.1016/j.compind.2017.05.004]
50. Muaremi A, Arnrich B, Tröster G. Towards measuring stress with smartphones and wearable devices during workday and sleep. Bionanoscience 2013 May;3:172-183 [FREE Full text] [doi: 10.1007/s12668-013-0089-2] [Medline: 25530929]
51. Oishi S. Culture and subjective well-being: conceptual and measurement issues. In: Handbook of Well-Being. Salt Lake City, UT: DEF; 2018.
52. Spielberger C, Vagg P, Wasala C. Occupational stress: job pressures and lack of support. In: Quick JC, Tetrick LE, editors. Handbook of Occupational Health Psychology. Washington, DC: American Psychological Association; 2003.

53. Chiang FF, Birtch TA, Kwan HK. The moderating roles of job control and work-life balance practices on employee stress in the hotel and catering industry. Int J Hosp Manag 2010 Mar;29(1):25-32. [doi: 10.1016/j.ijhm.2009.04.005]
54. Elovainio M, Heponiemi T, Kuusio H, Jokela M, Aalto A, Pekkarinen L, et al. Job demands and job strain as risk factors for employee wellbeing in elderly care: an instrumental-variables analysis. Eur J Public Health 2015 Feb 08;25(1):103-108. [doi: 10.1093/eurpub/cku115] [Medline: 25108118]
55. Cooke PJ, Melchert TP, Connor K. Measuring well-being: a review of instruments. Couns Psychol 2016 Aug 05;44(5):730-757. [doi: 10.1177/0011000016633507]
56. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385. [doi: 10.2307/2136404]
57. Huang Z, Kohler IV, Kämpfen F. A single-item Visual Analogue Scale (VAS) measure for assessing depression among college students. Community Ment Health J 2020 Feb 17;56(2):355-367. [doi: 10.1007/s10597-019-00469-7] [Medline: 31531784]
58. Bergkvist L, Rossiter JR. The predictive validity of multiple-item versus single-item measures of the same constructs. J Mark Res 2018 Oct 10;44(2):175-184. [doi: 10.1509/jmkr.44.2.175]
59. Liu J, Varghese BM, Hansen A, Xiang J, Zhang Y, Dear K, et al. Is there an association between hot weather and poor mental health outcomes? A systematic review and meta-analysis. Environ Int 2021 Aug;153:106533 [FREE Full text] [doi: 10.1016/j.envint.2021.106533] [Medline: 33799230]
60. Holick M. A perspective on the beneficial effects of moderate exposure to sunlight: bone health, cancer prevention, mental health and well-being. Compr Ser Photosci 2001:11-37. [doi: 10.1016/s1568-461x(01)80037-5]

Abbreviations

EMA: ecological momentary assessment
HR: heart rate
MWU: Mann-Whitney U
VAS: visual analog scale
VIF: variance inflation factor

Edited by A Mavragani; submitted 07.04.22; peer-reviewed by I Mohino-Herranz, K Villaume; comments to author 06.06.22; revised version received 21.07.22; accepted 22.09.22; published 10.11.22.

Please cite as:
Tump D, Narayan N, Verbiest V, Hermsen S, Goris A, Chiu CD, Van Stiphout R
Stressors and Destressors in Working From Home Based on Context and Physiology From Self-Reports and Smartwatch Measurements: International Observational Study Trial
JMIR Form Res 2022;6(11):e38562
URL: https://formative.jmir.org/2022/11/e38562
doi:10.2196/38562
PMID:36265030

Original Paper

Providers' Perspectives on Telemental Health Usage After the COVID-19 Pandemic: Retrospective Analysis

Hattie Wilczewski[1], BS; Samantha R Paige[1], MPH, PhD; Triton Ong[1], PhD; Hiral Soni[1], MBA, MS, PhD; Janelle F Barrera[1,2], MPH; Brandon M Welch[1,3], MS, PhD; Brian E Bunnell[1,2], PhD

[1]Doxy.me Research, Doxy.me Inc, Rochester, NY, United States

[2]Department of Psychiatry and Behavioral Neurosciences, University of South Florida, Tampa, FL, United States

[3]Biomedical Informatics Center, Medical University of South Carolina, Charleston, SC, United States

Corresponding Author:
Hattie Wilczewski, BS
Doxy.me Research
Doxy.me Inc
3445 Winton Pl, Suite 114
Rochester, NY, 14623
United States
Phone: 1 8444369963
Email: hattie.wilczewski@doxy.me

Abstract

Background: Mental health care pivoted to telemedicine during the COVID-19 pandemic, and there is uncertainty around the sustainability of this rapid shift.

Objective: This study examined how intentions to continue using telemedicine after the COVID-19 pandemic are influenced by provider perceptions of usefulness, ease of use, and professional social influence, facilitating organizational conditions.

Methods: We conducted a web-based, cross-sectional survey of 369 telemental health providers between February and March 2021. A hierarchical linear regression analysis was conducted to predict intentions to continue using telemedicine after the COVID-19 pandemic.

Results: Most providers began using telemedicine in March 2020 or later (257/369, 69.6%) and attended to ≥50% of their clients via telemedicine (299/369, 81.0%). Intention to continue using telemedicine after the COVID-19 pandemic was predicted by the telemedicine caseload (β=.10; *P*=.005), perceived usefulness in general (β=.10; *P*=.008), ease of use (β=.08; *P*=.04), social influence (β=.68; *P*<.001), and facilitating conditions (β=.08; *P*=.047).

Conclusions: Exploration of the predictors of telemedicine usage beyond the COVID-19 pandemic aids in surveillance of telemedicine usage, integration with future clinic workflows, and the shaping of public policy. It is important to consider telemedicine services as not only a response to a crisis but also an effective and useful solution for everyday life. Our results suggest widespread, sustainable telemedicine adoption.

(JMIR Form Res 2022;6(11):e39634) doi:10.2196/39634

KEYWORDS

telemedicine; telehealth; COVID-19; telemental health; mental health; pandemic; perception; use; usefulness; usage; workflow

Introduction

The COVID-19 pandemic significantly impacted mental health (MH) care systems, leading providers to transition rapidly to remote care delivery [1]. In early 2020, in-person mental health care decreased by 50%-70% [2,3] as telemedicine usage increased as much as 6500% [3,4]. This required MH providers to adjust to new technology and loss of in-person care [5] but proved satisfactory for patients owing to decreased wait time, travel time, and absenteeism from work [6,7]. Increased access to patients in rural regions and those with practical barriers to access mental health care may encourage providers' long-term integration of telemedicine into their practice [8]. Telemental health (TMH) emerged as a response to a crisis and has proven to be an effective, useful, and sustainable form of health care delivery [9]. However, widespread adoption of telemedicine was largely due to emergency regulations providing coverage during the pandemic. There is some concern of a telehealth cliff,

or sudden reversal of TMH usage once emergency regulations are lifted [10,11]. Toward this end, we aim to examine how TMH providers perceive the general and pandemic-specific usefulness of TMH and how this relates to their intention to use it in the future.

Behavioral intention is one of the strongest predictors of sustaining or changing a behavior [12]. Studies from both early 2019 [13] and August 2020 [14] found that TMH providers intended to use telemedicine more often in the future. In a study conducted several months before the COVID-19 pandemic, TMH providers voiced concerns about security and technological difficulties but indicated that the benefits to care and workflow strengthened their intentions to use it in the future [13]. Providers have reported other benefits such as improved work-life balance, more flexibility in scheduling, and being able to deliver innovative care [13,15]. During the pandemic, providers who served rural areas and were reimbursed through self-pay methods reported the greatest intentions to use telemedicine in the future, which may be attributed to greater comfort with using the technology as a result of more frequent use [14]. Further, one study found that psychologists anticipated a 5-fold increase in telemedicine usage from prepandemic rates [16]. There is a need to continue to understand MH providers' experiences and intentions as the pandemic evolves.

There are many barriers to integrating telemedicine services into regular care, including technology acceptance [17-19]. The Technology Acceptance Model [20] (TAM) and Unified Theory of Acceptance and Use of Technology [21] (UTAUT) are applied commonly to understand MH providers' use of telemedicine, but not all constructs in the TAM and UTAUT have equal predictive weight. Perceptions about the usefulness of telemedicine are grounded in how the technology can improve care [20,21]; that is, usefulness can be thought of in terms of the benefits the technology brings to providers (eg, reducing no-show rates, reducing costs or overhead, and improving work-life balance) [20,22]. Furthermore, perceptions about ease-of-use are generally based on design features and functionality to facilitate tasks such as coordinating care plans or effectively communicating with patients, patient families, and other health care providers [19]. Effort expectancy has been used synonymously with ease of use [23,24]. The effect of these perceptions on intentions to use telemedicine in the future have yet to be explored among mental health providers after COVID-19 uptake.

Previous studies have found that perceived usefulness was the strongest predictor of telemedicine adoption among MH providers along with perceived ease of use, social influence, and attitude [24,25]. Perceived usefulness from the telemedicine provider perspective refers to quality of care, diagnosis, and monitoring [24,26,27]. Social influence is the degree to which an individual's decision to use a technology is influenced by others' perceptions of the technology [24]. Social influence from other providers may shape intention to integrate a technology into one's own practice, especially when driven by the competitive desire to deliver cutting-edge, innovative care [28]. Thus, the effect of social influence from one's peers on TMH adoption warrants continued investigation [29].

MH providers report that they are still providing high-quality care and effectively communicating with their patients despite the abrupt transition to telemedicine during the COVID-19 pandemic [30]. However, the effect of facilitating conditions (ie, the amount of support, resources, and training available from one's organization to effectively practice their specialty via telemedicine) on this relation has received limited attention [24,31]. To optimize resources dedicated to increasing high-quality telemedicine use in MH care, there is a need to identify the "active ingredients" that predict providers' intentions to use telemedicine in the future.

Providers are the gatekeepers of telemedicine [32,33], warranting continued investigation into provider preferences of usage beyond the COVID-19 pandemic. Knowing the characteristics of MH providers who intend to use telemedicine after the COVID-19 pandemic is important for surveillance of telemedicine utilization and for efforts directed toward strengthening hesitant providers' intentions. Therefore, the purpose of this study is to further investigate determinants of TMH providers' intentions to continue using telemedicine beyond the COVID-19 pandemic.

Methods

Recruitment

Providers were invited to participate in a cross-sectional, web-based survey between February and March 2021, the primary results of which are published elsewhere [34]. Eligible providers included English-speaking, adult (≥18 years of age), mental health providers in the United States. Providers were registered with Doxy.me Inc [35], a commercial telemedicine company that offers secure telecommunications technology for providers to use in their own practice. In total, 495 providers agreed to participate in the study, 369 of whom had complete data for the purposes of this analysis (74.5% completion rate) [36]. The demographic characteristics of providers in the sample were consistent with those of providers in the mental health industry in the United States [13,14,30,34,37,38].

The survey was distributed to participants through email and administered via Qualtrics (Qualtrics Inc). The survey began with an electronic consent form detailing that deidentified data would be used for publication and that 1 free month of a Doxy.me professional membership account would be offered in compensation for time spent completing the survey.

Ethical Considerations

The study was reviewed and deemed exempt by the institutional review board of the University of South Florida (IRB#002053).

Survey and Measures

Overview

The research team iteratively refined and developed the survey exploring several aspects related to TMH practice before, during, and after the COVID-19 pandemic. Participants were asked to provide personal (eg, age, gender, race, and ethnicity) and clinical (eg, practice type, specialty, reimbursement, and treatment paradigm) demographics. They were then asked about their perceptions toward telemedicine relating to perceived

usefulness, perceived ease of use, social influence, facilitating conditions, and intention to use telemedicine after the pandemic. Multimedia Appendix 1 provides the survey questions and response frequencies.

Several demographic variables were recoded during analysis. Professional title and practice type variables were recoded to incorporate responses in the "other" category. Telemedicine caseload was dichotomized to "<50%" (including choices of <25% and 25%-49%) versus "≥50%" (including choices of 50%-75% and >75%). Further, the onset of telemedicine usage was dichotomized to "before March 2020" (combining "December 2019 or earlier" and "January or February 2020") versus "March 2020 or later."

Perceived Usefulness

Perceived usefulness was assessed using 2 scales. The first scale included the average of 3 items measuring provider perspectives about the *general benefits of telemedicine* (questions 27-29 in Multimedia Appendix 1). The second scale included the average of 3 items measuring provider perspectives about the *benefits of telemedicine specifically in relation to COVID-19* (questions 24-26 in Multimedia Appendix 1). Responses to all items were anchored on a 5-point Likert scale (1=*Not at all* to 5=*Extremely*). Both scales had adequate internal consistency (Cronbach =.70-.81).

Perceived Ease of Use

Perceived ease of use was assessed using the System Usability Scale (SUS) [39], which includes 10 items that are alternatingly worded positively (eg, "I think telemedicine is easy to use in my practice") and negatively (eg, "I find telemedicine unnecessarily complex"). Responses were anchored on a 5-point Likert Scale (1=*Strongly Disagree* to 5=*Strongly Agree*), and the SUS score was determined by rescaling the responses, resulting in a single score between 0 and 100. Internal consistency was adequate for the SUS scale (Cronbach =.77). The SUS items are detailed in Multimedia Appendix 1 (questions 14-23).

Social Influence

Social influence by other mental health providers was measured using 1 item (ie, "After the COVID-19 pandemic is resolved, I expect telemedicine to continue to be used by others in my profession"), which was anchored on a 5-point Likert scale from 1=*Much less* to 5=*Much more* (question 34 in Multimedia Appendix 1).

Facilitating Conditions

Facilitating conditions were measured by calculating the mean of 4 items asking about the extent to which providers felt comfortable, supported, trained, and adequately resourced by their practice relative to providing telemedicine services (questions 30-33 in Multimedia Appendix 1). Responses were anchored on a 5-point Likert scale from 1=*Strongly disagree* to 5=*Strongly agree*. Internal consistency was adequate for the facilitating conditions scale (Cronbach =.78).

Intentions to Continue Using Telemedicine After the COVID-19 Pandemic

Intentions to continue using telemedicine after the COVID-19 pandemic were measured using 1 item (ie, "After the COVID-19 pandemic is resolved, I expect to use telemedicine in my practice…"). Responses were anchored on a 5-point Likert scale from 1=*Much less* to 5=*Much more* (question 35 in Multimedia Appendix 1).

Statistical Analysis

SPSS (version 28; IBM Corp) was used for all analyses. Significance was determined with a 2-tailed α of <.05. We conducted independent samples *t* tests to examine differences in intentions to use telemedicine post the COVID-19 pandemic, based on telemedicine caseload (either <50% or ≥50% of patients served remotely), the onset of telemedicine usage in relation to the COVID-19 pandemic (either before March 2020 or during and after March 2020), gender (male or female), and ethnicity (Hispanic or non-Hispanic).

We conducted a 1-way ANOVA to determine differences in intentions to use telemedicine post the COVID-19 pandemic by race. Next, we conducted a correlation analysis to examine the relation between age and intentions to continue using telemedicine after the COVID-19 pandemic. Lastly, we conducted a 5-step hierarchical linear regression analysis to examine the predictive power of provider characteristics (ie, telemedicine caseload, onset of telemedicine usage in relation to the COVID-19 pandemic, and age) and perceptions toward telemedicine (ie, perceived usefulness, perceived ease-of-use, social influence, and facilitating conditions) with respect to intentions to continue using telemedicine after the COVID-19 pandemic. Each step in the regression analysis added a block of variables, determining incremental prediction above and beyond variables in the previous steps. Telemedicine caseload (0=<50%; 1=≥50%), onset of telemedicine usage (0=before March 2020; 1=March 2020 or later), and age were included in step 1 of the model as covariates because they were significantly associated with the intention to use telemedicine after the COVID-19 pandemic. The next step included perceived usefulness variables (eg, general perceived usefulness and perceived usefulness related to the COVID-19 pandemic), with step 3 adding ease of use, step 4 adding social influence, and step 5 adding facilitating conditions. ANOVAs were conducted to determine model comparisons.

Results

Sample Characteristics

The mean age of providers was 52 (SD 13.0) years. Most providers identified as female (300/369, 81.3%), White (298/369, 80.8%), and non-Hispanic (339/369, 91.9%). Professionally, most providers were mental health counselors (179/369, 48.5%), psychologists (108/369, 29.3%), or social workers (54/369, 14.6%) working in individual practice settings (279/369, 75.6%) and primarily treating adult patients (18-64 years old; 308/369, 83.5%). More details regarding provider and practice characteristics are reported elsewhere [34].

Characteristics of Telemedicine Practice and Measures

As shown in Table 1, a total of 257 of 369 (69.6%) providers began using telemedicine in March 2020 or later, and 299 (81.0%) reported that more than ≥50% of their caseload was attended to via telemedicine. With respect to intentions to continue using telemedicine after the COVID-19 pandemic, 173 (46.9%) intended to use it more, 66 (17.9%) intended to use it about the same, and 130 (35.2%) intended to use it less. Most providers (193/369, 52.3%) expected telemedicine to be used more by others in their profession, while fewer providers (115/369, 31.2%) expected their colleagues to use it about the same (61/369, 16.5%) or less.

Providers who attended to ≥50% of clients via telemedicine reported stronger intentions to continue using telemedicine after the COVID-19 pandemic (mean 3.4, SD 1.4) than those who attended to <50% of clients via telemedicine (mean 2.6, SD 1.3; t_{367}=–4.07; P<.001). Providers who began to use telemedicine before March 2020 also reported stronger intentions to continue using telemedicine after the COVID-19 pandemic (mean 3.5, SD 1.3) than those who began using it in March 2020 or later (mean 3.1, SD 1.4; t_{367}=2.46; P=.01). Provider age was negatively correlated with intentions to continue using telemedicine after the COVID-19 pandemic (r_{367}=–0.11; P=.03). There were no significant differences in intentions to use telemedicine after the COVID-19 pandemic by gender (t_{365}=–0.84; P=.40), race ($F_{6,360}$=0.431; P=.86), or ethnicity (t_{367}=0.26; P=.80). In subsequent analyses, telemedicine caseload, duration of usage, and age were included as covariates to be controlled for in the hierarchical regression analysis. Multimedia Appendix 2 shows the correlation table of covariates and measures.

Table 1. Characteristics of telemedicine practices and measures (N=369).

Demographics	Values
Onset of telemedicine usage, n (%)	
Before March 2020	112 (30.4)
March 2020 or later	257 (69.6)
Telemedicine caseload, n (%)	
<50%	70 (19.0)
≥50%	299 (81.0)
Measures, mean (SD)	
Perceived usefulness (general)	3.07 (1.11)
Perceived usefulness (COVID-19)	4.53 (0.67)
Perceived ease of use	75.26 (14.87)
Social influence	3.47 (1.36)
Facilitating conditions	4.11 (0.76)
Intentions to continue using telemedicine after the COVID-19 pandemic	3.24 (1.42)

Predicting Intentions to Continue Using Telemedicine After the COVID-19 Pandemic

Table 2 displays the results for each step of the hierarchical linear regression analysis and Table 3 shows the results for predictors in each step of the model. Step 1 of the model was significant and accounted for 7% of the variance in intentions to use telemedicine after the COVID-19 pandemic ($F_{3,365}$=9.12; P<.001; R^2=0.07). Telemedicine caseload (β=.21; P<.001) and onset of telemedicine usage (β=–.14; P=.007) significantly predicted intentions to continue using telemedicine after the COVID-19 pandemic. Age was not a significant predictor (P=.09).

Step 2 included the addition of perceived usefulness to the model. Controlling for the effects of telemedicine caseload, onset of usage, and age, the regression was significant and accounted for an additional 13% of the variance in intentions to continue using telemedicine after the COVID-19 pandemic ($F_{5,363}$=17.71; P<.001; R^2=0.20; ΔR^2=0.13). Onset of telemedicine usage (β=–.12; P=.009), perceived usefulness in general (β=.31; P<.001) and in relation to the COVID-19 pandemic (β=.11; P=.04) significantly predicted intentions to use telemedicine after the COVID-19 pandemic. Age and telemedicine caseload were not significant predictors (P>.05 for all).

Step 3 included the addition of perceived ease of use to the model. Controlling for the effects of covariates and perceived usefulness, the regression was significant and accounted for an additional 3% of the variance in intentions to continue using telemedicine after the COVID-19 pandemic ($F_{6,362}$=17.56; P<.001; R^2=0.23; ΔR^2=0.03). Onset of telemedicine usage (β=–.11; P=.02), perceived usefulness in general (β=.27; P<.001), and perceived ease of use (β=.18; P<.001) significantly predicted intentions to continue using telemedicine after the COVID-19 pandemic. Age, telemedicine caseload, and perceived usefulness in relation to the COVID-19 pandemic were not significant predictors (P>.05 for all).

Step 4 included the addition of social influence to the model. Controlling for the effects of covariates and perceived usefulness and ease of use, the regression was significant and accounted

for an additional 41% of the variance in intentions to continue using telemedicine after the COVID-19 pandemic ($F_{7,361}$=91.03; P<.001; R^2=0.64; ΔR^2=0.41). Onset of telemedicine usage (β=–.06; P=.02), telemedicine caseload (β=–.10; P=.005), perceived usefulness in general (β=.11; P=.004), ease of use (β=.11; P=.001), and social influence (β=.69; P<.001) significantly predicted intentions to use telemedicine after the COVID-19 pandemic. Age and perceived usefulness in relation to the COVID-19 pandemic were not significant predictors (P>.05 for all).

Step 5 included the addition of facilitating conditions to the model. Controlling for the effects of covariates, perceived usefulness and ease of use, and social influence, the regression was significant and accounted for an additional 0.4%, for a total of 64% of the variance in intentions to continue using telemedicine after the COVID-19 pandemic ($F_{8,360}$=80.80; P<.001; R^2=0.64; ΔR^2=0.004). Telemedicine caseload (β=.10; P=.005), perceived usefulness in general (β=.10; P=.008), ease of use (β=.08; P=.04), social influence (β=.68; P<.001), and facilitating conditions (β=.08; P=.047) significantly predicted intentions to use telemedicine after the COVID-19 pandemic. Age, onset of telemedicine usage, and perceived usefulness in relation to the COVID-19 pandemic were not significant predictors (P>.05 for all).

Table 2. Model comparisons for each step of the hierarchical regression analysis.

Step[a]	*F* test (*df*)	R^2	Adjusted R^2	ΔR^2
1	9.12 (3,365)	0.07	0.06	0.07
2	17.71 (5,363)	0.20	0.19	0.13
3	17.56 (6,362)	0.23	0.21	0.03
4	91.03 (7,361)	0.64	0.63	0.41
5	80.80 (8,360)	0.64	0.63	0.004

[a]P<.001 for all steps and ΔR^2.

Table 3. Predictors of intentions to use telemedicine after the COVID-19 pandemic.

Predictor	B (SE)	95% CI	β	*t* test (*df*)	*P* value
Step 1					
Constant	2.67 (0.48)	1.73 to 3.61	Reference	5.61 (365)	<.001
Telemedicine caseload	0.74 (0.18)	0.38 to 1.10	.21	4.04 (365)	<.001
Onset of telemedicine usage	–0.42 (0.16)	–0.73 to –0.12	–.14	–2.71 (365)	.007
Age	–0.01 (0.01)	–0.02 to –0.002	–.09	–1.68 (365)	.09
Step 2					
Constant	0.97 (0.61)	–0.22 to 2.17	Reference	1.60 (363)	.11
Telemedicine caseload	0.35 (0.18)	0.001 to 0.71	.10	1.96 (363)	.05
Onset of telemedicine usage	–0.38 (0.15)	–0.67 to –0.09	–.12	–2.61 (363)	.009
Age	–0.01 (0.01)	–0.02 to 0.002	–.07	–1.53 (363)	.13
Usefulness (general)	0.40 (0.07)	0.27 to 0.53	.31	5.93 (363)	<.001
Usefulness (COVID-19)	0.24 (0.11)	0.02 to 0.46	.11	2.11 (363)	.04
Step 3					
Constant	0.08 (0.64)	–1.19 to 1.35	Reference	0.12 (362)	.90
Telemedicine caseload	0.35 (0.18)	–0.002 to 0.70	.10	1.96 (362)	.05
Onset of telemedicine usage	–0.33 (0.14)	–0.61 to –0.05	–.11	–2.30 (362)	.02
Age	–0.01 (0.01)	–0.02 to 0.002	–.07	–1.51 (362)	.13
Usefulness (general)	0.34 (0.07)	–0.21 to 0.48	.27	5.10 (362)	<.001
Usefulness (COVID-19)	0.17 (0.11)	–0.05 to 0.40	.08	1.56 (362)	.12
Ease of use	0.02 (0.005)	0.01 to 0.03	.18	3.70 (362)	<.001
Step 4					
Constant	–1.15 (0.44)	–2.02 to –0.27	Reference	–2.58 (361)	.01
Telemedicine caseload	0.34 (0.12)	0.11 to 0.58	.10	2.84 (361)	.005
Onset of telemedicine usage	–0.20 (0.10)	–0.39 to –0.004	–.06	–2.01 (361)	.045
Age	–0.001 (0.004)	–0.01 to 0.01	–.01	–0.27 (361)	.79
Usefulness (general)	0.14 (0.05)	0.04 to 0.23	.11	2.90 (361)	.004
Usefulness (COVID-19)	0.05 (0.08)	–0.10 to 0.20	.03	0.69 (361)	.49
Ease of use	0.01 (0.003)	0.004 to 0.02	.11	3.27 (361)	.001
Social influence	0.72 (0.04)	0.65 to 0.79	.69	20.30 (361)	<.001
Step 5					
Constant	–1.39 (0.46)	–2.29 to –0.48	Reference	–3.02 (360)	.003
Telemedicine caseload	0.36 (0.12)	0.12 to 0.59	.10	2.94 (360)	.004
Onset of telemedicine usage	–0.18 (0.10)	–0.37 to 0.01	–.06	–1.82 (360)	.07
Age	–0.001 (0.004)	–0.01 to 0.01	–.01	–0.26 (360)	.79
Usefulness (general)	0.13 (0.05)	0.03 to 0.22	.10	2.65 (360)	.008
Usefulness (COVID-19)	0.03 (0.08)	–0.12 to 0.18	.02	0.42 (360)	.68
Ease of use	0.01 (0.004)	0.00 to 0.01	.08	2.08 (360)	.04
Social influence	0.71 (0.04)	0.64 to 0.78	.68	20.26 (360)	<.001
Facilitating conditions	0.14 (0.07)	0.002 to 0.28	.08	1.99 (360)	.047

Discussion

The findings of this study show that perceived usefulness in general, perceived ease of use, professional social influence, facilitating conditions, and telemedicine caseload predict MH providers' intentions to use telemedicine in the future. Age, onset of telemedicine usage, and perceived usefulness in relation to the COVID-19 pandemic were not found to be significant predictors in the final model.

Principal Findings

Social influence from other MH providers was the strongest predictor of intentions to continue using telemedicine after the COVID-19 pandemic. Pre–COVID-19 research found influence from organizational leadership to be less important for health care providers than perceived usefulness, reporting that a main driver of provider intentions was how useful they believed the technology would be for patients [24,40]. Other studies have examined social influence in general but not from others in one's profession [23,41]. Social context is an important factor for telemedicine acceptance, and researchers have started to identify a need for a "telehealth culture" to share experiences, opinions, and preferences among providers in the same profession [29,42]. Future research should aim to explore the development and impact of a professional telehealth culture since the onset of the COVID-19 pandemic.

Pre–COVID-19 research showed perceived usefulness to be the most important predictor of provider intentions to continue using telemedicine [24,41,43]. In contrast, our study found social influence from other MH providers to be most influential. This may be due to the somewhat mandatory adoption of telemedicine during the COVID-19 pandemic, compared to pre–COVID-19 studies where telemedicine use may have been more voluntary. The pandemic may have shifted MH providers' priorities from usefulness to other valuations (eg, social influence). Future research should investigate the role of patient preferences in MH providers' telemedicine usage.

It may be that providers with larger remote caseloads became more proficient with the use of telemedicine through experience. In previous research, increased experience using telemedicine was found to strengthen providers' positive attitudes toward remote care after acquiring practice, troubleshooting skills, and workflow integration [29]. Providers with more experience using telemedicine may have a better experience as a result of increased comfort and familiarity with telemedicine interfaces and features, and may experience less frustration owing to technological difficulties and incorporating new technology into their practice [44]. Notably, the onset of telemedicine usage was not a significant predictor in the final model, suggesting that the frequency of telemedicine use mattered more than the duration of use. Future research may focus on the relation between self-efficacy and telemedicine use (ie, workflow integration and technical disruptions) and quality of care (ie, therapeutic alliance and communication).

Perceived ease of use was not among the strongest predictors of intentions to use telemedicine. This finding is consistent with recent examinations of UTAUT in relation to the COVID-19 pandemic [18,24,43]. Providers in this sample reported high ease-of-use ratings, which may have been influenced by widespread adoption, larger telemedicine caseloads, and workflow integrations. Although not a particularly strong predictor, facilitating conditions predicted intentions for continued future use of telemedicine. This finding is consistent with telemedicine acceptance research among providers [41,43].

Notably, 75% of providers in this study practiced in individual practice settings and reported strong support, training, resources, and comfort using telemedicine in their practice. Our results suggest that these facilitating conditions may influence sustainability and practicality of long-term integration of telemedicine in one's practice. Future research should investigate the effect of different practice types on facilitating conditions.

Limitations

This study contains several limitations. First, data were sampled from users of one commercial telemedicine company, who were compensated with 1 month of free professional membership. This may not be representative of all telemedicine providers and may have biased sampling toward providers who are more interested in telemedicine. Future research should confirm the generalizability of our findings among TMH providers across platforms and in other contexts and countries, and should focus on negative opinions and experiences as well as positive ones. However, previous studies have reported that TMH users sampled from Doxy.me are representative of overall industry demographics, and other studies have reported similar findings regarding future telemedicine usage [15].

Furthermore, the secondary analysis of data excluded some constructs from the TAM and UTAUT models, and some provider demographics may be overrepresented in the sample (eg, individual practice, mental health counselor, cognitive behavioral therapy treatment paradigm, and primarily treats adults). A stratified sampling procedure would ensure equal representation of practice types and provider specialties. As COVID-19 regulations and norms evolve, future researchers should continue to investigate patterns in MH providers' telemedicine usage and intentions.

Conclusions

The purpose of this study was to examine TMH providers' intentions to continue using telemedicine after the COVID-19 pandemic, which changed the landscape of MH care by necessitating the need for service delivery via telemedicine. Most TMH providers reported intentions to continue using telemedicine in their profession between "about the same" or "more" after the COVID-19 pandemic. We speculate that this points toward widespread, sustained telemedicine adoption in the future. Stronger intention for future use was predicted by social influence, perceived usefulness, telemedicine caseload, perceived ease of use, and facilitating conditions. Social influence from others in one's profession was the strongest predictor of the continued use of telemedicine. Sustained high rates of telemedicine may lead to the development of a "telehealth culture" in which providers can depend on others in their profession for TMH training, resources, and workflow

systems. Rates of current telemedicine usage are expected to remain high even after the COVID-19 pandemic, especially for providers with a large percentage of caseloads seen via telemedicine. Telemedicine appears to be an important part of current and future MH care.

Acknowledgments

BEB was funded by the National Institute of Mental Health (grants K23MH118482 and R41MH126734) and BMW was funded by the National Cancer Institute (grant K07CA211786).

Conflicts of Interest

BMW is a shareholder, SRP is a former employee, and all other authors are employees of Doxy.me Inc, a commercial telemedicine company.

Multimedia Appendix 1
Survey questions and response frequencies.
[DOCX File , 32 KB - formative_v6i11e39634_app1.docx]

Multimedia Appendix 2
Correlation table (Pearson r and two-tailed *P* value) of practice characteristics and measures.
[DOCX File , 19 KB - formative_v6i11e39634_app2.docx]

References

1. 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]
2. Reilly SE, Zane KL, McCuddy WT, Soulliard ZA, Scarisbrick DM, Miller LE, et al. Mental health practitioners' immediate practical response during the COVID-19 pandemic: observational questionnaire study. JMIR Ment Health 2020 Oct 01;7(9):e21237 [FREE Full text] [doi: 10.2196/21237] [Medline: 32931440]
3. Humer E, Pieh C, Kuska M, Barke A, Doering BK, Gossmann K, et al. Provision of psychotherapy during the COVID-19 pandemic among Czech, German and Slovak Psychotherapists. Int J Environ Res Public Health 2020 Jul 04;17(13) [FREE Full text] [doi: 10.3390/ijerph17134811] [Medline: 32635422]
4. Patel SY, Mehrotra A, Huskamp HA, Uscher-Pines L, Ganguli I, Barnett ML. Trends in outpatient care delivery and telemedicine during the COVID-19 pandemic in the US. JAMA Intern Med 2021 Mar 01;181(3):388-391 [FREE Full text] [doi: 10.1001/jamainternmed.2020.5928] [Medline: 33196765]
5. Uscher-Pines L, Sousa J, Raja P, Mehrotra A, Barnett ML, Huskamp HA. Suddenly Becoming a "Virtual Doctor": Experiences of Psychiatrists Transitioning to Telemedicine During the COVID-19 Pandemic. Psychiatr Serv 2020 Nov 01;71(11):1143-1150 [FREE Full text] [doi: 10.1176/appi.ps.202000250] [Medline: 32933411]
6. Michaud TL, Zhou J, McCarthy MA, Siahpush M, Su D. Costs of home-based telemedicine programs: a systematic review. Int J Technol Assess Health Care 2018 Jan;34(4):410-418. [doi: 10.1017/S0266462318000454] [Medline: 30058505]
7. Atmojo JT, Sudaryanto WT, Widiyanto A, Ernawati E, Arradini D. Telemedicine, cost effectiveness, and patients satisfaction: a systematic review. J HEALTH POLICY MANAGE 2020;5(2):103-107. [doi: 10.26911/thejhpm.2020.05.02.02]
8. Nguyen M, Waller M, Pandya A, Portnoy J. A review of patient and provider satisfaction with telemedicine. Curr Allergy Asthma Rep 2020 Sep 22;20(11):72 [FREE Full text] [doi: 10.1007/s11882-020-00969-7] [Medline: 32959158]
9. Whaibeh E, Mahmoud H, Naal H. Telemental health in the context of a pandemic: the COVID-19 experience. Curr Treat Options Psychiatry 2020;7(2):198-202 [FREE Full text] [doi: 10.1007/s40501-020-00210-2] [Medline: 32292687]
10. Miliard M. As 'telehealth cliff' looms, hundreds of healthcare orgs urge Congress to act. HIMSS. 2021 Jul 27. URL: https://www.healthcareitnews.com/news/telehealth-cliff-looms-hundreds-healthcare-orgs-urge-congress-act [accessed 2021-08-17]
11. Jercich K. How we can avoid the 'telehealth cliff'. HIMSS. 2021 Jun 02. URL: https://www.healthcareitnews.com/news/how-we-can-avoid-telehealth-cliff [accessed 2021-09-14]
12. Davis F, Bagozzi R, Warshaw P. User acceptance of computer technology: a comparison of two theoretical models. Manag Sci 1989 Aug;35(8):982-1003 [FREE Full text] [doi: 10.1287/mnsc.35.8.982]
13. Bunnell BE, Barrera JF, Paige SR, Turner D, Welch BM. Acceptability of telemedicine features to promote its uptake in practice: a survey of community telemental health providers. Int J Environ Res Public Health 2020 Nov 17;17(22) [FREE Full text] [doi: 10.3390/ijerph17228525] [Medline: 33212979]
14. Zhu D, Paige SR, Slone H, Gutierrez A, Lutzky C, Hedriana H, et al. Exploring telemental health practice before, during, and after the COVID-19 pandemic. J Telemed Telecare 2021 Jul 09:1357633X211025943. [doi: 10.1177/1357633X211025943] [Medline: 34241545]

15. Malouff TD, TerKonda SP, Knight D, Abu Dabrh AM, Perlman AI, Munipalli B, et al. Physician satisfaction with telemedicine during the COVID-19 pandemic: The Mayo Clinic Florida experience. Mayo Clin Proc Innov Qual Outcomes 2021 Aug;5(4):771-782 [FREE Full text] [doi: 10.1016/j.mayocpiqo.2021.06.006] [Medline: 34226884]
16. Pierce BS, Perrin PB, Tyler CM, McKee GB, Watson JD. The COVID-19 telepsychology revolution: a national study of pandemic-based changes in U.S. mental health care delivery. Am Psychol 2021 Jan;76(1):14-25. [doi: 10.1037/amp0000722] [Medline: 32816503]
17. Hu PJ, Chau PY, Sheng ORL, Tam KY. Examining the technology acceptance model using physician acceptance of telemedicine technology. J Manag Inf Syst 2015 Dec 02;16(2):91-112. [doi: 10.1080/07421222.1999.11518247]
18. Chismar W, Wiley-Patton S. Does the extended technology acceptance model apply to physicians. 2003 Presented at: 36th Annual Hawaii International Conference on System Sciences; January 06-09, 2003; Big Island, HI. [doi: 10.1109/hicss.2003.1174354]
19. Kamal SA, Shafiq M, Kakria P. Investigating acceptance of telemedicine services through an extended technology acceptance model (TAM). Technol Soc 2020 Feb;60:101212. [doi: 10.1016/j.techsoc.2019.101212]
20. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 1989 Sep;13(3):319. [doi: 10.2307/249008]
21. Venkatesh V, Morris M, Davis G, Davis F. User acceptance of information technology: toward a unified view. MIS Quarterly 2003;27(3):425 [FREE Full text] [doi: 10.2307/30036540]
22. Hamzat S, Manawonku I. Influence of performance expectancy and facilitating conditions on use of digital library by engineering lecturers in universities in south-west, Nigeria. Libr Philos Pract 2018:1-16 [FREE Full text]
23. Rho MJ, Kim HS, Chung K, Choi IY. Factors influencing the acceptance of telemedicine for diabetes management. Cluster Comput 2014 Mar 12;18(1):321-331. [doi: 10.1007/s10586-014-0356-1]
24. 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]
25. McKee GB, Pierce BS, Donovan EK, Perrin PB. Examining models of psychologists' telepsychology use during the COVID-19 pandemic: a national cross-sectional study. J Clin Psychol 2021 Oct;77(10):2405-2423 [FREE Full text] [doi: 10.1002/jclp.23173] [Medline: 34028018]
26. Asua J, Orruño E, Reviriego E, Gagnon MP. Healthcare professional acceptance of telemonitoring for chronic care patients in primary care. BMC Med Inform Decis Mak 2012 Nov 30;12:139 [FREE Full text] [doi: 10.1186/1472-6947-12-139] [Medline: 23194420]
27. Orruño E, Gagnon MP, Asua J, Ben Abdeljelil A. Evaluation of teledermatology adoption by health-care professionals using a modified technology acceptance model. J Telemed Telecare 2011;17(6):303-307. [doi: 10.1258/jtt.2011.101101] [Medline: 21844171]
28. Lu J, Yao JE, Yu C. Personal innovativeness, social influences and adoption of wireless Internet services via mobile technology. J Strateg Inf Syst 2005 Sep;14(3):245-268. [doi: 10.1016/j.jsis.2005.07.003]
29. Connolly SL, Miller CJ, Lindsay JA, Bauer MS. A systematic review of providers' attitudes toward telemental health via videoconferencing. Clin Psychol (New York) 2020 Jan 06;27(2) [FREE Full text] [doi: 10.1111/cpsp.12311] [Medline: 35966216]
30. Slone H, Gutierrez A, Lutzky C, Zhu D, Hedriana H, Barrera JF, et al. Assessing the impact of COVID-19 on mental health providers in the southeastern United States. Psychiatry Res 2021 Aug;302:114055 [FREE Full text] [doi: 10.1016/j.psychres.2021.114055] [Medline: 34144509]
31. Venkatesh V, Brown S, Maruping L, Bala H. Predicting different conceptualizations of system use: the competing roles of behavioral intention, facilitating conditions, and behavioral expectation. MIS Quarterly 2008;32(3):483 [FREE Full text] [doi: 10.2307/25148853]
32. Whitten PS, Mackert MS. Addressing telehealth's foremost barrier: provider as initial gatekeeper. Int J Technol Assess Health Care 2005;21(4):517-521. [doi: 10.1017/S0266462305050725] [Medline: 16262977]
33. Cowan KE, McKean AJ, Gentry MT, Hilty DM. Barriers to use of telepsychiatry: clinicians as gatekeepers. Mayo Clin Proc 2019 Dec;94(12):2510-2523. [doi: 10.1016/j.mayocp.2019.04.018] [Medline: 31806104]
34. Wilczewski H, Paige SR, Ong T, Barrera JF, Soni H, Welch BM, et al. Perceptions of telemental health care delivery during COVID-19: a cross-sectional study with providers, February-March 2021. Front Psychiatry 2022 Apr 4;13:855138 [FREE Full text] [doi: 10.3389/fpsyt.2022.855138] [Medline: 35444579]
35. Doxy.me. URL: https://doxy.me/en/ [accessed 2022-03-01]
36. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
37. 2007-16: Demographics of the U.S. Psychology Workforce. American Psychological Association. 2018. URL: https://www.apa.org/workforce/publications/16-demographics [accessed 2022-11-04]
38. Bunnell BE, Kazantzis N, Paige SR, Barrera J, Thakkar RN, Turner D, et al. Provision of care by "Real World" telemental health providers. Front Psychol 2021;12:653652 [FREE Full text] [doi: 10.3389/fpsyg.2021.653652] [Medline: 34025519]
39. Bangor A, Kortum P, Miller J. An empirical evaluation of the system usability scale. Int J Hum-Comput Interact 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]

40. Monthuy-Blanc J, Bouchard S, Maïano C, Séguin M. Factors influencing mental health providers' intention to use telepsychotherapy in First Nations communities. Transcult Psychiatry 2013 Apr;50(2):323-343. [doi: 10.1177/1363461513487665] [Medline: 23666941]
41. Rouidi M, Elouadi A, Hamdoune A. Acceptance and use of telemedicine technology by health professionals: development of a conceptual model. Digit Health 2022;8:20552076221081693 [FREE Full text] [doi: 10.1177/20552076221081693] [Medline: 35223077]
42. Newman L, Bidargaddi N, Schrader G. Service providers' experiences of using a telehealth network 12 months after digitisation of a large Australian rural mental health service. Int J Med Inform 2016 Oct;94:8-20. [doi: 10.1016/j.ijmedinf.2016.05.006] [Medline: 27573307]
43. Garavand A, Aslani N, Nadri H, Abedini S, Dehghan S. Acceptance of telemedicine technology among physicians: a systematic review. Informatics in Medicine Unlocked 2022;30:100943. [doi: 10.1016/j.imu.2022.100943]
44. Lopez A, Schwenk S, Schneck CD, Griffin RJ, Mishkind MC. Technology-based mental health treatment and the impact on the therapeutic alliance. Curr Psychiatry Rep 2019 Jul 08;21(8):76. [doi: 10.1007/s11920-019-1055-7] [Medline: 31286280]

Abbreviations

MH: mental health
SUS: System Usability Scale
TAM: Technology Acceptance Model
TMH: telemental health
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by G Eysenbach; submitted 16.05.22; peer-reviewed by M Mishkind, B Li, T Vlavianos, N Dunn; comments to author 21.09.22; revised version received 12.10.22; accepted 31.10.22; published 11.11.22.

Please cite as:
Wilczewski H, Paige SR, Ong T, Soni H, Barrera JF, Welch BM, Bunnell BE
Providers' Perspectives on Telemental Health Usage After the COVID-19 Pandemic: Retrospective Analysis
JMIR Form Res 2022;6(11):e39634
URL: https://formative.jmir.org/2022/11/e39634
doi:10.2196/39634
PMID:36322787

Original Paper

The Emotional Anatomy of the Wuhan Lockdown: Sentiment Analysis Using Weibo Data

Xi Chen[1], BSSc; Michelle Yik[1], PhD

Division of Social Science, Hong Kong University of Science and Technology, Hong Kong, China

Corresponding Author:
Michelle Yik, PhD
Division of Social Science
Hong Kong University of Science and Technology
Clear Water Bay, Kowloon
Hong Kong
China
Phone: 852 23587815
Email: Michelle.Yik@ust.hk

Abstract

Background: On January 23, 2020, the city of Wuhan, China, was sealed off in response to the COVID-19 pandemic. Studies have found that the lockdown was associated with both positive and negative emotions, although their findings are not conclusive. In these studies, emotional responses to the Wuhan lockdown were identified using lexicons based on limited emotion types.

Objective: This study aims to map Chinese people's emotional responses to the Wuhan lockdown and compare Wuhan residents' emotions with those of people elsewhere in China by analyzing social media data from Weibo using a lexicon based on the circumplex model of affect.

Methods: Social media posts on Weibo from 2 weeks before to 2 weeks after the Wuhan lockdown was imposed (January 9, 2020, to February 6, 2020) were collected. Each post was coded using a valence score and an arousal score. To map emotional trajectories during the study period, we used a data set of 359,190 posts. To compare the immediate emotional responses to the lockdown and its longer-term emotional impact on Wuhan residents (n=1236) and non-Hubei residents (n=12,714), we used a second data set of 57,685 posts for multilevel modeling analyses.

Results: Most posts (248,757/359,190, 69.25%) made during the studied lockdown period indicated a pleasant mood with low arousal. A gradual increase in both valence and arousal before the lockdown was observed. The posts after the lockdown was imposed had higher valence and arousal than prelockdown posts. On the day of lockdown, the non-Hubei group had a temporarily boosted valence (γ_{20}=0.118; SE 0.021; *P*<.001) and arousal (γ_{30}=0.293; SE 0.022; *P*<.001). Compared with non-Hubei residents, the Wuhan group had smaller increases in valence (γ_{21}=−0.172; SE 0.052; *P*<.001) and arousal (γ_{31}=−0.262; SE 0.053; *P*<.001) on the day of lockdown. Weibo users' emotional valence (γ_{40}=0.000; SE 0.001; *P*=.71) and arousal (γ_{40}=0.001; SE 0.001; *P*=.56) remained stable over the 2 weeks after the lockdown was imposed regardless of geographical location (valence: γ_{41}=−0.004, SE 0.003, and *P*=.16; arousal: γ_{41}=0.003, SE 0.003, and *P*=.26).

Conclusions: During the early stages of the pandemic, most Weibo posts indicated a pleasant mood with low arousal. The overall increase in the posts' valence and arousal after the lockdown announcement might indicate collective cohesion and mutual support in web-based communities during a public health crisis. Compared with the temporary increases in valence and arousal of non-Hubei users on the day of lockdown, Wuhan residents' emotions were less affected by the announcement. Overall, our data suggest that Weibo users were not influenced by the lockdown measures in the 2 weeks after the lockdown announcement. Our findings offer policy makers insights into the usefulness of social connections in maintaining the psychological well-being of people affected by a lockdown.

(JMIR Form Res 2022;6(11):e37698) doi:10.2196/37698

KEYWORDS

Wuhan lockdown; COVID-19; public health emergency; emotion; circumplex model of affect; Weibo; jiayou

Introduction

Background

The COVID-19 pandemic broke out in China in December 2019, with the city of Wuhan as its epicenter. To contain the spread of the virus, the Chinese government put Wuhan into lockdown on January 23, 2020, by imposing restrictions on travel to and from the city [1]. More rigorous lockdown measures, including limits on outdoor activities and community containment, were gradually implemented in the following weeks [2]. Although lockdowns can effectively limit the spread of a virus [3], these measures seriously interfere with citizens' daily routines and social interactions [4]. Studies have found that travel restrictions and social distancing can lead to negative emotions such as distress, anger, and fear [5,6]. In this study, we sought to comprehensively map the emotional trajectories of Chinese people during the Wuhan lockdown using Weibo posts as our data source.

Understanding people's emotions during public health emergencies is important. At the individual level, emotional experiences are related to mental health. Negative emotions (eg, hostility and fear) during the pandemic were found to mediate the positive correlation between exposure to stressful events (eg, experiencing a lockdown and witnessing the death of loved ones) and symptoms of anxiety and depression [7]. Positive emotions, by contrast, were positively correlated with resilience, and the correlation was stronger among people who had experienced a higher level of negative emotions in the preceding week [8,9]. High-arousal emotions (eg, anger and fear) were found to be associated with poor sleep quality, stress, and anxiety during the pandemic [10,11]. Although emotions with low arousal (eg, calmness and boredom) make people feel relaxed, they have been found to be associated with depression [12] and a loss of meaning in life [13]. People's emotions may also contribute to behavioral outcomes in the context of public emergencies. For example, there is evidence that pleasant (eg, happiness) and unpleasant (eg, fear and anxiety) emotions were related to compliance with social distancing measures during the pandemic [14-17]. Heffner et al [18] found that pleasant emotions with higher arousal elicited by prosocial messages about COVID-19 were correlated with greater compliance with social distancing. The unrelenting and repetitive living situation was found to be associated with emotions with low arousal (eg, boredom) during the pandemic and stimulus-seeking behaviors such as impulsive buying after the pandemic [19]. Although it is clear that lockdown measures affected people's emotions, research findings are far from conclusive.

Why Study the Wuhan Lockdown?

Although many countries implemented lockdown measures during the pandemic [20], people's emotional responses following the Wuhan lockdown were unique for several reasons. First, the Wuhan lockdown occurred at the beginning of the pandemic. Owing to the public's lack of knowledge and preparation for the outbreak of a novel disease, the abrupt lockdown triggered particularly strong emotional reactions, including fear, anxiety, and hopelessness [21]. Second, the lockdown measures in China were more strictly implemented than anywhere else in the world because a wide range of local institutions (eg, residents' committees and property management companies) were mobilized [22]. Third, the Wuhan lockdown was implemented just before the Chinese New Year, a major holiday in Chinese communities. The mobility restrictions prevented individuals from participating in significant activities such as family reunions, leading to intense emotional responses (eg, anxiety, disappointment, anger, and hostility) and worsened psychological well-being [2,4,23,24].

Studying the emotional responses of Wuhan residents can be important as those who were physically restricted by the lockdown measures were more emotionally vulnerable to pandemic-related events. Studies have found that residents of Wuhan and the rest of the Hubei province were sensitive to milestone events (eg, the lockdown announcement on January 23, 2020, and the death of Dr Li Wenliang on February 7, 2020) and expressed more negative emotions after the lockdown was imposed than people from elsewhere in China [2,4,25,26]. In this study, we used sentiment analysis of Weibo data and the circumplex model of affect to examine the emotional responses of Chinese people in general and those living in Wuhan in particular.

Why Examine the Wuhan Lockdown Using Social Media Posts?

During the pandemic, social media became a major source of information about the public's emotions in response to lockdowns [27]. Unlike surveys, which rely on retrospective self-reports, social media data include real-time spontaneous feelings from a wide variety of people and are free from recall bias and demand characteristics [25,28,29]. Starting from the outbreak of COVID-19, a massive wave of discussions took place on Sina Weibo, one of China's most popular social media platforms [2,27]. In March 2022, Weibo had 582 million active users per month and 252 million active users per day [30]. Emotion information embedded in posts on Weibo provides a unique source of data for researchers to investigate emotions in response to the Wuhan lockdown. Therefore, in this study, we used Weibo posts to map the emotional trajectories of Chinese people before and after the lockdown was imposed.

Using social media data, recent studies have found an increase in negative emotions such as fear, panic, guilt, anger, and disappointment [2,4,25,26,28,31] as well as an increase in positive emotions such as feeling blessedness, hope, admiration, and encouragement after the Wuhan lockdown [2,26,32]. It appears that Chinese people experienced increases in both pleasant and unpleasant emotions after the lockdown was imposed. However, researchers have used several discrete emotion categories to code social media data. For example, Su et al [26] used the negative emotion and affect categories from the Simplified Chinese Linguistic Inquiry and Word Count lexicon [31]. Su et al [28] and Yu et al [25] adopted "depression," "like," "dislike," "anger," "sadness," "fear," "enjoyment," "disgust," and "surprise" from the Affective Lexicon Ontology [33]. Shen et al [2] manually coded emotion words from corpora as "anger," "fear," "neutral," "encouragement," and "hope." Cao et al [32] used the Ortony, Clore, and Collins model and the 6 emotions by Ekman to

categorize emotions as “joy,” “hope,” “distress,” “fear,” “admiration,” “reproach,” and “neutral.” The limited emotion types used in previous studies might not be sufficient to construct a complete picture of emotional responses during the Wuhan lockdown. Applying a sentiment analysis based on the circumplex model of affect, this study sought to map the comprehensive emotion trajectories of Chinese people during the Wuhan lockdown.

The Circumplex Model of Affect

The circumplex model of affect defines emotion in a Cartesian space with 2 independent dimensions and has received support in various cultural settings [34-37]. The horizontal axis, valence, denotes the degree of pleasantness and unpleasantness, and the vertical axis, arousal, denotes the degree of activation and deactivation. An emotion is composed of different levels of valence and arousal. For instance, although both *calmness* and *encouragement* are pleasant emotions, the former is less activated than the latter. Compared with categorical models in which emotions are restricted to a certain number of categories, the circumplex model allows for a potentially infinite number of emotions, thus enabling a more fine-grained analysis of emotion in sentiment analysis [38].

Although previous research findings have shown that both pleasant and unpleasant emotions were related to compliance with social distancing during the pandemic, the arousal dimension of the circumplex model of affect has been largely overlooked. For example, Shen et al [2] sorted emotions into “positive,” “negative,” and “neutral” categories, neglecting the arousal dimension. Researchers have mainly focused on high-arousal unpleasant emotions such as fear, panic, and anger [28] while neglecting low-arousal emotions such as calmness and glumness [36]. To obtain a full picture of the emotional trajectories of Chinese people in response to the Wuhan lockdown, we used a lexicon based on the circumplex model of affect for sentiment analysis in this study.

Objectives

Understanding citizens’ emotional responses to extreme social distancing measures such as a lockdown provides insights into how governments can effectively respond to a pandemic. The Wuhan lockdown certainly served as a catalyst for research on mapping the emotions of those affected by lockdown measures, but the findings remain far from conclusive. The aim of this study was to examine the emotional trajectories expressed on Weibo using a sentiment analysis approach. Specifically, we sought to answer three research questions (RQs): (1) What were the emotional trajectories of Chinese people during the Wuhan lockdown? (RQ 1); (2) Were there any differences in immediate emotional responses to the lockdown between people from Wuhan and those from non-Hubei areas? (RQ 2); and (3) Were there any differences in the longer-term impact of the lockdown on emotional responses between people from Wuhan and those from non-Hubei areas? (RQ 3).

Methods

Data Sets

Overview

To explore the emotional responses of Chinese people, we extracted 2 data sets from Weibo, each using a different set of keywords. From both data sets, we extracted posts created from 2 weeks before to 2 weeks after the lockdown announcement (ie, January 9, 2020, to February 6, 2020) and removed posts that yielded missing valence and arousal values. When the posts contained words not coded in the lexicon, missing values were assigned.

We included the 2 weeks before the lockdown in the analysis to create a full picture of prelockdown emotional trajectories. The limit of 2 weeks after the lockdown announcement was chosen to avoid the potential influence of the death of whistleblower Dr Li Wenliang on February 7, 2020 [25]. This event was accompanied by a dramatic increase in the number of posts and content characterized by negative emotions such as anger, fear, surprise, and depression [4,25,27], which could bias the emotional trajectories immediately following the lockdown announcement.

Data Set 1

The data in data set 1 were collected using Weibo’s built-in advanced search tool, through which the search can be optimized by specifying keywords, time frames, and categories (eg, original posts or reposts). A crawler was developed using the *httr* and *rvest* packages in R (R Foundation for Statistical Computing) [39,40]. The keyword “武汉” (“Wuhan”) was used in the search. The data set was created in June 2020 and contained 602,737 original posts made by 396,054 users from January 1, 2020 to February 15, 2020. After removing posts with missing values and extracting posts created from 2 weeks before to 2 weeks after the lockdown announcement, the final data set included 359,190 posts by 242,023 users.

Data Set 2

The data in data set 2 were extracted from a publicly available data set named Weibo-COV2 that contains >65 million Weibo posts made by 20 million active users from December 2019 to December 2020 [41]. The authors of Weibo-COV2 defined active users as those with >50 followers and 50 fans who had made a post within the 30 days before the data were collected. Inactive users were excluded as they did not make any posts during the pandemic. The authors first built a pool of active users whose posts were screened using 492 pandemic-related keywords. The 492 keywords used by the authors [41] covered events (eg, “武汉封城,” implying “Wuhan lockdown”), drugs and supplies (eg, “口罩,” implying “masks”), experts and physicians (eg, “钟南山,” implying “Zhong Nanshan”), and government policies (eg, “延迟开学,” implying “postponing school reopening”). In this study, we extracted Weibo-COV2 data from 2 weeks before to 2 weeks after the Wuhan lockdown announcement and removed posts with missing valence and arousal values. As we aimed to explore individuals’ emotional responses before and after the lockdown was imposed, data from users who created only 1 post were excluded. Ultimately,

data set 2 contained 974,317 posts—282,442 (28.99%) made before the lockdown announcement and 691,875 (71.01%) made after the lockdown announcement—by 219,446 unique users, each of whom made at least one post before and one post after the lockdown was imposed.

Comparing the 2 Data Sets

Each of our 2 data sets had strengths and limitations. Together, they provided an effective platform to address our 3 RQs.

First, different keywords were used to create each data set. The keyword "Wuhan" for data set 1 allowed us to understand the daily emotional trajectories before pandemic-related topics went viral on Weibo, providing a baseline measure. By contrast, the pandemic-related keywords for data set 2 allowed us to explore people's emotions regarding the pandemic. On January 20, 2020, when human-to-human transmission was confirmed, the number of daily posts on Weibo-COV2 increased dramatically [27,41]. To comprehensively map the emotional trajectory of Chinese people during the lockdown period, it is necessary to use both data sets 1 and 2.

Second, the 2 data sets had different structures. The data in data set 1 were cross-sectional in nature and included posts made by different people over time. Most user IDs (208,389/242,023, 86.1%) had only 1 corresponding post. Such cross-sectional data provide snapshots of Chinese people's emotions at different time points. By contrast, as data set 2 included all pandemic-related posts from each user, we were able to map their emotions before, during, and after the lockdown was imposed.

We used data set 1 to explore emotional trajectories before and after the Wuhan lockdown was imposed (RQ 1). Using data set 2, we compared the immediate emotional responses to the lockdown (RQ 2) and its longer-term emotional impact (RQ 3) for people from Wuhan and non-Hubei people.

Ethical Considerations

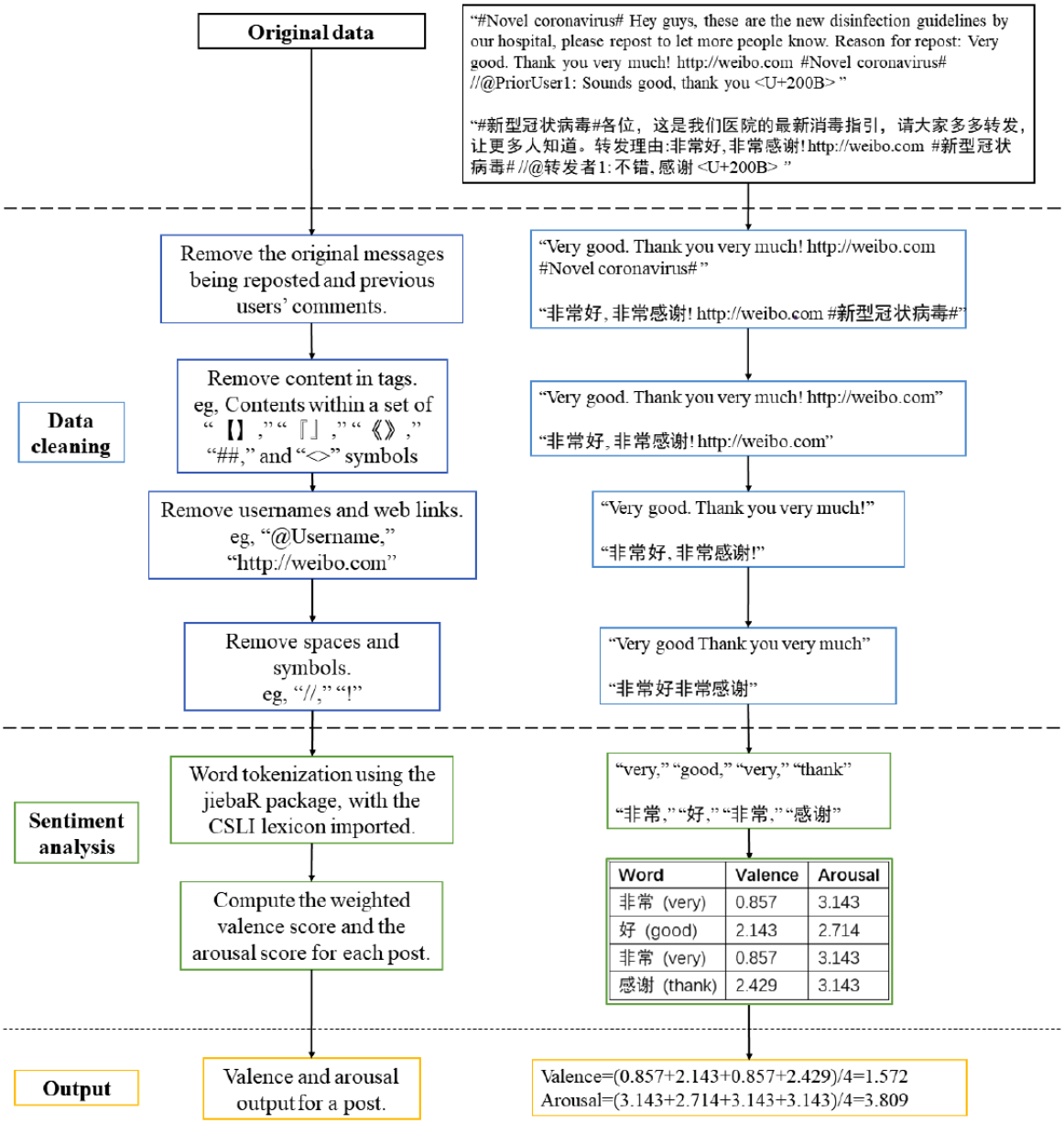
To protect human participants' privacy, all users' personal information was anonymized during our analysis. This study focused on human emotions at the aggregated level, meaning that individual-level information was averaged across users and no individual information was disclosed.

Data Cleaning

We cleaned the posts before assigning emotion scores. Figure 1 summarizes the process of data cleaning and sentiment analysis. Weibo posts have two major formats: original posts and reposts. An original post contains the original comments made by a user. A repost usually has three parts: the original message being reposted, other users' comments about the original post, and the reposting user's comments about the original post. As previous studies found that including the message being reposted in sentiment analysis significantly reduced the unpleasantness of emotions in Weibo posts, we excluded the original message being reposted and previous users' comments and analyzed the comments by the reposting user only [42]. For example, a repost might read, "Hey guys, these are the new disinfection guidelines by our hospital, please repost to let more people know. Reason for repost: Very good. Thank you very much! //@PriorUser1: Sounds good, thank you." In this repost, only the phrase "Very good. Thank you very much!" is an original comment made by the reposting user and would be subjected to our sentiment analysis. However, if the reposting user added no comments, we conducted sentiment analysis on all previous users' comments to represent the reposting user's comment as previous studies have found that Weibo users only repost other users' comments when they agree with them [43,44].

During data cleaning, we first retained the original posts and the comments made by the reposting user. We then removed tags, usernames, and web links from each post as such content would add noise to the analysis. The remaining content of each post was retained for sentiment analysis.

Figure 1. Process of data cleaning and sentiment analysis with post examples. CSLI: Chinese Sentiment Lexicon for Internet.



Sentiment Analysis

A lexicon-based sentiment analysis was conducted to assign each post a valence score and an arousal score based on the Chinese Sentiment Lexicon for Internet (CSLI) [45], which consists of 7088 Chinese words, each of which is annotated with a valence score ranging from −4 to 4 and an arousal score ranging from 0 to 8.

Our sentiment analysis procedure is summarized in Figure 1. The *JiebaR* package in R [46] was used for tokenization, which involved extracting the emotion words listed in the CSLI from each post. Following the procedures by Zhao et al [45], we calculated the valence and arousal scores of each post using the weighted average value. For each post, the valence score of each emotion word was multiplied by its frequency—the products of all emotion words in a post were summed and then divided by the total frequency of all emotion words in the post. The same method was applied to calculate the arousal score. Mean valence and arousal scores were assigned to each post to indicate its emotion. Using Weibo posts, Zhao et al [45] validated this sentiment analysis method by comparing valence and arousal scores rated by participants with the predicted CSLI scores for the same posts. The convergence was 0.70 for valence and 0.59 for arousal. As valence in the CSLI lexicon has a range of −4 to 4, we defined posts with valence scores <0 as *unpleasant* and scores >0 as *pleasant*. Similarly, as arousal has a range of 0 to 8, we defined posts with arousal scores <4 as *low-arousal* and scores >4 as *high-arousal*.

For subsequent analyses, we computed 2 types of daily emotion scores using the valence and arousal scores of posts yielded by the CSLI lexicon. One was daily emotion by person. For example, a daily valence score by person was calculated for a user by averaging the valence scores of all posts made by the user on that day. A daily arousal score by person was calculated in the same way. The other type of daily emotion score was emotion by day. A valence score by day was calculated by averaging all daily valence scores by person across users who

created posts on a particular day. An arousal score by day was calculated in the same way.

Geographical Location: Wuhan Versus Non-Hubei Residents

People who were physically restricted by the lockdown measures might have had different emotional responses from people who were not in the lockdown area. A previous study found that Hubei residents expressed more negative emotions than those living elsewhere in China during the pandemic [2]. To examine the differences in emotions in response to the Wuhan lockdown between Wuhan and non-Hubei areas, we grouped the Weibo users in data set 2 into a Wuhan group (ie, the lockdown area) and a non-Hubei group (ie, nonlockdown area) based on their postlockdown geographical location. The non-Hubei area, rather than the non-Wuhan area, was used as Wuhan was not the only city in the Hubei province that was locked down after January 23, 2020. A total of 15 other cities in Hubei gradually implemented similar restrictions up to January 27, 2020 [47]. Therefore, to compare the pre– and post–lockdown announcement periods, we decided to focus on the non-Hubei rather than the non-Wuhan area.

To differentiate between Wuhan and non-Hubei users, the IDs of users who made posts after the announcement that contained geotags (ie, the GPS coordinates of a user's device, including longitude and latitude) were filtered for geographical information [48]. As people could no longer enter or exit Wuhan after the lockdown was imposed, the users' GPS information enabled us to assess their actual geographical locations after the lockdown was imposed.

We used data set 2 to explore the emotional responses of the Wuhan and non-Hubei groups. We first used the geotags to separate user IDs into Wuhan and non-Hubei groups. All posts with these user IDs were then extracted, resulting in 6665 posts generated by 1236 users in Wuhan and 51,020 posts generated by 12,714 users outside Hubei. These users all made at least one post before and one post after the lockdown was imposed. It is important to note that Weibo users can select geotags from a list of locations, allowing them to misreport their actual locations. This function is an unfortunate drawback of Weibo data and should be treated as a source of error [48].

Multilevel Modeling

We used data set 2 to examine how people's everyday emotions evolved from 2 weeks before to 2 weeks after the lockdown was imposed. As data set 2 had a nested data structure in which repeated measures of emotion ratings were nested within each user, we used multilevel modeling [49]. A multilevel model separates the residual variances of a sample into within-person (level 1) and between-person (level 2) variances [50]. Multilevel modeling allows for data dependency [51,52], meaning that the measures of one person would be more similar to one another than to the measures of another person. In addition, social media data are not always complete, meaning that there would be missing observations if an individual did not post anything on certain days. Multilevel modeling is suitable for social media data as it can tolerate missing observations [51].

Piecewise multilevel models can capture changes in groups during different phases [53] and can therefore be used to compare the changes in emotional trajectories between the Wuhan and non-Hubei groups before and after the lockdown was imposed. The number of posts in data set 2 was small before January 20, 2020, as widespread discussion of pandemic-related topics had not yet started [27,41]. To ensure adequate observations of emotion ratings across users per day (ie, >50 users per day) [54], we used data set 2 data starting from January 20, 2020, for the multilevel modeling. The R package *lme4* [55] was used to estimate the multilevel models.

Results

Emotion Map

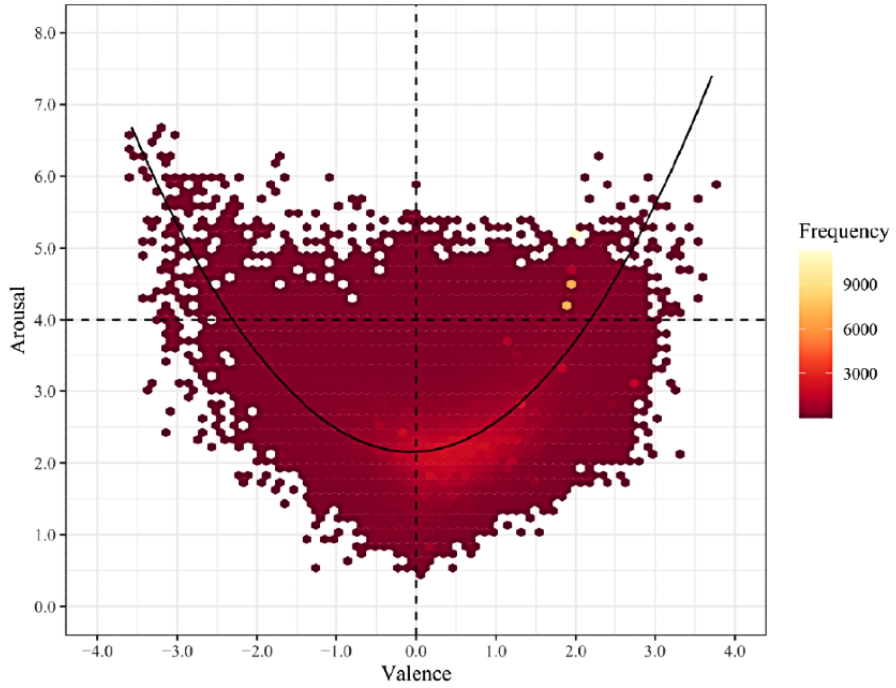
The data in data set 1 included 359,190 posts made by 242,023 users. To map the emotion distribution of Weibo posts, the valence and arousal scores of each post were used to draw the scatterplot shown in Figure 2. The valence and arousal combinations are depicted by their frequency. Posts were classified as *unpleasant* when their valence scores were <0 and as *pleasant* when their valence scores were >0. Posts were classified as *low-arousal* when their arousal scores were <4 and as *high-arousal* when their arousal scores were >4.

Overall, the emotions in the posts spread across the valence-arousal plane, covering both the pleasant and unpleasant areas. Highly frequent combinations of valence and arousal scores were gathered in low-arousal areas and leaned toward the pleasant half of the plane. Several highly frequent valence-arousal combinations were in the pleasant and high-arousal quadrants (valence >0; arousal >4).

Of the 359,190 posts, 248,757 (69.25%) fell in the pleasant and low-arousal quadrant (valence >0; arousal <4), 72,773 (20.26%) fell in the unpleasant and low-arousal quadrant (valence <0; arousal <4), 32,850 (9.15%) fell in the pleasant and high-arousal quadrant (valence >0; arousal >4), and 2023 (0.56%) fell in the unpleasant and high-arousal quadrant (valence <0; arousal >4). The remaining 0.78% (2787/359,190) of posts were neutral in valence or had a moderate arousal of 4. These results showed that, among all the combinations of valence and arousal, most posts (282,068/359,190, 78.53%) made from January 9, 2020, to February 6, 2020, were pleasant, among which there were more low- than high-arousal emotions.

To evaluate the relationship between the valence and arousal of posts on Weibo, a quadratic regression was fitted to the 359,190 posts. The results showed a significant quadratic relationship between valence and arousal ($F_{2,359,187}=1.554\times10^5$; $P<.001$; $R^2=0.464$). The regression equation was arousal = 2.156 + 0.046 × valence + 0.368 × valence^2. As shown in Figure 2, both pleasant and unpleasant emotions were accompanied by activation (refer to the study by Yik et al [37]).

Figure 2. Emotion map of 359,190 Weibo posts by 242,023 users from January 9, 2020, to February 6, 2020 (data set 1).



Daily Emotional Trajectories Before and After the Lockdown Was Imposed

To explore the emotional trajectories before and after the lockdown was imposed (RQ 1), the data in data set 1 were analyzed. We averaged the valence of all posts made on each day and plotted the daily valence scores over the 4 weeks. Figure 3A displays the daily trajectories in valence (−4 to +4). On January 20, 2020, when human-to-human transmission of COVID-19 was confirmed [25], a turning point emerged during which the valence score increased until January 23, 2020, and then reached a plateau and stabilized at a higher level. On February 3, 2020, a day after all patients infected or suspected to be infected with SARS-CoV-2 were required to be quarantined and isolated [25], valence rose to a higher level than the week after the lockdown announcement, reaching a second plateau. The results show that the emotions expressed on Weibo became more positive starting on January 20, 2020, 3 days before the lockdown. After the lockdown was imposed, a higher level of pleasant emotions was expressed, although valence was generally pleasant (valence >0) at a low level (arousal <4). Two valence plateaus were observed: one from January 23, 2020, to February 2, 2020, and one from February 3 to 6, 2020.

Similarly, we averaged the arousal of all posts made on each day and plotted the daily arousal scores over the 4 weeks. Figure 3B displays the daily trajectory of arousal scores (0 to 8). The plot shows that the turning point of arousal occurred on January 19, 2020, after which arousal increased until January 23, 2020, and then reached a first plateau and became stable. From February 3, 2020, arousal rose to a higher level than the week after the lockdown was imposed and reached a second plateau. These results show that the emotions expressed by Weibo users became more activated during the lockdown, although the daily arousal level was generally low. Two arousal plateaus were observed: one from January 23, 2020, to February 2, 2020, and one from February 3 to 6, 2020.

To investigate the reasons for the 2 valence and arousal plateaus during the lockdown, we conducted separate lexical analyses on the word frequencies across all words that were used during the 2 plateau periods. Word clouds, in which more frequently used words appear in larger font sizes, were used to visualize word frequency and demonstrate the main themes of the posts [56]. The 100 most frequently used words in the 2 weeks before the lockdown (January 9, 2020, to January 22, 2020) are illustrated in Figure 4A. The word 兄弟 (*brothers*) was the most commonly used word in the 2 weeks before the lockdown as fans of a music band named 摩登兄弟 (*Modern Brothers*) were doing promotion activities. This result provides a picture of Weibo topics before the lockdown, during which people on Weibo engaged in celebrity promotion activities and discussed topics other than the COVID-19 pandemic [57,58].

As shown in Figure 4B, of the 100 most frequently used words in the 2 weeks after the lockdown announcement (January 23, 2020, to February 6, 2020), 加油 (*jiayou* or *add oil*, which is a term of encouragement) had the highest frequency, accounting for 4.32% (81,969/1,895,476) of all emotion words found in the posts. The prevalence of *jiayou* increased to 6.48% (57,239/883,447) from February 3 to 6, 2020. Notably, *jiayou* accounted for only 0.53% (6734/1,277,449) of emotion words in the 2 weeks before the lockdown. In the CSLI lexicon, *jiayou* has a valence score of 2.000 (ranging from −4 to +4) and an arousal score of 5.143 (ranging from 0 to 8). It is likely that the escalating valence of Weibo posts during the lockdown can be attributed to an increased use of words of encouragement such as *jiayou*.

Figure 3. Cross-sectional emotion trajectories from January 9, 2020, to February 6, 2020 (data set 1).

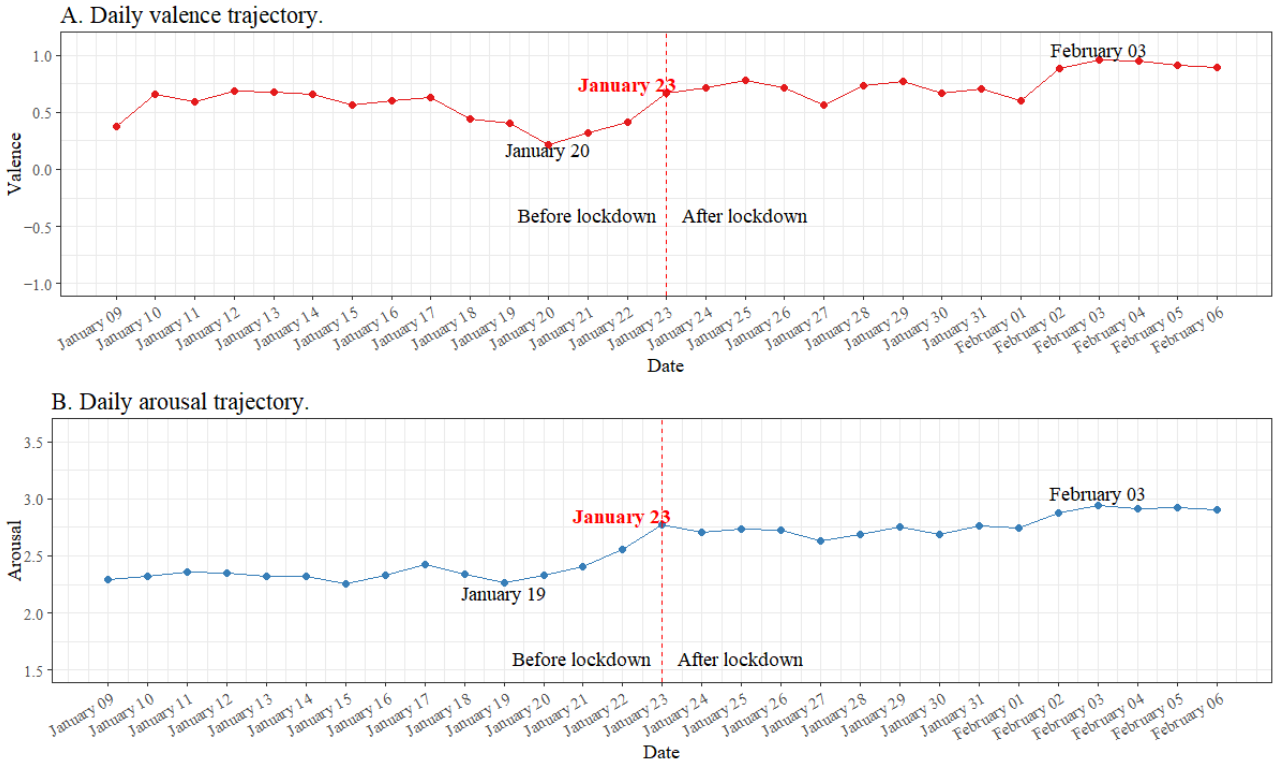


Figure 4. Word cloud of the (A) 100 most frequently used words in the 2 weeks before the lockdown and the (B) 100 most frequently used words in the 2 weeks after the lockdown was imposed (data set 1).

The Emotional Trajectories of Wuhan and Non-Hubei Residents

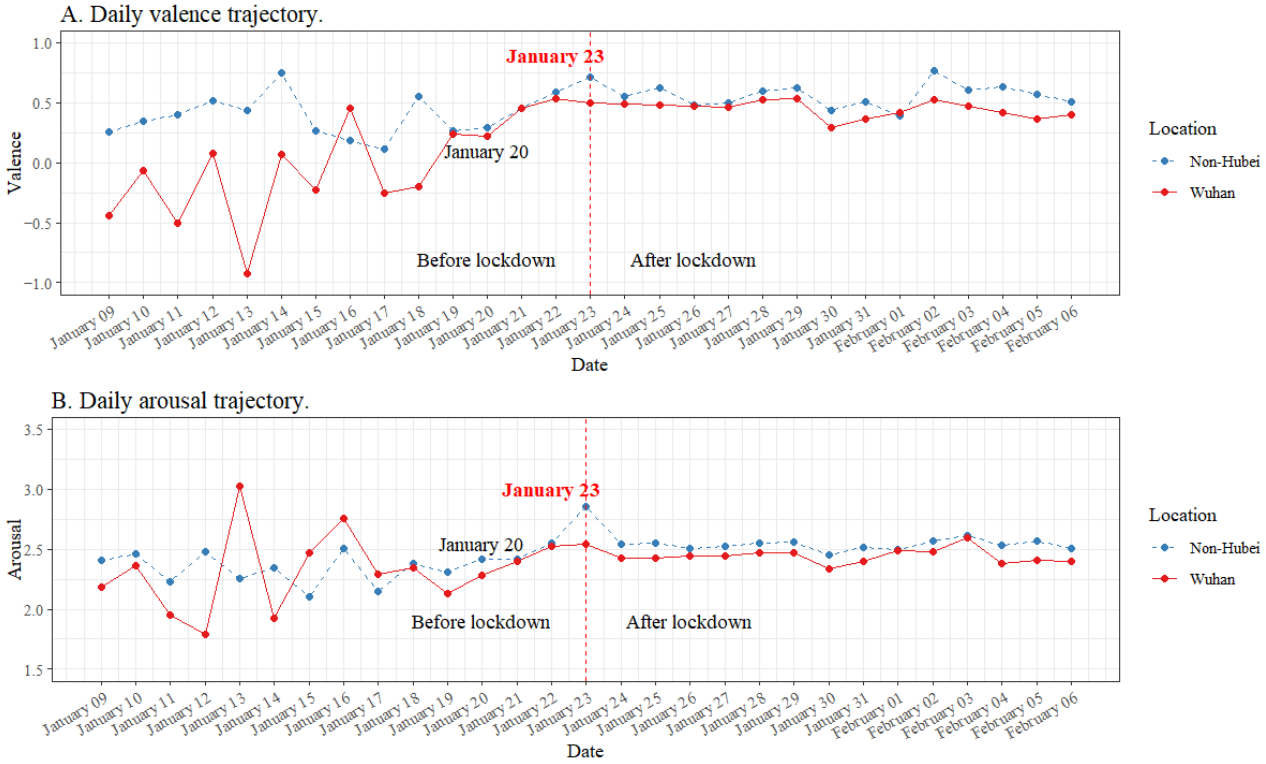
Data set 2 included 6665 posts made by 1236 Wuhan residents and 51,020 posts made by 12,714 non-Hubei residents from 2 weeks before to 2 weeks after the lockdown announcement. Each user made at least one post before and one post after the announcement. To examine the daily valence trajectories of Wuhan and non-Hubei residents, we averaged the daily valence across users on each day to generate daily valence scores for the Wuhan and non-Hubei groups. The daily arousal trajectories of Wuhan and non-Hubei residents were generated using the

same method. Figure 5A shows the daily valence trajectories of the Wuhan group and the non-Hubei group during the same period. Before human-to-human transmission of the virus was confirmed on January 20, 2020, the valence of the Wuhan group fluctuated more dramatically than that of the non-Hubei group, possibly because the data points were dominated by a small amount of data across individuals per day. Before January 20, 2020, the number of people who created posts per day in the Wuhan group ranged from 1 to 45, whereas the number of people who created posts per day in the non-Hubei group ranged from 11 to 91.

From January 20, 2020, the number of people who made posts per day dramatically increased in both the Wuhan and non-Hubei groups to a minimum of 162 individuals per day, which is consistent with the findings of Lu et al [27] and Hu et al [41], who reported a dramatic increase in Weibo-COV2 data after January 20, 2020. From January 20, 2020, the daily valence of the Wuhan and non-Hubei groups became parallel and increased steadily until the lockdown announcement on January 23, 2020, when the valence of the non-Hubei group peaked but that of the Wuhan group remained at a similar level to that of the previous day. After January 23, 2020, the daily valence of both groups increased and became similar across days. Notably, unlike the non-Hubei group, the Wuhan group did not express more pleasant emotions on the day of lockdown.

Figure 5B shows a plot of daily arousal trajectories of the Wuhan and non-Hubei groups during the same period. The trajectories of both groups became parallel and increased after January 20, 2020, before diverging on January 23, 2020, when arousal peaked in the non-Hubei group but remained steady in the Wuhan group. After the lockdown was imposed on January 23, 2020, the daily arousal trajectories of both groups stabilized at a higher level and became parallel across days. Unlike the non-Hubei group, the Wuhan group did not express more activated emotions on the day of lockdown.

Figure 5. Daily valence and arousal trajectories in the Wuhan and non-Hubei groups from January 9, 2020, to February 6, 2020 (data set 2).



Immediate Emotional Responses of Wuhan and Non-Hubei Weibo Users to the Wuhan Lockdown

Preliminary Analysis

To compare the immediate emotional responses of Wuhan and non-Hubei residents to the lockdown (RQ 2), multilevel modeling was conducted using data set 2. The descriptive statistics are provided in Table 1. The skewness of the valence scores ranged from −0.188 to 1.633, and the kurtosis ranged from 0.892 to 3.982, both of which are within the acceptable range for data normality (−2 to +2 for skewness and −7 to +7 for kurtosis) [59,60]. The skewness of the arousal scores ranged from 1.393 to 1.633, and the kurtosis ranged from 2.499 to 3.982, both of which are also within the acceptable range for data normality.

Table 1. Descriptive statistics of valence and arousal by person by day for the Wuhan and non-Hubei groups (data set 2).

Emotional response and group	Users, N	Mean (SD)	Skew	Kurtosis
Valence (−4 to +4)				
Non-Hubei	12,714	0.555 (0.913)	−0.188	0.892
Wuhan	1236	0.453 (0.845)	0.004	1.220
Arousal (0 to 8)				
Non-Hubei	12,714	2.544 (0.815)	1.393	2.499
Wuhan	1236	2.446 (0.718)	1.633	3.982

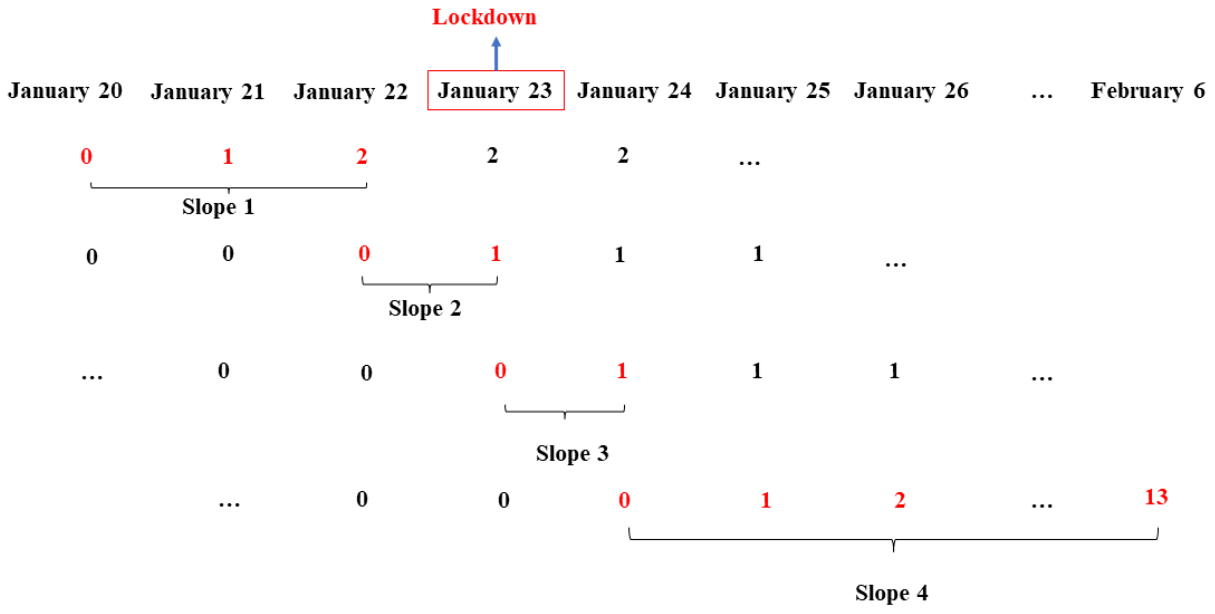
Multilevel Modeling

To examine the proportion of variance of valence accounted for by between- and within-person levels, we first used an empty multilevel model with random intercept only [51]. The results indicated a significant between-person variance of 0.332 (χ^2_4=1187.2; *P*<.001). The intraclass correlation coefficient was 0.134, implying that the between-person variance accounted for 13.40% of the variance. The same analysis was repeated for the arousal scores. The results indicated a significant between-person variance of 0.252 (χ^2_4=699.6; *P*<.001). The intraclass correlation coefficient was 0.098, implying that the between-person variance in arousal accounted for 9.77% of the variance. These results justified multilevel analysis at both level 1 (within-person) and level 2 (between-person).

To compare the emotional responses to the lockdown of Wuhan and non-Hubei residents (RQ 2), we introduced 4 slopes into the multilevel model. To ensure adequate observations of emotion ratings across users per day (ie, >50 users per day) [54], we used data after January 20, 2020, in data set 2. Slope 1 modeled the emotional trajectories between January 20, 2020, and January 22, 2020, just before the lockdown. To compare the immediate emotional responses to the lockdown of Wuhan and non-Hubei residents, slopes 2 and 3 were introduced into the model. Slope 2 tested whether the emotions on the day of lockdown (January 23, 2020) were different from the emotions on the day before (January 22, 2020); slope 3 tested whether the emotions on the day of lockdown (January 23, 2020) were different from the emotions the following day (January 24, 2020). To compare the longer-term emotional impact on Wuhan and non-Hubei residents after the lockdown was imposed, we used slope 4 to model emotional changes in the 2 weeks after the lockdown was imposed (ie, between January 24, 2020, and February 6, 2020). Figure 6 summarizes the coding scheme for the 4 slopes. The dates on or before the start date of a slope were coded as 0, and the dates after the start date of a slope were coded with a sequential integer depending on the distance from the start date. For example, as slope 4 started on January 24, 2020, the dates on or after January 24 were coded as 1 for January 25, 2 for January 26, and so on.

The outcomes of all models were tested using the daily emotion data (ie, either valence or arousal) of each person. Whenever location was included as a between-person (level 2) predictor, it was dummy coded as 1, denoting the Wuhan group, or 0, denoting the non-Hubei group.

Figure 6. Coding scheme for slopes 1 to 4 in multilevel modeling.



The random effects of the slopes were introduced into the models based on the RQs. To compare the immediate emotional responses to the lockdown of the Wuhan and non-Hubei groups from the day before to the day after the lockdown was imposed,

we estimated the random effects (ie, variability across users in Wuhan and non-Hubei groups) of slope 2 (the emotional trend from January 22, 2020, to January 23, 2020) and slope 3 (the emotional trend from January 23, 2020, to January 24, 2020). Location was introduced as the level-2 predictor. In this model, we examined the fixed effects of slope 1 (January 20, 2020, to January 22, 2020) and slope 4 (2 weeks after the lockdown was imposed) to test the overall changes in both the Wuhan and non-Hubei groups. In all the models, *i* denoted an individual, and *t* denoted a date. Multilevel model 1 is shown in Textbox 1.

Textbox 1. Multilevel model 1 for immediate emotional responses.

Multilevel model 1

- Level 1:
  - $Emotion_{it} = \beta_{0i} + \beta_{1i} \times slope\ 1 + \beta_{2i} \times slope\ 2 + \beta_{3i} \times slope\ 3 + \beta_{4i} \times slope\ 4 + e_{it}$
- Level 2:
  - $\beta_{0i} = \gamma_{00} + \gamma_{01} \times location + U_{0i}$
  - $\beta_{1i} = \gamma_{10}$
  - $\beta_{2i} = \gamma_{20} + \gamma_{21} \times location + U_{2i}$
  - $\beta_{3i} = \gamma_{30} + \gamma_{31} \times location + U_{3i}$
  - $\beta_{4i} = \gamma_{40}$

Valence

The results of the multilevel modeling with valence as the outcome variable are summarized in the upper panel of Table 2. Before the lockdown, the overall valence of Wuhan and non-Hubei residents yielded a positive coefficient of 0.123, implying an increase in average valence between January 20, 2020, and January 22, 2020 (SE 0.013; *P*<.001).

As location was introduced as a moderator for slopes 2 and 3 in multilevel model 1, the main effects of slopes 2 and 3 indicated the emotion trends conditional at a location of 0 (ie, the non-Hubei group). To test the emotional changes in response to the lockdown, we relied on slope 2. The main effect of slope 2 was positive (γ_{20}=0.118; SE 0.021; *P*<.001), implying that the valence score of the non-Hubei group increased from the day before. The random effect of slope 2 was not significant (SD 0.259; χ^2_3=3.8; *P*=.28), implying that there were no significant variations in slope 2 across individuals in either the Wuhan or non-Hubei group. However, the interaction between slope 2 and location yielded a significant result of −0.172 (SE 0.052; *P*=.001). This significant moderation of location with an insignificant random effect can be interpreted as a systematic difference between the Wuhan and the non-Hubei groups [53]. Although there was no difference in the valence trend of the individuals in the Wuhan and non-Hubei groups, the mean valence increase was higher in the non-Hubei group than in the Wuhan group. In other words, the non-Hubei residents reacted to the lockdown more positively than the Wuhan residents.

To test the emotional change immediately after the lockdown was imposed, we relied on slope 3. The main effect of slope 3 was negative (γ_{30}=−0.146; SE 0.020; *P*<.001), indicating that the valence score of the non-Hubei residents decreased after the day of lockdown. The random effect of slope 3 was significant (SD 0.260; χ^2_3=24.7; *P*<.001), implying that the slope varied across individuals in both the Wuhan and non-Hubei groups. Location significantly moderated slope 3 and, therefore, the valence score of the Wuhan users did not decrease as much as that of the non-Hubei users (γ_{31}=0.109; SE 0.047; *P*=.02). In other words, the valence score of the non-Hubei residents decreased more than that of the Wuhan residents.

Table 2. Immediate emotional responses based on a multilevel model predicting valence or arousal with slope 2 and slope 3 at level 1 by location at level 2 (data set 2).

Variable	Fixed effect			Random effect		
	Coefficient	SE	*P* value	SD	Chi-square (*df*)[a]	*P* value
Valence						
Intercept (γ_{00})	0.337	0.022	<.001	0.452	656.8 (3)	<.001
Slope 1 (γ_{10})[b]	0.123	0.013	<.001	—[c]	—	—
Slope 2 (γ_{20})[d]	0.118	0.021	<.001	0.259	3.8 (3)	.28
Slope 3 (γ_{30})[e]	−0.146	0.020	<.001	0.260	24.7 (3)	<.001
Slope 4 (γ_{30})[f]	0.000	0.001	.99	—	—	—
Location (γ_{01})[g]	−0.037	0.032	.24	—	—	—
Slope 2×location (γ_{21})[d,g]	−0.172	0.052	.001	—	—	—
Slope 3×location (γ_{31})[e,g]	0.109	0.047	.02	—	—	—
Arousal						
Intercept (γ_{00})	2.365	0.019	<.001	0.265	183.3 (3)	<.001
Slope 1 (γ_{10})[b]	0.089	0.011	<.001	—	—	—
Slope 2 (γ_{20})[d]	0.293	0.022	<.001	0.569	333.2 (3)	<.001
Slope 3 (γ_{30})[e]	−0.315	0.021	<.001	0.607	403.4 (3)	<.001
Slope 4 (γ_{40})[f]	0.000	0.001	.30	—	—	—
Location (γ_{01})[g]	−0.042	0.026	.11	—	—	—
Slope 2×location (γ_{21})[d,g]	−0.262	0.053	<.001	—	—	—
Slope 3×location (γ_{31})[e,g]	0.218	0.050	<.001	—	—	—

[a]The chi-square value is calculated as the difference in –2loglikelihood between a model with and a model without the random effect.

[b]Slope 1: emotional trend before the lockdown (between January 20, 2020, and January 22, 2020).

[c]The cells were kept empty when the random effect of a variable was not included in the model.

[d]Slope 2: difference in emotion between January 23, 2020 (the day of lockdown) and January 22, 2020.

[e]Slope 3: difference in emotion between January 23, 2020 (the day of lockdown) and January 24, 2020.

[f]Slope 4: longer-term emotional trend in the 2 weeks after the lockdown was imposed.

[g]Location: dummy coded as 1 denoting the Wuhan group and 0 denoting the non-Hubei group.

Arousal

The results of the multilevel modeling with arousal as the outcome variable are summarized in the lower panel of Table 2. From January 20, 2020, to January 22, 2020, the overall arousal trend of the Wuhan and non-Hubei groups yielded a positive coefficient (γ_{10}=0.089; SE 0.011; *P*<.001), indicating an increase in average arousal.

The main effect of slope 2 was positive (γ_{20}=0.293; SE 0.022; *P*<.001), indicating that the arousal score of the non-Hubei group on the day of lockdown increased from the day before. The random effect of slope 2 was significant (SD 0.569; χ^2_3=333.2; *P*<.001), implying that slope 2 varied across individuals in both the Wuhan and non-Hubei groups. Location significantly moderated slope 2 (γ_{21}=−0.262; SE 0.053; *P*<.001), indicating that the non-Hubei group reacted to the lockdown with a greater increase in arousal than that of the Wuhan users.

The main effect of slope 3 had a coefficient of −0.315 (SE 0.021; *P*<.001), indicating that the arousal score of the non-Hubei residents decreased after the day of lockdown. The random effect of slope 3 was significant (SD 0.607; χ^2_3=403.4; *P*<.001), indicating that the slope varied across individuals in both the Wuhan and non-Hubei groups. Location significantly moderated slope 3 (γ_{31}=0.218; SE 0.050; *P*<.001), indicating that the non-Hubei group reacted to the lockdown with a greater drop in arousal than that of the Wuhan users.

Taken together, the results indicate that the non-Hubei users' emotions were influenced by the Wuhan lockdown and became more pleasant and aroused. However, the effect was temporary.

Both the valence and arousal of the non-Hubei group decreased immediately after the day of lockdown, and the magnitude of the decreases in the non-Hubei group was significantly greater than that in the Wuhan group. In other words, the emotions of Wuhan residents were less affected by the lockdown.

Longer-term Emotional Impact of the Wuhan Lockdown on Wuhan and Non-Hubei Weibo Users

Multilevel Modeling

To examine the longer-term emotional impact of the lockdown (RQ 3), we compared the emotional trajectories in the 2 weeks after the lockdown was imposed (ie, between January 24, 2020, and February 6, 2020) of the Wuhan and non-Hubei groups. We included all 4 slopes tested in the previous model. We tested the random effect of slope 4 (ie, variability across users in the Wuhan and non-Hubei groups). Location was introduced as the level-2 predictor. We also tested the fixed effects of the remaining slopes. Multilevel model 2 is shown in Textbox 2.

Textbox 2. Multilevel model 2 for longer-term emotional impact.

Multilevel model 2

- Level 1:
  - $\text{Emotionit} = \beta_{0i} + \beta_{1i} \times \text{slope 1} + \beta_{2i} \times \text{slope 2} + \beta_{3i} \times \text{slope 3} + \beta_{4i} \times \text{slope 4} + e_{it}$
- Level 2:
  - $\beta_{0i} = \gamma_{00} + \gamma_{01} \times \text{locationi} + U_{0i}$
  - $\beta_{1i} = \gamma_{10}$
  - $\beta_{2i} = \gamma_{20}$
  - $\beta_{3i} = \gamma_{30}$
  - $\beta_{4i} = \gamma_{40} + \gamma_{41} \times \text{locationi} + U_{4i}$

Valence

The results of multilevel model 2 with valence as the outcome variable are summarized in the upper panel of Table 3. As location was introduced as the moderator for slope 4 in multilevel model 2, the main effects of slope 4 indicated emotional trends conditional on a location of 0 (ie, the non-Hubei group). From January 24, 2020, the main effect of slope 4 for valence was 0.000 (SE 0.001; *P*=.71), indicating no valence change on average for the non-Hubei group in the 2 weeks after the lockdown was imposed. The random effect of slope 4 was significant (SD 0.021; χ^2_2=15.8; *P*<.001), indicating that the slope varied across individuals in both the Wuhan and non-Hubei groups. However, location did not moderate slope 4 (γ_{41}=−0.004; SE 0.003; *P*=.16), indicating that, although people differed in the longer-term valence trend, the difference was not explained by the location difference between the Wuhan and non-Hubei groups.

Table 3. Longer-term emotional impact based on a multilevel model predicting valence or arousal with slope 4 at level 1 by location at level 2 (data set 2).

Variable	Fixed effect			Random effect		
	Coefficient	SE	*P* value	SD	Chi-square (*df*)[a]	*P* value
Valence						
Intercept (γ_{00})	0.339	0.021	<.001	0.356	946.8 (2)	<.001
Slope 1 (γ_{10})[b]	0.124	0.013	<.001	—[c]	—	—
Slope 2 (γ_{20})[d]	0.094	0.019	<.001	—	—	—
Slope 3 (γ_{30})[e]	−0.128	0.018	<.001	—	—	—
Slope 4 (γ_{40})[f]	0.000	0.001	.71	0.021	15.8 (2)	<.001
Location (γ_{01})[g]	−0.080	0.020	<.001	—	—	—
Slope 4×location (γ_{41})[f,g]	−0.004	0.003	.16	—	—	—
Arousal						
Intercept (γ_{00})	2.378	0.018	<.001	0.257	500.2 (2)	<.001
Slope 1 (γ_{10})[b]	0.085	0.011	<.001	—	—	—
Slope 2 (γ_{20})[d]	0.261	0.017	<.001	—	—	—
Slope 3 (γ_{30})[e]	−0.284	0.016	<.001	—	—	—
Slope 4 (γ_{40})[f]	0.001	0.001	.56	0.016	5.9 (2)	.05
Location (γ_{01})[g]	−0.109	0.017	<.001	—	—	—
Slope 4×location (γ_{41})[f,g]	0.003	0.003	.26	—	—	—

[a]The chi-square value is calculated as the difference in –2loglikelihood between a model with and a model without the random effect.

[b]Slope 1: emotional trend before the lockdown (between January 20, 2020, and January 22, 2020).

[c]The cells were kept empty when the random effect of a variable was not included in the model.

[d]Slope 2: difference in emotion between January 23, 2020 (the day of lockdown) and January 22, 2020.

[e]Slope 3: difference in emotion between January 23, 2020 (the day of lockdown) and January 24, 2020.

[f]Slope 4: longer-term emotional trend in the 2 weeks after the lockdown was imposed.

[g]Location: dummy coded as 1 denoting the Wuhan group and 0 denoting the non-Hubei group.

Arousal

The results of the multilevel model with arousal as the outcome variable are summarized in the lower panel of Table 3. From January 24, 2020, the main effect of slope 4 was 0.001 (SE 0.001; *P*=.56), indicating no significant arousal change in the non-Hubei group in the 2 weeks after the lockdown was imposed. The random effect of slope 4 was insignificant (SD 0.016; χ^2_2=5.9; *P*=.05), implying that there was not enough evidence to suggest that the slope varied across individuals in both the Wuhan and non-Hubei groups. Location did not significantly moderate slope 4 (γ_{41}=0.003; SE 0.003; *P*=.26), indicating that there was no difference in longer-term arousal changes between the Wuhan and non-Hubei groups.

Overall, there is not enough evidence to suggest that longer-term changes in emotional valence and arousal occurred in the 2 weeks after the lockdown was imposed, regardless of geographical area.

Discussion

Principal Findings

Using the Wuhan lockdown as our context, we sought to understand the interplay between lockdown measures and the emotions of residents in different areas of China. Our results showed that, during the lockdown, most posts on Weibo were pleasant, with low arousal. Compared with posts before the lockdown, posts after the lockdown was imposed had higher valence and arousal, indicating more pleasant and activated emotions. Word cloud analysis revealed an increased use of the encouragement word *jiayou* after the lockdown was imposed, which might account for the increase in valence and arousal. Non-Hubei users' valence and arousal were significantly higher on the day of lockdown than on the day before and after the lockdown was imposed. In contrast, Wuhan residents showed little immediate change in emotion in response to and after the Wuhan lockdown was imposed. Overall, in the 2 weeks after the lockdown was imposed, the valence and arousal scores of

Weibo users remained constant regardless of their geographical location.

In contrast to studies that observed more negative emotions during the lockdown, our overall emotion trajectories suggest that Weibo users demonstrated more pleasant and activated emotions after the lockdown was imposed. A possible explanation for the discrepancies between our results and those of previous studies is the differences in word choice when coding the data. Studies that reported more negative emotions after the lockdown was imposed tended to use emotion types such as "stress," "hostility," "disappointment," "surprise," "fear," "guilt," and "blame" [2,4,25,28,32], and those that reported positive emotions after the lockdown was imposed tended to include emotion types such as "encouragement," "admiration," "hope," and "blessedness" [2,28,32]. The word choices could have biased their results toward conclusions that were determined by the word categories chosen. Instead of focusing on a few selected emotion categories, we chose a lexicon based on the circumplex model of affect in which its two dimensions (valence and arousal) underlie most—if not all—emotions [34-36]. In this model, emotions are composed of valence and arousal, meaning that both pleasant and unpleasant emotions with different levels of arousal are covered.

Unlike non-Hubei people, who were significantly affected, people in Wuhan experienced smaller immediate changes in emotions in response to the lockdown event. Unlike the temporary peak in valence and arousal found in non-Hubei areas on the day of lockdown, there was a minimal change in valence and arousal for Wuhan people. This result appears to support the *psychological typhoon eye effect*, which describes the phenomenon of people closer to the epicenter of a devastating event having less intense psychological reactions to it [61]. Zhang et al [62] concluded that people closer to the outbreak displayed fewer mental health problems in response to the COVID-19 pandemic as they had a more accurate estimation of the situation and were desensitized by repeated exposure to stress.

Although the emotions of Wuhan people were more stable than those of non-Hubei residents on the day of lockdown, their emotional changes were similar to those of non-Hubei people over the longer term. Overall, Chinese Weibo users, regardless of whether they were in lockdown or not, expressed stable emotions with higher valence and arousal in the 2 weeks after the lockdown was imposed than before the lockdown. This result is in stark contrast to that of the study by Zhao et al [4], who found fewer positive emotion words (eg, "happiness") and more negative emotion words (eg, "fear" and "pain") in Wuhan than in non-Hubei areas after the lockdown was imposed. The difference may be rooted in the fact that, unlike our study, which selected data from up to 2 weeks after the lockdown was imposed (ie, January 23, 2020, to February 6, 2020), their data covered a longer postlockdown period (January 23, 2020, to February 16, 2020) that included the significant event of the death of Dr Li Wenliang on February 7, 2020. Li's death was accompanied by dramatic increases in the number of posts and expression of negative emotions such as "anger," "fear," "surprise," and "depression" [4,25,27], and the inclusion of this event may have negatively biased the emotional trajectories following the lockdown announcement.

We did not observe a systematic difference in valence and arousal scores between the Wuhan and non-Hubei groups after controlling for the moderating effect of location for slopes 2 and 3. That is, the average score of valence and arousal by person by day did not differ between people in or out of lockdown. This finding is different from a recent study by Meock et al [63] that compared the emotional responses to the Australian lockdown of people who were in and out of lockdown. The participants (both in and not in lockdown) in that study recorded their emotions 7 to 9 times per day for 1 week. Those in lockdown had slightly more negative emotions and slightly fewer positive emotions than those not in lockdown. A possible explanation for the discrepancy in results is the timing of data collection as Meock et al [63] collected data at a later stage of the Australian lockdown. Our study focused on the first 2 weeks of the Wuhan lockdown, during which the emotional consequences might not yet have emerged.

It is reasonable to expect negative emotions during a public health emergency such as the COVID-19 pandemic, but the power of positive emotions should not be overlooked. Our results highlight collective positive emotional responses during a public health emergency. Although Weibo users made both pleasant and unpleasant posts, their average daily emotions were pleasant and of low arousal throughout the 4 weeks under scrutiny. Unpleasant posts only accounted for 20.82% (74,796/359,190) of all posts during the lockdown. These findings are similar to those of Pan et al [64], who found that >50% of comments on Weibo during the pandemic were positive, and Arora et al [29], who found that happiness was the most prevalent emotion expressed by Indian Twitter users during the lockdown, accounting for 40% of users' posts. Maintaining a positive emotional state may contribute to better mental health and lead to better coping when facing stressful events [62,65]. Positive feelings such as "calmness," "hope," "love," and "gratitude" during the COVID-19 pandemic have been found to be positively correlated with a resilience mindset, and the relationship is stronger among people with a higher level of negative emotions (eg, "anger," "anxiety," "regret," and "sadness") than among people with low levels of negative emotions [9].

Why Was the Lockdown Positively Received?

Why did Weibo users in China respond positively to the adverse situation after the lockdown was imposed? The cognitive processes of people in Eastern cultures are driven by Eastern naïve dialecticism, suggesting a tendency to see good in the bad and bad in the good [66,67]. Therefore, from a cross-cultural perspective, Chinese people are more likely than their Western counterparts to embrace adversity during a pandemic. Studies have found cultural differences in emotional responses during the pandemic. Yap et al [68] collected data on affect, optimism, well-being, and meaning in life from China and Canada in March 2020 and April 2020. Compared with the Canadian participants, the Chinese participants reported more positive and fewer negative emotions, greater optimism, greater psychological well-being, and greater sense of meaning during

the pandemic. Our findings of higher valence and arousal after the lockdown was imposed, along with an abundance of encouragement words such as *jiayou*, add new evidence that Chinese people reacted more optimistically to the outbreak of COVID-19. Future studies could collect social media data from different countries to compare how people with different cultural backgrounds react to public health emergencies.

People tend to bond with strangers after an exogenous crisis, creating a sense of belonging to emergent groups [69,70]. Such collective cohesion and shared identity enable mutual support in adverse situations such as pandemics [71]. Research findings have also shown that a collective experience of positive emotions contributed to more resilience and better mental health during the COVID-19 pandemic [8]. Similar collective behavior was found during the Wuhan lockdown. In in-depth interviews by Qian and Hanser [21], Wuhan residents reported that they soon adapted to reality upon the abrupt announcement of the lockdown, with mutual support among neighbors on WeChat groups and in communities. Our results add further evidence for such collective behavior, with people on Weibo shouting *jiayou* together to support each other through the unfolding uncertainties of the pandemic. This finding is in line with the suggestion of Elcheroth and Drury [71] that the continuity of social ties might help people manage negative feelings during social crises such as pandemics and lockdowns.

Implications and Limitations

Although the emotion trajectories of the Wuhan lockdown have been examined in previous studies, they have focused on limited types of emotions. Instead of coding social media data using limited types of emotions, we explored emotional responses to lockdown using a lexicon based on the valence and arousal dimensions of the circumplex model of affect [34-36]. Hence, our coding was not limited to a certain number of emotions. Rather, we sought to cover most if not all emotions.

The current study offers significant practical implications. In addition to providing an emotion map describing the valence and arousal of Weibo users, we offer explanations for the changes in emotional response patterns. We highlight the power of web-based social cohesion, which may explain the stability of Wuhan residents' emotional states when the lockdown was imposed and the 2 weeks afterward. Such findings provide insights for policy makers to consider the reasonable length of lockdown measures.

However, our study is not without limitations. The first is the potential influence of censorship on our data sets. Posts that violate Weibo regulations are usually removed within 24 hours of their creation, meaning that real-time data scraping is crucial to understand the complete emotional picture on the web [27]. The observation that censored posts from the day of lockdown had a higher degree of criticism than of support whereas uncensored posts had similar amounts of criticism and support [27] might suggest that the real valence after the lockdown was imposed was lower than that detected in our study. To overcome this issue, future studies could consider using Weibo data collected in real time or within 24 hours after an event. Furthermore, as some posts contained words that were not coded in the CSLI lexicon, 7.38% (28,601/387,791) of the posts in data set 1 were excluded from the sentiment analysis. These posts may have contained emotion information. In the future, researchers could use machine learning and deep learning algorithms to capture the emotions embedded in the words that were not coded in the lexicon and map out emotional trajectories [2,32].

Conclusions

In this study, we applied a circumplex model of affect on 2 data sets of Weibo posts to explore emotional responses before and after the Wuhan lockdown was imposed. Our results suggest that, even while facing the adverse situation of a lockdown, the overall aggregated emotion expressed on Weibo was pleasant and of low arousal. Over time, the web-based discussion included more encouraging messages. In addition, compared with people from more distant areas, Wuhan residents had more stable emotions on the day of lockdown, but they showed similar changes to non-Hubei residents in the 2 weeks after the lockdown was imposed. Interestingly, the lockdown did not seem to subdue the feelings of Wuhan residents despite the fact that they were physically constrained. The increased valence and arousal of posts after the lockdown was imposed might suggest a collective cohesion that provides mutual support for people under adverse situations. Policy makers could consider various approaches to forming and maintaining social connections for mutual encouragement and better adjustment during lockdowns and public health emergencies.

Acknowledgments

The preparation of the manuscript was facilitated by the General Research Fund of the Hong Kong Research Grants Council (projects 16601818 and AoE/E-603/18). The authors thank Zoe Franklin, Felity Kwok, Ines Lin, and James Pho for their help in preparing this manuscript.

Data Availability

The raw data collected for this study are not publicly available as they contain Weibo users' personal information.

Conflicts of Interest

None declared.

References

1. Lau H, Khosrawipour V, Kocbach P, Mikolajczyk A, Schubert J, Bania J, et al. The positive impact of lockdown in Wuhan on containing the COVID-19 outbreak in China. J Travel Med 2020 May 18;27(3):taaa037 [FREE Full text] [doi: 10.1093/jtm/taaa037] [Medline: 32181488]
2. Shen L, Yao R, Zhang W, Evans R, Cao G, Zhang Z. Emotional attitudes of Chinese citizens on social distancing during the COVID-19 outbreak: analysis of social media data. JMIR Med Inform 2021 Mar 16;9(3):e27079 [FREE Full text] [doi: 10.2196/27079] [Medline: 33724200]
3. Atalan A. Is the lockdown important to prevent the COVID-19 pandemic? Effects on psychology, environment and economy-perspective. Ann Med Surg (Lond) 2020 Aug;56:38-42 [FREE Full text] [doi: 10.1016/j.amsu.2020.06.010] [Medline: 32562476]
4. Zhao N, Wang Y, Li S, Liu X, Wu P, Zhu T. Psychological and behavioral impact of Wuhan lockdown and suggestions. Bull Chinese Acad Sci 2020;35(3):264-272 [FREE Full text]
5. Xiang YT, Yang Y, Li W, Zhang L, Zhang Q, Cheung T, et al. Timely mental health care for the 2019 novel coronavirus outbreak is urgently needed. Lancet Psychiatry 2020 Mar;7(3):228-229 [FREE Full text] [doi: 10.1016/S2215-0366(20)30046-8] [Medline: 32032543]
6. Mukhtar S. Psychological health during the coronavirus disease 2019 pandemic outbreak. Int J Soc Psychiatry 2020 Aug;66(5):512-516 [FREE Full text] [doi: 10.1177/0020764020925835] [Medline: 32434402]
7. Zhao N, Zhou G. Social media use and mental health during the COVID-19 pandemic: moderator role of disaster stressor and mediator role of negative affect. Appl Psychol Health Well Being 2020 Dec;12(4):1019-1038 [FREE Full text] [doi: 10.1111/aphw.12226] [Medline: 32945123]
8. Waters L, Algoe SB, Dutton J, Emmons R, Fredrickson BL, Heaphy E, et al. Positive psychology in a pandemic: buffering, bolstering, and building mental health. J Posit Psychol 2022;17(3):303-323. [doi: 10.1080/17439760.2021.1871945]
9. Israelashvili J. More positive emotions during the COVID-19 pandemic are associated with better resilience, especially for those experiencing more negative emotions. Front Psychol 2021 May 5;12:648112 [FREE Full text] [doi: 10.3389/fpsyg.2021.648112] [Medline: 34054649]
10. Basishvili T, Oniani N, Sakhelashvili I, Eliozishvili M, Khizanashvili M, Arabidze M, et al. Insomnia, pre-sleep arousal, psychosocial factors and changes in sleep pattern during the second wave lockdown of the COVID-19 pandemic in Georgia. Brain Sci 2021 Dec 24;12(1):17 [FREE Full text] [doi: 10.3390/brainsci12010017] [Medline: 35053761]
11. Gorgoni M, Scarpelli S, Mangiaruga A, Alfonsi V, Bonsignore MR, Fanfulla F, Board of the Italian Association of Sleep Medicine (AIMS). Pre-sleep arousal and sleep quality during the COVID-19 lockdown in Italy. Sleep Med 2021 Dec;88:46-57 [FREE Full text] [doi: 10.1016/j.sleep.2021.10.006] [Medline: 34731828]
12. Picard RW, Fedor S, Ayzenberg Y. Multiple arousal theory and daily-life electrodermal activity asymmetry. Emot Rev 2015 Mar 02;8(1):62-75. [doi: 10.1177/1754073914565517]
13. Laurensia Hartoyo A, Wijaya E. Correlation between boredom and meaning of life in late adolescence during the Covid-19 pandemic. In: Proceedings of the 3rd Tarumanagara International Conference on the Applications of Social Sciences and Humanities. 2021 Presented at: TICASH '21; August 5-6, 2021; Jakarta, Indonesia. [doi: 10.2991/assehr.k.220404.240]
14. Krekel C, Swanke S, De Neve J, Fancourt D. Are happier people more compliant? Global evidence from three large-scale surveys during Covid-19 lockdowns. SSRN J 2020 Sep 15:13690. [doi: 10.2139/ssrn.3691403]
15. Min C, Shen F, Yu W, Chu Y. The relationship between government trust and preventive behaviors during the COVID-19 pandemic in China: exploring the roles of knowledge and negative emotion. Prev Med 2020 Dec;141:106288 [FREE Full text] [doi: 10.1016/j.ypmed.2020.106288] [Medline: 33091414]
16. Harper CA, Satchell LP, Fido D, Latzman RD. Functional fear predicts public health compliance in the COVID-19 pandemic. Int J Ment Health Addict 2021;19(5):1875-1888 [FREE Full text] [doi: 10.1007/s11469-020-00281-5] [Medline: 32346359]
17. Leung GM, Ho LM, Chan SK, Ho SY, Bacon-Shone J, Choy RY, et al. Longitudinal assessment of community psychobehavioral responses during and after the 2003 outbreak of severe acute respiratory syndrome in Hong Kong. Clin Infect Dis 2005 Jun 15;40(12):1713-1720. [doi: 10.1086/429923] [Medline: 15909256]
18. Heffner J, Vives ML, FeldmanHall O. Emotional responses to prosocial messages increase willingness to self-isolate during the COVID-19 pandemic. Pers Individ Dif 2021 Feb 15;170:110420 [FREE Full text] [doi: 10.1016/j.paid.2020.110420] [Medline: 33082614]
19. Deng S, Wang W, Xie P, Chao Y, Zhu J. Perceived severity of COVID-19 and post-pandemic consumption willingness: the roles of boredom and sensation-seeking. Front Psychol 2020 Sep 16;11:567784 [FREE Full text] [doi: 10.3389/fpsyg.2020.567784] [Medline: 33041933]
20. Imran AS, Daudpota SM, Kastrati Z, Batra R. Cross-cultural polarity and emotion detection using sentiment analysis and deep learning on COVID-19 related tweets. IEEE Access 2020 Sep 28;8:181074-181090 [FREE Full text] [doi: 10.1109/ACCESS.2020.3027350] [Medline: 34812358]
21. Qian Y, Hanser A. How did Wuhan residents cope with a 76-day lockdown? Chinese Sociol Rev 2021;53(1):55-86. [doi: 10.1080/21620555.2020.1820319]
22. Ren X. Pandemic and lockdown: a territorial approach to COVID-19 in China, Italy and the United States. Eurasian Geogr Econ 2020 May 08;61(4-5):423-434. [doi: 10.1080/15387216.2020.1762103]

23. Birditt KS, Turkelson A, Fingerman KL, Polenick CA, Oya A. Age differences in stress, life changes, and social ties during the COVID-19 pandemic: implications for psychological well-being. Gerontologist 2021 Feb 23;61(2):205-216 [FREE Full text] [doi: 10.1093/geront/gnaa204] [Medline: 33346806]
24. Marchetti M, Gatti D, Inguscio L, Mazzoni G. Psychological well-being and lockdown: a study on an Italian sample during the first COVID-19 wave. Psychology Hub 2022 Apr 5;39(1):17-24. [doi: 10.31234/osf.io/mnu7e]
25. Yu S, Eisenman D, Han Z. Temporal dynamics of public emotions during the COVID-19 pandemic at the epicenter of the outbreak: sentiment analysis of weibo posts from Wuhan. J Med Internet Res 2021 Mar 18;23(3):e27078 [FREE Full text] [doi: 10.2196/27078] [Medline: 33661755]
26. Su Y, Xue J, Liu X, Wu P, Chen J, Chen C, et al. Examining the impact of COVID-19 lockdown in Wuhan and Lombardy: a psycholinguistic analysis on weibo and Twitter. Int J Environ Res Public Health 2020 Jun 24;17(12):4552 [FREE Full text] [doi: 10.3390/ijerph17124552] [Medline: 32599811]
27. Lu Y, Pan J, Xu Y. Public sentiment on Chinese social media during the emergence of COVID19. J Qual Description 2021 Apr 26;1(2021):1-47. [doi: 10.51685/jqd.2021.013]
28. Su Y, Wu P, Li S, Xue J, Zhu T. Public emotion responses during COVID-19 in China on social media: an observational study. Human Behav Emerg Tech 2021 Jan;3(1):127-136. [doi: 10.1002/hbe2.239]
29. Arora A, Chakraborty P, Bhatia MP, Mittal P. Role of emotion in excessive use of Twitter during COVID-19 imposed lockdown in India. J Technol Behav Sci 2021;6(2):370-377 [FREE Full text] [doi: 10.1007/s41347-020-00174-3] [Medline: 33102690]
30. Weibo Corporation. Weibo reports first quarter 2020 unaudited financial results. Cision Distribution. 2022 Jun 1. URL: https://www.prnewswire.com/news-releases/weibo-reports-first-quarter-2022-unaudited-financial-results-301558699.html [accessed 2022-08-04]
31. Gao R, Hao B, Li H, Gao Y, Zhu T. Developing simplified chinese psychological linguistic analysis dictionary for microblog. In: Proceedings of the 2013 International Conference on Brain and Health Informatics. 2013 Presented at: BHI '13; October 29-31, 2013; Maebashi, Japan p. 359-368. [doi: 10.1007/978-3-319-02753-1_36]
32. Cao G, Shen L, Evans R, Zhang Z, Bi Q, Huang W, et al. Analysis of social media data for public emotion on the Wuhan lockdown event during the COVID-19 pandemic. Comput Methods Programs Biomed 2021 Nov;212:106468 [FREE Full text] [doi: 10.1016/j.cmpb.2021.106468] [Medline: 34715513]
33. Xu L, Lin H, Pan Y, Ren H, Chen J. Constructing the affective lexicon ontology. J China Soc Scientific Tech Inf 2008;27(2):180-185.
34. Yik M. Studying affect among the Chinese: the circular way. J Pers Assess 2009 Sep;91(5):416-428. [doi: 10.1080/00223890903087596] [Medline: 19672748]
35. Russell JA. A circumplex model of affect. J Pers Soc Psychol 1980 Dec;39(6):1161-1178. [doi: 10.1037/h0077714]
36. Yik M, Russell JA, Steiger JH. A 12-point circumplex structure of core affect. Emotion 2011 Aug;11(4):705-731. [doi: 10.1037/a0023980] [Medline: 21707162]
37. Yik M, Mues C, Sze IN, Kuppens P, Tuerlinckx F, De Roover K, et al. On the relationship between valence and arousal in samples across the globe. Emotion (forthcoming) 2022 Apr 21. [doi: 10.1037/emo0001095] [Medline: 35446055]
38. Yu LC, Lee LH, Hao S, Wang J, He Y, Hu J, et al. Building chinese affective resources in valence-arousal dimensions. In: Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2016 Presented at: NAACL '16; June 12-17, 2016; San Diego, CA, USA p. 540-545. [doi: 10.18653/v1/n16-1066]
39. Wickham H. Package ‘rvest’. The Comprehensive R Archive Network. 2022 Aug 19. URL: https://cran.r-project.org/web/packages/rvest/rvest.pdf [accessed 2022-08-04]
40. Wickham H. Package 'httr'. The Comprehensive R Archive Network. 2022 Aug 17. URL: https://cran.r-project.org/web/packages/httr/index.html [accessed 2022-08-04]
41. Hu Y, Huang H, Chen A, Mao XL. Weibo-COV: a large-scale COVID-19 social media dataset from Weibo. In: Proceedings of the 1st Workshop on NLP for COVID-19 (Part 2) at EMNLP 2020. 2020 Dec Presented at: EMNLP '20; December 2020; Online. [doi: 10.18653/v1/2020.nlpcovid19-2.34]
42. Li Y, Gao X, Du M, He R, Yang S, Xiong J. What causes different sentiment classification on social network services? Evidence from weibo with genetically modified food in China. Sustainability 2020 Feb 12;12(4):1345. [doi: 10.3390/su12041345]
43. Liu W, Zhang M. Semi-supervised sentiment classification method based on weibo social relationship. In: Proceedings of the 16th International Conference on Web Information Systems and Applications. 2019 Presented at: WISA '19; September 20-22, 2019; Qingdao, China p. 480-491. [doi: 10.1007/978-3-030-30952-7_47]
44. West R, Paskov HS, Leskovec J, Potts C. Exploiting social network structure for person-to-person sentiment analysis. Trans Assoc Comput Linguist 2014 Dec;2:297-310. [doi: 10.1162/tacl_a_00184]
45. Zhao JL, Li MZ, Yao J, Qin GH. The development of the Chinese sentiment lexicon for Internet. Front Psychol 2019 Nov 5;10:2473 [FREE Full text] [doi: 10.3389/fpsyg.2019.02473] [Medline: 31749746]
46. Qin W, Wu Y. Package ‘jiebaR’. The Comprehensive R Archive Network. 2019 Dec 13. URL: https://cran.r-project.org/web/packages/jiebaR/jiebaR.pdf [accessed 2022-08-04]

47. The last prefectural-level city in Hubei province is locked down when the Xiangyang train station is closed. International Finance News. 2020. URL: https://www.thepaper.cn/newsDetail_forward_5671283 [accessed 2022-08-04]
48. Jendryke M, Balz T, Liao M. Big location-based social media messages from China's Sina Weibo network: collection, storage, visualization, and potential ways of analysis. Trans GIS 2017 Mar 10;21(4):825-834. [doi: 10.1111/tgis.12266]
49. Lafit G, Adolf J, Dejonckheere E, Myin-Germeys I, Viechtbauer W, Ceulemans E. Selection of the number of participants in intensive longitudinal studies: a user-friendly shiny app and tutorial to perform power analysis in multilevel regression models that account for temporal dependencies. Adv Methods Pract Psychol Sci 2021 Mar 23;4(1):2515245920978738. [doi: 10.31234/osf.io/dq6ky]
50. Reis HT, Gable SL, Maniaci MR. Methods for studying everyday experience in its natural context. In: Reis HT, Judd CM, editors. Handbook of Research Methods in Social and Personality Psychology. Cambridge, UK: Cambridge University Press; 2014:373-403.
51. Snijders TA, Bosker RJ. Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling. Thousand Oaks, CA, USA: Sage Publications; 2011.
52. West SG, Ryu E, Kwok OM, Cham H. Multilevel modeling: current and future applications in personality research. J Pers 2011 Feb;79(1):2-50. [doi: 10.1111/j.1467-6494.2010.00681.x] [Medline: 21223263]
53. Hoffman L. Longitudinal Analysis: Modeling Within-Person Fluctuation and Change. London, UK: Routledge; 2015.
54. Maas CJ, Hox JJ. Sufficient Sample Sizes for Multilevel Modeling. Methodology 2005 Jan;1(3):86-92. [doi: 10.1027/1614-2241.1.3.86]
55. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J Stat Soft 2015 Oct 7;67(1):1-48. [doi: 10.18637/jss.v067.i01]
56. Xu J, Tao Y, Lin H. Semantic word cloud generation based on word embeddings. In: Proceedings of the 2016 IEEE Pacific Visualization Symposium. 2016 Presented at: PacificVis '16; April 19-22, 2016; Taipei, Taiwan p. 239-243. [doi: 10.1109/pacificvis.2016.7465278]
57. Jia X, Hung K, Zhang K. Celebrity fans in China: motives, characteristics, and marketing impacts. In: Wang CL, editor. Handbook of Research on the Impact of Fandom in Society and Consumerism. Hershey, PA, USA: IGI Global; 2020:104-126.
58. Wang H, Lei K, Xu K. Profiling the followers of the most influential and verified users on Sina Weibo. In: Proceedings of the 2015 IEEE International Conference on Communications. 2015 Presented at: ICC '15; June 8-12, 2015; London, UK p. 1158-1163. [doi: 10.1109/icc.2015.7248479]
59. Curran PJ, West SG, Finch JF. The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. Psychol Methods 1996 Mar;1(1):16-29. [doi: 10.1037/1082-989x.1.1.16]
60. West SG, Finch JF, Curran PJ. Structural equation models with nonnormal variables: problems and remedies. In: Hoyle RH, editor. Structural Equation Modeling: Concepts, Issues, and Applications. Thousand Oaks, CA, USA: Sage Publications; 1995:56-75.
61. Li S, Rao LL, Ren XP, Bai XW, Zheng R, Li JZ, et al. Psychological typhoon eye in the 2008 Wenchuan earthquake. PLoS One 2009;4(3):e4964 [FREE Full text] [doi: 10.1371/journal.pone.0004964] [Medline: 19305501]
62. Zhang L, Ma M, Li D, Xin Z. The psychological typhoon eye effect during the COVID-19 outbreak in China: the role of coping efficacy and perceived threat. Global Health 2020 Oct 27;16(1):105 [FREE Full text] [doi: 10.1186/s12992-020-00626-8] [Medline: 33109228]
63. Moeck E, Freeman-Robinson R, O'Brien ST, Woods J, Grewal K, Bagnara L, et al. The emotional anatomy of lockdown. PsyArXiv 2022 Jun 17. [doi: 10.31234/osf.io/hxatc]
64. Pan W, Wang RJ, Dai WQ, Huang G, Hu C, Pan WL, et al. China public psychology analysis about COVID-19 under considering Sina Weibo data. Front Psychol 2021 Sep 8;12:713597 [FREE Full text] [doi: 10.3389/fpsyg.2021.713597] [Medline: 34566790]
65. Moskowitz JT, Cheung EO, Freedman M, Fernando C, Zhang MW, Huffman JC, et al. Measuring positive emotion outcomes in positive psychology interventions: a literature review. Emot Rev 2020 Aug 30;13(1):60-73. [doi: 10.1177/1754073920950811]
66. Peng K, Nisbett RE. Culture, dialectics, and reasoning about contradiction. Am Psychol 1999 Sep;54(9):741-754. [doi: 10.1037/0003-066x.54.9.741]
67. Yik M. Culture, gender, and the bipolarity of momentary affect. Cogn Emot 2007 Apr;21(3):664-680. [doi: 10.1080/02699930600823702]
68. Yap S, Lee A, Ji LJ, Li Y, Dong Y. Cultural differences in people's psychological response to COVID-19. Front Psychol 2021 Jun 12;12:636062 [FREE Full text] [doi: 10.3389/fpsyg.2021.636062] [Medline: 34322051]
69. Drury J, Brown R, González R, Miranda D. Emergent social identity and observing social support predict social support provided by survivors in a disaster: solidarity in the 2010 Chile earthquake. Eur J Soc Psychol 2016 Mar;46(2):209-223. [doi: 10.1002/ejsp.2146]
70. Drury J, Carter H, Cocking C, Ntontis E, Tekin Guven S, Amlôt R. Facilitating collective psychosocial resilience in the public in emergencies: twelve recommendations based on the social identity approach. Front Public Health 2019 Jun 4;7:141 [FREE Full text] [doi: 10.3389/fpubh.2019.00141] [Medline: 31214561]

71. Elcheroth G, Drury J. Collective resilience in times of crisis: lessons from the literature for socially effective responses to the pandemic. Br J Soc Psychol 2020 Jul;59(3):703-713 [FREE Full text] [doi: 10.1111/bjso.12403] [Medline: 32584480]

Abbreviations

CSLI: Chinese Sentiment Lexicon for Internet
RQ: research question

Edited by A Mavragani; submitted 03.03.22; peer-reviewed by L Shen, A Pal; comments to author 09.04.22; revised version received 21.08.22; accepted 19.09.22; published 14.11.22.

Please cite as:
Chen X, Yik M
The Emotional Anatomy of the Wuhan Lockdown: Sentiment Analysis Using Weibo Data
JMIR Form Res 2022;6(11):e37698
URL: https://formative.jmir.org/2022/11/e37698
doi:10.2196/37698
PMID:36166650

Original Paper

Students' Emotional Well-being and Academic Functioning Before, During, and After Lockdown in Germany: Cohort Study

Tania R Nuñez[1], MA; Nina Pallasch[2], BSc; Theda Radtke[1], PhD

[1]Department of Health Psychology and Applied Diagnostics, Institute of Psychology, University of Wuppertal, Wuppertal, Germany

[2]Department of Psychology, Lund University, Lund, Sweden

Corresponding Author:
Tania R Nuñez, MA
Department of Health Psychology and Applied Diagnostics
Institute of Psychology
University of Wuppertal
Gaußstraße 20
Wuppertal, 42119
Germany
Phone: 49 0202 439 2301
Email: nunez@uni-wuppertal.de

Abstract

Background: The COVID-19 lockdowns have led to social detriments and altered learning environments among university students. Recent research indicates that such ramifications may engender various impairments to students' mental health. However, such research has major limitations, such as the lack of a prepandemic control measure, the focus on singular well-being parameters, or the investigation of only the early phases of the pandemic.

Objective: To address these research gaps, this comprehensive and nationwide study compared 3 student cohorts (aged 17-48 years) in Germany: a *prepandemic cohort* (January-February 2020), a *postlockdown cohort* (May 2020-July 2020), and an *intralockdown cohort* (January-February 2021) regarding students' general emotional well-being and academic functioning. It was hypothesized that, because of rigorous lockdown-related restrictions, students in the intralockdown cohort would report diminished general emotional well-being compared with the other cohorts. Furthermore, because of ongoing remote learning since the beginning of the pandemic, it was expected that students' academic functioning would decrease across all 3 cohorts.

Methods: The data collection was performed over 3 consecutive semesters (fall semester 2019-2020, spring semester 2020, and fall semester 2020-2021). Students were surveyed on the web on various aspects regarding their general emotional well-being (eg, stress and general well-being) and academic functioning (eg, concentration and study-related flow). Data analyses were performed using multivariate ANOVAs.

Results: A total of 787 students participated in this study. Results indicated higher general well-being in the postlockdown cohort than in the intralockdown cohort (P=.02). As for students' academic functioning, our results revealed that students in the prepandemic cohort reported higher study-related flow (P=.007) and concentration (P=.001) than those in the intralockdown cohort. In addition, students reported higher flow (P=.04) and concentration (P=.04) in the postlockdown cohort than those in the intralockdown cohort. No cohort effects were revealed for other aspects of general emotional well-being (eg, perceived stress) and academic functioning (eg, procrastination).

Conclusions: This study indicates that students' general emotional well-being as well as motivational and attentional components of academic functioning can be impaired owing to the COVID-19 lockdowns and ongoing remote learning formats. The necessity and design of interventional programs remedying such effects in light of the ongoing crisis need to be addressed.

(JMIR Form Res 2022;6(11):e34388) doi:10.2196/34388

KEYWORDS

self-efficacy; academic self-concept; test anxiety; achievement motivation; positive and negative affect; mobile phone; COVID-19

Introduction

Background

The novel COVID-19 outbreak, which was declared a global pandemic on March 11, 2020, has altered people's everyday lives in an unparalleled way. With a total of 240,940,937 confirmed cases and 4,903,911 COVID-19–related deaths worldwide reported by the *World Health Organization* [1] at the time of writing this paper, it becomes abundantly clear that the disease is a serious threat to people's physical health [2-4]. However, research on large-scale health crises and quarantine-like situations [5] as well as recent work concerning COVID-19 [6-8] suggest that the pandemic and concomitant restrictions may have led to a multitude of ramifications similarly affecting people's mental health.

For university students, such ramifications have been manifold. Nationwide lockdowns have not only led to incisive limitations regarding general social contact and peer relationships but also to altered learning environments because of a rapid shift from in-person to remote learning [9]. Although coinciding with multiple other lockdown-related stressors (eg, student job loss, insecurity about mandatory internships and future employment, and fears about contracting the disease) [9,10], such social and academic ramifications can engender ample mental health detriments affecting students' both general emotional well-being [7,8,11,12] and academic functioning [13-15].

Students' Emotional Well-being

Cross-sectional studies during the early phases of the pandemic suggest that COVID-19 lockdowns have led to a high prevalence of sadness and frustration [16], depressive symptoms [17-19], anxiety and stress [17,20-26], and sleep disturbances [16] among students. Furthermore, adverse changes in such lockdown-related mental health aspects have been found in studies using both retrospective [27] as well as actual prepandemic control measures [7,8,12]. For instance, a study comparing 3 student cohorts tested in fall 2019 (prepandemic measure), spring 2020 (when lockdown provisions were initiated), and fall 2020 (when lockdown provisions were eased) revealed increases in depression, anger, stress, and mania between the fall 2019 cohort and the spring 2020 cohort [7]. However, results showed no differences between the prepandemic and the fall 2020 cohort. Another study comparing students' depressive symptoms before the lockdown in Italy (October 2019 and December 2019), during the lockdown (April 2020), and after the lockdown was lifted (May 2020 and June 2020) showed similar results [8].

In summary, research indicates that COVID-19 lockdowns may have an adverse effect on students' general emotional well-being. However, studies have also shown that students appear to recover quickly from well-being–related detriments once lockdown provisions are eased [7] or lifted [8]. Still, it should be noted that, despite the possibility of rapid recovery, severe periodic emotional impairments can come with long-term consequences for students, who have been shown across studies to be a vulnerable group regarding mental health problems [12,28-30]. Furthermore, such impairments may also create momentary as well as future costs regarding academic functioning [13,14,31]. Consequently, more research is needed to gain comprehensive insights into lockdown-related and possible long-term consequences.

Students' Academic Functioning

Stressful situations coupled with social isolation and web-based learning can similarly compromise not only general emotional well-being but also academic functioning. On the one hand, web-based teaching during the COVID-19 pandemic had to be implemented rapidly, which may have led to a lack in adequate design and organization of web-based teaching formats as well as uncertainty and anxiety regarding course work and exam preparations [32]. Such shortcomings can result in heightened self-doubt and difficulties in information processing on the part of the students [9,20,25,33]. By contrast, research has shown that because of COVID-19 lockdown provisions, people have vastly increased their overall technology use [34]. Although technology-mediated communication has been an important way of staying socially and academically connected during the COVID-19 lockdowns [35], increased screen time and the simultaneous use of technologies for different everyday tasks (eg, staying connected with friends and family, schoolwork, and entertainment) can engender cognitive (eg, inability to concentrate) and performance-related (eg, worse grades) detriments [36,37].

Research on the effects of COVID-19 lockdowns and remote learning arrangements on students' academic functioning is limited. However, qualitative data derived from interviews conducted with students in the United States showed that 81.5% (159/195) of interviewees stated concerns about their general academic performance because of the COVID-19 pandemic [25]. Furthermore, 88.7% (173/195) of interviewees indicated that they were negatively affected in their ability to concentrate on schoolwork. Other qualitative [38] and quantitative yet observational studies [23,33] corroborated these findings. Other studies revealed that students, compared with a retrospectively assessed control measure, reported declines in their motivation [14,31] as well as their behavioral and emotional academic engagement [31]. Moreover, it has been found that students experienced decreased attention and heightened externalizing problems [39] as well as increased study-related stress during the early phase of the pandemic compared with the time before [13].

In summary, lockdown-related social detriments as well as (emergency) remote learning may have taken a toll on students' academic functioning, including factors such as motivation, concentration, and study-related stress. However, existing studies using prepandemic control measures are extremely scarce. Thus, more research is needed to gain a deeper insight into the changes in students' academic functioning in the context of COVID-19 lockdowns and investigate which aspects of academic functioning may be particularly affected by the crisis.

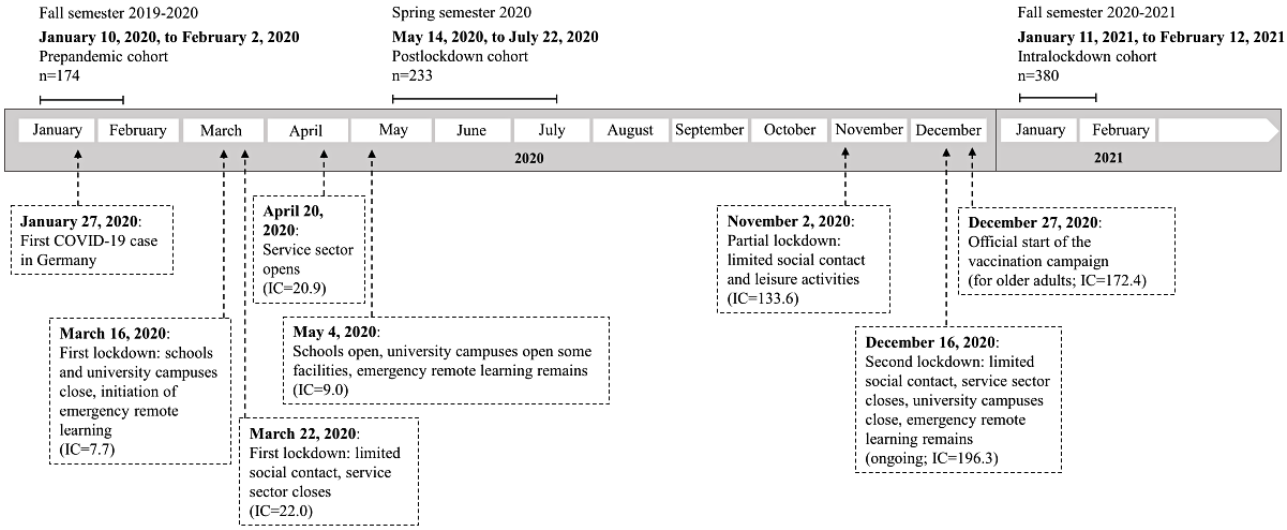
Aims of This Study

To develop interventional programs remedying emotional and academic complications among the vulnerable population of students [12,28-30], more research on the effects of the COVID-19 lockdowns on the student population is needed.

However, recent research has major limitations. First, only few studies incorporate a control measure (ie, a prepandemic measure) into their investigations [7,8,39]. Second, many studies are limited to 1 university [7,13,40] or 1 specific study department of a singular institution [12], limiting the generalizability of the results obtained. Third, most studies have investigated the effects of the COVID-19 crisis on students during its early onset but not in later phases of the pandemic [8,12-14]. Fourth, recent studies have mainly focused on few student characteristics such as stress [13,27,41], depression [42], or acceptance of web-based formats [43], thus neglecting a comprehensive examination of students' well-being.

To address these research gaps, this study aimed to compare between-subject data gathered in a comprehensive research project on students' mental health and academic functioning encompassing 3 student cohorts: a *prepandemic cohort* (ie, a student cohort tested before the pandemic), a *postlockdown cohort* (ie, a student cohort tested after the first lockdown in Germany, when social and everyday life restrictions were eased but remote teaching at universities continued), and an *intralockdown cohort* (ie, a student cohort tested during the second lockdown in Germany, when social and everyday life restrictions were reinstated in addition to remote teaching and web-based examinations; Figure 1).

Figure 1. Visualization of the data collection process in relation to the most relevant pandemic-related events in Germany between January 2020 and February 2021, including incidence values. IC: new COVID-19 infections during the last 7 days per inhabitant×100,000.



Hypotheses

Students' General Emotional Well-being

Research indicates that general social detriments [44] and COVID-19–related lockdown provisions [19,22,45] can engender affective detriments in students (eg, negative emotions, stress, and depression). However, research also shows that students' general emotional well-being appears to rapidly improve to prepandemic levels when lockdown-related restrictions are eased [7] or lifted [8]. In line with this, for students' general emotional well-being, hypothesis 1 is given in Textbox 1.

Textbox 1. Hypothesis 1.

Hypothesis 1.1

- The intralockdown cohort will report (a) less positive and (b) more negative affect, (c) less general well-being, and (d) higher perceived stress than the prepandemic cohort.

Hypothesis 1.2

- The intralockdown cohort will report (a) less positive and (b) more negative affect, (c) less general well-being, and (d) higher perceived stress than the postlockdown cohort.

Hypothesis 1.3

- The postlockdown cohort will report similar levels of (a) positive and (b) negative affect, (c) general well-being, and (d) perceived stress compared with the prepandemic cohort.

Students' Academic Functioning

Research also indicates that general social deprivation [15] and COVID-19–related lockdowns [14,23,33] can impair students' academic functioning [46,47]. Therefore, we assume that the intralockdown cohort will exhibit a severe decline in their academic functioning compared with the prepandemic cohort. However, the postlockdown cohort should also be affected in their academic functioning owing to remaining emergency remote learning, albeit not as pronounced as those students assessed during the second lockdown.

Therefore, as for students' academic functioning encompassing study-related emotional well-being (hypotheses 2a-2b), academic self-perception (hypotheses 2c-2d), motivation (hypotheses 2e-2f), and self-regulation (hypotheses 2g-2i), we have given hypothesis 2 in Textbox 2.

Textbox 2. Hypothesis 2.

Hypothesis 2.1

- The intralockdown cohort will report (a) more study-related stress, (b) more test anxiety, (c) a lower academic self-concept, (d) less study-related self-efficacy, (e) adverse achievement motivation, (f) less study-related flow, (g) less concentration, (h) lower frequency of study activities, and (i) more procrastination than the prepandemic cohort.

Hypothesis 2.2

- The intralockdown cohort will report (a) more study-related stress, (b) more test anxiety, (c) a lower academic self-concept, (d) less study-related self-efficacy, (e) adverse achievement motivation, (f) less study-related flow, (g) less concentration, (h) lower frequency of study activities, and (i) more procrastination than the postlockdown cohort.

Hypothesis 2.3

- The postlockdown cohort will report (a) more study-related stress, (b) more test anxiety, (c) a lower academic self-concept, (d) less study-related self-efficacy, (e) adverse achievement motivation, (f) less study-related flow, (g) less concentration, (h) lower frequency of study activities, and (i) more procrastination than the prepandemic cohort.

Methods

Study Design

This study is part of a large longitudinal randomized controlled study investigating the effectiveness of a planning intervention of reduced smartphone interference on students' mental health. In the longitudinal study, 3 student cohorts were investigated on various mental health aspects as well as smartphone use behaviors over 5 time points. In this study, preinterventional baseline data related to self-reported general emotional well-being and academic functioning from all 3 student cohorts (ie, prepandemic cohort, postlockdown cohort, and intralockdown cohort) were used (trial registration: ClinicalTrials.gov NCT04550286).

Recruitment

Participants were recruited nationwide through on- (only for the prepandemic cohort) and off-campus advertisements, social media platforms, and universities' listservs. The inclusion criteria for the study were participants who were students, were aged ≥16 years, had at least one graded examination during the exam period of the respective semester, possessed sound knowledge of the German language, and used an Android smartphone regularly. Individuals who failed to meet these inclusion criteria were excluded from study participation (Multimedia Appendix 1).

Procedure

As illustrated in Figure 1, between-subject data were collected via the platform *SoSci Survey* (SoSci Survey GmbH) in 3 consecutive semesters. The prepandemic cohort was tested between January 10, 2020, and February 2, 2020; the postlockdown cohort was tested between May 14, 2020, and July 22, 2020; and the intralockdown cohort was tested between January 11, 2021, and February 12, 2021. During the first 2 lockdowns that were instated in Germany, the population faced vast restrictions regarding social contact and the closure of non–system-relevant service sector industries as well as primary, secondary, and tertiary institutions of education. Restaurants, cafés, clubs, and other gastronomic and entertainment establishments were closed. Sports centers had to discontinue classes and other activities. Parks were locked or limited to a certain number of individuals, and pedestrians were not allowed to remain in larger groups. Visits to hospitals and homes for older adults were prohibited. Many began to work remotely, and schools as well university campuses were shut down. As a consequence, many universities transitioned from in-person to emergency remote learning. In May 2020, restrictions related to the first lockdown were eased, allowing for small social gatherings and the opening of the service sector. However, although universities alleviated access limitations to campuses and allowed for some in-person examinations under rigorous hygiene concepts, they continued with remote teaching formats. Remote teaching was continued during the second lockdown (Figure 1).

Participation in this study was voluntary; students of the institution responsible for the study's conduction received course credit. All participants were given the opportunity to take part in a raffle of various prizes, including adventure activity gift cards and other vouchers (worth US $830).

Ethics Approval

Participants had to give their informed consent to take part in the study and were treated in accordance with the ethical standards of the Helsinki Declaration [48]. The study was approved by the Witten and Herdecke University Ethics Commission (215/2019).

Measures

General Emotional Well-being

Positive and Negative Affect

Students' positive and negative affect was investigated using a shortened German version of the *Positive and Negative Affect Schedule* [49,50]. Participants were to indicate how they had felt during the last 7 days on 5 items covering positive emotions

(eg, "attentive"; ω=0.77) and 5 items covering negative emotions and emotional expressions (eg, "upset"; ω=0.80). Responses were to be given on a 5-point Likert scale ranging from 1=*not at all* to 5=*extremely*.

General Well-being

Students' general well-being was measured using a German version of the *World Health Organization-5 Well-being Index* [51]. Regarding the last 7 days, participants were to appraise their well-being on 5 items (eg, "I felt calm and relaxed"; ω=0.83) on a 6-point Likert scale ranging from 0=*never* to 5=*all the time*. A well-being index score of 13 represents low well-being and can be used as a screening marker for the presence of depressive symptoms.

Perceived Stress

Perceived stress was measured using a German version of the *Perceived Stress Scale* [52,53]. Participants were to indicate how often they had perceived stress on 9 items (eg, "How often during the last 7 days...have you felt nervous and stressed"; ω=0.85) on a 5-point Likert scale ranging from 1=*never* to 5=*very often*.

Academic Functioning

Study-Related Stress

Students' study-related stress was investigated using an adaptation of the stress measure used by Schmidt et al [54]. This measure consists of 4 items (eg, "During the last 7 days...I felt nervous and stressed due to the preparations for my exam(s)"; ω=0.88) that were to be answered on a 5-point Likert scale ranging from 1=*not at all* to 5=*extremely*.

Test Anxiety

Test anxiety was assessed using items adapted from the German version of the *Test Anxiety Inventory* [55,56]. The measure used consists of 10 items; 5 items address *test-related agitation* (*Test Anxiety Inventory-Emotionality*; eg, "My heart is in my mouth"; ω=0.89), whereas the other 5 items assess *test-related worries* (*Test Anxiety Inventory-Worry*; eg, "I wonder whether my performance will suffice"; ω=0.83). Participants were to respond on a 6-point Likert scale ranging from 1=*completely disagree* to 6=*completely agree*.

Academic Self-concept

Students' academic self-concept was assessed using the subscale *academic self-concept* of the *German Scales regarding Academic Self-Concept* [57]. Academic self-concept encompasses students' general perceptions and beliefs about their own academic capabilities [58]. The subscale used consists of 5 items (eg, "I think I am very intelligent"; ω=0.88) that were to be answered on a 7-point Likert scale ranging from 1=*completely disagree* to 7=*completely agree*.

Study-Related Self-efficacy

Students' study-related self-efficacy was measured using the *German Self-efficacy Scale* [56], adapted to the student context. Study-related self-efficacy encompasses students' expectations regarding possible achievements in a given context, which is thought to be a precursor of students' academic self-concept [58]. This instrument comprises 7 items (eg, "During the last 7 days...I was sure that I can solve even the difficult tasks and texts for the exam if I make an effort"; ω=0.86). Participants were to respond on a 6-point Likert scale ranging from 1=*completely disagree* to 6=*completely agree*.

Achievement Motivation

Students' achievement motivation was assessed using the *German Questionnaire for the Assessment of Current Motivation in Learning and Performance Situations* [59]. In this study, the subscales *probability of success* (4 items; eg, "I think everybody can pass this exam"; ω=0.81) and *probability of failure* (4 items; eg, "When thinking about the upcoming exam, I am somewhat worried"; ω=0.79) were used. Probability of success encompasses students' expectations regarding the probability of performing well in an upcoming exam. Probability of failure, by contrast, encompasses students' expectations regarding the probability of not being able to deal with exam pressures and, consequently, performing badly. Participants were to indicate their achievement motivation regarding the upcoming exams and their study engagement on a 5-point Likert scale ranging from 1=*completely disagree* to 5=*completely agree*.

Study-Related Flow

Study-related flow was assessed using the *German Measure for Flow Experience* [60]. Flow experience can be understood as an action-related (as opposed to target-related) motivator in the academic context. The flow measure used includes 13 items (eg, "During the last 7 days...I was completely absorbed in what I was studying"; ω=0.82). Participants were to respond on a 5-point Likert scale ranging from 1=*completely disagree* to 5=*completely agree*.

Concentration

Students' concentration while studying was measured using the subscale *concentration* of the *German Study-Related Learning Strategy Scales* [61], which consists of 6 items (eg, "During the last 7 days...I was unconcentrated"; ω=0.94). Participants were to respond on a 5-point Likert scale ranging from 1=*completely disagree* to 5=*completely agree*.

Frequency of Study Activities

Students' frequency of study activities was assessed using a self-developed measure. Students were to indicate on how many days during the last 7 days they had studied, at least for a little while (eg, memorizing information, reading, and doing exercises). Participants were to choose whether they had studied on *1 day*, *2 days*, *3 days*, *4 days*, *5 days*, *6 days*, *7 days*, or *no days*.

Procrastination

Students' tendency to procrastinate was assessed using the German version of the *Aitken Procrastination Scale* [62]. Procrastination encompasses the tendency to postpone tasks, which may affect a broad range of activities and is generally independent of specific situational stimuli. This measure encompasses 13 items (eg, "I often need a long time to get things started"; ω=0.91) that participants were to answer on a 5-point Likert scale ranging from 1=*completely disagree* to 5=*completely agree*.

Statistical Analysis

All data analyses were performed with SPSS Statistics (version 27; IBM Corp). Across variables, <1% of missing data (item nonresponses) were identified; thus, analyses were performed with the available information. To test our hypotheses, we performed multivariate analysis of variances (MANOVAs). For that, all variables assessed as indicators of general emotional well-being and all variables of the respective aspects regarding academic functioning (ie, study-related well-being, academic self-perception, academic motivation, and academic self-regulation; cf Measures) were conjointly analyzed.

Univariate outliers were identified using visual inspection of box plots and analyses of *z* scores. Multivariate outliers were identified using visual inspection of *Q*-*Q* plots and analyses of Mahalanobis distance. We identified 5 univariate outliers ($|z|$=3.29) and 2 multivariate outliers as assessed based on critical chi-square values (P<.001). Univariate outliers were winsorized. Multivariate outliers were not omitted as analyses yielded similar results whether extreme cases were excluded or not. Normality of outcomes was confirmed as assessed by visual inspection of histograms and Normal *Q*-*Q* Plots. Homogeneity of variances was tested using the Levene test, showing that, for most variables, variances were equal for all student cohorts, with significance values ranging from P=.05 (positive affect) to P=.99 (negative affect). The Levene test showed significant results only for study-related stress (P=.02). For multivariate analyses, homogeneity of covariance matrices was confirmed by the Box test (P>.05). Visual inspection of Normal *P*-*P* Plots and histograms plotting standardized predicted values against standardized residuals confirmed linearity and homoscedasticity. Furthermore, we detected no evidence of multicollinearity as assessed by the Pearson correlation ($|r|$<0.9) and the variance inflation index (variance inflation factor <2).

As recent COVID-19 studies have shown that demographic factors may influence general emotional well-being and academic functioning [6,17,22,26], we performed additional hypotheses tests including control variables (ie, gender, age, semester, and number of exams in the respective semester; Multimedia Appendices 2-6).

Results

Participants

A total of 787 participants were recruited for this study. The final sample of the prepandemic cohort consisted of 174 students (n=127, 73% women; n=46, 26.4% men; and n=1, 0.6% nonbinary) with a mean age of 22.89 (SD 3.58) years ranging from 17 to 38 years. The postlockdown cohort consisted of 233 students (n=172, 73.8% women; n=60, 25.8% men; and n=1, 0.4% nonbinary) with a mean age of 23.32 (SD 4.40) years ranging from 18 to 48 years. The intralockdown cohort encompassed 380 students (n=270, 71.1% women; n=106, 27.9% men; and n=4, 1.1% nonbinary) with a mean age of 22.47 (SD 3.31) years ranging from 17 to 43 years. Participants in the postlockdown cohort were, on average, slightly older than those in the other 2 cohorts ($F_{2,784}$=3.80; P=.02; η_p^2=0.010). The distribution of men and women was equal across the 3 cohorts (N=781, χ^2_2=0.4; P=.81). The cohorts consisted of students from >200 German universities from all 16 federal states. Most participants were enrolled in higher semesters (579/787, 73.6%). First-year students were unequally distributed across cohorts, with relatively fewer first-year students in the postlockdown cohort (22/233, 9.4%) than in the prepandemic (53/174, 30.5%) and intralockdown (133/389, 35%) cohorts (N=787, χ^2_2=50.4; P<.001). Participants were enrolled in different study programs: arts and design (16/787, 2%), education studies (21/787, 2.7%), agricultural studies (25/787, 3.2%), cultural and language studies (29/787, 3.7%), psychology (62/787, 7.9%), social sciences (83/787, 10.5%), medicine and health (104/787, 13.2%), economy and law (139/787, 17.7%), natural sciences (140/787, 17.8%), and technology (165/787, 21%). On average, participants prepared for 4 (SD 1.51) exams during the respective semesters. The exam count was equal across cohorts ($F_{2,784}$=2.89; P=.28; η_p^2=0.003). An overview of the sample demographics is provided in Table 1.

For the longitudinal study, the sample size was calculated a priori using G*Power [63]. The necessary sample size for identifying a small effect (Cohen d=0.17) regarding the smartphone interference reduction intervention in relation to a control group with a power of 1–β=.95 at a P value of .05 was 116. For this study, a post hoc power analysis was performed using G*Power. The analysis for the sample size of 787 resulted in a power of 1–β=.91 given a small cohort effect of f=0.13 and an error probability of α=.05.

Table 1. Demographics regarding all student cohorts (N=787).

	Prepandemic cohort (n=174)	Postlockdown cohort (n=233)	Intralockdown cohort (n=380)	Total
Gender, n (%)				
Men	46 (26.4)	60 (25.8)	106 (27.9)	212 (26.9)
Women	127 (72.9)	172 (73.8)	270 (71.1)	569 (72.3)
Nonbinary	1 (0.6)	1 (0.4)	4 (1.1)	6 (0.8)
Semester information, n (%)				
First study	136 (78.2)	184 (79)	291 (76.6)	611 (77.6)
Second study	38 (21.8)	49 (21)	89 (23.4)	176 (22.4)
First-semester students	53 (30.5)	22 (9.4)	133 (35)	208 (26.4)
Higher-semester students	121 (69.5)	211 (90.6)	247 (65)	579 (73.6)
Study program, n (%)				
Arts and design	2 (1.1)	4 (1.7)	10 (2.6)	16 (2)
Education studies	4 (2.3)	7 (3)	10 (2.6)	21 (2.7)
Agricultural studies	4 (2.3)	4 (1.7)	17 (4.5)	25 (3.2)
Cultural and language studies	9 (5.2)	7 (3)	13 (3.4)	29 (3.7)
Psychology	27 (15.5)	12 (5.2)	23 (6.1)	62 (7.9)
Social sciences	11 (6.3)	31 (13.3)	41 (10.8)	83 (10.5)
Medicine and health	28 (16.1)	30 (12.9)	46 (12.1)	104 (13.2)
Economy and law	35 (20.1)	42 (18)	62 (16.3)	139 (17.7)
Natural sciences	21 (12.1)	43 (18.5)	76 (20)	140 (17.8)
Technology	31 (17.8)	53 (22.7)	81 (21.3)	165 (21)
Other	1 (0.6)	0 (0)	1 (0.3)	1 (0.1)
Federal state, n (%)				
Baden-Württemberg	38 (21.8)	40 (17.2)	54 (14.2)	132 (16.8)
Bavaria	13 (7.5)	58 (24.9)	80 (21.1)	151 (19.2)
Berlin	10 (5.7)	7 (3)	14 (3.7)	31 (3.9)
Brandenburg	0 (0)	11 (4.7)	5 (1.3)	16 (2)
Bremen	0 (0)	12 (5.2)	8 (2.1)	20 (2.5)
Hamburg	3 (1.7)	6 (2.6)	0 (0)	9 (1.1)
Hesse	17 (9.8)	8 (3.4)	57 (15)	82 (10.4)
Lower Saxony	3 (1.7)	8 (3.4)	33 (8.7)	44 (5.6)
Mecklenburg-Western Pomerania	14 (8)	24 (10.3)	4 (1.1)	42 (5.3)
North Rhine-Westphalia	40 (23)	4 (1.7)	67 (17.6)	111 (14.1)
Rhineland-Palatinate	4 (2.3)	17 (7.3)	6 (1.6)	27 (3.4)
Saarland	0 (0)	1 (0.4)	0 (0)	1 (0.1)
Saxony	8 (4.6)	18 (7.7)	10 (2.6)	36 (4.6)
Saxony-Anhalt	9 (5.2)	8 (3.4)	14 (3.7)	31 (3.9)
Schleswig-Holstein	0 (0)	5 (2.1)	4 (1.1)	9 (1.1)
Thuringia	14 (8)	6 (2.6)	24 (6.3)	44 (5.6)
Age (years), mean (SD)	22.89 (3.58)	23.32 (4.40)	22.47 (3.31)	22.82 (3.73)
Exam count, mean (SD)	3.99 (1.57)	4.23 (1.56)	4.17 (1.45)	4.15 (1.51)

Descriptive Outcome Analysis

Table 2 shows the mean values and SDs of all outcome variables for each student cohort. As for general emotional well-being, the descriptive data show that students reported positive and negative affect close to the average of the scale, indicating a moderate level. Students' general well-being scores were below the cutoff value of 13, indicating low well-being in all student cohorts. Students in all cohorts also reported a perceived level of stress above the scale average (range 9-45). In line with this, students reported high study-related stress and increased worries regarding their upcoming exams (mean values were above the scale average). As for students' self-perception, the data revealed scores above the average of the scale regarding academic self-concept and study-related self-efficacy. Furthermore, students reported high achievement motivation (mean values were above the scale average) with regard to their perceived probability of both success and failure. However, study-related flow was only moderate (mean values were close to the scale average). Similarly, students' ability to concentrate was moderate, whereas the time that students reportedly engaged in their exam preparations was rather high (ie, >4.5 days per week on average). Finally, students' tendency to procrastinate was also above the average of the scale.

Almost all outcome variables were correlated with one another, indicating that general emotional well-being and the different components of academic functioning were interconnected (see Table 3 for Pearson correlations of all outcome variables across cohorts). Noticeably, frequency of study activities was positively correlated only with positive affect (r_{787}=0.18; P<.001), study-related flow (r_{787}=0.15; P<.001), and concentration (r_{787}=0.15; P<.001). All other outcome variables were not correlated with students' study engagement.

Table 2. Descriptive data of outcomes, including mean scores, score range, and the effect size (η_p^2) for the differences among all cohorts.

	Prepandemic cohort, mean (SD)	Postlockdown cohort, mean (SD)	Intralockdown cohort, mean (SD)	Range	η_p^2
General emotional well-being					
Positive affect	15.22 (3.75)	15.21 (3.14)	14.64 (3.53)	5-25	0.007
Negative affect	10.83 (4.05)	10.88 (4.22)	11.23 (4.10)	5-25	0.002
General well-being	11.51 (4.92)	11.89 (4.67)	10.92 (4.45)	0-25	0.008
Perceived stress	25.56 (6.44)	26.57 (6.40)	26.88 (6.30)	9-45	0.007
Study-related emotional well-being					
Study-related stress	3.12 (1.15)	3.14 (1.06)	3.14 (1.03)	1-4	<0.001
Test anxiety (agitation)	3.34 (1.28)	3.25 (1.22)	3.36 (1.25)	1-6	0.001
Test anxiety (worry)	4.46 (1.04)	4.40 (0.99)	4.40 (1.06)	1-6	<0.001
Self-perception					
Academic self-concept	4.76 (1.03)	4.77 (0.91)	4.62 (0.98)	1-7	0.006
Study-related self-efficacy	4.06 (0.76)	4.05 (0.79)	3.95 (0.88)	1-6	0.004
Motivation					
Motivation (success)	3.51 (0.80)	3.48 (0.82)	3.40 (0.85)	1-5	0.003
Motivation (failure)	3.32 (1.01)	3.18 (0.90)	3.35 (0.94)	1-5	0.007
Study-related flow	2.69 (0.60)	2.65 (0.53)	2.55 (0.54)	1-5	0.012
Self-regulation					
Concentration	2.54 (0.97)	2.42 (0.90)	2.25 (0.92)	1-5	0.017
Frequency of study activities	4.64 (1.96)	4.89 (1.93)	4.83 (2.01)	0-7	0.002
Procrastination	3.05 (0.81)	2.99 (0.83)	3.12 (0.78)	1-5	0.005

Table 3. Pearson correlations of outcome variables across cohorts.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Positive affect	1	—[a]	—	—	—	—	—	—	—	—	—	—	—	—	—
2. Negative affect	−0.38[b]	1	—	—	—	—	—	—	—	—	—	—	—	—	—
3. General well-being	0.68[b]	−0.53[b]	1	—	—	—	—	—	—	—	—	—	—	—	—
4. Perceived stress	−0.51[b]	0.69[b]	−0.60[b]	1	—	—	—	—	—	—	—	—	—	—	—
5. Study-related stress	−0.32[b]	−0.58[b]	−0.50[b]	0.66[b]	1	—	—	—	—	—	—	—	—	—	—
6. Test anxiety (agitation)	−0.31[b]	0.67[b]	−0.43[b]	0.63[b]	0.66[b]	1	—	—	—	—	—	—	—	—	—
7. Test anxiety (worry)	−0.24[b]	0.52[b]	−0.37[b]	0.55[b]	0.59[b]	0.66[b]	1	—	—	—	—	—	—	—	—
8. Academic self-concept	0.39[b]	−0.31[b]	0.34[b]	−0.43[b]	−0.34[b]	−0.30[b]	−0.28[b]	1	—	—	—	—	—	—	—
9. Study-related self-efficacy	−0.43[b]	−0.39[b]	0.43[b]	−0.54[b]	−0.49[b]	−0.40[b]	−0.43[b]	0.64[b]	1	—	—	—	—	—	—
10. Achievement motivation (success)	0.34[b]	−0.40[b]	0.37[b]	−0.52[b]	−0.52[b]	−0.48[b]	−0.50[b]	0.50[b]	0.66[b]	1	—	—	—	—	—
11. Achievement motivation (failure)	−0.28[b]	0.52[b]	−0.39[b]	0.51[b]	0.50[b]	0.59[b]	0.67[b]	−0.24[b]	−0.37[b]	−0.40[b]	1	—	—	—	—
12. Study-related flow	0.55[b]	−0.29[b]	0.44[b]	−0.45[b]	−0.31[b]	−0.27[b]	−0.27[b]	0.36[b]	0.41[b]	0.37[b]	−0.28[b]	1	—	—	—
13. Concentration	0.42[b]	−0.36[b]	−0.34[b]	−0.40[b]	−0.32[b]	−0.34[b]	−0.32[b]	0.24[b]	0.23[b]	0.25[b]	−0.34[b]	0.58[b]	1	—	—
14. Frequency of study activities	0.18[b]	−0.012	0.029	−0.023	0.065	−0.010	0.036	0.028	0.001	0.033	−0.012	0.15[b]	0.15[b]	1	—
15. Procrastination	−0.43[b]	0.31[b]	−0.35[b]	0.40[b]	−0.33[b]	0.30[b]	0.25[b]	−0.27[b]	−0.26[b]	−0.29[b]	0.27[b]	−0.44[b]	−0.45[b]	−0.30[b]	1

[a]Not applicable.

[b]$P<.001$.

Primary Outcome Analysis

General Emotional Well-being

MANOVA results for positive and negative affect, general well-being, and perceived stress with student cohort as independent factor indicated differences among the prepandemic, postlockdown, and intralockdown cohorts (V=0.020; $F_{8,1562}$=2.00; P=.04; η_p^2=0.010). Univariate analyses indicated that the student cohort had an effect on general well-being ($F_{2,783}$=3.32; P=.04; η_p^2=0.008) but neither on positive affect ($F_{2,783}$=2.64; P=.07; η_p^2=0.007), negative affect ($F_{2,783}$=0.81; P=.45; η_p^2=0.002), nor perceived stress ($F_{2,783}$=2.63; P=.07; η_p^2=0.007).

Bonferroni-corrected post hoc comparisons for general well-being demonstrated that, although students in the intralockdown cohort perceived similar levels of well-being as those in the prepandemic cohort (P=.25), those in the intralockdown cohort perceived lower general well-being than those in the postlockdown cohort (P=.02). No differences in students' general well-being were found between the postlockdown and prepandemic cohorts (P=.50).

Study-Related Emotional Well-being

MANOVA results for study-related stress and the 2 test anxiety subscales—agitation and worry—with student cohort as independent factor did not reveal any differences among the cohorts for the combined dependent variables (V=0.004; $F_{6,1566}$=0.56; P=.76; η_p^2=0.002).

Academic Self-perception

MANOVA results for academic self-concept and study-related self-efficacy with student cohort as independent factor showed no cohort effect on the combined dependent variables (V=0.006; $F_{4,1568}$=1.16; P=.33; η_p^2=0.003).

Motivation

MANOVA results for the 2 achievement motivation subscales—probability of success and probability of failure—as well as study-related flow with student cohort as independent factor revealed differences among the student cohorts (V=0.018; $F_{6,1566}$=2.40; P=.03; η_p^2=0.009). Univariate analyses demonstrated a cohort effect on study-related flow ($F_{2,784}$=4.90; P=.008; η_p^2=0.012) but neither on probability of success ($F_{2,784}$=1.27; P=.28; η_p^2=0.003) nor on probability of failure ($F_{2,784}$=2.65; P=.07; η_p^2=0.007).

As for students' study-related flow, Bonferroni-corrected post hoc analyses indicated that students in the intralockdown cohort reported lower flow levels than those in the prepandemic cohort (P=.007). Furthermore, students in the intralockdown cohort reported less study-related flow than those in the postlockdown

cohort (P=.04). No differences were found between the postlockdown and prepandemic cohorts (P=.50).

Academic Self-regulation

MANOVA results for students' concentration, frequency of study activities, and procrastination with student cohort as independent factor showed a cohort effect on the combined dependent variables (V=0.022; $F_{6,1566}$=2.94; P=.008; η_p^2=0.011). Univariate analyses revealed that students' concentration differed among the cohorts ($F_{2,784}$=6.61; P<.001; η_p^2=0.017). However, no differences were found for students' frequency of study activities ($F_{2,784}$=0.84; P=.43; η_p^2=0.002) or for students' procrastination ($F_{2,784}$=1.95; P=.14; η_p^2=0.005).

Bonferroni-corrected post hoc comparisons for students' concentration showed that students in the intralockdown cohort had more difficulties concentrating on their study activities than students in the prepandemic cohort (P=.001). Students in the intralockdown cohort also reported lower concentration levels than those in the postlockdown cohort (P=.04). No difference in concentration was found between the postlockdown and prepandemic cohorts (P=.31).

Discussion

Principal Findings

This study is one of few comprehensive investigations into the effects of 2 of the COVID-19 lockdowns in Germany on students' mental health incorporating a prepandemic control measure. Specifically, this study compared 3 student cohorts (a prepandemic, postlockdown, and intralockdown cohort) regarding various factors concerning students' emotional well-being and academic functioning.

General Emotional Well-being

As for students' general emotional well-being, including positive and negative affect, general well-being, and perceived stress, the results of our study indicated support for our hypotheses only for general well-being. Particularly, students assessed during the lockdown in Germany (ie, intralockdown cohort) experienced decreased general well-being (hypothesis 1c) compared with students assessed after lockdown-related restrictions were lifted (ie, postlockdown cohort). Furthermore, in line with our hypotheses, no differences regarding students' well-being were observed between the postlockdown and prepandemic cohorts. These results are in accordance with existing empirical evidence showing that well-being–related detriments can emerge as a consequence of COVID-19 lockdown provisions but may be rapidly attenuated once restrictions are eased [7] or lifted [8]. However, it should be noted that all student cohorts reported concerningly low well-being scores, which were only exacerbated during the lockdown. These findings call attention to the general vulnerability of individuals in tertiary education [12,28-30] and the necessity of mental health programs particularly designed to help students during the COVID-19 crisis and beyond.

Contrary to our hypotheses, students' positive (hypothesis 1a) and negative affect (hypothesis 1b) as well as stress (hypothesis 1d) did not differ among the 3 cohorts. However, for all cohorts, positive affect was only moderate, and stress presented itself as rather high, which appears to be complementary to the low well-being scores. It should be noted that, in the current literature, there is indication for an increase in negative affect and stress in students during the COVID-19 lockdowns [7,8,12]. However, there are various possible explanations for the absence of the expected differences regarding these outcomes in this research.

For instance, it may be that affect and stress were less affected by the lockdown-related restrictions as many students relocated to their caregivers' place. Recent research has shown that relocating can reduce some of the experienced material and psychological burden during the pandemic and, thereby, attenuate perceived stress [27].

However, it is also possible that the measures used did not detect adverse changes that emerged among the student cohorts. As for positive and negative affect, it may be that, owing to low general well-being [64] and the social deprivation [44,65] precipitated by the COVID-19 lockdowns, students have experienced dampened emotionality. Thus, the mood measures used assessing high-activity emotional expressions (ie, anger, hostility, wakefulness, and determination) may have been inadequate to capture lockdown-related detriments. In other words, comprehensive measures encompassing low-activity emotions similar to those used in the applied well-being measure (eg, calm and relaxed, interested, and in good spirits) may allow for deeper insights into the lockdown effects on students' mental health. Similarly, the stress measure used may not be an adequate indicator of lockdown-related detriments. A recent longitudinal study comparing students' stress before and after the onset of the COVID-19 pandemic in Germany corroborates the absence of stress differences because of the lockdown [66]. However, the study revealed shifts in students' behavior and experience patterns such as that healthy (ie, effective stress coping and positive study-related behaviors) and overexertion (ie, ineffective stress coping and study-related overcommitment) tendencies decreased, whereas unambitious (ie, effective stress coping and low study-related commitment) and burnout (ie, ineffective stress coping and low study-related ambition) patterns increased between the prepandemic and intrapandemic measure. Thus, although the COVID-19 lockdowns may not have affected reported stress, underlying stress perceptions and coping behaviors may have changed.

In line with this, it is also noteworthy that, unlike most existing studies [8,12-14], our study investigated well-being–related detriments both in the earlier phase (ie, the postlockdown cohort) and in a later phase (ie, the intralockdown cohort) of the pandemic. Consequently, students had more experience with lockdown-related restrictions, which may have led to the development of various adaptive and maladaptive coping strategies as well as behavior and experience patterns. With this, our results indicate that more research comparing different phases of the pandemic and different emotional expressions as well as coping patterns with regard to students' general emotional well-being is needed. Previous research has started to investigate possible coping strategies used by students during

the COVID-19 crisis [67-69]. However, this research is still extremely limited.

Academic Functioning

As for students' academic functioning, our results showed support for some but not all outcome variables. Specifically, in line with our hypotheses, we found that study-related flow (hypothesis 2f) and concentration (hypothesis 2g) were lowest in the intralockdown cohort compared with the prepandemic and postlockdown cohorts.

These findings are in line with recent research demonstrating detriments to students' motivation [14,31] and attention [39] associated with lockdown-precipitated social detriments and the shortcomings of remote learning formats. As for study-related flow in particular, the results obtained are in line with research suggesting adverse effects brought about by uncertainty regarding study and exam requirements, insufficient interstudent and student-teacher communication, and enhanced and often undefined workload [9]. Important components of study-related flow are clear definitions of study goals, frequent feedback, and high perceived control [60]. Consequently, the lack of such aspects may hinder study-related flow. These findings are not just of relevance as diminished study-related flow and concentration may directly impair students' academic functioning but also as they are associated with the frequency of study activities. Thus, detriments in these academic areas may also worsen study engagement in general. It should be noted that, contrary to our hypotheses, we did not identify any differences in study-related flow and concentration between the postlockdown and prepandemic cohorts. This result indicates that despite potential problems arising in the context of remote learning, the remote learning format alone may not lead to academic detriments as long as it does not co-occur with other pandemic-related stressors such as social detriments. However, more research is needed to support this claim.

In contrast to our hypotheses, we did not identify any cohort effects on students' study-related stress (hypothesis 2a), test anxiety (hypothesis 2b), academic self-concept (hypothesis 2c), self-efficacy (hypothesis 2d), achievement motivation (hypothesis 2e), frequency of study activities (hypothesis 2h), and procrastination (hypothesis 2i).

There are multiple possible explanations for the lack of observed differences. Most of these factors concerning academic functioning represent rather stable person characteristics; something that is also reflected in the measures used to assess the respective variables. It is possible that COVID-19 lockdown measures and remote learning—despite being incisive life events—simply do not unsettle stable person characteristics as easily as more variable aspects of academic functioning. Research also suggests that factors such as general self-efficacy may act as protective factors buffering the individual from potential COVID-19–related detriments [70]. Similarly, it is possible that adverse person characteristics such as the tendency to procrastinate act as risk factors enhancing the adverse consequences arising from COVID-19–related academic stressors. In fact, in our study, academic self-concept and study-related self-efficacy scores were rather high; however, students still experienced some general emotional well-being– and academic functioning–related impairments, albeit the effects were generally small. Consequently, more investigations into protective and risk factors are needed to identify potential implications for interventional programs on students' academic functioning during the crisis.

Another explanation for the lack of differences concerning the aforementioned factors could be that some aspects of academic functioning related to the exam phase of the semester, which was still over a month away at the time students participated in the web-based survey. Consequently, such measures (ie, study-related stress, test anxiety, study-related self-efficacy, and achievement motivation) may have not been able to identify strains that would have occurred shortly before the exam phase.

Finally, it is important to note that, as part of a larger longitudinal study, this study only included individuals who were highly versed in the use of mobile technologies, particularly the smartphone. Earlier, we referred to prior work suggesting that, besides social detriments and remote learning in general, heightened screen time [34,36,37] may further engender academic impairments. However, students in our study possessed sound knowledge of their devices as—for study participation—they had to use their smartphones on a daily basis. Consequently, these students may have encountered few complications (eg, technical problems and internet access) with regard to remote learning formats. Moreover, as students participated in our study with the goal of reducing their smartphone interference while studying, it is highly probable that most of them had been exhibiting smartphone and technology overuse tendencies already before the pandemic. Thus, the potential academic detriments of heightened screen time because of COVID-19 lockdowns may have simply not become apparent in our student sample.

Still, in light of the revealed impairments to students' concentration and study-related flow as well as the uncertainty regarding the development of the COVID-19 crisis, it appears reasonable to develop academic programs promoting students' academic functioning. This is especially important as remote learning formats are likely to be continued in one way or another.

Limitations

Despite its contributions, this study also has limitations. First, it must be noted that we investigated between-subject data; thus, students in the different cohorts were not the same individuals and, therefore, may differ in the general emotional well-being and academic functioning aspects that were assessed in this study. Moreover, even though we drew from a nationwide sample, some sociodemographic factors were not equally distributed across cohorts. Future comprehensive research with representative study samples incorporating multiple measurement points for one or several cohorts as has been done in few existing studies would be beneficial to better understand the effects of the COVID-19 crisis on the student population. Alternatively, it is possible that students who were particularly impaired because of COVID-19 lockdown measures and emergency remote learning discontinued their studies; thus, cohorts may not be easily comparable because of unknown student fluctuation. This possibility should be regarded in future

research. Furthermore, we do not know how students felt during the first lockdown. Clearly, this insight would be particularly valuable to answer questions on whether coping has been, at least to some degree, successful over time and whether interventional programs are needed. In addition, our study can only provide information regarding the first 2 lockdowns in Germany. Thus, we want to emphasize the importance of continuing the investigation of the ongoing effects of the COVID-19 pandemic on students and the general population. We want to encourage future research to address these issues by analyzing available data from different time points during the pandemic with regard to students' general emotional well-being and academic functioning.

Conclusions

This study contributes to the COVID-19 literature regarding students' mental health. It is one of few studies incorporating a prepandemic control measure and a measure during the later phase of the pandemic while also providing broad insights into various mental health aspects relating to students' general well-being and academic functioning. In summary, this study showed that students experience both general emotional well-being–related detriments (ie, worsened general well-being) as well as some impairments to their academic functioning (ie, decreased concentration and study-related flow). Thus, in light of the ongoing crisis, possible interventional approaches implementable within educational institutions should be addressed.

Acknowledgments

The authors would like to thank Alexandra Pongs, Lina Brockmeier, and Christopher Grolla for their support during the recruiting phase of this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Flow of participants.
[PNG File , 534 KB - formative_v6i11e34388_app1.png]

Multimedia Appendix 2
Multivariate analysis of covariance results for the combined dependent variables positive affect, negative affect, general well-being, and perceived stress.
[PDF File (Adobe PDF File), 138 KB - formative_v6i11e34388_app2.pdf]

Multimedia Appendix 3
Multivariate analysis of covariance results for the combined dependent variables study-related stress, test anxiety (agitation), and test anxiety (worry).
[PDF File (Adobe PDF File), 134 KB - formative_v6i11e34388_app3.pdf]

Multimedia Appendix 4
Multivariate analysis of covariance results for the combined dependent variables academic self-concept and self-efficacy.
[PDF File (Adobe PDF File), 127 KB - formative_v6i11e34388_app4.pdf]

Multimedia Appendix 5
Multivariate analysis of covariance results for the combined dependent variables achievement motivation (probability of success), achievement motivation (probability of failure), and study-related flow.
[PDF File (Adobe PDF File), 133 KB - formative_v6i11e34388_app5.pdf]

Multimedia Appendix 6
Multivariate analysis of covariance results for the combined dependent variables concentration, frequency of study activities, and procrastination.
[PDF File (Adobe PDF File), 131 KB - formative_v6i11e34388_app6.pdf]

References

1. Coronavirus disease (COVID-19) pandemic. World Health Organization. 2021. URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019 [accessed 2021-10-20]

2. Krishnan A, Hamilton JP, Alqahtani SA, Woreta TA. A narrative review of coronavirus disease 2019 (COVID-19): clinical, epidemiological characteristics, and systemic manifestations. Intern Emerg Med 2021 Jun;16(4):815-830 [FREE Full text] [doi: 10.1007/s11739-020-02616-5] [Medline: 33453010]
3. Mehta OP, Bhandari P, Raut A, Kacimi SE, Huy NT. Coronavirus disease (COVID-19): comprehensive review of clinical presentation. Front Public Health 2021 Jan 15;8:582932 [FREE Full text] [doi: 10.3389/fpubh.2020.582932] [Medline: 33520910]
4. Cabrera Martimbianco AL, Pacheco RL, Bagattini Â, Riera R. Frequency, signs and symptoms, and criteria adopted for long COVID-19: a systematic review. Int J Clin Pract 2021 Oct;75(10):e14357 [FREE Full text] [doi: 10.1111/ijcp.14357] [Medline: 33977626]
5. 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]
6. Rodríguez-Rey R, Garrido-Hernansaiz H, Collado S. Psychological impact of COVID-19 in Spain: early data report. Psychol Trauma 2020 Jul;12(5):550-552. [doi: 10.1037/tra0000943] [Medline: 32538657]
7. Charles NE, Strong SJ, Burns LC, Bullerjahn MR, Serafine KM. Increased mood disorder symptoms, perceived stress, and alcohol use among college students during the COVID-19 pandemic. Psychiatry Res 2021 Feb;296:113706 [FREE Full text] [doi: 10.1016/j.psychres.2021.113706] [Medline: 33482422]
8. Meda N, Pardini S, Slongo I, Bodini L, Zordan MA, Rigobello P, et al. Students' mental health problems before, during, and after COVID-19 lockdown in Italy. J Psychiatr Res 2021 Feb;134:69-77. [doi: 10.1016/j.jpsychires.2020.12.045] [Medline: 33360865]
9. Aristovnik A, Keržič D, Ravšelj D, Tomaževič N, Umek L. Impacts of the COVID-19 pandemic on life of higher education students: a global perspective. Sustainability 2020 Oct 13;12(20):8438. [doi: 10.3390/su12208438]
10. Hagedorn RL, Wattick RA, Olfert MD. "My entire world stopped": college students' psychosocial and academic frustrations during the COVID-19 pandemic. Appl Res Qual Life 2022;17(2):1069-1090 [FREE Full text] [doi: 10.1007/s11482-021-09948-0] [Medline: 33995688]
11. Leigh-Hunt N, Bagguley D, Bash K, Turner V, Turnbull S, Valtorta N, et al. An overview of systematic reviews on the public health consequences of social isolation and loneliness. Public Health 2017 Nov;152:157-171. [doi: 10.1016/j.puhe.2017.07.035] [Medline: 28915435]
12. Elmer T, Mepham K, Stadtfeld C. Students under lockdown: comparisons of students' social networks and mental health before and during the COVID-19 crisis in Switzerland. PLoS One 2020 Jul 23;15(7):e0236337 [FREE Full text] [doi: 10.1371/journal.pone.0236337] [Medline: 32702065]
13. von Keyserlingk L, Yamaguchi-Pedroza K, Arum R, Eccles JS. Stress of university students before and after campus closure in response to COVID-19. J Community Psychol 2022 Jan;50(1):285-301 [FREE Full text] [doi: 10.1002/jcop.22561] [Medline: 33786864]
14. Santamaría-Vázquez M, Del Líbano M, Martínez-Lezaun I, Ortiz-Huerta JH. Self-regulation of motivation and confinement by COVID-19: a study in Spanish University students. Sustainability 2021 May 13;13(10):5435. [doi: 10.3390/su13105435]
15. Baumeister RF, DeWall CN, Ciarocco NJ, Twenge JM. Social exclusion impairs self-regulation. J Pers Soc Psychol 2005 Apr;88(4):589-604. [doi: 10.1037/0022-3514.88.4.589] [Medline: 15796662]
16. Tasso AF, Hisli Sahin N, San Roman GJ. COVID-19 disruption on college students: academic and socioemotional implications. Psychol Trauma 2021 Jan;13(1):9-15. [doi: 10.1037/tra0000996] [Medline: 33382329]
17. Kecojevic A, Basch CH, Sullivan M, Davi NK. The impact of the COVID-19 epidemic on mental health of undergraduate students in New Jersey, cross-sectional study. PLoS One 2020 Sep 30;15(9):e0239696 [FREE Full text] [doi: 10.1371/journal.pone.0239696] [Medline: 32997683]
18. Khan AH, Sultana MS, Hossain S, Hasan MT, Ahmed HU, Sikder MT. The impact of COVID-19 pandemic on mental health and wellbeing among home-quarantined Bangladeshi students: a cross-sectional pilot study. J Affect Disord 2020 Dec 01;277:121-128 [FREE Full text] [doi: 10.1016/j.jad.2020.07.135] [Medline: 32818775]
19. Kohls E, Baldofski S, Moeller R, Klemm SL, Rummel-Kluge C. Mental health, social and emotional well-being, and perceived burdens of university students during COVID-19 pandemic lockdown in Germany. Front Psychiatry 2021 Apr 6;12:643957 [FREE Full text] [doi: 10.3389/fpsyt.2021.643957] [Medline: 33889102]
20. Alemany-Arrebola I, Rojas-Ruiz G, Granda-Vera J, Mingorance-Estrada Á. Influence of COVID-19 on the perception of academic self-efficacy, state anxiety, and trait anxiety in college students. Front Psychol 2020 Oct 9;11:570017 [FREE Full text] [doi: 10.3389/fpsyg.2020.570017] [Medline: 33154727]
21. Aslan I, Ochnik D, Çınar O. Exploring perceived stress among students in Turkey during the COVID-19 pandemic. Int J Environ Res Public Health 2020 Dec 02;17(23):8961 [FREE Full text] [doi: 10.3390/ijerph17238961] [Medline: 33276520]
22. Bourion-Bédès S, Tarquinio C, Batt M, Tarquinio P, Lebreuilly R, Sorsana C, et al. Stress and associated factors among French university students under the COVID-19 lockdown: the results of the PIMS-CoV 19 study. J Affect Disord 2021 Mar 15;283:108-114 [FREE Full text] [doi: 10.1016/j.jad.2021.01.041] [Medline: 33540333]

23. Deshetler L, Gangadhar M, Battepati D, Koffman E, Mukherjee RK, Menon B. Learning on lockdown: a study on medical student wellness, coping mechanisms and motivation during the COVID-19 pandemic. MedEdPublish 2021;10(1):50. [doi: 10.15694/mep.2021.000050.1]
24. Dhar BK, Ayittey FK, Sarkar SM. Impact of COVID-19 on psychology among the university students. Glob Chall 2020 Sep 28;4(11):2000038 [FREE Full text] [doi: 10.1002/gch2.202000038] [Medline: 33042575]
25. Son C, Hegde S, Smith A, Wang X, Sasangohar F. Effects of COVID-19 on college students' mental health in the united states: interview survey study. J Med Internet Res 2020 Sep 03;22(9):e21279 [FREE Full text] [doi: 10.2196/21279] [Medline: 32805704]
26. Schlichtiger J, Brunner S, Steffen J, Huber BC. Mental health impairment triggered by the COVID-19 pandemic in a sample population of German students. J Investig Med 2020 Dec;68(8):1394-1396. [doi: 10.1136/jim-2020-001553] [Medline: 33087426]
27. Husky MM, Kovess-Masfety V, Swendsen JD. Stress and anxiety among university students in France during Covid-19 mandatory confinement. Compr Psychiatry 2020 Oct;102:152191 [FREE Full text] [doi: 10.1016/j.comppsych.2020.152191] [Medline: 32688023]
28. Robinson E, Sutin AR, Daly M, Jones A. A systematic review and meta-analysis of longitudinal cohort studies comparing mental health before versus during the COVID-19 pandemic in 2020. J Affect Disord 2022 Jan 01;296:567-576 [FREE Full text] [doi: 10.1016/j.jad.2021.09.098] [Medline: 34600966]
29. Xiong J, Lipsitz O, Nasri F, Lui LM, Gill H, Phan L, et al. Impact of COVID-19 pandemic on mental health in the general population: a systematic review. J Affect Disord 2020 Dec 01;277:55-64 [FREE Full text] [doi: 10.1016/j.jad.2020.08.001] [Medline: 32799105]
30. Saleh D, Camart N, Romo L. Predictors of stress in college students. Front Psychol 2017 Jan 25;8:19 [FREE Full text] [doi: 10.3389/fpsyg.2017.00019] [Medline: 28179889]
31. Daniels LM, Goegan LD, Parker PC. The impact of COVID-19 triggered changes to instruction and assessment on university students' self-reported motivation, engagement and perceptions. Soc Psychol Educ 2021;24(1):299-318 [FREE Full text] [doi: 10.1007/s11218-021-09612-3] [Medline: 33613084]
32. Petillion RJ, McNeil WS. Student experiences of emergency remote teaching: impacts of instructor practice on student learning, engagement, and well-being. J Chem Educ 2020 Aug 06;97(9):2486-2493. [doi: 10.1021/acs.jchemed.0c00733]
33. Quintiliani L, Sisto A, Vicinanza F, Curcio G, Tambone V. Resilience and psychological impact on Italian university students during COVID-19 pandemic. Distance learning and health. Psychol Health Med 2022 Jan;27(1):69-80. [doi: 10.1080/13548506.2021.1891266] [Medline: 33602027]
34. Sultana A, Tasnim S, Hossain MM, Bhattacharya S, Purohit N. Digital screen time during the COVID-19 pandemic: a public health concern. F1000Res 2021 Feb 8;10:81. [doi: 10.12688/f1000research.50880.1]
35. David ME, Roberts JA. Smartphone use during the COVID-19 pandemic: social versus physical distancing. Int J Environ Res Public Health 2021 Jan 25;18(3):1034 [FREE Full text] [doi: 10.3390/ijerph18031034] [Medline: 33503907]
36. Montagni I, Guichard E, Kurth T. Association of screen time with self-perceived attention problems and hyperactivity levels in French students: a cross-sectional study. BMJ Open 2016 Feb 26;6(2):e009089 [FREE Full text] [doi: 10.1136/bmjopen-2015-009089] [Medline: 26920440]
37. Amez S, Baert S. Smartphone use and academic performance: a literature review. Int J Educ Res 2020;103:101618. [doi: 10.1016/j.ijer.2020.101618]
38. Pappa S, Yada T, Perälä-Littunen S. International master's degree students' well-being at a Finnish University during COVID-19. Open Educ Stud 2020 Nov 28;2(1):240-251. [doi: 10.1515/edu-2020-0128]
39. Copeland WE, McGinnis E, Bai Y, Adams Z, Nardone H, Devadanam V, et al. Impact of COVID-19 pandemic on college student mental health and wellness. J Am Acad Child Adolesc Psychiatry 2021 Jan;60(1):134-41.e2 [FREE Full text] [doi: 10.1016/j.jaac.2020.08.466] [Medline: 33091568]
40. Magni D, Sestino A. Students learning outcomes and satisfaction. An investigation of knowledge transfer during social distancing policies. Int J Learn Intellect Capital 2021;18(4):339-351. [doi: 10.1504/ijlic.2021.10035251]
41. Hoyt LT, Cohen AK, Dull B, Maker Castro E, Yazdani N. "Constant stress has become the new normal": stress and anxiety inequalities among U.S. college students in the time of COVID-19. J Adolesc Health 2021 Feb;68(2):270-276. [doi: 10.1016/j.jadohealth.2020.10.030] [Medline: 33288459]
42. Juchnowicz D, Baj J, Forma A, Karakuła K, Sitarz R, Bogucki J, et al. The outbreak of SARS-CoV-2 pandemic and the well-being of polish students: the risk factors of the emotional distress during COVID-19 lockdown. J Clin Med 2021 Mar 01;10(5):944 [FREE Full text] [doi: 10.3390/jcm10050944] [Medline: 33804368]
43. Shim TE, Lee SY. College students' experience of emergency remote teaching due to COVID-19. Child Youth Serv Rev 2020 Dec;119:105578 [FREE Full text] [doi: 10.1016/j.childyouth.2020.105578] [Medline: 33071405]
44. Williams KD. Chapter 6 Ostracism: a temporal need - threat model. In: Zanna MP, editor. Advances in Experimental Social Psychology. Volume 41. Amsterdam, The Netherlands: Elsevier; 2009:275-314.
45. Zacher H, Rudolph CW. Individual differences and changes in subjective wellbeing during the early stages of the COVID-19 pandemic. Am Psychol 2021 Jan;76(1):50-62. [doi: 10.1037/amp0000702] [Medline: 32700938]

46. Oliveira L, Mesquita A, Sequeira A, Oliveira A. Emergency remote learning during COVID-19: socio-educational impacts on Portuguese students. In: Proceedings of the 23rd International Conference on Educating Engineers for Future Industrial Revolutions. 2020 Presented at: ICL '20; September 23-25, 2020; Tallinn, Estonia p. 303-314. [doi: 10.1007/978-3-030-68198-2_28]
47. Parker SW, Hansen MA, Bernadowski C. COVID-19 campus closures in the United States: American student perceptions of forced transition to remote learning. Soc Sci 2021 Feb 09;10(2):62. [doi: 10.3390/socsci10020062]
48. Rickham PP. Human experimentation. Code of ethics of the world medical association. Declaration of Helsinki. Br Med J 1964 Jul 18;2(5402):177 [FREE Full text] [doi: 10.1136/bmj.2.5402.177] [Medline: 14150898]
49. Krohne HW, Egloff B, Kohlmann CW, Tausch A. Positive and Negative Affect Schedule--German Version. Washington, DC, USA: American Psychological Association; 1996.
50. Thompson ER. Development and validation of an internationally reliable short-form of the positive and negative affect schedule (PANAS). J Cross Cult Psychol 2007 Mar;38(2):227-242. [doi: 10.1177/0022022106297301]
51. Brähler E, Mühlan H, Albani C, Schmidt S. Teststatistische Prüfung und Normierung der deutschen Versionen des EUROHIS-QOL Lebensqualität-Index und des WHO-5 Wohlbefindens-Index. Diagnostica 2007 Apr;53(2):83-96. [doi: 10.1026/0012-1924.53.2.83]
52. Cohen S, Kamarck T, Mermelstein R. Perceived stress scale. In: Cohen S, Kessler RC, Gordon LU, editors. Measuring Stress: A Guide for Health and Social Scientists. Oxford, UK: Oxford University Press; 1994.
53. Klein EM, Brähler E, Dreier M, Reinecke L, Müller KW, Schmutzer G, et al. The German version of the Perceived Stress Scale - psychometric characteristics in a representative German community sample. BMC Psychiatry 2016 May 23;16:159 [FREE Full text] [doi: 10.1186/s12888-016-0875-9] [Medline: 27216151]
54. vorgelegt von Schmidt-Gürtler LS. Stressbewältigung und Lebensqualität im Bachelorstudiengang Psychologie. Justus-Liebig-Universität Gießen. 2013 Jul. URL: https://d-nb.info/1069065439/34 [accessed 2021-10-20]
55. Hodapp V, Laux L, Spielberger CD. Theorie und Messung der emotionalen und kognitiven Komponente der Prüfungsangst. Zeitschrift Differentielle Diagnostische Psychol 1982;3(3):169-184.
56. Schwarzer R, Jerusalem M. Skalen zur Erfassung von Lehrer-und Schülermerkmalen: Dokumentation der psychometrischen Verfahren im Rahmen der Wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen. Freie Universität Berlin. 1999. URL: https://www.psyc.de/skalendoku.pdf [accessed 2021-10-20]
57. Dickhäuser O, Schöne C, Spinath B, Stiensmeier-Pelster J. Die Skalen zum akademischen Selbstkonzept: Konstruktion und Überprüfung eines neuen Instrumentes. Z Differentielle Diagnostische Psychol 2002 Dec;23(4):393-405. [doi: 10.1024//0170-1789.23.4.393]
58. Bong M, Skaalvik EM. Academic self-concept and self-efficacy: how different are they really? Educ Psychol Rev 2003;15(1):1-40. [doi: 10.1023/A:1021302408382]
59. Rheinberg F, Vollmeyer R, Burns BD. FAM: Ein Fragebogen zur Erfassung aktueller Motivation in Lern- und Leistungssituationen. Diagnostica 2001 Apr;47(2):57-66. [doi: 10.1026//0012-1924.47.2.57]
60. Rheinberg F, Vollmeyer R, Engeser S. Die Erfassung des flow-Erlebens. In: Stiensmeier-Pelster J, Rheinberg F, editors. Diagnostik von Selbstkonzept, Lernmotivation und Selbstregulation. Göttingen, Germany: Hogrefe; 2003:261-279.
61. Schiefele U, Wild KP. Lernstrategien im Studium: Ergebnisse zur Faktorenstruktur und Reliabilität eines neuen Fragebogens. Z Differentielle Diagnostische Psychol 1994;15(4):185-200.
62. Patzelt J, Opitz I. Deutsche Version des academic procrastination state-inventory (APSI-d). Zusammenstellung sozialwissenschaftlicher Items Skalen 2005. [doi: 10.6102/zis139]
63. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007 May;39(2):175-191. [doi: 10.3758/bf03193146] [Medline: 17695343]
64. Beblo T, Fernando S, Klocke S, Griepenstroh J, Aschenbrenner S, Driessen M. Increased suppression of negative and positive emotions in major depression. J Affect Disord 2012 Dec 10;141(2-3):474-479. [doi: 10.1016/j.jad.2012.03.019] [Medline: 22483953]
65. Williams KD. Ostracism. Annu Rev Psychol 2007;58:425-452. [doi: 10.1146/annurev.psych.58.110405.085641] [Medline: 16968209]
66. Voltmer E, Köslich-Strumann S, Walther A, Kasem M, Obst K, Kötter T. The impact of the COVID-19 pandemic on stress, mental health and coping behavior in German University students - a longitudinal study before and after the onset of the pandemic. BMC Public Health 2021 Jul 13;21(1):1385 [FREE Full text] [doi: 10.1186/s12889-021-11295-6] [Medline: 34256717]
67. Wasil AR, Franzen RE, Gillespie S, Steinberg JS, Malhotra T, DeRubeis RJ. Commonly reported problems and coping strategies during the COVID-19 crisis: a survey of graduate and professional students. Front Psychol 2021 Feb 25;12:598557 [FREE Full text] [doi: 10.3389/fpsyg.2021.598557] [Medline: 33716864]
68. Rogowska AM, Kuśnierz C, Bokszczanin A. Examining anxiety, life satisfaction, general health, stress and coping styles during COVID-19 pandemic in Polish sample of university students. Psychol Res Behav Manag 2020 Sep 28;13:797-811 [FREE Full text] [doi: 10.2147/PRBM.S266511] [Medline: 33061695]

69. Gaeta ML, Gaeta L, Rodriguez MD. The impact of COVID-19 home confinement on Mexican university students: emotions, coping strategies, and self-regulated learning. Front Psychol 2021 Apr 28;12:642823 [FREE Full text] [doi: 10.3389/fpsyg.2021.642823] [Medline: 33995197]
70. Bendau A, Plag J, Kunas S, Wyka S, Ströhle A, Petzold MB. Longitudinal changes in anxiety and psychological distress, and associated risk and protective factors during the first three months of the COVID-19 pandemic in Germany. Brain Behav 2021 Feb;11(2):e01964 [FREE Full text] [doi: 10.1002/brb3.1964] [Medline: 33230969]

Abbreviations

MANOVA: multivariate analysis of variance

Edited by A Mavragani; submitted 20.10.21; peer-reviewed by L Günthner, R Ricard, N Marotta; comments to author 08.04.22; revised version received 22.07.22; accepted 23.07.22; published 15.11.22.

Please cite as:
Nuñez TR, Pallasch N, Radtke T
Students' Emotional Well-being and Academic Functioning Before, During, and After Lockdown in Germany: Cohort Study
JMIR Form Res 2022;6(11):e34388
URL: https://formative.jmir.org/2022/11/e34388
doi:10.2196/34388
PMID:36228133

Original Paper

Fear of COVID-19 and Prevention Behaviors: Cross-Lagged Panel Analysis

Katherine M Anderson[1,2], MPH; Jamila K Stockman[1], MPH, PhD

[1]Department of Medicine, School of Medicine, University of California San Diego, La Jolla, CA, United States

[2]Department of Behavioral, Social, and Health Education Sciences, Rollins School of Public Health, Emory University, Atlanta, GA, United States

Corresponding Author:
Katherine M Anderson, MPH
Department of Medicine
School of Medicine
University of California San Diego
9500 Gilman Dr
MC 0507
La Jolla, CA, 92093
United States
Phone: 1 7209036196
Email: k4anderson@health.ucsd.edu

Abstract

Background: The ongoing COVID-19 pandemic has brought forth conversations about effective behavior change models for increasing prevention behavior, ranging from wearing masks in public to physical distancing. Among the considered behavior change techniques is the use of fear appeals, through which a negative possible outcome is emphasized to invoke fear, which in turn may promote prevention behaviors to counter the likelihood of the negative outcome. Although fear is hypothesized as health promoting in some theories of health behavior, little research has rigorously assessed the relationship.

Objective: In our exploratory analyses, we aim to examine the association, including directionality of the association between fear of COVID-19 and COVID-19 prevention behaviors across 2 time points during the early COVID-19 pandemic among a sample of US women.

Methods: The COPE study, a web-based survey of US women's COVID-19 experiences, was deployed in May-June 2020 (time 1) with follow-up in December 2020-January 2021 (time 2; n=200). Demographic characteristics as well as fear of COVID-19 and COVID-19 prevention behaviors (eg, staying home except for essential activities, physical distancing in public, and masking in public) were measured. Descriptive and bivariate analyses were used to characterize COVID-19 prevention behaviors and fear of COVID-19 among participants. Cross-lagged panel analysis, a type of structural equation modeling that assesses directionality of temporal associations, was used to understand relationships, if any, between variables of interest.

Results: We found cross-sectional associations between fear of COVID-19 and staying home and physical distancing, as well as temporal associations between fear at time 1 and time 2 and prevention behaviors at time 1 and time 2. However, results of the cross-lagged panel analysis indicated no cross-lagged temporal relationships between fear of COVID-19 and COVID-19 prevention behaviors 6 months apart.

Conclusions: Fear of COVID-19 did not appear to predict COVID-19 prevention behaviors 6 months after initial measurements among the sample of women recruited for our study. Future research should rigorously test these associations longitudinally, and alternative methods of public health prevention promotion should be considered.

(JMIR Form Res 2022;6(11):e35730) doi:10.2196/35730

KEYWORDS

fear appeals; structural equation modelling; cross-lagged model; prevention behavior; COVID-19; fear; women; behavior; change; health; physical distance; relationships; pandemic; research; association; prevention; experience; panel; interest; public; distancing

Introduction

Amid the ongoing COVID-19 pandemic, behavioral prevention is vital to reduce viral transmission. Uptake of prevention behaviors, including masking in public, social distancing, and staying home except for essential activities [1], has been variable [2,3], underscoring the need for identification of mechanisms of behavior change to encourage uptake of behaviors appropriate for the current level of risk and dominant variants. Concerningly, it has been suggested that those who have previously been infected with COVID-19 are even less likely to use prevention behaviors [4], despite risk of reinfection and continued evolution of variants. As a mechanism for behavior change, fear appeals have long been used in public health [5] and justified with behavioral theory constructs of perceived risk or perceived severity, such as within the Health Belief Model [6-8]. Results are mixed on the effectiveness [9] and ethics [10] of fear appeals. Significant literature has emerged on the role of fear of COVID-19 [2,11,12], though somewhat less on fear and its association with prevention behavior [3,8]. The available literature has identified positive cross-sectional associations between anxiety or fear of infection related to COVID-19 and willingness to vaccinate [13], with odds of vaccine hesitancy approximately 5.5 times greater among those with no fear of COVID-19 compared to those with a great extent of fear [14]. Fear has also been demonstrated to mediate the relationship between COVID-19 exposure and intent to be vaccinated [15] and the relationship between COVID-19 information seeking and prevention behavior performance [16]. Further, COVID-19 fear is predictive of behavioral intention to perform prevention behaviors [17,18], and fear of contamination is predictive of obsessive-compulsive hand washing [19]. In contrast to this evidence, longitudinal studies have found that as the pandemic progressed, fear diminished over time while prevention behaviors increased, even as uncertainty related to the pandemic remained stable [20].

Arguments for fear appeals are based on assumed directionality from fear of an outcome to behavior preventing that outcome. This does not take into account the potential of promotion and use of prevention behaviors increasing anxiety or fear related to COVID-19 nor does it consider competing fear, such as the negative impacts of social isolation resulting from prevention behaviors [21,22]. However, most studies of fear appeals do not thoroughly assess directionality or further interactions such as mediation or moderation [23,24], with few exceptions [25]. As such, this analysis attempts to answer the following research question: what is the directionality of the relationship between fear of COVID-19 and practice of mask wearing, physical distancing in public, and staying home except for essential activities across 2 time points 6 months apart among adult US women enrolled in an internet-based study?

Methods

Ethical Considerations

All procedures were approved by the University of California San Diego's (project 200663) institutional review board. Participants provided documented informed consent prior to completing surveys at each time point.

Procedures

Participants were recruited for The COPE Study baseline survey from May to June 2020 (time 1 [t_1]), using the Facebook advertising algorithm. Women aged ≥18 years were targeted for advertisements on Facebook (83.1%) and other non-Facebook–owned websites on which the program anticipates reaching the demographic of interest (Facebook Audience Network: 16.5%). The aim of The COPE Study was to understand US adult women's experiences with COVID-19, service access, and violence during the first months of the COVID-19 pandemic. Of 682 potential participants, 633 (92.8%) provided consent and responded to eligibility questions. Eligible participants were ≥18 years of age, self-identified as women, lived in the United States, and could speak and understand English; of 626 eligible participants, 491 (78%) completed the internet-based survey at t_1. For the follow-up survey, administered between December 2020 and January 2021 (time 2 [t_2]), we conducted 2-stage stamping. In stage 1, all non-White participants were purposively sampled for overrepresentation of racial or ethnic minorities. Due to the underrepresentation of racial and ethnic minority individuals in research, this tactic helped to bolster the diversity of our sample. In stage 2, we used simple random sampling of the remaining participants, not inclusive of nonresponders from the first stage of sampling, until achieving a sample size of N=333, due to funding constraints. Each participant was emailed an invitation to participate in the follow-up survey at an email address provided at t_1, along with permission to recontact them. All data were collected using REDCap [26]. Participants were compensated with Amazon e-gift cards (t_1: US $20; t_2: US $15).

Participants

Of 333 participants invited, 296 completed the survey at t_2 (88.9%), and 200 (60%) passed response validity tests. Validity testing, which was only performed at t_2, took the form of an attention check, wherein participants were instructed to select a specific response to a question at 2 independent points in the survey. Of the remaining cases, 4/200 (2%) had missing data. Not accounting for data removed due to failed validity testing, 200/237 (84%) of invited participants completed the survey (200 of 237 participants not known failed validity checks). Median age of participants was 33 (IQR 18-69); 118/200 (59%) were White; 27/200 (13.5%) were Black; 16/200 (8%) were of Asian descent; and 9/200 (4.5%) identified as being multiracial; 27/200 (13.5%) were Hispanic or Latinx. Most participants had completed college (n=74/200, 37%) or had at least some graduate school education (n=62/200, 31%); 123/200 (62.5%) were employed, with 47/200 (23.5%) self-identified as essential workers.

Measures

Four variables of interest were considered. Fear of COVID-19 was measured using the 7-item Fear of COVID-19 Scale [27]. Items included being afraid of dying from COVID-19 and experiencing physical symptoms of fear. Response options were on a 5-point Likert scale from "strongly disagree" to "strongly

agree." Responses were summed, with scores ranging from 7 to 35 (Cronbach α=.897 at t_1; Cronbach α=.904 at t_2). The following 3 prevention behaviors were measured: (1) staying home except for essential activities, (2) physical distancing of 6 feet from nonhousehold members, and (3) using a face mask in public. Participants were asked, "which of the following prevention behaviors have you been using?" with response options of "yes" or "no" for each.

Statistical Analyses

Descriptive statistics were calculated, and 2-tailed independent samples *t* tests were run using IBM SPSS (version 24; IBM Corp). Given multiple comparisons, a Bonferroni adjustment and significance was set at *P*=.008 for *t* tests. Cross-lagged panel analysis [28] was conducted in Mplus (version 8) [29]. All responses were retained through maximum likelihood estimation with robust standard errors, using the weighted least squares mean and variance adjusted estimator in MPlus [30]; significance was set at *P*<.05. In total, 3 relationships of interest were assessed over 2 time points, as follows: (1) practice of staying home except for essential activities during the COVID-19 pandemic and fear of COVID-19, (2) practice of physical distancing in public during the COVID-10 pandemic and fear of COVID-19, and (3) practice of wearing a mask in public during the COVID-19 pandemic and fear of COVID-19. Age, formal educational attainment, and parental status (t_1), as well as essential worker status (t_2) were entered as time-invariant covariates in adjusted models. As a saturated cross-lagged model with 2 time points, goodness-of-fit indices are not used for model interpretation [31].

Results

Demographic characteristics of the sample and distribution of the predictor and outcome variables are presented in Table 1 and Table 2; bivariate results are presented in Table 3. Prevention behaviors were practiced by most participants across time points, though the proportion of 'staying home except for essential activities' declined from t_1 to t_2 (93.5% to 78.5%), and the proportion of 'wearing masks in public' increased (83.5% to 95.5%), whereas the proportion of 'physically distancing in public' remained approximately the same (90% at t_1 and 89% at t_2). Fear of COVID-19 at t_1 and t_2 were not significantly different (mean difference 0.460; *P*=.25, not depicted) and were significantly correlated (*r*=0.662; *P*<.001; Table 4). Mean fear of COVID-19 was higher among women who stayed home except for essential activities at t_2; however, upon correction for multiple comparisons, it was not significant (20.87 vs 19.44; *P*=.047), and it was significantly higher among women who maintained physical distance of 6 feet in public at both t_1 (21.23 vs 17.0; *P*=.008) and t_2 (20.82 vs 16.10; *P*=.005). Across other prevention behaviors and time points, fear and prevention behaviors were not statistically significantly associated.

Beta estimates for each cross-lagged panel model are presented and depicted in Figure 1, and beta estimates and correlations as well as significance of associations at a level of *P*<.05 are presented in Table 4. Results of the cross-lagged models indicate that fear of COVID-19 (t_1) does not predict practicing prevention behaviors 6 months later (t_2), including staying home except for essential activities (adjusted model: β=.022; *P*=.11), physically distancing in public (adjusted model: β=.005; *P*=.74), or wearing a mask in public (adjusted model: β=.003; *P*=.86). Relatedly, practicing of prevention behaviors (t_1) did not predict fear of COVID-19 6 months later (t_2) for staying home except for essential activities (adjusted model: β=1.577; *P*=.31), physically distancing in public (adjusted model: β=2.001; *P*=.08), or wearing a mask in public (adjusted model: β=.823; *P*=.31). Fear of COVID-19 was strongly and significantly associated at t_1 and t_2 across all models (*P*<.001 for all), and prevention behavior at t_1 was significantly associated with prevention behavior at t_2 across all models (staying home: adjusted *P*=.02; distancing in public and wearing a mask in public: adjusted *P*<.001). Finally, physical distancing in public at t_1 was statistically significantly associated with fear of COVID-19 at t_1 (adjusted model: β=.380; *P*=.008).

Table 1. Demographic characteristics among a sample of US adult women (N=200).

Variables	Values
Age (years), mean (SD)	34.89 (11.1)
Race, N (%)[a]	
White	118 (59)
Black	27 (13.5)
Asian	16 (8)
American Indian and Alaska Native	5 (2.5)
Multiple races	9 (4.5)
Middle Eastern and North African	3 (1.5)
Pacific Islander	0 (0)
Ethnicity (Hispanic or Latinx), n (%)	27 (13.5)
Education, n (%)	
High school diploma, GED[b], or less	28 (14)
Some college, or some (or completed) trade or vocational school	34 (17)
Completed college	74 (37)
Some (or completed) graduate school	62 (31)
Employed, n (%)[c]	125 (62.5)
In a relationship, n (%)	128 (64)
Parent to children of any age, n (%)	100 (50)
Essential worker, n (%)	47 (23.5)

[a]Participants selected all applicable races; some participants did not provide a race, identifying only as Hispanic or Latinx.

[b]GED: Graduate Educational Development.

[b]Includes full-time and part-time employees as well as self-employed.

Table 2. Variables of interest among a sample of US adult women (N=200).

Variables of interest	Time 1 (baseline)	Time 2 (6-month follow-up)
Staying home, n (%)	187 (93.5)	157 (78.5)
Physical distancing, n (%)	180 (90)	178 (89)
Masking in public, n (%)	167 (83.5)	191 (95.5)
Fear of COVID-19, mean (SD)	20.81 (6.76)	20.35 (7.11)

Table 3. Bivariate analyses of prevention behavior and fear of COVID-19 among a sample of US adult women (N=200). Italicized *P* values are significant.

Prevention behaviors	Fear of COVID-19			
	Time 1 (baseline), mean (SD)	*P* value	Time 2 (6-month follow-up), mean (SD)	*P* value
Staying home				
Time 1				
Yes	20.92 (6.7)	0.39	20.5 (7.06)	0.25
No	19.23 (7.64)		18.15 (7.86)	
Time 2				
Yes	21.18 (6.66)	.13	20.87 (6.97)	.05
No	19.44 (7.00)		19.44 (7.39)	
Physical distancing				
Time 1				
Yes	21.23 (7.48)	*.008*	20.82 (6.93)	*.005*
No	17.0 (7.97)		16.10 (7.48)	
Time 2				
Yes	20.92 (6.75)	.53	20.58 (7.06)	.19
No	19.95 (6.93)		18.45 (7.45)	
Masking in public				
Time 1				
Yes	20.95 (6.51)	.51	20.52 (6.77)	.52
No	20.09 (7.97)		19.48 (8.72)	
Time 2				
Yes	20.86 (6.59)	.64	20.40 (7.08)	.63
No	19.78 (10.07)		19.22 (8.18)	

Table 4. Estimated betas and correlations for cross-lagged models of prevention behavior and fear of COVID-19 among a sample of US adult women (N=200). Italicized values are significant.

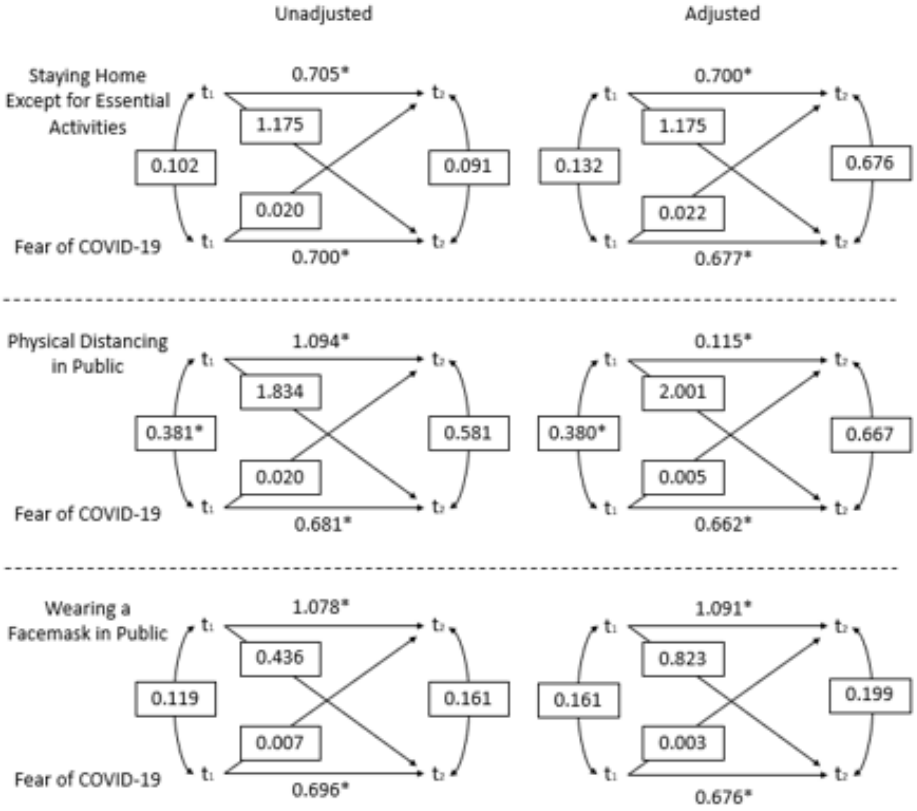
Regression models	Unadjusted		Adjusted[a]	
	Estimate	*P* value	Estimate	*P* value
Model 1				
Staying home (t_1[b] to t_2[c])	*.705*	*.05*	*.700*	*.02*
Fear (t_1 to t_2)	*.700*	*<.001*	*.677*	*<.001*
Staying home (t_1) to fear (t_2)	1.175	.54	1.577	.317
Fear (t_1) to staying home (t_2)	.020	.15	.022	.13
Staying home (t_1) with fear (t_1)	.102	.36	.132	.22
Staying home (t_2) with fear (t_2)	.646	.18	.676	.17
Model 2				
Distancing (t_1 to t_2)	*1.094*	*<.001*	*1.15*	*<.001*
Fear (t_1 to t_2)	*.681*	*<.001*	*.662*	*<.001*
Distancing (t_1) to fear (t_2)	1.834	.68	2.001	.08
Fear (t_1) to distancing (t_2)	.002	.91	.005	.74
Distancing (t_1) with fear (t_1)	*.381*	*.01*	*.380*	*.008*
Distancing (t_2 with fear (t_2)	.581	.31	.677	.25
Model 3				
Masking (t_1 to t_2)	*1.078*	*<.001*	*1.091*	*<.001*
Fear (t_1 to t_2)	*.696*	*<.001*	*.676*	*<.001*
Masking (t_1) to fear (t_2)	.436	.60	.823	.31
Fear (t_1) to masking (t_2)	.007	.67	.003	.86
Masking (t_1) with fear (t_1)	.119	.45	.161	.30
Masking (t_2) with fear (t_2)	.161	.81	.199	.75

[a]Models are adjusted for age, formal educational attainment, and parental status (time 1) as well as essential worker status (time 2).

[b]t_1: time 1 (baseline).

[c]t_2: time 2 (6-month follow-up).

Figure 1. Cross-lagged panel models of prevention behavior and fear of COVID-19 among a sample of US adult women. *Denotes statistical significance at a value of .05. Adjusted models are adjusted for age, education, parental status, and essential worker status. t_1: time 1. t_2: time 2.



Discussion

Principal Findings

Findings from this exploratory analysis indicate that although fear of COVID-19 and practice of prevention behaviors may be self-predictive over 6 months, they are not cross-predictive among the cohort of US women enrolled in The COPE Study. In bivariate analyses, significant associations were found between fear and staying home except for essential activities at t_1 and fear and physical distancing in public at both time points; however, further exploration of these relationships using cross-lagged panel analysis suggests these may not be temporal associations.

Comparison With Prior Work

Previous literature has examined the use of fear appeals in behavioral health interventions, largely with mixed or inconclusive findings. Conflicting meta-analyses have suggested that fear appeals are not sufficiently effective [6] and that 'strong' fear appeals are very effective [9]. In the context of COVID-19, studies have found that fear is cross-sectionally associated with compliance with behavioral prevention [23] as well as willingness and intention to be vaccinated [13,14], and it is longitudinally associated with intention to perform prevention behaviors [17,18]. However, longitudinal findings suggest that fear and practice of prevention behaviors have had inverse trajectories throughout the pandemic, and therefore, are not positively associated [20]. Our findings support the latter of these studies and add to the literature suggesting that fear is not an effective predictor of prevention behavior over a span of 6 months among US women included in the sample.

Limitations

There were a number of limitations to consider in this study. The COPE Study used an internet-based sample recruited through Facebook advertising, enabling broad reach, as most women have access to internet and use Facebook [32]; however, this resulted in variable data completion rates and quality, despite the presence of validity checks within the survey. Additionally, this is a secondary, exploratory analysis of data intended to capture women's interpersonal experiences; assessing prevention behavior and fear of COVID-19 was not the primary focus, and therefore, the data captured are not ideal for this application. Measures used for prevention behavior were captured using a dichotomous variable, restricting the range of responses and prohibiting the exploration of more nuanced dynamics of prevention behavior frequency. A validated scale for fear of COVID-19 was used in this study, but the scale was developed rapidly in the midst COVID-19 pandemic; therefore, development may not have been as rigorous, possibly jeopardizing validity; furthermore, recent findings have documented issues with measurement invariance across countries [33]; however, reliability of the scale was strong at both time points. Participant data were only available for 2 time points, underscoring the need for caution in causal interpretation; additional time points would strengthen causal inference in future research. Maximum likelihood estimation was used in order to use all available data without listwise deletion. Upon removal of participants from the denominator who failed validity checks, an 84% (200/237) completion rate indicates that there

is the small possibility of some response bias. However, validity checks were only performed at t_2, and the high rate of failure of validity tests suggests that t_1 data may have faced similar challenges in terms of invalid responses; however, only data from participants who passed t_2 validity checks were included in the analysis, ensuring at least one layer of assurance and increasing the likelihood of only valid data being included. It is possible that a causal relationship exists between fear and prevention behaviors that was not identified in this study due to the length of time between assessments; we were not able to assess if fear had a more proximal but not cross-sectional impact on prevention behaviors. Finally, although the average age and racial or ethnic distribution of participants is similar to that of the US population of women, this sample is not representative of US women and may not adequately represent women without access to the internet or regular use of social media, potentially underrepresenting low-income or older individuals who may be most at risk for COVID-19. Each of these should be taken into consideration to weigh against analysis findings.

Conclusions

Despite these limitations, this exploratory analysis uses innovative methods to examine the directionality of an important relationship hypothesized in the public health sphere—that between fear of a health outcome and prevention behaviors related to the outcome—in the highly relevant context of COVID-19. These findings have implications for public health educational and communication efforts; particularly, it may not be effective to emphasize fear as a public health tactic to promote COVID-19 prevention behaviors. The ethicality of fear appeals, particularly given the potential lack of effectiveness, should continue to be discussed; this is particularly true under circumstances of identified effectiveness, wherein there should be sustained conversation of what is effective versus what is acceptable and appropriate. Formative research on the effectiveness of fear-provoking public health campaigns should be rigorously conducted to ensure cost-effective distribution of funding to effective educational and behavior change campaigns. Further, practitioners and researchers may want to consider the nuanced dynamics of fear, when fear may not be equally applied to others (particularly those who are immunocompromised or otherwise at risk) and oneself, limiting its influence on personal behaviors. Alternative methods, such as changing negative attitudes about prevention behaviors and improving subjective norms [34], should be explored for their feasibility and effectiveness in altering prevention behaviors in the context of COVID-19. Enlisting opinion leaders within communities disproportionately affected by COVID-19 is critical in this effort, as opinion leaders can act as gatekeepers for prevention efforts, help change social norms, and accelerate behavior change [35]. Simultaneously, community outreach and education programs can be used to maximize uptake and adherence to prevention behaviors.

Acknowledgments

The authors would like to thank Regine Haardoerfer for her support in the development of this manuscript.

Conflicts of Interest

None declared.

References

1. West R, Michie S, Rubin GJ, Amlôt R. Applying principles of behaviour change to reduce SARS-CoV-2 transmission. Nat Hum Behav 2020 May;4(5):451-459. [doi: 10.1038/s41562-020-0887-9] [Medline: 32377018]
2. Stockman JK, Wood BA, Anderson KM. Racial and Ethnic Differences in COVID-19 Outcomes, Stressors, Fear, and Prevention Behaviors Among US Women: Web-Based Cross-sectional Study. J Med Internet Res 2021 Jul 12;23(7):e26296 [FREE Full text] [doi: 10.2196/26296] [Medline: 33819909]
3. Anderson KM, Stockman JK. Staying Home, Distancing, and Face Masks: COVID-19 Prevention among U.S. Women in The COPE Study. Int J Environ Res Public Health 2020 Dec 29;18(1):180 [FREE Full text] [doi: 10.3390/ijerph18010180] [Medline: 33383745]
4. Kaim A, Siman-Tov M, Jaffe E, Adini B. From isolation to containment: perceived fear of infectivity and protective behavioral changes during the COVID-19 vaccination campaign. Int J Environ Res Public Health 2021 Jun 16;18(12):6503 [FREE Full text] [doi: 10.3390/ijerph18126503] [Medline: 34208729]
5. Leventhal H. Fear appeals and persuasion: the differentiation of a motivational construct. Am J Public Health 1971 Jun;61(6):1208-1224. [doi: 10.2105/ajph.61.6.1208] [Medline: 4110702]
6. Ruiter R, Kessels L, Peters G, Kok G. Sixty years of fear appeal research: current state of the evidence. Int J Psychol 2014 Apr;49(2):63-70 [FREE Full text] [doi: 10.1002/ijop.12042] [Medline: 24811876]
7. Glanz K, Rimer B, Viswanath K. Health Behavior and Health Education: Theory, Research, and Practice. Hoboken, NJ: John Wiley & Sons; 2008.
8. Park T, Ju I, Ohs JE, Hinsley A. Optimistic bias and preventive behavioral engagement in the context of COVID-19. Res Social Adm Pharm 2021 Jan;17(1):1859-1866 [FREE Full text] [doi: 10.1016/j.sapharm.2020.06.004] [Medline: 33317765]
9. Witte K, Allen M. A meta-analysis of fear appeals: implications for effective public health campaigns. Health Educ Behav 2000 Oct;27(5):591-615. [doi: 10.1177/109019810002700506] [Medline: 11009129]

10. Guttman N, Salmon CT. Guilt, fear, stigma and knowledge gaps: ethical issues in public health communication interventions. Bioethics 2004 Nov;18(6):531-552. [doi: 10.1111/j.1467-8519.2004.00415.x] [Medline: 15580723]
11. Salehi L, Rahimzadeh M, Molaei E, Zaheri H, Esmaelzadeh-Saeieh S. The relationship among fear and anxiety of COVID-19, pregnancy experience, and mental health disorder in pregnant women: a structural equation model. Brain Behav 2020 Nov;10(11):e01835 [FREE Full text] [doi: 10.1002/brb3.1835] [Medline: 32969190]
12. Heiat M, Heiat F, Halaji M, Ranjbar R, Tavangar Marvasti Z, Yaali-Jahromi E, et al. Phobia and Fear of COVID-19: origins, complications and management, a narrative review. Ann Ig 2021;33(4):360-370 [FREE Full text] [doi: 10.7416/ai.2021.2446] [Medline: 33908601]
13. Bendau A, Plag J, Petzold MB, Ströhle A. COVID-19 vaccine hesitancy and related fears and anxiety. Int Immunopharmacol 2021 Aug;97:107724 [FREE Full text] [doi: 10.1016/j.intimp.2021.107724] [Medline: 33951558]
14. Willis DE, Andersen JA, Bryant-Moore K, Selig JP, Long CR, Felix HC, et al. COVID-19 vaccine hesitancy: Race/ethnicity, trust, and fear. Clin Transl Sci 2021 Jul 02:2200-2207 [FREE Full text] [doi: 10.1111/cts.13077] [Medline: 34213073]
15. Chu H, Liu S. Integrating health behavior theories to predict American's intention to receive a COVID-19 vaccine. Patient Educ Couns 2021 Aug;104(8):1878-1886 [FREE Full text] [doi: 10.1016/j.pec.2021.02.031] [Medline: 33632632]
16. Kim K, Yang J, Jeon YJ, Lee YJ, Lee Y, Kim HC, et al. The effects of information-seeking behaviours on prevention behaviours during the COVID-19 pandemic: the mediating effects of anxiety and fear in Korea. Epidemiol Health 2021;43:e2021085 [FREE Full text] [doi: 10.4178/epih.e2021085] [Medline: 34696572]
17. Frounfelker RL, Santavicca T, Li ZY, Miconi D, Venkatesh V, Rousseau C. COVID-19 experiences and social distancing: insights from the theory of planned behavior. Am J Health Promot 2021 Nov;35(8):1095-1104 [FREE Full text] [doi: 10.1177/08901171211020997] [Medline: 34074154]
18. Roberto AJ, Zhou X, Lu AH. The effects of perceived threat and efficacy on college students' social distancing behavior during the COVID-19 pandemic. J Health Commun 2021 Apr 03;26(4):264-271. [doi: 10.1080/10810730.2021.1903628] [Medline: 34086533]
19. Knowles KA, Olatunji BO. Anxiety and safety behavior usage during the COVID-19 pandemic: The prospective role of contamination fear. J Anxiety Disord 2021 Jan;77:102323 [FREE Full text] [doi: 10.1016/j.janxdis.2020.102323] [Medline: 33137593]
20. Peng S, Yang XY, Yang T, Zhang W, Cottrell RR. Uncertainty stress, and its impact on disease fear and prevention behavior during the COVID-19 epidemic in China: a panel study. am j health behav 2021 Mar 01;45(2):334-341. [doi: 10.5993/ajhb.45.2.12]
21. Flanagan EW, Beyl RA, Fearnbach SN, Altazan AD, Martin CK, Redman LM. The impact of COVID-19 stay-at-home orders on health behaviors in adults. Obesity (Silver Spring) 2021 Feb;29(2):438-445 [FREE Full text] [doi: 10.1002/oby.23066] [Medline: 33043562]
22. Singh S, Roy D, Sinha K, Parveen S, Sharma G, Joshi G. Impact of COVID-19 and lockdown on mental health of children and adolescents: a narrative review with recommendations. Psychiatry Res 2020 Nov;293:113429 [FREE Full text] [doi: 10.1016/j.psychres.2020.113429] [Medline: 32882598]
23. Harper CA, Satchell LP, Fido D, Latzman RD. Functional fear predicts public health compliance in the COVID-19 pandemic. Int J Ment Health Addict 2020 Apr 27:1-14 [FREE Full text] [doi: 10.1007/s11469-020-00281-5] [Medline: 32346359]
24. Mallinas S, Maner J, Ashby Plant E. What factors underlie attitudes regarding protective mask use during the COVID-19 pandemic? Pers Individ Dif 2021 Oct;181:111038 [FREE Full text] [doi: 10.1016/j.paid.2021.111038] [Medline: 34092834]
25. Ayandele O, Ramos-Vera CA, Iorfa SK, Chovwen CO, Olapegba PO. Exploring the complex pathways between the fear of COVID-19 and preventive health behavior among Nigerians: mediation and moderation analyses. Am J Trop Med Hyg 2021 Jul 08;105(3):701-707 [FREE Full text] [doi: 10.4269/ajtmh.20-0994] [Medline: 34237014]
26. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
27. Ahorsu DK, Lin C, Imani V, Saffari M, Griffiths MD, Pakpour AH. The fear of COVID-19 scale: development and initial validation. Int J Ment Health Addict 2020 Mar 27:1-9 [FREE Full text] [doi: 10.1007/s11469-020-00270-8] [Medline: 32226353]
28. Finkel S. Causal Analysis with Panel Data. New York, NY: Sage Publications; 1995.
29. Muthén L, Muthén B. Mplus. Statistical analysis with latent variables User's guide. 2009. URL: https://www.statmodel.com/download/usersguide/MplusUserGuideVer_8.pdf [accessed 2021-01-05]
30. Bandalos DL. Relative performance of categorical diagonally weighted least squares and robust maximum likelihood estimation. Struct Equ Modeling 2014 Jan 31;21(1):102-116. [doi: 10.1080/10705511.2014.859510]
31. Hayashi K, Bentler P, Yuan K. 7 - Structural equation modeling. In: Essential Statistical Methods for Medical Statistics. Boston: North-Holland; 2011:202-234.
32. Internet/broadband fact sheet. Pew Research Center. 2019. URL: https://www.pewresearch.org/internet/fact-sheet/internet-broadband/ [accessed 2022-11-09]

33. Lin C, Hou W, Mamun MA, Aparecido da Silva J, Broche-Pérez Y, Ullah I, et al. Fear of COVID-19 Scale (FCV-19S) across countries: measurement invariance issues. Nurs Open 2021 Jul;8(4):1892-1908 [FREE Full text] [doi: 10.1002/nop2.855] [Medline: 33745219]
34. Mao Y, Chen H, Wang Y, Chen S, Gao J, Dai J, et al. How can the uptake of preventive behaviour during the COVID-19 outbreak be improved? An online survey of 4827 Chinese residents. BMJ Open 2021 Feb 16;11(2):e042954 [FREE Full text] [doi: 10.1136/bmjopen-2020-042954] [Medline: 33593779]
35. Valente TW, Pumpuang P. Identifying opinion leaders to promote behavior change. Health Educ Behav 2007 Dec;34(6):881-896. [doi: 10.1177/1090198106297855] [Medline: 17602096]

Abbreviations

t_1: time 1
t_2: time 2

Edited by A Mavragani; submitted 15.12.21; peer-reviewed by CY Lin, R Shahriarirad, K Ross, J Agley; comments to author 18.07.22; revised version received 17.09.22; accepted 26.09.22; published 17.11.22.

Please cite as:
Anderson KM, Stockman JK
Fear of COVID-19 and Prevention Behaviors: Cross-Lagged Panel Analysis
JMIR Form Res 2022;6(11):e35730
URL: https://formative.jmir.org/2022/11/e35730
doi:10.2196/35730
PMID:36346895

Original Paper

Prevalence and Correlates of COVID-19 Vaccine Information on Family Medicine Practices' Websites in the United States: Cross-sectional Website Content Analysis

Jonathan V Ackleh-Tingle[1], MSc; Natalie M Jordan[1], MPH; Udodirim N Onwubiko[1], MBBS, MPH; Christina Chandra[1], MPH; Paige E Harton[1], MPH; Shelby T Rentmeester[1], MPH; Allison T Chamberlain[1], PhD, MS

Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA, United States

Corresponding Author:
Jonathan V Ackleh-Tingle, MSc
Department of Epidemiology
Rollins School of Public Health
Emory University
1518 Clifton Road
Atlanta, GA, 30322
United States
Phone: 1 404 727 6159
Email: jonathan.tingle@emory.edu

Abstract

Background: Primary care providers are regarded as trustworthy sources of information about COVID-19 vaccines. Although primary care practices often provide information about common medical and public health topics on their practice websites, little is known about whether they also provide information about COVID-19 vaccines on their practice websites.

Objective: This study aimed to investigate the prevalence and correlates of COVID-19 vaccine information on family medicine practices' website home pages in the United States.

Methods: We used the Centers for Medicare and Medicaid National Provider Identifier records to create a sampling frame of all family medicine providers based in the United States, from which we constructed a nationally representative random sample of 964 family medicine providers. Between September 20 and October 8, 2021, we manually examined the practice websites of these providers and extracted data on the availability of COVID-19 vaccine information, and we implemented a 10% cross-review quality control measure to resolve discordances in data abstraction. We estimated the prevalence of COVID-19 vaccine information on practice websites and website home pages and used Poisson regression with robust error variances to estimate crude and adjusted prevalence ratios for correlates of COVID-19 vaccine information, including practice size, practice region, university affiliation, and presence of information about seasonal influenza vaccines. Additionally, we performed sensitivity analyses to account for multiple comparisons.

Results: Of the 964 included family medicine practices, most (n=509, 52.8%) had ≥10 distinct locations, were unaffiliated with a university (n=838, 87.2%), and mentioned seasonal influenza vaccines on their websites (n=540, 56.1%). In total, 550 (57.1%) practices mentioned COVID-19 vaccines on their practices' website home page, specifically, and 726 (75.3%) mentioned COVID-19 vaccines anywhere on their practice website. As practice size increased, the likelihood of finding COVID-19 vaccine information on the home page increased (n=66, 27.7% among single-location practices, n=114, 52.5% among practices with 2-9 locations, n=66, 56.4% among practices with 10-19 locations, and n=304, 77.6% among practices with 20 or more locations, $P<.001$ for trend). Compared to clinics in the Northeast, those in the West and Midwest United States had a similar prevalence of COVID-19 vaccine information on website home pages, but clinics in the south had a lower prevalence (adjusted prevalence ratio 0.8, 95% CI 0.7 to 1.0; $P=.02$). Our results were largely unchanged in sensitivity analyses accounting for multiple comparisons.

Conclusions: Given the ongoing COVID-19 pandemic, primary care practitioners who promote and provide vaccines should strongly consider utilizing their existing practice websites to share COVID-19 vaccine information. These existing platforms have the potential to serve as an extension of providers' influence on established and prospective patients who search the internet for information about COVID-19 vaccines.

(JMIR Form Res 2022;6(11):e38425) doi:10.2196/38425

KEYWORDS

primary care; vaccine hesitancy; COVID-19; health communications; health information; health website; family practice; primary care; vaccine information; online health; health platform; online information

Introduction

As of July 2022, at least 1 million Americans have died from COVID-19, although official counts may underestimate total attributable deaths [1,2]. While 67% of the total population has been fully vaccinated against COVID-19 [1], achieving community protection requires a coverage of at least 80% [3]. However, gains in vaccination coverage within the United States have slowed, in part owing to vaccine hesitancy, which is defined as a "delay in acceptance or refusal of [COVID-19] vaccination despite availability of vaccination services" [4,5].

COVID-19 vaccine hesitancy is a well-established infodemiological problem in public health, resulting from a complex interplay among many factors, including, but not limited to, systemic racism, political ideology, intentional and unintentional spread of misinformation, social media networks and behaviors, and skepticism of the scientific community [6-14]. While longer-term solutions to COVID-19 vaccine hesitancy, specifically—as well as hesitancy toward other vaccines, more generally—will require targeting these systems-level causes, shorter-term interventions can also improve COVID-19 vaccine uptake by addressing downstream factors that influence vaccine intent at the level of clinics and individuals, such as individuals' health literacy and vaccine proponents' trustworthiness [15-17]. In the early stages of the COVID-19 pandemic, the US Centers for Disease Control and Prevention identified many of these factors as targets of their national campaign, "Vaccinate with Confidence: Strategy to Reinforce Confidence in COVID-19 Vaccines" [17]. In particular, the Vaccinate with Confidence campaign emphasizes the importance of "effective messages delivered by trusted messengers," in which physicians and other health care professionals tailor narratives about COVID-19 vaccines to the unique needs of their particular patient populations [17].

Primary care providers (PCPs) are consistently regarded as trusted messengers for vaccine-related information and are well-qualified to communicate effectively to overcome hesitancy [17-19]. Enhancing PCPs' web-based promotion of COVID-19 vaccines through their existing practice websites may allow them to reach individuals seeking COVID-19 vaccine information on the internet. Patients already visit these websites to access logistical information such as practice address, telephone number, and patient portals. However, it is unclear whether PCPs use their websites to communicate information about COVID-19 vaccines. The aim of this paper was to describe the prevalence and correlates of the provision of COVID-19 vaccine information on family medicine practices' websites in the United States.

Methods

Overview

Using the National Provider Identifier records of the Centers for Medicare and Medicaid Services' National Plan and Provider Enumeration System, we created a sampling frame of 136,531 providers who indicated "family medicine" as their primary taxonomy code and were practicing in the United States as of September 12, 2021. Using this sampling frame, we selected an overall target sample size of 1000 unique websites, maintaining convention with nationwide surveys and polls [20], as well as similar exploratory analyses for which no a priori power calculations exist [21]. Ensuring a nationally representative sample, the state-level target sample size was determined on the basis of the proportion of the total number of family medicine providers listed for that state. For example, if a state had 1365 providers listed in the sampling frame (1% of 136,531), that state's target sample size in our sample would be 10 websites (1% of 1000). For each state, a random number generator assigned each practitioner a number, and then the team sorted the observations by the random number and selected the top *n* observations (*n*=state-level target sample size).

Five data abstractors searched internet browsers from September 20 to October 8, 2021, to identify sampled providers' practice websites. Websites were excluded if the practice (1) did not have an apparent focus on family medicine or primary care, (2) was located in a different state than that listed in the National Provider Identifier database, (3) was a duplicate, (4) had a nonfunctional website link, (5) used only social media pages (eg, Facebook), or (6) was permanently closed. Websites affiliated with the military or court-mandated health care systems were also excluded owing to their varying approaches to vaccine policies, which may have included messaging that the general public was not exposed to stringent vaccine mandates that were not applicable to the general public, or limited autonomy to shape vaccine messaging at the level of individual clinics and providers. For data quality control, the team cross-reviewed 10% of the websites, resolving any discordances in data extraction.

Our primary outcome was a mention of COVID-19 vaccines on the practices' website home page, specifically. Secondary outcomes included a mention of COVID-19 vaccines anywhere on the website, evidence that a clinic provided COVID-19 vaccines on site, provision of tailored vaccine information, explicit mention of full Food and Drug Administration (FDA) approval of the Pfizer-BioNTech COVID-19 vaccine (the only fully FDA-approved COVID-19 vaccine at the time of data collection [22,23]), and provision of a frequently asked question (FAQ) section about COVID-19 vaccines. "Tailored vaccine information" was defined as COVID-19 vaccine information thoughtfully curated or presented in a patient-centered manner or from the perspective of the providers in that clinic specifically. For example, an explicit mention of the patient

populations served by the clinic or explicit acknowledgment of common needs and concerns of the clinic's population would illustrate tailoring of standard vaccine information. Exploratory independent variables included the number of clinic locations, affiliation with a large hospital or managed health care system (conglomerate affiliation), university affiliation, mention of seasonal influenza vaccination, and US Census Bureau–defined geographic region within the United States [24]. We specifically explored the promotion of seasonal influenza vaccines because we sought to verify whether providers adopted similar policies of promoting another vaccine with well-established vaccine hesitancy in the general public [25].

We estimated the period prevalence of all outcomes. We estimated period prevalence, rather than point prevalence, because our data collection process was carried out over a period of several weeks rather than at a single time point. We reported unstratified results, as well as results stratified by the number of clinic locations. Cochran-Armitage tests for trend were used to identify statistically significant differences across strata of the number of clinic locations. The Cochran-Armitage test for trend is appropriate for assessing whether data in a contingency table of dimensions $2\times C$, where C is an ordinal variable with >2 levels, differ between the 2 groups [26,27]. We verified that all expected cell counts in these $2\times C$ tables were $n>5$.

Independent variables associated with mentioning COVID-19 vaccines on the website home page were assessed using modified Poisson regression models with robust error variances. In these analyses, our primary and secondary outcomes were each modeled as binary indicator outcome variables, which were set to 0 when a practice website did not have the attribute to 1 when the practice did. Modified Poisson regression is appropriate when estimating the prevalence or risk (including prevalence and risk ratios) of a high-prevalence, binary outcome [28,29]. We assumed that all observations were independent. Given the exploratory nature of our regression analyses, we also assumed no interaction between our independent variables, and we assumed that each variable had a linear association with the outcome on the log scale. Finally, we conducted a sensitivity analysis for these Poisson regression models in which P values were adjusted for multiple comparisons [30-33].

Statistical analyses were performed using SAS (version 9.4; SAS Institute Inc). All data analyzed in this study are included in Multimedia Appendices 1 and 2.

Ethical Considerations

There were no human participants in this study, and all data were collected via publicly available internet searches. Therefore, this study did not require ethics approval.

Results

After applying the exclusion criteria, 964 unique practice websites were included in the analytic sample. Practices typically had ≥10 locations (n=509, 52.8%), were affiliated with a conglomerate (n=663, 68.2%), were unaffiliated with a university (n=838, 87.2%), and mentioned seasonal influenza vaccines on their website (n=540, 56.1%; Table 1).

Between September 20 and October 8, 2021, overall, 550 (57.1%) websites mentioned COVID-19 vaccines on their home page, and 726 (75.3%) mentioned them somewhere on their website (Table 2). Additionally, 580 (60.2%) tailored the content provided, 426 (44.2%) provided a COVID-19 vaccine–focused FAQ section, and 199 (20.6%) explicitly mentioned the Pfizer BioNTech vaccine's full FDA approval.

Mentioning COVID-19 vaccines on home pages ranged from 28% among single-location practices to 78% among ≥20-location practices ($P<.001$ for trend). Similar statistically and clinically significant trends were observed between the number of clinic locations and all other outcomes (Table 2).

After adjusting for university affiliation, mention of influenza vaccines, and geographic region, the prevalence of mentioning COVID-19 vaccines on the home page was 1.7 (95% CI 1.4 to 2.2), 1.7 (95% CI 1.3 to 2.2), and 2.3 (95% CI 1.8 to 2.8) times higher among clinics with 2-9, 10-19, or ≥20 locations, respectively, than for clinics with only 1 location (Table 3). Compared to websites from Northeastern practices, those from the West and Midwest United States had comparable outcome prevalence; practices in the south had a lower prevalence (adjusted prevalence ratio 0.8, 95% CI 0.7 to 1.0). Sensitivity analyses, in which Poisson regression P values were adjusted for multiple comparisons, yielded similar statistical results (Multimedia Appendix 2, Table S1).

Table 1. Characteristics of family medicine practices in the United States included in our fall 2021 COVID-19 website content review.

Practice characteristics	Frequency (n=964), n (%)
Practice locations	
1	238 (24.7)
2-9	217 (22.5)
10-19	117 (12.1)
≥20	392 (40.7)
Practice affiliated with a large hospital or managed health care system (conglomerate)	
Yes	663 (68.2)
No	301 (31.2)
Practice affiliated with a university	
Yes	123 (12.8)
No	838 (87.2)
Practice website mentions seasonal influenza vaccination	
Yes	540 (56.1)
No	423 (43.9)
Translations to other languages available on website	
Yes	266 (27.6)
No	698 (72.4)
US region[a]	
Midwest	243 (25.2)
Northeast	145 (15.0)
South	333 (34.5)
West	243 (25.2)

[a]US regions: Midwest (Iowa, Illinois, Indiana, Kansas, Michigan, Minnesota, Missouri, North Dakota, Nebraska, Ohio, South Dakota, and Wisconsin), northeast (Connecticut, Massachusetts, Maine, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont), south (Alabama, Arkansas, District of Columbia, Delaware, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, and West Virginia), and west (Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, New Mexico, Nevada, Oregon, Utah, Washington, and Wyoming).

Table 2. COVID-19 vaccine–related information on family medicine practice websites in the United States included in our fall 2021 COVID-19 website content review.

	Total, n (%)	Practice locations, n (%)				*P* value[a]
		1	2-9	10-19	≥20	
Mentions the COVID-19 vaccine on the home page						<.001
Yes	550 (57.1)	66 (27.7)	114 (52.5)	66 (56.4)	304 (77.6)	
No	414 (42.9)	172 (72.3)	103 (47.5)	51 (43.6)	88 (22.4)	
Mentions the COVID-19 vaccine mentioned anywhere on the website						<.001
Yes	726 (75.3)	98 (41.2)	158 (72.8)	97 (82.9)	373 (95.2)	
No	238 (24.7)	140 (58.8)	59 (27.2)	20 (17.1)	19 (4.9)	
Provides tailored information on COVID-19 vaccines[b]						<.001
Yes	580 (60.2)	59 (24.8)	114 (52.5)	71 (60.7)	336 (85.7)	
No	384 (39.8)	179 (75.2)	103 (47.5)	46 (39.3)	56 (14.3)	
Mentions Food and Drug Administration approval of the COVID-19 vaccine[c]						<.001
Yes	199 (20.6)	15 (6.3)	22 (10.1)	24 (20.5)	138 (35.2)	
No	765 (79.4)	223 (93.7)	195 (89.9)	93 (79.5)	254 (64.8)	
Mentions the booster or third dose of the vaccine						<.001
Yes	400 (41.5)	28 (11.8)	67 (30.9)	55 (47.0)	250 (63.8)	
No	564 (58.5)	210 (88.2)	150 (69.1)	62 (53.0)	142 (36.2)	
Has information on COVID-19 vaccine eligibility criteria						<.001
Yes	530 (55.0)	51 (21.4)	101 (46.5)	66 (56.4)	312 (79.6)	
No	434 (45.0)	187 (78.6)	116 (53.5)	51 (43.6)	80 (20.4)	
Has a "frequently asked questions" section on COVID-19 vaccines on the website						<.001
Yes	426 (44.2)	31 (13.0)	62 (28.6)	48 (41.0)	285 (72.7)	
No	538 (55.8)	207 (87.0)	155 (71.4)	69 (59.0)	107 (27.3)	
Nature of the website's language regarding COVID-19 vaccination						<.001
No mention	245 (25.4)	142 (59.7)	64 (29.5)	22 (19.0)	17 (4.3)	
Neutral	172 (17.9)	37 (15.5)	51 (23.5)	25 (21.6)	59 (15.1)	
Supportive	546 (56.7)	59 (24.8)	102 (47.0)	69 (59.5)	316 (80.6)	

[a]Cochran-Armitage trend test.

[b]Tailored information described as curated information indicating an effort from the practice to explain and endorse the COVID-19 vaccines.

[c]The website explicitly mentioned the approval of the COVID-19 vaccine (Pfizer's Comirnaty approved on August 23, 2021) by the US Food and Drug Administration anywhere on the website.

Table 3. Unadjusted and adjusted prevalence ratios of whether a family medicine practice mentioned the COVID-19 vaccine on their website home page.

	Mentioned the COVID-19 vaccine on the home page, n (%)	Unadjusted prevalence ratio		Adjusted prevalence ratio[a]	
		Ratio (95% CI)	*P* value	Ratio (95% CI)	*P* value
Number of locations					
1	66 (27.7)	Reference	—[b]	Reference	—
2-9	114 (52.5)	1.89 (1.49-2.41)	<.001	1.73 (1.37-2.20)	<.001
10-19	66 (56.4)	2.03 (1.57-2.64)	<.001	1.73 (1.34-2.23)	<.001
≥20	304 (77.6)	2.80 (2.26-3.46)	<.001	2.25 (1.81-2.79)	<.001
US region[c]					
Northeast	93 (64.1)	Reference	—	Reference	—
Midwest	148 (60.9)	0.95 (0.81-1.11)	.52	0.96 (0.84-1.10)	.57
South	166 (49.9)	0.78 (0.66-0.91)	.002	0.84 (0.73-0.97)	.02
West	143 (58.9)	0.92 (0.78-1.08)	.29	0.99 (0.86-1.14)	.86
University-affiliated[d]	94 (76.4)	1.41 (1.26-1.59)	<.001	1.15 (1.03-1.29)	.01
Mentioned influenza vaccination[d]	394 (73.0)	1.99 (1.74-2.28)	<.001	1.62 (1.41-1.86)	<.001

[a]Models adjusted for the number of locations, region, university affiliation, and mention of seasonal influenza vaccines.

[b]—: not applicable.

[c]US regions: Midwest (Iowa, Illinois, Indiana, Kansas, Michigan, Minnesota, Missouri, North Dakota, Nebraska, Ohio, South Dakota, and Wisconsin), northeast (Connecticut, Massachusetts, Maine, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont), south (Alabama, Arkansas, District of Columbia, Delaware, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, and West Virginia), and west (Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, New Mexico, Nevada, Oregon, Utah, Washington, and Wyoming).

[d]Comparing "Yes" versus "No" responses (reference).

Discussion

Principal Findings

As revealed in this analysis, at the time of data collection, not all family medicine practices were maximizing their existing practice websites for COVID-19 vaccine promotion. As of October 2021, nearly half of the websites examined by our research team did not mention COVID-19 vaccines on their home page. In addition, only one-fifth of them mentioned full FDA approval of the Pfizer vaccine, a regulatory accomplishment that many postulated would increase vaccine confidence and uptake [34]. Although medical professionals may have missed the opportunity to amplify this milestone on their practices' websites, future milestones may present similar opportunities, including full FDA approval of vaccines for young children and approval of additional strain-specific booster vaccines.

It is critical for PCPs, as trusted sources of vaccine-related information, to optimize communication with the general public [18,19]. PCPs can accomplish medical and public health goals concurrently by considering how their website can serve as a web-based extension of the tailored messaging they often use inside the clinic. Furthermore, given the potential for novel SARS-CoV-2 variants (eg, the B1 variant) and future infection surges, it is increasingly important for vaccine providers to be cognizant of where and how to make thoughtful use of their existing practice websites for additional vaccine promotion.

How vaccine content is described is also important. With only 60% of included websites providing tailored content, there is additional room to translate the tailored, values-based messaging often used inside clinics to the clinic's internet domain [35,36]. If curated appropriately on the home page, a practice website can quickly convey the priorities and values of the practice as they pertain to vaccines.

Our analysis also revealed that as practice size increased, the presence of COVID-19 vaccine information also increased. Although our analysis was not designed to determine the cause of any association between practice size and web-based vaccine promotion, we hypothesize that this finding reflects that larger, well-resourced practices have greater means to update their websites. This would signal an opportunity for entities such as national professional societies to develop guidance to help less-resourced practices provide accurate and specialty-specific information on the internet.

Finally, we observed that compared to northeastern clinics, southern clinics were less likely to mention COVID-19 vaccines on their practices' website home pages. This finding emerges amid evidence that southern states continue to have COVID-19 vaccination coverage below the national average [37]. Given the lower baseline rates of COVID-19 information on the home pages of websites of southern clinics, we hypothesize that in

the future, concerted efforts to increase COVID-19 vaccine information on practices' websites may have the highest overall impact in this region.

Strengths and Limitations

The strengths of our study include a large, nationwide sampling frame of family medicine practitioners to yield a sample distributed proportionally across all 50 states. Furthermore, our team implemented a 10% cross-checking measure to reduce data entry errors and improve internal validity.

Our study also has limitations. As our study is exploratory in nature, it is not designed to assess the causal impact of web-based information on the COVID-19 vaccine on the population uptake of these vaccines. In addition, this study's cross-sectional design does not reflect whether and how clinics change their website content over time. Finally, our analytic sample did not include other relevant provider types (eg, pediatricians) who may differ in approaches to provision of vaccine content on their websites.

Conclusions

During a pandemic, it is vital for medical providers to utilize all means to promote vaccination when safe and effective vaccines are available. With many PCPs already hosting websites, there is little reason not to utilize this platform to reach patients who are eager to understand their trusted provider's views.

Conflicts of Interest

None declared.

Multimedia Appendix 1
This includes our raw and analysis-ready datasets.
[ZIP File (Zip Archive), 5149 KB - formative_v6i11e38425_app1.zip]

Multimedia Appendix 2
This is our sensitivity analysis to account for multiple testing.
[DOCX File , 40 KB - formative_v6i11e38425_app2.docx]

References

1. COVID data tracker weekly review. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/coronavirus/2019-ncov/covid-data/covidview/index.html [accessed 2022-07-25]
2. Adam D. COVID's true death toll: much higher than official records. Nature 2022 Mar 10;603(7902):562-562. [doi: 10.1038/d41586-022-00708-0] [Medline: 35277684]
3. Mandavilli A. Reaching 'herd immunity'is unlikely in the US, experts now believe. The New York Times. 2021. URL: https://www.nytimes.com/2021/05/03/health/covid-herd-immunity-vaccine.html [accessed 2021-11-15]
4. Engber D. Vaccination in America might have only one tragic path forward. The Atlantic. URL: https://www.theatlantic.com/health/archive/2021/07/america-covid-19-vaccine-decline/619474/ [accessed 2021-12-17]
5. MacDonald NE, SAGE Working Group on Vaccine Hesitancy. Vaccine hesitancy: definition, scope and determinants. Vaccine 2015 Aug 14;33(34):4161-4164 [FREE Full text] [doi: 10.1016/j.vaccine.2015.04.036] [Medline: 25896383]
6. Troiano G, Nardi A. Vaccine hesitancy in the era of COVID-19. Public Health 2021 May;194:245-251 [FREE Full text] [doi: 10.1016/j.puhe.2021.02.025] [Medline: 33965796]
7. Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. J Med Internet Res 2009 Mar 27;11(1):e11 [FREE Full text] [doi: 10.2196/jmir.1157] [Medline: 19329408]
8. Hildreth JEK, Alcendor DJ. Targeting COVID-19 vaccine hesitancy in minority populations in the US: implications for herd immunity. Vaccines (Basel) 2021 May 11;9(5) [FREE Full text] [doi: 10.3390/vaccines9050489] [Medline: 34064726]
9. Chandrasekaran R, Desai R, Shah H, Kumar V, Moustakas E. Examining public sentiments and attitudes toward COVID-19 vaccination: infoveillance study using Twitter posts. JMIR Infodemiology 2022;2(1):e33909 [FREE Full text] [doi: 10.2196/33909] [Medline: 35462735]
10. Calac AJ, Haupt MR, Li Z, Mackey T. Spread of COVID-19 vaccine misinformation in the ninth inning: retrospective observational infodemic study. JMIR Infodemiology 2022 Mar 16;2(1):e33587 [FREE Full text] [doi: 10.2196/33587] [Medline: 35320982]
11. Skafle I, Nordahl-Hansen A, Quintana DS, Wynn R, Gabarron E. Misinformation about COVID-19 vaccines on social media: rapid review. J Med Internet Res 2022 Aug 04;24(8):e37367 [FREE Full text] [doi: 10.2196/37367] [Medline: 35816685]
12. Yasmin F, Najeeb H, Moeed A, Naeem U, Asghar MS, Chughtai NU, et al. COVID-19 vaccine hesitancy in the United States: a systematic review. Front Public Health 2021;9:770985 [FREE Full text] [doi: 10.3389/fpubh.2021.770985] [Medline: 34888288]

13. Padamsee TJ, Bond RM, Dixon GN, Hovick SR, Na K, Nisbet EC, et al. Changes in COVID-19 vaccine hesitancy among Black and White individuals in the US. JAMA Netw Open 2022 Jan 04;5(1):e2144470 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.44470] [Medline: 35061038]
14. Lazarus JV, Wyka K, White TM, Picchio CA, Rabin K, Ratzan SC, et al. Revisiting COVID-19 vaccine hesitancy around the world using data from 23 countries in 2021. Nat Commun 2022 Jul 01;13(1):3801 [FREE Full text] [doi: 10.1038/s41467-022-31441-x] [Medline: 35778396]
15. Weerakoon SM, Henson-Garcia M, Valerio-Shewmaker MA, Messiah SE, Knell G. Contributions of trustworthiness, health literacy, and self-efficacy in communicating with COVID-19 vaccine-hesitant audiences: web-based survey study. JMIR Form Res 2022 Aug 01;6(8):e38076 [FREE Full text] [doi: 10.2196/38076] [Medline: 35878123]
16. Kasting M, Macy J, Grannis S, Wiensch A, Lavista Ferres JM, Dixon B. Factors associated with intention to receive the COVID-19 vaccine: a cross-sectional national study. JMIR Public Health Surveill 2022 Oct 09:2022 [FREE Full text] [doi: 10.2196/37203] [Medline: 36219842]
17. 12 COVID-19 vaccination strategies for your community. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/vaccines/covid-19/vaccinate-with-confidence/community.html [accessed 2021-11-03]
18. KFF COVID-19 Vaccine Monitor. Kaiser Family Foundation. URL: https://www.kff.org/coronavirus-covid-19/dashboard/kff-covid-19-vaccine-monitor-dashboard/ [accessed 2021-11-17]
19. Latkin CA, Dayton L, Miller JR, Yi G, Jaleel A, Nwosu CC, et al. Behavioral and attitudinal correlates of trusted sources of COVID-19 vaccine information in the US. Behav Sci (Basel) 2021 Apr 20;11(4):56 [FREE Full text] [doi: 10.3390/bs11040056] [Medline: 33924118]
20. Frequently asked questions. Pew Research Center. URL: https://www.pewresearch.org/our-methods/u-s-surveys/frequently-asked-questions/ [accessed 2021-12-30]
21. Lehnert J, Ellingson M, Goryoka G, Kasturi R, Maier E, Chamberlain A. Use of obstetric practice web sites to distribute zika virus information to pregnant women during a zika virus outbreak. J Public Health Manag Pract 2017;23(6):608-613 [FREE Full text] [doi: 10.1097/PHH.0000000000000537] [Medline: 28125540]
22. FDA. FDA Approves First COVID-19 Vaccine. U.S. Food and Drug Administration. URL: https://www.fda.gov/news-events/press-announcements/fda-approves-first-covid-19-vaccine [accessed 2022-07-15]
23. Coronavirus (COVID-19) update: FDA takes key action by approving second COVID-19 vaccine. U.S. Food and Drug Administration. URL: https://www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-fda-takes-key-action-approving-second-covid-19-vaccine [accessed 2022-07-15]
24. Geographic levels. United States Census Bureau. URL: https://www.census.gov/programs-surveys/economic-census/guidance-geographies/levels.html [accessed 2022-01-06]
25. Schmid P, Rauber D, Betsch C, Lidolt G, Denker M. Barriers of influenza vaccination intention and behavior - a systematic review of influenza vaccine hesitancy, 2005 - 2016. PLoS One 2017 Jan 26;12(1):e0170550 [FREE Full text] [doi: 10.1371/journal.pone.0170550] [Medline: 28125629]
26. Agresti A. Categorical data analysis. John Wiley & Sons. 2012. URL: http://ebookcentral.proquest.com/lib/emory/detail.action?docID=1168529 [accessed 2022-07-26]
27. The FREQ Procedure. SAS Help Center. URL: https://documentation.sas.com/doc/en/pgmsascdc/9.4_3.3/statug/statug_freq_details74.htm [accessed 2022-07-26]
28. Zou G. A modified poisson regression approach to prospective studies with binary data. Am J Epidemiol 2004 Apr 01;159(7):702-706. [doi: 10.1093/aje/kwh090] [Medline: 15033648]
29. Barros AJD, Hirakata VN. Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio. BMC Med Res Methodol 2003 Oct 20;3:21 [FREE Full text] [doi: 10.1186/1471-2288-3-21] [Medline: 14567763]
30. Feise RJ. Do multiple outcome measures require p-value adjustment? BMC Med Res Methodol 2002 Jun 17;2(1):8 [FREE Full text] [doi: 10.1186/1471-2288-2-8] [Medline: 12069695]
31. The MULTTEST Procedure: PROC MULTTEST Statement. SAS Help Center. URL: https://documentation.sas.com/doc/en/pgmsascdc/9.4_3.3/statug/statug_multtest_syntax01.htm [accessed 2022-08-02]
32. The MULTTEST procedure: p-value adjustments. SAS Help Center. URL: https://documentation.sas.com/doc/en/pgmsascdc/9.4_3.3/statug/statug_multtest_details11.htm [accessed 2022-08-02]
33. Glickman ME, Rao SR, Schultz MR. False discovery rate control is a recommended alternative to Bonferroni-type adjustments in health studies. J Clin Epidemiol 2014 Aug;67(8):850-857. [doi: 10.1016/j.jclinepi.2014.03.012] [Medline: 24831050]
34. Cohen E. Full FDA approval of Pfizer's Covid-19 vaccine had only a modest impact on uptake. Here's what mattered more. CNN Health. URL: https://www.cnn.com/2021/10/04/health/pfizer-covid-19-vaccine-approval-bump/index.html [accessed 2021-12-17]
35. McClure CC, Cataldi JR, O'Leary ST. Vaccine hesitancy: where we are and where we are going. Clin Ther 2017 Aug;39(8):1550-1562. [doi: 10.1016/j.clinthera.2017.07.003] [Medline: 28774498]
36. Amin AB, Bednarczyk RA, Ray CE, Melchiori KJ, Graham J, Huntsinger JR, et al. Association of moral values with vaccine hesitancy. Nat Hum Behav 2017 Dec;1(12):873-880. [doi: 10.1038/s41562-017-0256-5] [Medline: 31024188]

37. Oladipo G. US fight against Covid threatened by growing vaccine gap in the south. The Guardian. URL: https://www.theguardian.com/world/2021/jun/27/us-covid-vaccine-disparities-south [accessed 2021-12-17]

Abbreviations

FDA: Food and Drug Administration
PCP: primary care provider

Edited by A Mavragani; submitted 31.03.22; peer-reviewed by A Rovetta, G Davlyatov; comments to author 13.07.22; revised version received 03.08.22; accepted 29.10.22; published 17.11.22.

Please cite as:
Ackleh-Tingle JV, Jordan NM, Onwubiko UN, Chandra C, Harton PE, Rentmeester ST, Chamberlain AT
Prevalence and Correlates of COVID-19 Vaccine Information on Family Medicine Practices' Websites in the United States: Cross-sectional Website Content Analysis
JMIR Form Res 2022;6(11):e38425
URL: https://formative.jmir.org/2022/11/e38425
doi:10.2196/38425
PMID:36343211

Original Paper

The Impact of the COVID-19 Pandemic on eHealth Use in the Daily Practice and Life of Dutch-Speaking General Practitioners in Belgium: Qualitative Study With Semistructured Interviews

Jan Ismail Yagiz[1*], MD; Geert Goderis[1*], MD, PhD

Leuven Institute for Healthcare Policy, Medical Faculty, KU Leuven, Leuven, Belgium

[*]all authors contributed equally

Corresponding Author:
Jan Ismail Yagiz, MD
Leuven Institute for Healthcare Policy
Medical Faculty
KU Leuven
Kapucijnenvoer 35
blok d - bus 7001 - 3rd floor
Leuven, 3000
Belgium
Phone: 32 16 32 12 59
Email: yagizmd@msn.com

Abstract

Background: The COVID-19 crisis has led to rapid and far-reaching changes in digital health care, but little is known about what, why, and how changes occurred in eHealth use in Flemish general practice during the pandemic.

Objective: This study aims to understand how general practitioners (GPs) perceive and evaluate eHealth solutions and their eHealth experience during the COVID-19 pandemic.

Methods: This qualitative study was conducted using in-depth 1-on-1 semistructured interviews with the help of an interview guide. Several areas were identified beforehand to help assess the impact of the COVID-19 pandemic: perceptions of digital technologies in GP practices; changes in the use of these technologies during and after the COVID-19 pandemic; GPs' adaptation to digitalization, benefits, risks, and challenges of eHealth; GPs motivations to change practice; and future perspectives on eHealth. In this study, purposive sampling and snowballing methods were used. Between October 2021 and April 2022, we interviewed 15 Dutch-speaking GPs in the Flemish region via the Zoom online conferencing tool.

Results: GPs indicated that eHealth was used more frequently during the COVID-19 pandemic than before, a change that helped them reduce their workload, enabling greater accessibility to health care services and the complementary use of digital and physical consultations. Our findings suggest that physicians underwent a significant cognitive shift in their perceptions, causing them to be more open and prepared to adopt eHealth solutions. However, there remains significant doubt and uncertainty about digital literacy for certain groups, privacy, data security, reimbursement, and the burden of technical information and communication technologies (ICT) issues.

Conclusions: The COVID-19 pandemic seems to have been a turning point for eHealth by Flemish GPs. eHealth is an essential complementary health care service that can reduce pressure on health care as well as increase health care accessibility. Sensitive aspects, such as privacy, data security, digital literacy, reimbursement, and the burden of technical ICT issues, are particularly emphasized. With our results, we can offer recommendations to health IT policymakers and developers that will help maintain the continuity of eHealth solutions beyond the COVID-19 pandemic, considering the expectations and sensitivities presented in the study.

(JMIR Form Res 2022;6(11):e41847) doi:10.2196/41847

KEYWORDS

COVID-19; impact; eHealth; GPs; Flemish; practice; qualitative study; semistructured interviews

Introduction

eHealth refers to tools that use information and communication technologies (ICT) to prevent, diagnose, treat, and monitor health-related problems. The term “eHealth” is inclusive, and what falls under eHealth is also constantly changing under technological developments, such as mobile health (mHealth), telemonitoring, telemedicine, cloud platforms for data storage, and artificial intelligence (AI). The term “mHealth” refers to the use of mobile communication and devices that provide care or improve health outcomes, for example, with the help of an app on smartphones [1]. Telemonitoring is described as daily remote, wireless, and ambulatory monitoring of various medical and technical data and parameters via sensors, cameras, and devices implanted in the patient or placed on the body or in the patient's clothing, such as blood pressure, weight, heart rate, and body temperature [2]. The term “telemedicine” relates to health care services that permit patients to receive care in their everyday life and overpass the distance between health care professionals and patients through ICT [3].

eHealth solutions are represented as complex interventions, as several interacting components pose some additional problems for raters and already theoretical and methodological issues [4]. However, over the past decade, the growing use of eHealth has increased pressure on health care [5,6], playing an increasingly important role in the sustainability of future health care systems and increasing passion in patient empowerment [7,8]. An aging population, rising chronic diseases, and the COVID-19 pandemic are increasing pressure on health care [9-12]. Therefore, innovations, such as eHealth, are needed to maintain the accessibility and quality of care [13-16]. Meanwhile, digital health technologies have greatly accelerated patient engagement [17-20]. In line with these developments, medical institutions have intensively integrated eHealth into traditional face-to-face counseling [21]. The combination of eHealth and face-to-face consultation can be defined as hybrid health care [22,23]. Sometime after the COVID-19 pandemic, it was shown that the pandemic has accelerated the implementation of eHealth solutions, and even though it is a crisis situation, many health care providers and institutions have quickly embraced digital medicine. Adopting more digital health is a phase with the potential to improve the quality of care [24,25].

General practitioners (GPs) are an essential part of primary health care in Belgium's health system. They usually work in solo, dual, or group practices, and their main tasks include preventive care, diagnosis, and treatment for a wide range of health problems. Since 2007, digitization systems have been used in health care in Belgium [26].

According to a Belgian 2019 study, the adoption of eHealth by GPs in Belgium—as in other countries—remains modest [27]. These results are in line with other studies showing that the health care sector is a laggard in adopting digital services [28]. According to these studies, several factors play an active role in the sector's slowness to adopt digital practices, such as the strict regulations of the sector, the sensitivity surrounding personally identifiable information, the resistance of health care providers to digital apps, the lack of prioritization of the patient experience, and the cost of investment [29,30].

The first COVID-19 death in Belgium occurred on March 10, 2020, and the first quarantine period was imposed on the population from March 14 to May 5, 2020, to prevent the spread of the COVID-19 virus [31,32]. Rapid and far-reaching changes have been made so that health care providers can ensure continuity of care while minimizing the risk of the virus spreading. Digital health tools, such as social follow-up, texting, getting photos from patients, and e-consultations, have been used more than ever [33]. In Belgium, temporary remote consultations (by phone or video calls) were introduced. According to the COVID-19 monitoring report of the National Institute for Health and Disability Insurance (RIZIV), 3.8 million consultations were billed between March and May 2020, and these consultations were performed mainly by GPs. Teleconsultations were commonly used to obtain prescriptions, follow up on a chronic or existing condition, deal with new coronavirus-related complaints, and discuss a sickness report or a new complaint unrelated to coronavirus [34].

Although other studies have previously shown that the COVID-19 pandemic significantly impacted GPs’ use of eHealth, little is known about the impact of the COVID-19 crisis on the daily use of eHealth technologies among GPs in the Flemish region of Belgium. In our study, we investigated the following research questions: What has changed in the mind and practice of GPs toward the use of eHealth as a result of the COVID-19 crisis? What other changes were experienced by GPs during this period, why, and how? Moreover, we aimed to gain deep insight into the perceived impact of the crisis.

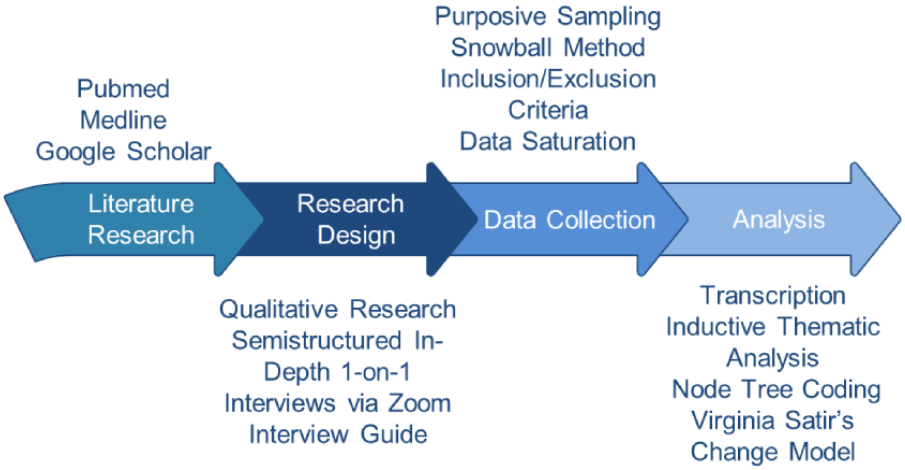
This paper describes the GPs’ perception, appreciation, and use of various eHealth tools during the pandemic; their expectations; concerns regarding eHealth use; the perceived impact of the COVID-19 pandemic on eHealth use; and perceived barriers, enablers, potential gains, and success points in eHealth use.

Methods

Study Design and Population

We conducted qualitative research using semistructured interviews of 15 Dutch-speaking GPs in Belgium. One-on-one interviews were chosen as the research design, which allowed us to obtain in-depth information from the respondents. Purposive sampling and the snowball method were used during the study. The study population was assumed to be around 20-30 GPs, but when data saturation provided enough knowledge for a deep understanding of the subject, the study finished before the targeted population numbers were reached. The study was conducted between October 2021 and April 2022. The study design is shown in Figure 1.

Figure 1. Study design.

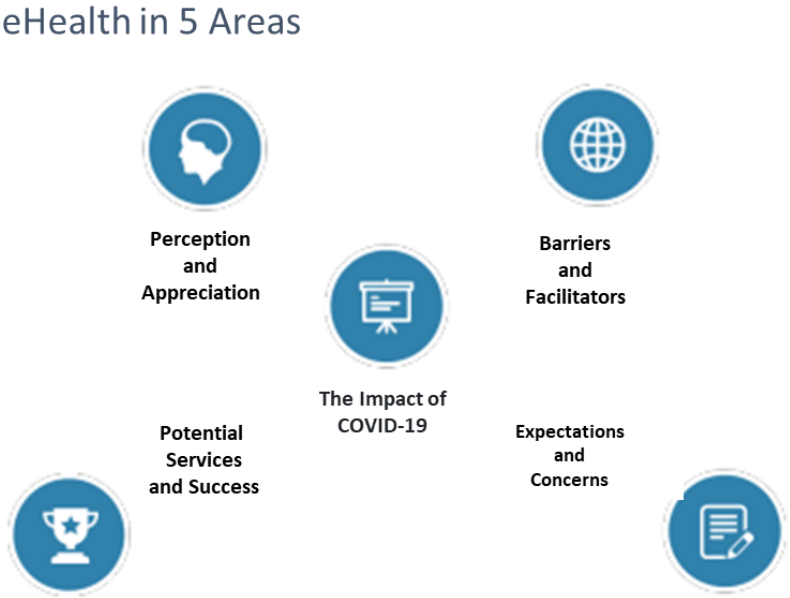


Interview Procedures, Data Collection, and Analysis

Due to the COVID-19 pandemic, interviews were scheduled and conducted through the Zoom online videoconferencing tool. The interview guide (Multimedia Appendix 1) was developed in close collaboration with the promotor. Before the interviews, informed consent was collected via email from the interviewees. The subinvestigator conducted the semistructured interviews using the following thematic blocks: demographics; the perception and appreciation of digital technologies in GP practices; changes in the use of these technologies during and after the COVID-19 pandemic; the adaptation of GPs to digitalization; eHealth's benefits, risks, and challenges; and motivations for and future perspectives on eHealth. The interviews lasted approximately 30 minutes. All interviews were recorded, transcribed verbatim, and analyzed thematically using Microsoft Excel 2010. Braun and Clarke's [35] 6-step guide was used. Notes were grouped into 5 domains (with themes and subthemes): the perception and appreciation of eHealth, the impact of COVID-19 on eHealth use, the barriers to and facilitators of eHealth use, the potential services and success points of eHealth, and concerns and expectations for the future (see Figure 2). All quotes were translated from Dutch to English, considered representative, and are reported in the Results section.

All data were treated confidentially and pseudonymized with due care during the project. Depending on our research questions, the recorded interviews were only watched and listened to and typed out by the researcher supervised by the promotor. The interviews were transcribed in an anonymous manner. The sound recordings and transcripts were recorded in a safe and secure way, and these files were saved on a personal computer that is password-protected. The study was conducted on a volunteer basis, and there was no compensation given.

Figure 2. Defined themes.



Ethical Considerations

Ethical committee approval and informed consent letters were asked for and obtained beforehand from the KU Leuven Ethics Committee (reference MP016788, July 20, 2021).

Results

Characteristics of Study Participants

We interviewed 15 GPs from 5 different regions of Flanders: Antwerp (n=7, 47%), Limburg (n=3, 20%), Flemish Brabant (n=2, 13%), East Flanders (n=2, 13%), and West Flanders (n=1, 7%). Of the 15 interviewees, 10 (67%) reported practicing medicine for 0-20 years and 5 (33%) reported practicing medicine for more than 20 years. On average, they had 14.8 years of experience in primary care, from 3 to 40 years. Most interviewees reported working in a group (n=7, 47%) or alone (n=5, 33%). The gender distribution was 8 (53%) males and 7 (47%) females. The demographic characteristics of the participants are shown in Table 1.

Based on the thematic analysis of the interviews, we collected all the data on 5 important domains with subdomains, as shown in Tables 2-6.

Table 1. Demographics of participants (N=15).

Experience (years)	Solo practice, n (%)	Duo practice, n (%)	Group practice, n (%)	Community, n (%)	Total GPs[a], n (%)
<5	0	0	2 (13): M[b], Antwerp, GP7 and GP8	0	2 (13)
5-10	1 (7): F[c], Antwerp, GP9	0	2 (13): M, East Flanders and Antwerp, GP10 and GP11 2 (13): F, Antwerp, GP5 and GP6	2 (13): M, Brussels, GP1 and GP2	7 (47)
10-20	0	1 (7): F, Limburg, GP14	0	0	1 (7)
>20	2 (13): F, Limburg, GP3 and GP4 2 (13): M, West Flanders, GP13; F, Antwerp, GP12	0	1 (7): M, East Flanders GP15	0	5 (33)

[a]GP: general practitioner.

[b]M: male.

[c]F: female.

Table 2. Themes and subthemes of the "impact of the COVID-19 pandemic on eHealth use" domain.

Themes and subthemes	Quotes from the interviews
Changes and shifts in service	
There is greater eHealth uptake.	"COVID created an extra training for us, especially for older GP[a] generations who were against the digitalization but came to understand the benefits of eHealth." [GP2]
GPs reduce strain with eHealth.	"Teleconsultation is not a pleasant thing, but it was an important tool in the intensity of the pandemic." [GP3]
Teleconsultations	
GPs provide triage for suspicious infections.	"We did triage for COVID infection via teleconsultations." [GP5]
Workload	
GPs' workload increased and sped up the process of digitalization.	"COVID sped up and also increased our work via digitalization. It also changed our way of working." [GP6]
Digital administrator	
GPs provide e-certificates, e-receipts, and reimbursement papers.	"COVID has changed the way we work and boosted digital solutions." [GP10]
Extra investment in IT solutions	
New investments increase in eHealth, mHealth[b], and video consults.	"We adopted an application on the mobile phone so that we could also work from home." [GP7]

[a]GP: general practitioner.

[b]mHealth: mobile health.

Table 3. Themes and subthemes of the "eHealth perception and appreciation" domain.

Themes and subthemes	Quotes from the interviews
eHealth perception	
eHealth is a digital solution for GP[a] work.	"eHealth is a portal to digital to use all the capabilities of information, EMD[b], tele-consult, medication…" [GP1] "Digital solutions increased the quality of and communication in health care." [GP8]
The future is digital.	"Digital apps will be used more because GPs are retired, and the time spent on physical contact will be less than now." [GP4]
eHealth appreciation	
eHealth has a lot of advantages.	"There are many positive things." [GP1] "Our clinic is 90% digital now." [GP6] "I find it very good to have. I hope that will get better, be adopted more quickly, and implemented at a faster pace. I found it very slow, I mean, the evolution over the two years." [GP13]

[a]GP: general practitioner.

[b]EMD: Electronic Medical Dossier.

Table 4. Themes and subthemes of the "potential services and success points" domain.

Themes and subthemes	Quotes from the interviews
Potential services	
GPs[a] are prepared for the unknown.	"I think we are more ready for the unknown…the crisis made us stronger and able to provide health care in difficult situations." [GP10]
GPs have more digital literacy.	"Software is new, and I am not fully integrated with the software. I need more training so that I can use it more efficiently." [GP14]
GPs have sight for mHealth and video consultations in the future.	"The role of the GP will be changed to include controlling data from applications. We could follow up closely chronic diseases with the applications, especially with chronic illnesses such as heart failure" [GP15] "I want to implement video consultations in the future, especially so that I can see the patient." [GP9]
Success points	
There is increased accessibility.	"I am in urgency mode for the patients, reachable 24/7." [GP12] "Quick connecting with people to monitor them at home, especially older people" [GP3]
GPs are aware of the mental shift to remote work.	"One of the most important things we learned from the coronavirus crisis is that eHealth solutions shifted our mindset from classical role of the physician to an integrated role using more remote solutions." [GP11]
Physician-patient satisfaction with telemedicine	"In general, the patients are satisfied with the speed of work. They do not wait long and do not come here to take medications." [GP4]

[a]GP: general practitioner.

Table 5. Themes and subthemes of the "barriers and facilitators to using eHealth" domain.

Themes and subthemes	Quotes from the interviews
Barriers	
There is an ICT[a] burden.	"Lack of IT support to solve the problems such as older software or blockages." [GP[b]4]
Digital literacy is needed.	"To use the digital platforms, you need to have an educated population." [GP1]
Facilitators	
COVID-19 pandemic measures	"Due to the measures, the disallowing of physical contact made us facilitate creative solutions, such as using more technology and remote options." [GP5]
Government and European Union (EU) policies	"Remuneration for telehealth was useful to stimulate complementary remote and physical care." [GP9]
Patient's role	"The patient is central now. We have to listen to them, and most patients are comfortable and are doing well with the technology and happy to use it." [GP12]

[a]ICT: information and communication technologies.

[b]GP: general practitioner.

Table 6. Themes and subthemes of the "future expectations and concerns" domain.

Themes and subthemes	Quotes from the interviews
Expectations	
GPs[a] were seeking for integrated first and second service lines.	"More stable eHealth, technically strong and more integrated solutions. Currently, digital integration between the 1 and 2 service lines is terrible." [GP13]
GPs expect more efficient IT solutions.	"I would like to see more user-friendly solutions for all population levels, as well as more education on digital eHealth solutions." [GP10]
GPs demand remuneration for telemedicine services.	"We would like to have further remuneration for teleconsultations." [GP9]
Concerns	
GPs are aware of safety concerns related to data.	"We often use e-mails, text messaging, and WhatsApp. These are not fully safe, not fully protected. These must be protected from cyberattacks." [GP8]
GPs are concerned about fragile groups.	"There are subgroups of patients who are not aware of or capable enough to use eHealth tools." [GP15]
GPs have concerns about losing physical contact and follow-ups.	"I do not want to lose my physical contact with the patient…Why? Problems with follow-up. Why? Because even though it is a prescription, I am doing more than that. Are there any other issues to follow up on? For example, screening for colorectal cancer, mammography, vaccinations, blood pressure measurements, advice, conversations with the patient." [GP1]

[a]GP: general practitioner.

Domain 1: Impact of the COVID-19 Pandemic on eHealth

Changes and Shifts in Service

According to the GPs surveyed, eHealth use showed a tremendous difference before and after the COVID-19 pandemic. All the participants agreed on the differentiation and increased use of eHealth tools.

Before the coronavirus crisis, they noted that they were using eHealth frameworks infrequently or in a limited number of cases, and they used phone calls and emails mainly for simple investigations. However, they had to find new arrangements during the pandemic due to limited in-person contact. Teleconsultations and eHealth apps, such as e-receipts, e-certificates, digital reimbursement, and online appointment modules, were widely used.

They also stated that a few apps that had been in use before the pandemic worked exceptionally well during the pandemic, such as e-receipts. They generally summarized the impact of the coronavirus crisis on eHealth use in 3 domains: an expanded uptake of eHealth in their practice, decreased strain on work with eHealth arrangements, and differentiation of the use of eHealth solutions.

Increased eHealth Uptake

> *Digitalization increased after COVID in our practice from 40% to 75%...The quality of digitalization is also better, and it is easier than before.* [GP14]

> *We were using telephone contacts with our initiative for talking about patient results before COVID, but now we use it for teleconsultations, triage, and prescriptions.* [GP15]

GPs Reduced Their Work Strain With eHealth

> *COVID increased the use of e-certificates, e-receipts, and other online solutions that helped us and patients to get out of this crisis in a healthy way.* [GP10]

Workload

All the interviewees agreed on differentiation in workload due to the COVID-19 pandemic. They said their work increased, was boosted, and sped up with the impact of the COVID-19 crisis on eHealth use.

Meaningful feedback from GP1 was:

> *COVID increased our workload, sped up digital solutions, differentiated the use of tools, and implemented teleconsultations to solve urgent problems.* [GP1]

Teleconsultations: The New Way of Working

All the GPs have implemented teleconsultations since the start of the pandemic. The GPs said they used teleconsultations as a complementary tool almost daily in their practices. They blocked extra hours for teleconsultations between physical consultations. They found it critical for continuity of care, and they used it as an exit strategy during the enormous workload of the pandemic. Teleconsultations were used for working safely, triaging, and following patients.

> *Teleconsultation is not a pleasant thing, but it was a way for us to overcome the intensity of the pandemic.* [GP3]

Creating a Safe Working Environment and Triaging via Teleconsultations

> *Via teleconsultations, we protected ourselves from COVID infection, we could work safely, and we guaranteed the continuity of care during lockdowns.* [GP12]

GP as a Digital Administrator

All the interviewees noted general displeasure about digital administration for different reasons. They all pointed out that the pandemic brought many extra administration duties at the expense of patient care.

> *The impact of COVID was huge. Suddenly, we had e-certificates that we had to deliver to patients, but there was the practical problem that the digital signature was not accepted, so every patient had to come to us again, but they could not because they had to stay home…So, we had to sign certificates for these people, which is absurd; if someone had a positive test and got a digital result, why should I have to write a homestay certificate again?* [GP13]

Increased Investments in IT

During the pandemic, GPs were aware of eHealth benefits. They described how they tended to invest in IT solutions from the start of the pandemic and noted that they searched for and invested in more stable software, online modules, and providers.

GP12 described how the COVID-19 pandemic boosted the investment in her practice:

> *I had been thinking about switching to an online appointment program for many years. The COVID-19 pandemic led me to invest in this program with an extra telemedicine module.* [GP12]

GP7, who worked in a group practice, added that with the new investments, all the colleagues worked more flexibly and remotely:

> *Now, everyone at the practice has laptops with new software that we can take home so that we have access to our medical records at home.* [GP7]

Domain 2: Perception and Appreciation of eHealth

Digital Communication Tool

Across multiple interviews, all the GPs defined eHealth as a digital hub of instruments and administrations that provides different health-related tools. The primary eHealth solutions were private software for patient dossiers (eg, Care-connect), online appointment modules, online secretary tools, telemodules, and mHealth apps (eg, the Collaborative Care Platform [CoZo], e-receipts, e-certificates).

The interviewees stated that eHealth apps were critical during the COVID-19 pandemic and that their perception of eHealth, which they previously associated with ICT problems, has changed positively due to the pandemic because they use this system more efficiently.

> *Digital solutions increased the quality and continuity of health care.* [GP1]

GP2 shared vital feedback about how the old GP generations appreciated eHealth during the COVID-19 pandemic:

> *COVID gave us extra training, especially older GP generations who were against the digitalization. It allowed them to understand the benefits of it because before that; they were not using any digital solutions even though we had them.* [GP2]

The Future Is Digital

GPs stated that they work predominantly in the digital environment and that eHealth is the future.

> *The more, the better, I am happy.* [GP13]
>
> *This is an inevitable future.* [GP15]

Advantages of eHealth

GPs noted that providing primary care via eHealth gave them immense satisfaction in their work. The participants noted 3 main advantages of eHealth. First, eHealth collected all data easily, quickly, and efficiently. Second, it saved time. Third, it sped up their work, especially in urgent conditions.

> *Thanks to eHealth, I save time and I spend this extra time with my family. Before I had to take all the documents home, which was extra work, but now, it is effortless.* [GP14]

GP1 added that eHealth has ”a lot of positive things…easy to find, collect, and deliver data, easy to work paperless, easy to send medicine, easy to work quickly, easy to send invoices…“

Disadvantages of eHealth

All the GPs added that there were some disadvantages that they had to overcome, such as technical problems, sitting for too long in front of the computer, a lack of physical contact with the patients, and being perceived as a ”24–7 online doctor.“

> *Mmm…I think we see more patients in a short time and have shorter contact with the patient, which is an advantage, but at the same time also a disadvantage because the patient has the idea that we are always accessible and he can say, “I have a skin problem, I will take a picture of it and forward it to the GP by e-mail,” and he thinks that he will immediately get an answer. Unfortunately, it usually does not work, so a bit of a fine-tuning is still needed, not only for us but also for the patients, to figure out how we should deal with that.* [GP12]

Potential Services and Success Points Related to eHealth Use

The GPs noted several possible services and success points during the pandemic regarding eHealth use. First, they all stated that the COVID-19 pandemic trained them to cope with the unknown. They were forced to deliver care during the lockdowns and had to find more innovative care delivery methods.

Potential Services

Preparedness for the Unknown

The interviewees acknowledged that eHealth use and new experiences prepared them for other possible pandemics or health crashes. Some of them described how they implemented their initiatives.

> *After the first shock, we were urged to deliver care only for emergencies and COVID infections. All other*

> *things were canceled. We created a common Excel document with our patients' data categorized with colors (eg, red meant urgent, yellow meant less urgent, green meant solved), and we had to check that frequently. This was our adaptation to keep track of our patients. In addition, we moved physical consultations with COVID-like illnesses to very early in the morning and had to disinfect all the practice rooms after. These were different solutions we had to find. We met continuously with other colleagues to learn and share new things.* [GP5]

Digital Literacy

All the participants said they were busy improving their digital literacy to use eHealth effectively during the lockdowns.

> *As a practice, we asked for extra pieces of training and workshops from the providers.* [GP14]

mHealth and Video Consultations

The responders indicated that they used mHealth apps and video consultations after telephone consultations.

> *I do not use applications to follow the patients, but I think that would be interesting. For example, diabetes follow-ups. This would be exciting if people take a test at home and integrate the results with their medical files. I also think there are applications for it.* [GP6]

A few of them tried to use video consultations for selected patients. They all acknowledged that a stable internet connection and technical competency were needed for successful telehealth practices:

> *The next step we were thinking of was implementing video consultations. We saw it during a congress in the United States, and we thought about it before the COVID-19 pandemic, but now we are busy with investment for that.* [GP2]

Success Points

Increased Accessibility

The convenience of remote care made it easier for patients to reach doctors in the first months of the COVID-19 pandemic when the quarantine instructions were applied. Using eHealth solutions allowed GPs to coordinate primary health care and safely communicate with patients.

> *Using eHealth is helpful because people can have greater access to care during the pandemic even outside working hours. We are available at night and on the weekends to provide the urgent needs. eHealth solution makes people reach our services and increase accessibility, which is critical during the pandemic.* [GP8]

The Mental Shift to Remote Work

GPs collaborated through physical and remote work and achieved good results from their remote work. A mind shift occurred during remote consultations. This mind shift was 1 of the vital breaking points to fastening eHealth solutions' uptake. The GPs said they had to find answers to solve problems urgently and remotely. In this way, they changed their work and this allowed them to collaborate physically and via teleconsultations.

> *The COVID-19 pandemic forced us to solve problems remotely, which required a mental shift to combine physical and teleconsultations.* [GP10]

> *One of the most important things we learned from the COVID crisis is that eHealth solutions shifted our mindset…from relying on the traditional doctor's role to considering an integrated role using more remote solutions.* [GP11]

Physician-Patient Satisfaction With Telemedicine

As doctor-patient relationships evolved over the past decades, communication became more diverse during the pandemic. The satisfaction of patients became more crucial than ever according to the surveyed GPs. The interviewees added that young adults used all the digital technologies efficiently, and adaptation to eHealth services during the pandemic was high. Using eHealth allowed patients not to replace their selves, to find quick solutions remotely, and to be satisfied with the health service.

> *This was a comfort for the patient, no transport, no parking, no time wasted on coming to the doctor.* [GP4]

> *Most patients are comfortable and are doing well with eHealth solutions. They are happy to use them.* [GP12]

Domain 4: Barriers to and Facilitators of eHealth

Barriers

According to GPs, there were 2 critical barriers to eHealth uptake. Although eHealth barriers diminished compared with previous studies, the ICT burden and patients' digital literacy remained.

ICT Burden

> *eHealth information does not work on Monday, works on Tuesday, Wednesday, Thursday, and then stops working on Friday. This is such a long weekend, and things are not reachable, insurance and citizen numbers cannot be seen…then, we are blocked.* [GP4]

Digital Literacy

The responders commented on another significant barrier: patients' digital literacy. They found that this was critical for providing equal access to services during the pandemic. The patients who did not understand how to use digital health tools could not access those tools in their language, which disadvantaged them.

> *If you work in the suburbs of Brussels, you know that some populations cannot speak Dutch or French. They do not understand the invitation briefs from the government, and they come to ask what these papers say. Sometimes, the language used in official papers is very complex, not simple. These groups are also not eligible to use digital solutions. Education of these*

fragile groups is essential because they feel like they are outside of the community. [GP1]

Facilitators

The GPs in our study also identified 3 main domains as facilitators to the uptake of eHealth. The biggest was COVID-19 measures, followed by government and European Union (EU) policies and the patients' active role during the pandemic.

COVID-19 Pandemic Measures

The COVID-19 measures applied for preventing the spread of the virus also reduced access to care. According to GPs, measures were a change moment for transforming ideas and discovering the benefits of eHealth solutions.

We had to find a solution to maintain care, which was difficult because everything was canceled...There was no more physical contact because of the measures, which forced us to devise creative solutions, use more technology, and get used to it… [GP5]

Government and EU Policies

Remuneration for teleservices from the government and developed digital apps to provide PCR (polymerase chain reaction) tests and e-certificates accelerated the uptake for eHealth according to GPs.

We grew more accustomed to eHealth tools through providing for our patients by creating and sending PCR tests, e-certificates, e-receipts, and reimbursement papers. [GP7]

The Role of the Patient

Patient demand for telemedicine outstripped the ability of health care providers to supply it. As our interviewees noted, there was considerable satisfaction on the part of the patient about reaching their doctor quickly and getting solutions during the pandemic, which stimulated greater eHealth uptake in health care.

Communication with the patient was online, quick, and necessary…They were more satisfied than us, I think. Patients using these tools efficiently definitely increased our adaptation to eHealth. [GP11]

Domain 5: Future Expectations and Concerns

Expectations

All the GPs were satisfied with the information circle between the first line of medical services, but they noted that there was no efficient information flow between the first and the second lines.

More Integrated First and Second Service Lines

Responders emphasized that they also had difficulties transferring information to the second and third medicine lines. They complained about a lack of integration between health care services in eHealth.

Every hospital has a different eHealth system. In 2022, this cannot happen. [GP13]

We are working on the same patient, but there is a lot of missing data between the service lines. What I do cannot be seen by the specialist… [GP11]

More Efficient IT Solutions

Another issue was stable IT support. Although IT support has significantly improved since the beginning of the COVID-19 pandemic, almost all participants declared that they expected more efficient IT solutions.

Stable, technically strong, and well-integrated IT solutions are expected in the future. [GP4]

Remuneration for Telemedicine Services

All the interviewees agreed that receiving reimbursement was beneficial during the pandemic. One physician added that she and her colleagues were upset about the reimbursement department's decision regarding the cancellation of reimbursement for teleservices.

I am despondent about this decision. Remuneration for teleservices was first decreased and then canceled. We are busy collecting signatures to get back the teleconsultation reimbursement. I think this is very beneficial for us now and in the future. Remuneration for teleconsultations must be well regulated if the policymakers want us to cope with unknown circumstances… [GP9]

Concerns

The interviewees were worried about some issues. The main concerns were the security of patients' data and fragile groups who were not eligible to use digital solutions and the loss of in-person contact.

Awareness of Data Security

First, the increase in digitalization made the GPs concerned about data regulation and safety. During the pandemic, there was an extensive unprotected circulation of patient data on different digital platforms. Emails, text messages, and WhatsApp messages were not protected enough according to GPs.

One thing we mostly forget is that the patients' data circulate on unprotected platforms like emails and WhatsApp...I find it personally very dangerous and open to cyber-attacks. [GP8]

Fragile Groups

Second, they were sensitive to the patients with socioeconomic problems or language barriers and subgroups that were not eligible to use digital solutions. The GPs mentioned that fewer controls and in-person contact would probably lead to substantial health problems for these people. They expected to experience more distress from COVID-19 and the lockdowns for these fragile groups.

Cultural differences and language barriers were the main reasons patients did not adopt eHealth tools. [GP13]

Some people were unaware of their health problems, and others were not eligible to use digital systems and needed in-person visits. These people are fragile,

and access to health care could be more difficult if we only use digital platforms. [GP1]

Losing In-Person Contact and Follow-Ups

Even though the shift from in-person to remote care during the COVID-19 pandemic was supposedly advantageous for GPs, there was a big concern about losing in-person contact with the patients. Some added that they deliberately minimized teleconsultations in order to avoid losing this contact with the patients, and others asked patients to come to the clinic after a few digital consultations.

We do now have a limit of 3 months for e-prescriptions. After that, they have to come physically to see us. We are limiting teleconsultations. Things are now returning to normal. We do not want to lose in-person contact with the patients because we are responsible for the quality of care. [GP2]

Only Teleconsultations Are Not Enough

The GPs noticed that although teleconsultations were critical for continuity of care and used as an exit strategy during the enormous workload of the pandemic, pure digital consultations are unrealistic and cannot replace natural, physical interaction with the patient. Some GPs added that fully digital solutions are not enough to provide holistic care.

You cannot do everything digital…Some patients need physical contact. Others must be followed in-person contact because they are not fully ready to control their health and are unaware of what is going on… [GP13]

We are holistic health care providers; we evaluate all the aspects of the patient, physiological, social, physical…Furthermore, unfortunately, in teleconsultation, we only make up for what it lacks. Therefore, we always say we leave the door open for in-person visits. [GP15]

Implementing Results on Virginia Satir's Model

Why this eHealth uptake rapidly changed after years of slowness in adopting digital practices and why and how the COVID-19 pandemic became a tipping point are not easy questions to answer, because health care organizations are complex. It is not always possible to predict changes or the effects of interventions, because of the significant, interdependent, hierarchical, top-down, and fragmented characteristics of health care organizations [36,37]. Therefore, to implement a change in complex health care contexts, change management methodologies are often used as guiding principles [38,39]. In this study, we used Virginia Satir's change model to understand this change [40]; see Figure 3. This model was created by Virginia Satir, a family therapist, to aid individuals or organizations recover the way they deal with noteworthy, unpredicted change. Using this model in health care may help explain the changes because it includes hierarchical and growth models. The hierarchical model attempts to simplify life and allows for a complicated structure of layered responsibilities, whereas the growth model is geared toward the complexity of human interactions [41]. This model includes 5 stages: old status quo, resistance, chaos, integration, and new status quo.

In the old status quo phase, we can see how things were shortly before we realized the significant and disruptive change. Before the COVID-19 pandemic, eHealth uptake was low due to different barriers. There was less investment in eHealth services, and eHealth was perceived as time-consuming. Physical contact with patients was the cornerstone of care.

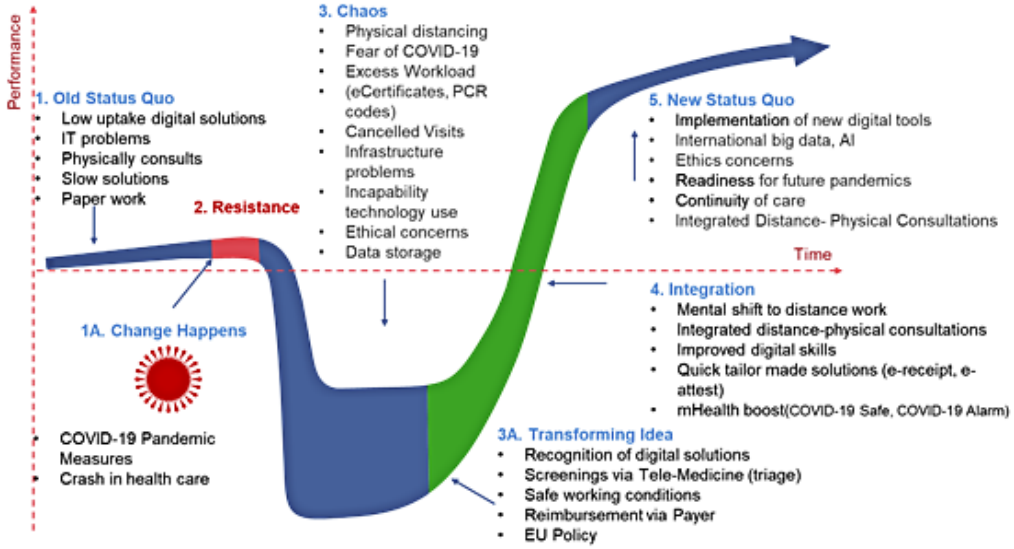
The resistance phase describes resistance to the foreign element that disturbs the comfort of the status quo. In this case, the pandemic caused a significant shock to the health care system. Continuity of care was disrupted, and there was a service shift to critical COVID-19–related issues. Teleservices were used extensively but were insufficient due to the loss of nonverbal language. Concerns were raised about the continuity of regular care from GPs.

In the chaos and idea transformation phase, the foreign element gains critical weight and changes work. The old way of working is not adopted. Performance declines, and a sense of urgency emerges—feelings of stress, confusion, vulnerability, fear, and panic arise. Nevertheless, these feelings can generate new ideas, which will be a path to transformation. Crucial to this stage is to see how the external element can benefit through a transformative idea. Trying new ways of working and acquiring new skills can significantly improve performance. During the chaos phase of the COVID-19 pandemic, enormous changes took place. Fear of COVID-19 infection and measures (physical distancing, curfews) changed the way of daily practice. eHealth solutions became unavailable, infrastructure collapsed, and feelings of anxiety, stress, and uncertainty about the future emerged. During this phase, we noticed that GPs took the initiative to create safe working conditions. They realized the new use of digital solutions and consulted via telemedicine. eHealth enabled them to reduce the burden on their daily practice work. They recovered their labor fees from the authorities and followed EU policies.

In the integration phase, integration occurs through the reinforcement of numerous practices and the new state of change. In the final stages of the COVID-19 pandemic, GPs routinely used teleservices as an integral part of their daily care to ensure continuity of care. They improved their digital capabilities through many apps. They integrated physical consultations and teleworking in this phase, allowing the mental shift to the new integration.

In the last phase, the new status quo, new working methods begin to be applied. New skills become second nature. New norms are formed as part of this. Work performance starts to align with the new skills, and a new status quo is formed. At this stage, we noticed that GPs were ready for unknown conditions and worked using the new digital services. They prioritized patient satisfaction. There was more sensitivity to ethical and privacy issues. Continuity of care has provided preparedness for future pandemics.

Figure 3. Adapted from Virginia Satir's change model [40]. AI: artificial intelligence; EU: European Union. PCR: polymerase chain reaction.



Discussion

Principal Findings

Our study indicates that the COVID-19 crisis was a critical turning point in adopting eHealth tools by GP practices in Flanders. Since the pandemic's start, the use of eHealth has rapidly increased and evolved. There was an increase in not only telemedicine but also mHealth, remote monitoring, and direct communication between health providers and patients [34].

According to the European Public Health Alliance (EPHA) briefing from November 2021, there has also been a tremendous increase in the deployment of digital health tools in European countries, such as Sweden and Italy, from doubling to a 30.1 times increase, respectively [42]. Tracking apps and Digital Green Certificates were developed to support the resumption of international mobility for tourism, leisure, and business, both internationally and in Europe.

GPs have been more flexible than ever before, working remotely, while maintaining continuity of care. Our respondents reported 3 factors facilitating eHealth uptake: COVID-19 pandemic measures, government and EU policies, and changing patient roles. Implementing teleservices for triage, charging for remote administration, and using digital applications (eg, COVID Safe, CoZo, and e-receipts). Interestingly, the GPs considered patient satisfaction with eHealth services essential in our study. This observation may support the hypothesis that the role of the patient has been actively changed, centered, and prioritized. During the pandemic, the participation of patients in eHealth use and teleservices was frequent, forcing the GPs to be more flexible with technological solutions. All these enablers have been documented in other studies.

This momentum was significant because health care was previously identified as 1 of the sectors lagging the furthest in adopting digital services. According to research, there were different reasons, such as the strict regulations of the sector, the sensitivity surrounding personally identifiable information, the resistance of health care providers to digital solutions, the lack of prioritization of patient experience, and the cost of investments [3-5]. Despite the potential benefits of eHealth in Belgium and other EU countries, uptake was slower than expected before the coronavirus crisis [28,43,44]. Research by the European Commission has shown that eHealth adoption in all European countries is much more complex and time-consuming than initially envisaged [45,46].

A Belgium study conducted between 2018 and 2019 identified infrastructure as the most significant barrier to eHealth adoption by GPs [27]. Respondents argued that although there had been improvement recently, there were continuity problems, crashes, and software updates when they were busy with patients and their electronic medical files. eHealth is seen as time-consuming, requiring significant investment and ICT services, leading to information overload, needing data security, and feeling dependent on external factors. However, the findings of this study differ from the previous research on eHealth perception and appreciation. Interestingly, there was a vast difference in the perception and appreciation of eHealth adoption before and after the COVID-19 crisis. In our study, the interviewed GPs agreed that the pandemic increased, boosted, and sped up their daily work, making eHealth essential. The GPs are aware of the benefits of eHealth and feel that eHealth is an important complementary part of health care, can reduce the burden of chaos, and increase access to health care. Contrary to what the earlier studies have found, the interviewees no longer see eHealth as a time-consuming tool; they see it as a time-saving instrument. All the GPs emphasized that eHealth is part of their work, stating that they will use it more in the future and noting that they are voluntarily investing in eHealth solutions.

In addition to the many advantages of eHealth use, the GPs noted some disadvantages, such as decreased privacy with the patients perceiving 24–7 online access to their doctor and increased busyness with the digital tools. In addition, their dissatisfaction was due to constantly being the digital administrator and prescribing more e-certificates than needed. These results have not previously been described in other studies.

One of the most important findings of our study is the mental shift of GPs to be more optimistic about teleworking. Although all the participants stated that physical contact with the patient would never lose its importance, the interviewees have become accustomed to working remotely and routinely practicing telemedicine. Telehealth is an appropriate and satisfactory modality for receiving care. It was essential to providing remote medical assistance, reducing GP workload, and helping increase patient satisfaction. This shift in mindset is evident because some GPs have stated that there will never again be only physical examinations and that a rational combination of telemedicine and physical examinations is vital. The interviewees also mentioned that the reimbursement conditions for telemedicine should be improved and continued. After the coronavirus crisis, teleconsultation became a complementary tool in the daily practice of Flemish GPs. These findings are similar to the results of previous quantitative and qualitative studies regarding the acceptance of eHealth during the COVID-19 pandemic [47,48].

In line with other studies [27,49], our study exposes patient digital literacy and the ICT burden as the main barriers to eHealth use. These are the primary obstacles noted by the GPs. The GPs stated that digital literacy is crucial for the elderly and socially deprived groups. They pointed to harmful discrimination perpetrated against these groups if only eHealth solutions were used. Further, blocking software and system crashes during work adversely affected eHealth uptake. Although the interviewees are satisfied with the massive development in IT solutions, more work still needs to be done. The GPs expect a stable and effective ICT health care strategy in the future. Confidentiality and security of patient data are no longer barriers to using eHealth but remain a concern among GPs. Other areas of concern include losing physical contact with the patient and missing follow-ups.

It is, therefore, likely that eHealth has become part of the daily routine in GP practices as a result of the pandemic. The practices demand a more stable IT infrastructure in the future to work efficiently, and more care communications between the first and second lines within the eHealth services are needed. This interorganizational coordination is crucial in order to create integrated services.

Finally, as eHealth becomes more implemented, we find that GPs are more open and flexible to using eHealth solutions, especially after the advent of COVID-19, and they feel more prepared for uncertain conditions, such as a pandemic. GPs' digital literacy increased, and they intend to use more eHealth solutions, such as video consultations and mHealth apps, in the future.

Limitations

There were some limitations to this study. First, practices in only a few districts in the Flemish region of Belgium with Dutch-speaking GPs were surveyed, and the experiences of other GPs in different regions were not considered. The sample size was quite small, and the use of the snowball method could have led to overrepresentation of GPs more "enthusiastic" toward eHealth. Nevertheless, our GPs also clearly indicated the risks, dangers, and drawbacks of eHealth and telemedicine. Further, the interviews were conducted between October 2021 and April 2022, and the GPs were asked about their experience of eHealth use since March 2020, when the COVID-19 pandemic started. This might have led to recall bias. In contrast, interviews were conducted in a period that eHealth use became already more integrated as a relatively sustainable routine in daily practice. As a result, eHealth use and the intention to continue using it may have changed marginally between the October 2021 and April 2022 interviews due to the normalization of GPs' daily practices.

Furthermore, we did not interview patients or other health care providers for our study. Therefore, our findings are limited to what the doctors detailed. In addition, we used semistructured interviews, which are highly accepted in terms of deep discussion, adaptability, locking into productive nature, empowering modern thoughts, and seeing interviews taken in their natural forms by counting nonverbal communication. However, we used Zoom to conduct online interviews due to the COVID-19 crisis, which could have resulted in a loss of information because the conditions are not the same as face-to-face interviews.

Further research will be beneficial for understanding which eHealth solutions could be implemented in the future, such as video consultations and mHealth apps, and their remuneration possibilities. Another important research topic is how to boost digital literacy for doctors and patients to sustain eHealth in the future. Finally, our qualitative research suggests that eHealth services have been adopted by GPs, but this hypothesis as well as the reach of implementation should be tested by quantitative research methods.

We recommend that health IT policymakers and developers maintain the continuity of eHealth solutions beyond the COVID-19 pandemic, considering the expectations and vulnerabilities presented in this study.

Conclusion

With this study presented in the Flemish community of Belgium, we tried to acquire deep insights from the GPs into the perceived effects of the COVID-19 crisis on eHealth use and why and how things changed in their daily practices.

Previously defined areas of research were thoroughly analyzed, and we showed that the COVID-19 pandemic was a critical situation that provided significant and unstoppable changes in the uptake of eHealth by Flemish GPs. The coronavirus crisis was an accelerator for digital health care, which was previously described as "backward." It made a positive mental shift toward eHealth adoption.

According to GPs, eHealth became essential to their daily practice during the COVID-19 crisis, reducing the burden on health care services and increasing health care accessibility. The perception of and appreciation for eHealth have changed positively, and eHealth has become an integral part of daily care. Patient satisfaction has increased and been prioritized significantly, and GPs have become more open and ready to implement more eHealth and mHealth solutions in their daily practice.

Although there has been a positive cognitive shift toward eHealth adoption, this study shows that there is still a significant amount of skepticism and uncertainty around privacy, security of patient data, digital literacy, and remuneration. Future expectations were addressed as more integrated first and second services lines, a more stable IT infrastructure, and remuneration for digital services.

Acknowledgments

The authors thank the interviewees. Without their help, this research would not have been possible. The authors received no specific funding for this work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide.
[DOCX File , 20 KB - formative_v6i11e41847_app1.docx]

References

1. European Public Health Alliance. Mobile Health (mHealth). 2013. URL: https://epha.org/epha-briefing-mobile-health-mhealth/ [accessed 2022-11-20]
2. Callens S, Peers J. Organisatie van de gezondheidszorg. Belgium: Intersentia; 2015.
3. Jansen-Kosterink S. The Added Value of Telemedicine Services for Physical Rehabilitation. 2014. URL: https://ris.utwente.nl/ws/portalfiles/portal/6030964 [accessed 2022-09-18]
4. Bonten TN, Rauwerdink A, Wyatt JC, Kasteleyn MJ, Witkamp L, Riper H, EHealth Evaluation Research Group. Correction: online guide for electronic health evaluation approaches: systematic scoping review and concept mapping study. J Med Internet Res 2020 Aug 21;22(8):e23642 [FREE Full text] [doi: 10.2196/23642] [Medline: 32822315]
5. van der Kleij RM, Kasteleyn MJ, Meijer E, Bonten TN, Houwink EJ, Teichert M, et al. SERIES: eHealth in primary care. Part 1: concepts, conditions and challenges. Eur J Gen Pract 2019 Oct 10;25(4):179-189 [FREE Full text] [doi: 10.1080/13814788.2019.1658190] [Medline: 31597502]
6. Nijland N. Grounding eHealth: Towards a Holistic Framework for Sustainable eHealth Technologies. 2011. URL: https://research.utwente.nl/en/publications/grounding-ehealth-towards-a-holistic-framework-for-sustainable-eh [accessed 2022-09-18]
7. Hibbard JH, Mahoney ER, Stock R, Tusler M. Do increases in patient activation result in improved self-management behaviors? Health Serv Res 2007 Aug;42(4):1443-1463 [FREE Full text] [doi: 10.1111/j.1475-6773.2006.00669.x] [Medline: 17610432]
8. Boers S, Jongsma K, Lucivero F, Aardoom J, Büchner FL, de Vries M, et al. SERIES: eHealth in primary care. Part 2: exploring the ethical implications of its application in primary care practice. Eur J Gen Pract 2020 Dec;26(1):26-32 [FREE Full text] [doi: 10.1080/13814788.2019.1678958] [Medline: 31663394]
9. van Hattem NE, Silven A, Bonten T, Chavannes N. COVID-19's impact on the future of digital health technology in primary care. Fam Pract 2021 Nov 24;38(6):845-847 [FREE Full text] [doi: 10.1093/fampra/cmab081] [Medline: 34268563]
10. Golinelli D, Boetto E, Carullo G, Nuzzolese AG, Landini MP, Fantini MP. Adoption of digital technologies in health care during the COVID-19 pandemic: systematic review of early scientific literature. J Med Internet Res 2020 Nov 06;22(11):e22280 [FREE Full text] [doi: 10.2196/22280] [Medline: 33079693]
11. Thulesius H. Increased importance of digital medicine and eHealth during the COVID-19 pandemic. Scand J Prim Health Care 2020 Jun;38(2):105-106 [FREE Full text] [doi: 10.1080/02813432.2020.1770466] [Medline: 32484725]
12. Organisation for Economic Cooperation and Development. Health at a Glance 2021. 2021. URL: https://tinyurl.com/bdh7u75r [accessed 2022-11-20]
13. Centraal Planbureau. Zorgkeuzes in Kaart 2020: Analyse van beleidsopties van politieke partijen voor de zorg. 2022 Jul 24. URL: https://www.cpb.nl/zorgkeuzes-in-kaart-2020[accessed [accessed 2022-09-18]
14. Mann DM, Chokshi SK, Kushniruk A. Bridging the gap between academic research and pragmatic needs in usability: a hybrid approach to usability evaluation of health care information systems. JMIR Hum Factors 2018 Nov 28;5(4):e10721 [FREE Full text] [doi: 10.2196/10721] [Medline: 30487119]
15. Heckemann B, Wolf A, Ali L, Sonntag SM, Ekman I. Discovering untapped relationship potential with patients in telehealth: a qualitative interview study. BMJ Open 2016 Mar 02;6(3):e009750 [FREE Full text] [doi: 10.1136/bmjopen-2015-009750] [Medline: 26936904]

16. Aardoom J, van DL, Rompelberg C, Standaar L, Suijkerbuijk A, van TL. Indicatoren E-healthmonitor 2021-2023 en doelstellingen voor e-health. Rijksinstituut voor Volksgezondheid en Milieu. 2021. URL: https://rivm.openrepository.com/handle/10029/624864[accessed [accessed 2022-09-18]
17. Snyder H, Engström J. The antecedents, forms and consequences of patient involvement: a narrative review of the literature. Int J Nurs Stud 2016 Jan;53:351-378 [FREE Full text] [doi: 10.1016/j.ijnurstu.2015.09.008] [Medline: 26602069]
18. Bentvelsen R, van der Vaart R, Veldkamp K, Chavannes N. Systematic development of an mHealth app to prevent healthcare-associated infections by involving patients: 'Participatient'. Clinical eHealth 2021;4:37-44 [FREE Full text] [doi: 10.1016/j.ceh.2021.03.001]
19. Bruce CR, Harrison P, Nisar T, Giammattei C, Tan NM, Bliven C, et al. Assessing the impact of patient-facing mobile health technology on patient outcomes: retrospective observational cohort study. JMIR Mhealth Uhealth 2020 Jun 26;8(6):e19333 [FREE Full text] [doi: 10.2196/19333] [Medline: 32589161]
20. Matamala-Gomez M, Maisto M, Montana J, Mavrodiev P, Baglio F, Rossetto F, et al. The role of engagement in teleneurorehabilitation: a systematic review. Front Neurol 2020;11:354 [FREE Full text] [doi: 10.3389/fneur.2020.00354] [Medline: 32435227]
21. Tossaint-Schoenmakers R, Versluis A, Chavannes N, Talboom-Kamp E, Kasteleyn M. The challenge of integrating eHealth into health care: systematic literature review of the Donabedian model of structure, process, and outcome. J Med Internet Res 2021 May 10;23(5):e27180 [FREE Full text] [doi: 10.2196/27180] [Medline: 33970123]
22. Chan S, Torous J, Hinton L, Yellowlees P. Mobile tele-mental health: increasing applications and a move to hybrid models of care. Healthcare (Basel) 2014 May 06;2(2):220-233 [FREE Full text] [doi: 10.3390/healthcare2020220] [Medline: 27429272]
23. Hughes M, Gorman J, Ren Y, Khalid S, Clayton C. Increasing access to rural mental health care using hybrid care that includes telepsychiatry. J Rural Ment Health 2019 Jan;43(1):30-37 [FREE Full text] [doi: 10.1037/rmh0000110]
24. 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]
25. Barkai G, Gadot M, Amir H, Menashe M, Shvimer-Rothschild L, Zimlichman E. Patient and clinician experience with a rapidly implemented large-scale video consultation program during COVID-19. Int J Qual Health Care 2021 Feb 20;33(1):mzaa165 [FREE Full text] [doi: 10.1093/intqhc/mzaa165] [Medline: 33313891]
26. HealthManagement.org. Healthcare IT in Belgium. 2009. URL: https://ftp.healthmanagement.org/c/it/issuearticle/healthcare-it-in-belgium [accessed 2022-11-20]
27. Detollenaere J, Christiaens W, Dossche D, Camberlin C, Lefèvre M, Devriese S. Barriers and Facilitators for eHealth Adoption by General Practitioners in Belgium. Analysis Based on the Integrated Allowance for GP Practices. 2021. URL: https://kce.fgov.be/sites/default/files/2021-11/KCE_337_eHealth_adoption_in_Belgium_Report_V2.pdf [accessed 2022-11-20]
28. Hennemann S, Beutel ME, Zwerenz R. Ready for eHealth? Health professionals' acceptance and adoption of eHealth interventions in inpatient routine care. J Health Commun 2017 Mar;22(3):274-284. [doi: 10.1080/10810730.2017.1284286] [Medline: 28248626]
29. Hulme G. Majority of U.S Healthcare Providers Rank Themselves as Digital Laggards. Digiruptio Newsletter. 2019. URL: https://digirupt.io/majority-of-u-s-healthcare-providers-rank-themselves-as-digital-laggards/ [accessed 2022-09-18]
30. Kruse CS, Kristof C, Jones B, Mitchell E, Martinez A. Barriers to electronic health record adoption: a systematic literature review. J Med Syst 2016 Dec;40(12):252 [FREE Full text] [doi: 10.1007/s10916-016-0628-9] [Medline: 27714560]
31. COVID-19 Mortality. 2021. URL: https://www.healthybelgium.be/en/health-status/covid-19-crisis/covid-19-mortality [accessed 2022-09-18]
32. Brzozowski A. Belgium Activates Emergency Measures to Contain Coronavirus Outbreak. 2020. URL: https://www.euractiv.com/section/coronavirus/news/belgium-activates-emergency-measures-to-contain-coronavirus-outbreak/ [accessed 2022-09-18]
33. Academic Health Science Network for the North East and North Cumbria (AHSN NENC). Rapid Insights into Digital GP Solutions during the COVID-19 Pandemic: June 2020 NENC Daa. URL: https://www.ahsn-nenc.org.uk/wp-content/uploads/2020/06/NENC-ICS.docx [accessed 2022-11-20]
34. Avalosse H, Bruyneel L, Hens E. Intermutualistische Enquete Raadplegingen Op Afstand. 2020. URL: https://www.riziv.fgov.be/SiteCollectionDocuments/teleconsultatie_resultaten_rapport.pdf [accessed 2022-09-18]
35. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
36. Pype P, Krystallidou D, Deveugele M, Mertens F, Rubinelli S, Devisch I. Healthcare teams as complex adaptive systems: focus on interpersonal interaction. Patient Educ Couns 2017 Nov;100(11):2028-2034. [doi: 10.1016/j.pec.2017.06.029] [Medline: 28687278]
37. Stange KC. The problem of fragmentation and the need for integrative solutions. Ann Fam Med 2009 Mar 01;7(2):100-103 [FREE Full text] [doi: 10.1370/afm.971] [Medline: 19273863]

38. Ratnapalan S, Lang D. Health care organizations as complex adaptive systems. Health Care Manag 2019 Dec 26;39(1):18-23. [doi: 10.1097/hcm.0000000000000284]
39. Harrison R, Fischer S, Walpola RL, Chauhan A, Babalola T, Mears S, et al. Where do models for change management, improvement and implementation meet? A systematic review of the applications of change management models in healthcare. J Healthc Leadersh 2021 Mar;13:85-108. [doi: 10.2147/jhl.s289176]
40. Satir V, Banmen J, Gerber J. Perceiving the world. In: The Satir Model: Family Therapy and Beyond. Palo Alto: Science and Behaviour Books; 1991.
41. Hutchinson T. Whole Person Care. Cham: Springer; 2017.
42. European Public Health Alliance. Case Study Collection: The Impact of COVID-19 on the Digital Transformation of Health and Care. 2021 Nov. URL: https://epha.org/wp-content/uploads/2021/11/case-studies-digital-transformations-2021.pdf [accessed 2022-11-20]
43. Ajami S, Ketabi S, Isfahani S, Heidari A. Readiness assessment of electronic health records implementation. Acta Inform Med 2011 Dec;19(4):224-227 [FREE Full text] [doi: 10.5455/aim.2011.19.224-227] [Medline: 23407861]
44. Wachter R. Work: Harnessing the Power Health Information Technology to Improve Care in England. Report of the National Advisory Group on Health Information Technology in England. 2016. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/550866/Wachter_Review_Accessible.pdf [accessed 2022-09-18]
45. Stroetmann K, Artmann J, Stroetmann V. 1. European Countries on their Journey Towards National eHealth Infrastructures. 2011. URL: http://www.ehealth-strategies.eu/report/eHealth_Strategies_Final_Report_Web.pdf [accessed 2022-09-18]
46. Watson R. European Union leads way on e-health, but obstacles remain. BMJ 2010 Sep 21;341(sep21 2):c5195. [doi: 10.1136/bmj.c5195] [Medline: 20858645]
47. Srinivasan M, Asch S, Vilendrer S, Thomas SC, Bajra R, Barman L, et al. Qualitative assessment of rapid system transformation to primary care video visits at an academic medical center. Ann Int Med 2020 Oct 06;173(7):527-535. [doi: 10.7326/m20-1814]
48. Malouff TD, TerKonda SP, Knight D, Abu Dabrh AM, Perlman AI, Munipalli B, et al. Physician satisfaction with telemedicine during the COVID-19 pandemic: The Mayo Clinic Florida Experience. Mayo Clin Proc Innov Qual Outcomes 2021 Aug;5(4):771-782 [FREE Full text] [doi: 10.1016/j.mayocpiqo.2021.06.006] [Medline: 34226884]
49. Zolbin M, Nikou S. Health literacy improvement and use of digital health services in aged people: a systematic literature review. 2021 Presented at: Conference: International Telecommunications Society (ITS); June 21-23, 2021; Gothenburg, Sweden p. 21-32.

Abbreviations

AI: artificial intelligence
CoZo: Collaborative Care Platform
EU: European Union
GP: general practitioner
ICT: information and communication technologies
mHealth: mobile health

Edited by A Mavragani; submitted 11.08.22; peer-reviewed by AS Alaqra, M Kapsetaki, T Parikh, A Das Gupta; comments to author 30.08.22; revised version received 19.09.22; accepted 04.10.22; published 28.11.22.

Please cite as:
Yagiz JI, Goderis G
The Impact of the COVID-19 Pandemic on eHealth Use in the Daily Practice and Life of Dutch-Speaking General Practitioners in Belgium: Qualitative Study With Semistructured Interviews
JMIR Form Res 2022;6(11):e41847
URL: https://formative.jmir.org/2022/11/e41847
doi:10.2196/41847
PMID:36399650

Original Paper

Enrollment and Completion Characteristics for Novel Remote Delivery Modes of the Self-management Programs During the COVID-19 Pandemic: Exploratory Analysis

Kristin Pullyblank[1,2], MS; Serdar Atav[2], PhD

[1]Center for Rural Community Health, Bassett Research Institute, Bassett Medical Center, Cooperstown, NY, United States

[2]Decker College of Nursing and Health Sciences, Binghamton University, State University of New York, Binghamton, NY, United States

Corresponding Author:
Kristin Pullyblank, MS
Center for Rural Community Health
Bassett Research Institute
Bassett Medical Center
One Atwell Road
Cooperstown, NY, 13326
United States
Phone: 1 6075476711
Email: kristin.pullyblank@bassett.org

Abstract

Background: In-person, evidence-based, peer-facilitated chronic disease self-management programs have been shown to be effective for individuals from a variety of backgrounds, including rural and minority populations and those with lower socioeconomic status. Based in social learning theory, these programs use group processes to help participants better manage their chronic disease symptoms and improve their quality of life. During the pandemic, these in-person programs were forced to rapidly transition to remote delivery platforms, and it was unclear whether doing so increased disparities within our rural population.

Objective: The objectives of this analysis were to ascertain self-management program enrollment and completion characteristics between 2 remote delivery platforms, as well as determine the individual level characteristics that drove enrollment and completion across delivery modes.

Methods: We analyzed enrollment and completion characteristics of 183 individuals who either enrolled in a self-management workshop delivered through a web-based videoconference (VC) system or through a traditional, audio-only conference call (CC) offered by our health care network between April and December 2020. Chi-square tests of association were used to describe the characteristics of and differences between groups. Logistic regression analysis was used to determine significant predictors of enrollment and completion.

Results: Those who enrolled in the VC platform were significantly likelier to be younger and college educated than those who enrolled in the CC platform. Those who completed a program, regardless of delivery mode, were likelier to be older and college educated than those who did not complete a program. Multivariate analyses indicated that of those enrolled in the CC platform, completers were likelier to not be enrolled in Medicaid. Among those enrolled in the VC platform, completers were older, college graduates, female, and likelier to have reported poorer health than those who did not complete the program.

Conclusions: The transition of self-management programs to remote delivery modes, particularly to those that rely on VC platforms, revealed that certain demographic groups may no longer be able or willing to access the service. Efforts need to be made to increase engagement in remote self-management workshops. In addition, equivalent quality services that do not rely on a digital platform must continue to be offered in order to promote health equity.

(JMIR Form Res 2022;6(11):e38357) doi:10.2196/38357

KEYWORDS

self-management programs; self management; internet-based intervention; health promotion; COVID-19; health equity; socioeconomic status; remote healthcare; health delivery; virtual care; remote care; remote delivery; videoconference; videoconferencing; adherence; attrition; completion; virtual health

Introduction

The COVID-19 pandemic has resulted in skyrocketing utilization of digital platforms for health-related services, such as provider visits, support groups, and wellness classes [1-3]. Digital platforms provided a safer alternative to in-person meetings during the peak of the COVID-19 pandemic, and in some cases, improved access by facilitating more timely appointments or eliminating the need to travel [4-7]. However, in the rush to implement remote health-related services, we inadvertently risk increasing inequities to accessing care when we fail to consider the consumer groups who are unable or unwilling to use remote modes of delivery. When chronic disease self-management education (CDSME) programs switched to remote delivery modes in 2020, our team had the opportunity to explore how the delivery mechanism of the program affected participant engagement.

The evidence-based suite of chronic disease self-management programs, originally developed at Stanford University and now licensed through the Self-Management Resource Center, is a nationally disseminated community-based intervention. The program process and content are based on Bandura's [8] social learning theory and focus on improving self-efficacy so that people are better able to manage their chronic conditions.

Typically, 2 trained peer leaders facilitate the small group workshops, which consist of 6 weekly 2.5-hour sessions. Even though leaders follow a scripted curriculum, the programs are designed to be highly participatory with group participants tailoring the content through brainstorming, pairing and sharing, and problem-solving activities (eg, if a participant states that he/she has had trouble communicating effectively with a health care provider and would like some ideas, the group will embark on a structured problem solve). In the United States, workshop delivery is usually funded by community-based organizations (eg, area agency on aging and senior center) or governmental agency (eg, local health department) and, thus, are generally offered at no cost to the participant.

The suite of CDSME workshops, which includes the Chronic Disease Self-Management Program, the Diabetes Self-Management Program, and the Chronic Pain Self-Management Program among others, has demonstrated effectiveness at improving a variety of health-related outcomes such as increased medication adherence, decreased depression, improved diabetes self-efficacy, improved pain self-efficacy, better communication with the health care team, and a reduction in hemoglobin A1c levels [9-13]. Additionally, participation in the program may reduce overall cost of care for those with chronic conditions [14,15].

The in-person workshops have been shown to be effective for individuals from a variety of backgrounds [11,16-21] including rural and minority populations and those with lower socioeconomic status (SES). However, recruitment of these populations is often challenging [22-24] owing to barriers associated with transportation, work and family obligations, and cultural beliefs [18,25,26].

Prior to the COVID-19 pandemic, our health care system had offered the Chronic Disease Self-Management Program, the Diabetes Self-Management Program, and the Chronic Pain Self-Management Program in small, in-person groups. In response to the pandemic, the Self-Management Resource Center developed two additional delivery modes: digital "in-person" sessions via videoconference (VC) using a web-based platform and a mailed toolkit with weekly small group telephone conference calls (CC). The VC format is similar in time and attention to the in-person program—that is, a 2.5-hour meeting once a week for 6 weeks with 2 peer leaders facilitating the workshop. The small group CC with a mailed toolkit is facilitated by a peer leader for 30-60 minutes per week for 6 weeks with the expectation that participants will review tool kit resources independently.

This analysis has 2 purposes. First, we wanted to ascertain enrollment and completion characteristics among the different remote delivery modes of CDSME workshops offered in our health care network's region during the first year of the COVID-19 pandemic. Second, we wanted to determine the individual level characteristics that drove enrollment and completion across the different delivery modes. The findings from these analyses may identify groups that are at risk for inequitable access to remote health education services during and after the COVID-19 pandemic. An evidence-based program that does not reach the target population or has high rates of attrition jeopardizes the overall impact of the intervention and can potentially exacerbate existing health disparities.

Methods

Settings and Participants

The remotely delivered workshops were offered throughout an 8-county region in upstate New York. The region is largely rural, with an average population density of 55 residents per square mile (ie, 21 residents/km^2), compared to 238 residents per square mile (ie, 92 residents/km^2) in New York, excluding New York City [27]. Similar to other rural areas in the northeast, the population is predominantly White, non-Hispanic. Nearly 30% of the general population is at least 60 years old.

Participants who enrolled in 1 of 2 workshop delivery modes (VC or CC) between April and December 2020 and consented to have their data deidentified and shared for research purposes are included in this analysis. To be considered "enrolled," an individual had to complete the baseline survey and register and receive materials for the workshop.

Ethical Considerations

The implementation, delivery, and evaluation of the CDSME workshops within our health care system's service region was originally developed as a quality improvement project and was determined to be exempt from ongoing oversight by the Mary Imogene Bassett Hospital's institutional review board. Participants provided written consent to attend the workshop and to have their deidentified data collected for research and evaluation purposes. Participants received a US $30 gift card to a local grocery store as compensation for completing data collection activities. Individuals who did not consent to have

their data collected and hence not included in the analyses, were still able to participate in the workshop.

Measures

Sociodemographic data (age, gender, education, marriage status, and insurance status) were collected either through paper or electronic questionnaires at baseline. In addition to sociodemographic questions, participants were asked to complete a single-item, self-rated health question [28]. All variables were dichotomized after data collection. Process measures included enrollment information (including workshop delivery mode), attendance, completion status, and reasons for not being able to attend all sessions of the workshop.

Statistical Analyses

The analyses described below were used to answer the following questions: what are the characteristics of participants enrolled in each mode of remote workshop? What are the differences between completers and noncompleters across workshop delivery modes? What are the differences between completers and noncompleters within each delivery mode?

Descriptive statistics were used to characterize the entire sample, as well as for enrollment in each delivery mode. Bivariate analyses were conducted using chi-square tests of association to determine which characteristics were significant between groups for enrollment, an overall comparison of completers versus noncompleters, and which factors influence the likelihood of completion within each delivery mode. Based on these findings, multivariate analyses using logistic regression were conducted to explore which characteristics were independent predictors of enrolling in or completing a particular delivery mode. In all analyses, noncompleters included individuals who failed to show up for the workshop (eg, “no shows”) as well as those who showed up for fewer than 4 sessions.

Data were analyzed with SPSS (version 27; IBM Corp). Four outliers were identified in SPSS and were removed from the data set prior to analyses.

Results

Enrollment

Between April and December 2020, a total of 183 individuals who consented to share their deidentified data enrolled in a VC or small group CC self-management program workshop (2 additional individuals enrolled in a workshop but did not consent to share their data). The majority of the participants across workshops were older (mean age 58.6 years, median 60.0 years), female (n=143, 78.1%), married or partnered (n=93, 52.5%), and self-rated their health as good or better (n=155, 85.6%). Participant demographic characteristics are summarized in Table 1. There were significant differences in age and education level among users of the 2 different delivery modes at enrollment. Those who enrolled in the VC workshop were 3.44 times more likely to be aged 60 years or younger and 2.24 times more likely to be a college graduate (Table 1). Gender, marriage status, self-reported health, and Medicaid status were not significant indicators of program mode enrollment (*P*=.19, .76, .45, and .11, respectively).

In the multivariate analysis, age and education remained significant predictors of enrollment. The logistic regression model was significant overall (χ^2_6=22.3, *P*<.001). Those who were 60 years old or younger were significantly likelier to enroll in VC than those older than 60 years (odds ratio [OR] 3.16, 95% CI 6.26; *P*<.001). College graduates were also significantly likelier to enroll in VC than those without college education (OR 2.02, 95% CI 1.00-4.09; *P*=.05).

Table 1. Demographic characteristics of the sample (N=183) by program delivery mode (videoconference [VC] or conference call [CC]) and differences between groups.

Characteristics	Total, n (%)	VC (n=99), n (%)	CC (n=84), n (%)	Chi-square (*df*)	Odds ratio (95% CI)	*P* value
Age (years)				16.3 (1)	3.44 (1.87-6.33)	<.001
≤60	95 (51.9)	65 (65.7)	30 (35.7)			
>60	88 (48.1)	34 (34.3)	54 (64.3)			
Gender				1.8 (1)	1.64 (0.79-3.43)	.19
Male	38 (21)	24 (24.7)	14 (16.7)			
Female	143 (79)	73 (75.3)	70 (83.3)			
Marital status				0.1 (1)	1.10 (0.61-1.99)	.76
Married	93 (52.5)	52 (53.6)	41 (51.2)			
Not married	84 (47.5)	45 (46.4)	39 (48.8)			
Education level				6.4 (1)	2.24 (1.19-4.19)	.01
College graduate	67 (37.2)	45 (33.3)	22 (27.2)			
Less than college education	113 (62.8)	54 (66.7)	59 (72.8)			
Self-reported health status				0.6 (1)	1.39 (0.59-3.25)	.45
Fair or poor health	26 (14.4)	16 (16.2)	10 (12.2)			
At least good health	155 (85.6)	83 (83.8)	72 (87.8)			
Medicaid status				2.7 (1)	0.58 (0.30-1.12)	.11
Medicaid	53 (30.3)	62 (64.6)	60 (75.9)			
No Medicaid	122 (69.7)	34 (35.4)	19 (24.1)			

Completion

There were significant differences between program completers and noncompleters in terms of age, education, and self-reported health (Table 2). Completers likelier to be older than 60 years (OR 2.76, 95% CI 1.51-5.02), have a college education (OR 2.50, 95% CI 1.34-4.68), and report poorer general health (OR 4.60, 95% CI 1.65-12.81) than noncompleters.

These differences remained significant in the multivariate analyses. The logistic regression model was significant overall (χ^2_6=22.5, *P*<.001). Workshop completers were likelier be older than 60 years (OR 3.10, 95% CI 1.53-6.31; *P*=.002), be college graduates (OR 2.52, 95% CI 1.22-5.21; *P*=.01), and report poorer health (OR 4.26, CI 1.30-13.99; *P*=.02). In addition, gender emerged as a significant predictor, with females being likelier to complete a workshop than males (OR 2.40, 95% CI 1.03-5.62; *P*=.04).

When the data were stratified by workshop type (CC and VC), the CC logistic regression model (χ^2_6=16.7, *P*=.01) revealed that Medicaid status was the only independent predictor of completing a CC workshop. Those not enrolled in Medicaid were likelier to complete the program than those enrolled in Medicaid (OR 4.17, 95% CI 1.16-15.07). For the VC logistic regression model (χ^2_6=24.8, *P*<.001), age, gender, education, and self-reported health remained independent predictors of program completion, in that, those who were older than 60 years (OR 5.56, 95% CI 1.81-17.10), female (OR 4.24, 95% CI 1.16-15.43), college graduates (OR 2.77, 95% CI 1.01-7.56), and those who reported poor or fair health (OR 7.00, 95% CI 1.44, 34.07) were likelier to complete the workshop (Table 3).

Table 2. Differences between program completers and noncompleters.

Characteristics	Completers (n=95), n (%)	Noncompleters (n=88), n (%)	Chi-square (*df*)	Odds ratio (95% CI)	*P* value
Age (years)			11.2 (1)	2.76 (1.51-5.02)	<.001
>60	57 (60.0)	31 (35.2)			
≤60	38 (40.0)	57 (64.8)			
Gender			1.2 (1)	0.71 (0.34-1.45)	.28
Male	17 (17.9)	21 (24.4)			
Female	78 (82.1)	65 (75.6)			
Marital status			0.6 (1)	1.27 (0.70-2.29)	.43
Married	52 (55.3)	41 (49.4)			
Not married	42 (44.7)	42 (50.6)			
Education level			8.4 (1)	2.50 (1.34-4.68)	.004
College graduate	44 (47.3)	23 (26.4)			
Less than college education	49 (52.7)	64 (73.6)			
Health status			9.7 (1)	4.60 (1.65-12.81)	.002
Fair or poor health	21 (22.1)	5 (5.8)			
At least good health	74 (77.9)	81 (94.2)			
Medicaid status			3.4 (1)	1.83 (0.95-3.53)	.07
No Medicaid	69 (75.8)	53 (63.1)			
Medicaid	22 (24.2)	31 (36.9)			

Table 3. Logistic regression analysis for predictors of completion split by delivery mode.

Predictors	*B* (SE)	Wald chi-square test (df)	*P* value	Odds ratio (95% CI)
Workshop delivery mode: conference call				
Age>60 years	0.617 (0.563)	1.202 (1)	.27	1.85 (0.62-5.59)
Female gender	0.271 (0.673)	0.162 (1)	.69	1.31 (0.35-4.90)
Not married or partnered	0.415 (0.570)	0.530 (1)	.47	1.52 (0.50-4.63)
College graduate	1.258 (0.672)	3.446 (1)	.06	3.48 (0.93-13.00)
Self-reported less than good health	1.484 (1.137)	1.703 (1)	.19	4.41 (0.48-41.00)
Not enrolled in Medicaid	1.429 (0.655)	4.754 (1)	.03	4.17 (1.16-15.07)
Workshop delivery mode: videoconference				
Age>60 years	1.716 (0.573)	8.959 (1)	.003	5.56 (1.81-17.10)
Female gender	1.443 (0.659)	4.791 (1)	.03	4.24 (1.16-15.43)
Not married or partnered	–0.537 (0.521)	1.066 (1)	.30	0.58 (0.21-1.62)
College graduate	1.019 (0.515)	3.920 (1)	.048	2.77 (1.01-7.59)
Self-reported less than good health	1.945 (0.808)	5.700 (1)	.02	6.99 (1.44-34.06)
Not enrolled in Medicaid	–0.538 (0.551)	0.953 (1)	.33	0.58 (0.20-1.72)

Discussion

Principal Findings

The purpose of these analyses was to gain a better understanding of the characteristics of those who enroll in and complete novel remote delivery modes of CDSME workshops that were developed in response to the onset of the COVID-19 pandemic. Our findings suggest that younger individuals were likelier to enroll in the VC platform than older individuals. However, older individuals were likelier to successfully complete any self-management workshop, regardless of delivery mode. When controlling for all variables within the multivariate models, not being enrolled in Medicaid significantly predicted CC workshop completion, while being older, female, a college graduate and rating one's health as fair or poor were all independent predictors for completing a VC workshop.

The lower rate of digital technology use among older populations has been well documented [29], and the disparity has been highlighted during the COVID-19 pandemic [30,31]. However, of those who did enroll, those who were older than 60 years were likelier to complete the program, suggesting that once engaged with the remote delivery platform, older adults are able and willing to complete the workshops. Other researchers have demonstrated that self-management workshop completion rates are much higher among older cohorts [20]. Thus, while the VC delivery mode may initially engage those who are younger, we found it difficult to retain younger individuals in this program. A possible explanation for this finding is that younger and middle-aged adults are more likely engaged in the workforce and with family obligations and are hence unable to commit to a 6-week program, even if the program is offered in the evening [7,32]. This could have been particularly true during the first months of the pandemic when many parents suddenly found themselves essentially home-schooling their children. Another possible reason is that participants may be likelier to drop out if they have difficulty identifying with others in the group (eg, most participants are female, older, and retired) [33,34].

Self-management programs have typically struggled with recruiting and retaining men [35]. Our analyses are in agreement with nationwide studies, in that, nearly 80% of the participants were female. Women were also likelier to complete the program overall. However, it is interesting to note that there was no difference in completion rates between genders for the CC delivery mode. A possible explanation is that the CC mode allows for some level of anonymity; therefore, men may not feel they have to conform as much to masculine help-seeking behaviors. The CC group size is also smaller, and fewer participants per session have been shown to increase retention rates among men enrolled in the self-management programs [35].

The analyses revealed that education level was a significant factor for enrolling in and completing the VC workshop. This finding raises questions about the utility of the VC workshop for individuals who do not have a college education, particularly if one of the delivery modes demonstrates superior health-related outcomes in future randomized controlled trials. The availability of a non–internet-based remote option for those who either did not have internet access or who prefer not to use the internet is important from a health equity standpoint. However, at this time, it is unclear if the CC delivery mode has similar outcomes as the VC platform, particularly since it is shorter in time and attention.

Education level and Medicaid status are often used as proxies for socioeconomic status (SES) [36]. The digital divide is well documented with those of lower SES having less access to and lower usage of web-based services [37,38]; therefore, the abovementioned finding is not surprising. While a nationwide evaluation of the in-person CDSME workshops did not examine differences in education between completers and noncompleters, a Canadian study by McGowan found that education level was not a significant factor for completing in-person self-management workshops [13]. Our in-person workshop findings also indicated no significant difference in education level between completers and noncompleters [39,40].

A mixed methods investigation on access to, use of, and benefits from digital health services also found that being older and having less education affected access to digital health services, as well as living in a lower-SES area. Having less education was also associated with less use of digital health services [41]. However, the largest factors that influenced use were trust in digital health services, eHealth literacy, and confidence in using them. While further research is needed to understand the determinants of trusting digital health services, it is clear that a trusted provider's recommendation to utilize the service may increase the level of trust in digital services, especially among underserved populations [41].

In this analysis, not being enrolled in Medicaid predicted completion of the program through the CC delivery mode, but did not predict program completion through the VC delivery mode or enrollment through either program delivery mode, which was an unexpected finding. Recently, investigators have attempted to understand the relationship between SES and engagement in self-management programs. A review by Hardman et al [24] on the moderating effect of SES on self-management interventions found that there was some influence of SES on attrition rates, but the lack of high-quality research made it difficult to draw conclusions. Hardman et al [24] recognized the heterogeneity of low-SES groups, and that appropriate interventions or recruitment methods depend largely on the context of that group.

A scoping review of scoping reviews also illustrated the complexities of engaging individuals in web-based health services, recognizing how an individual's culture and perceived effectiveness of telehealth technologies is intertwined not only with digital literacy but also the social and structural determinants of health [42]. It is likely that there are other factors that we did not measure, which are stronger predictors of completion of the VC delivery mode (eg, digital literacy and trust in technology). From a health equity standpoint, we recommend further research to explore the complex relationships among education, insurance status, SES, the social and structural determinants of health, and remote workshop completion status.

Finally, it is important to recognize the effect that self-reported health status may have on the odds of workshop completion. The overwhelming majority of participants rated their health as good or better; nonetheless, over 80% of those reporting fair or poor health completed the program, compared to 47% of those reporting good or better health. While our sample size was small, leading to very large CIs, this finding does illustrate 2 important points. First, it was difficult to recruit those who rate their health as fair or poor. This finding is in line with that of other research indicating that individuals with poorer self-reported health perceive fewer benefits from digital health services [43]. However, our data show that once these individuals are engaged, they are likely to complete the program. Therefore, exploring ways to engage those who are facing poor health is critical. To do this, program implementers must acknowledge and account for the many social determinants of

health, including cultural and environmental determinants, which can impact health outcomes [24,41,43,44].

Limitations

There are several limitations that must be acknowledged in this analysis. First, the analysis was a service evaluation and was not based on an experimental design. Therefore, it is impossible to determine any causal relationships. Second, external constraints on program delivery dictated how often each mode of workshop was offered, which could have added additional bias into our findings. We also had a very small sample of individuals who reported less than good health, which may affect the accuracy of those findings. In addition, our sample population was from one particular region of New York, and the results are not generalizable to other areas. Finally, the evaluation took place during the first 10 months of the pandemic, which likely influenced health-seeking behaviors. As individuals have settled into the peripandemic era, remote workshop usage patterns may have shifted.

Conclusions

Despite these limitations, our analysis demonstrates strong signals that certain demographic groups are less likely to initially engage with web-based CDSME workshops, particularly those who are older and less educated. In addition, completion of web-based workshops is likelier for those who have a college education, as well as those who are older and female. The traditional CC workshop appeared more accessible to participants. As health care resources rapidly become more digitized, health educators and practitioners must remain aware of this potential disparity, and future work needs to focus on understanding the nuances underlying these disparities and how we can effectively work with our patients and communities to overcome them. Remote program offerings, including those that are independent or asynchronous, can remove barriers that prevent individuals with chronic disease from engaging with in-person self-management workshops, such as transportation issues, childcare concerns, fatigue, and pain [7,45]. However, equivalent, high-quality service that does not rely on a digital platform must continue to be offered to promote health care equity.

Acknowledgments

This work was supported in part by the US Department of Health and Human Services Healthcare Resources and Services Administration (D04RH31785-01-00). The authors would like to acknowledge the Living Well peer leaders and community organizations who made the implementation of the self-management programs possible.

Data Availability

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

Conflicts of Interest

None declared.

References

1. Mackenzie SC, Cumming KM, Garrell D, Brodie D, Wilson L, Mehar S, et al. Massive open online course for type 2 diabetes self-management: adapting education in the COVID-19 era. BMJ Innov 2020 Nov 16;7(1):141-147. [doi: 10.1136/bmjinnov-2020-000526]
2. McElroy JA, Day TM, Becevic M. The influence of telehealth for better health across communities. Prev Chronic Dis 2020 Jul 16;17:E64 [FREE Full text] [doi: 10.5888/pcd17.200254] [Medline: 32678060]
3. Smithson R, Roche E, Wicker C. Virtual models of chronic disease management: lessons from the experiences of virtual care during the COVID-19 response. Aust. Health Review 2021;45(3):311. [doi: 10.1071/ah20190]
4. Zanaboni P, Fagerlund AJ. Patients' use and experiences with e-consultation and other digital health services with their general practitioner in Norway: results from an online survey. BMJ Open 2020 Jun 17;10(6):e034773 [FREE Full text] [doi: 10.1136/bmjopen-2019-034773] [Medline: 32554721]
5. Zolot J. Virtual care takes over during COVID-19 pandemic. Am J Nurs 2021;121(5):15. [doi: 10.1097/01.NAJ.0000751040.50837.c0]
6. LeBlanc M, Petrie S, Paskaran S, Carson D, Peters P. Patient and provider perspectives on eHealth interventions in Canada and Australia: a scoping review. RRH 2020 Sep 19. [doi: 10.22605/rrh5754]
7. Allen JO, Concha JB, Mejía Ruiz MJ, Rapp A, Montgomery J, Smith J, et al. Engaging underserved community members in diabetes self-management: evidence from the YMCA of Greater Richmond diabetes control program. Diabetes Educ 2020 Apr 04;46(2):169-180 [FREE Full text] [doi: 10.1177/0145721720907059] [Medline: 32129157]
8. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977;84(2):191-215. [doi: 10.1037/0033-295x.84.2.191]
9. Lorig K, Ritter PL, Turner RM, English K, Laurent DD, Greenberg J. A diabetes self-management program: 12-month outcome sustainability from a nonreinforced pragmatic trial. J Med Internet Res 2016 Dec 15;18(12):e322 [FREE Full text] [doi: 10.2196/jmir.6484] [Medline: 27979790]

10. Lorig KR, Sobel DS, Ritter PL, Laurent D, Hobbs M. Effect of a self-management program on patients with chronic disease. Eff Clin Pract 2001;4(6):256-262. [Medline: 11769298]
11. Ory M, Ahn S, Jiang L, Smith ML, Ritter PL, Whitelaw N, et al. Successes of a national study of the Chronic Disease Self-Management Program: meeting the triple aim of health care reform. Med Care 2013 Nov;51(11):992-998 [FREE Full text] [doi: 10.1097/MLR.0b013e3182a95dd1] [Medline: 24113813]
12. Hevey D, Wilson O'Raghallaigh J, O'Doherty V, Lonergan K, Heffernan M, Lunt V, et al. Pre-post effectiveness evaluation of Chronic Disease Self-Management Program (CDSMP) participation on health, well-being and health service utilization. Chronic Illn 2020 Jun 08;16(2):146-158. [doi: 10.1177/1742395318792063] [Medline: 30089405]
13. McGowan P. The relative effectiveness of self-management programs for type 2 diabetes. Can J Diabetes 2015 Oct;39(5):411-419. [doi: 10.1016/j.jcjd.2015.04.005] [Medline: 26165582]
14. Ahn S, Basu R, Smith ML, Jiang L, Lorig K, Whitelaw N, et al. The impact of chronic disease self-management programs: healthcare savings through a community-based intervention. BMC Public Health 2013 Dec 06;13(1):1141 [FREE Full text] [doi: 10.1186/1471-2458-13-1141] [Medline: 24314032]
15. Turner RM, Ma Q, Lorig K, Greenberg J, DeVries AR. Evaluation of a Diabetes Self-Management Program: claims analysis on comorbid illnesses, health care utilization, and cost. J Med Internet Res 2018 Jun 22;20(6):e207 [FREE Full text] [doi: 10.2196/jmir.9225]
16. Horrell LN, Kneipp SM, Ahn S, Towne SD, Mingo CA, Ory MG, et al. Chronic disease self-management education courses: utilization by low-income, middle-aged participants. Int J Equity Health 2017 Jun 27;16(1):114 [FREE Full text] [doi: 10.1186/s12939-017-0604-0] [Medline: 28655319]
17. Goldfinger JZ, Kronish IM, Fei K, Graciani A, Rosenfeld P, Lorig K, et al. Peer education for secondary stroke prevention in inner-city minorities: design and methods of the prevent recurrence of all inner-city strokes through education randomized controlled trial. Contemp Clin Trials 2012 Sep;33(5):1065-1073 [FREE Full text] [doi: 10.1016/j.cct.2012.06.003] [Medline: 22710563]
18. Bobitt J, Aguayo L, Payne L, Jansen T, Schwingel A. Geographic and social factors associated with chronic disease self-management program participation: going the "Extra-Mile" for disease prevention. Prev Chronic Dis 2019 Mar 07;16:E25 [FREE Full text] [doi: 10.5888/pcd16.180385] [Medline: 30844360]
19. Horowitz CR, Eckhardt S, Talavera S, Goytia C, Lorig K. Effectively translating diabetes prevention: a successful model in a historically underserved community. Transl Behav Med 2011 Sep 16;1(3):443-452 [FREE Full text] [doi: 10.1007/s13142-011-0067-6] [Medline: 22189897]
20. Erdem E, Korda H. Self-management program participation by older adults with diabetes. Fam Community Health 2014;37(2):134-146. [doi: 10.1097/fch.0000000000000025]
21. Ory MG, Smith ML. Research, practice, and policy perspectives on evidence-based programing for older adults. Front Public Health 2015 Apr 27;3:136 [FREE Full text] [doi: 10.3389/fpubh.2015.00136] [Medline: 25973417]
22. Towne SD, Smith ML, Ahn S, Altpeter M, Belza B, Kulinski KP, et al. National dissemination of multiple evidence-based disease prevention programs: reach to vulnerable older adults. Front Public Health 2014 Apr 27;2:156 [FREE Full text] [doi: 10.3389/fpubh.2014.00156] [Medline: 25964901]
23. Smith ML, Ory MG, Jiang L, Lorig K, Kulinski KP, Ahn S. Workshop characteristics related to chronic disease self-management education program attendance. Front Public Health 2015 Apr 27;3:19 [FREE Full text] [doi: 10.3389/fpubh.2015.00019] [Medline: 25964943]
24. Hardman R, Begg S, Spelten E. What impact do chronic disease self-management support interventions have on health inequity gaps related to socioeconomic status: a systematic review. BMC Health Serv Res 2020 Feb 27;20(1):150 [FREE Full text] [doi: 10.1186/s12913-020-5010-4] [Medline: 32106889]
25. Smith M, Towne Jr S, Bergeron C, Zhang D, McCord C, Mier N, et al. Geographic disparities associated with travel to medical care and attendance in programs to prevent/manage chronic illness among middle-aged and older adults in Texas. Rural Remote Health 2019 Nov 9. [doi: 10.22605/rrh5147]
26. Lee H, McDonagh M. Updating the rural nursing theory base. In: Rural Nursing: Concepts, Theory, and Practice (5th edition). New York, NY: Springer; 2018.
27. Vital Statistics of New York State. New York State Department of Health. URL: https://www.health.ny.gov/statistics/vital_statistics/ [accessed 2022-11-22]
28. Idler EL, Angel RJ. Self-rated health and mortality in the NHANES-I epidemiologic follow-up study. Am J Public Health 1990 Apr;80(4):446-452. [doi: 10.2105/ajph.80.4.446] [Medline: 2316767]
29. Levine DM, Lipsitz SR, Linder JA. Trends in seniors' use of digital health technology in the United States, 2011-2014. JAMA 2016 Aug 02;316(5):538-540. [doi: 10.1001/jama.2016.9124] [Medline: 27483069]
30. Eberly LA, Kallan MJ, Julien HM, Haynes N, Khatana SAM, Nathan AS, et al. Patient characteristics associated with telemedicine access for primary and specialty ambulatory care during the COVID-19 pandemic. JAMA Netw Open 2020 Dec 01;3(12):e2031640 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.31640] [Medline: 33372974]
31. Kalicki AV, Moody KA, Franzosa E, Gliatto PM, Ornstein KA. Barriers to telehealth access among homebound older adults. J Am Geriatr Soc 2021 Sep 13;69(9):2404-2411 [FREE Full text] [doi: 10.1111/jgs.17163] [Medline: 33848360]

32. Gucciardi E, Demelo M, Offenheim A, Stewart DE. Factors contributing to attrition behavior in diabetes self-management programs: a mixed method approach. BMC Health Serv Res 2008 Feb 04;8(1):33 [FREE Full text] [doi: 10.1186/1472-6963-8-33] [Medline: 18248673]
33. Verevkina N, Shi Y, Fuentes-Caceres VA, Scanlon DP. Attrition in chronic disease self-management programs and self-efficacy at enrollment. Health Educ Behav 2014 Dec 25;41(6):590-598. [doi: 10.1177/1090198114529590] [Medline: 24771128]
34. Borek AJ, Abraham C. How do small groups promote behaviour change? An integrative conceptual review of explanatory mechanisms. Appl Psychol Health Well Being 2018 Mar;10(1):30-61. [doi: 10.1111/aphw.12120] [Medline: 29446250]
35. Smith ML, Bergeron CD, Ahn S, Towne SD, Mingo CA, Robinson KT, et al. Engaging the underrepresented sex: male participation in chronic disease self-management education (CDSME) programs. Am J Mens Health 2018 Jul;12(4):935-943 [FREE Full text] [doi: 10.1177/1557988317750943] [Medline: 29355070]
36. Casey JA, Pollak J, Glymour MM, Mayeda ER, Hirsch AG, Schwartz BS. Measures of SES for electronic health record-based research. Am J Prev Med 2018 Mar;54(3):430-439 [FREE Full text] [doi: 10.1016/j.amepre.2017.10.004] [Medline: 29241724]
37. Shaw J, Brewer LC, Veinot T. Recommendations for health equity and virtual care arising from the COVID-19 pandemic: narrative review. JMIR Form Res 2021 Apr 05;5(4):e23233 [FREE Full text] [doi: 10.2196/23233] [Medline: 33739931]
38. Crawford A, Serhal E. Digital health equity and COVID-19: the innovation curve cannot reinforce the social gradient of health. J Med Internet Res 2020 Jun 02;22(6):e19361 [FREE Full text] [doi: 10.2196/19361] [Medline: 32452816]
39. Pullyblank K, Brunner W, Wyckoff L, Krupa N, Scribani M, Strogatz D. Implementation of evidence-based disease self-management programs in a rural region: leveraging and linking community and health care system assets. Health Educ Behav 2022 Feb 18:10901981221078516. [doi: 10.1177/10901981221078516] [Medline: 35179055]
40. Pullyblank K, Brunner W, Scribani M, Krupa N, Wyckoff L, Strogatz D. Evaluation of a peer led chronic pain self-management program in a rural population. J Prim Care Community Health 2022 Sep 12;13:21501319221121464 [FREE Full text] [doi: 10.1177/21501319221121464] [Medline: 36112865]
41. Foley K, Freeman T, Ward P, Lawler A, Osborne R, Fisher M. Exploring access to, use of and benefits from population-oriented digital health services in Australia. Health Promot Int 2021 Aug 30;36(4):1105-1115. [doi: 10.1093/heapro/daaa145] [Medline: 33367568]
42. Budhwani S, Fujioka J, Thomas-Jacques T, De Vera K, Challa P, De Silva R, et al. Challenges and strategies for promoting health equity in virtual care: findings and policy directions from a scoping review of reviews. J Am Med Inform Assoc 2022 Apr 13;29(5):990-999. [doi: 10.1093/jamia/ocac022] [Medline: 35187571]
43. Heponiemi T, Jormanainen V, Leemann L, Manderbacka K, Aalto A, Hyppönen H. Digital divide in perceived benefits of online health care and social welfare services: national cross-sectional survey study. J Med Internet Res 2020 Jul 07;22(7):e17616 [FREE Full text] [doi: 10.2196/17616] [Medline: 32673218]
44. Braveman P, Gottlieb L. The social determinants of health: it's time to consider the causes of the causes. Public Health Rep 2014 Jan 01;129(1_suppl2):19-31. [doi: 10.1177/00333549141291s206]
45. Christensen NI, Drejer S, Burns K, Lundstrøm SL, Hempler NF. A qualitative exploration of facilitators and barriers for diabetes self-management behaviors among persons with type 2 diabetes from a socially disadvantaged area. Patient Prefer Adherence 2020;14:569-580 [FREE Full text] [doi: 10.2147/PPA.S237631] [Medline: 32210542]

Abbreviations

CC: conference call
CDSME: chronic disease self-management education
OR: odds ratio
SES: socioeconomic status
VC: videoconference

Edited by A Mavragani; submitted 04.04.22; peer-reviewed by F Balli, J Wherton; comments to author 11.10.22; revised version received 01.11.22; accepted 21.11.22; published 30.11.22.

Please cite as:
Pullyblank K, Atav S
Enrollment and Completion Characteristics for Novel Remote Delivery Modes of the Self-management Programs During the COVID-19 Pandemic: Exploratory Analysis
JMIR Form Res 2022;6(11):e38357
URL: https://formative.jmir.org/2022/11/e38357
doi:10.2196/38357
PMID:36413622

Original Paper

User Engagement Within an Online Peer Support Community (Depression Connect) and Recovery-Related Changes in Empowerment: Longitudinal User Survey

Dorien Smit[1,2], MSc; Janna N Vrijsen[1,3], PhD; Theo Broekman[4], MSc; Bart Groeneweg[5], PhD; Jan Spijker[1,2], MD, PhD

[1]Pro Persona Mental Health Care, Pro Persona Research, Depression Expertise Centre, Nijmegen, Netherlands

[2]Behavioural Science Institute, Radboud University Nijmegen, Nijmegen, Netherlands

[3]Department of Psychiatry, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Centre, Nijmegen, Netherlands

[4]Bureau Bêta, Nijmegen, Netherlands

[5]Dutch Depression (Patient) Association, Amersfoort, Netherlands

Corresponding Author:
Dorien Smit, MSc
Pro Persona Mental Health Care
Pro Persona Research
Depression Expertise Centre
Nijmeegsebaan 61
Nijmegen, 6525 DX
Netherlands
Phone: 31 647074551
Email: d.smit@propersona.nl

Abstract

Background: The chronic nature of depression and limited availability of evidence-based treatments emphasize the need for complementary recovery-oriented services, such as peer support interventions (PSIs). Peer support is associated with positive effects on clinical and personal recovery from mental illness, but little is known about the processes of engagement that foster change, and studies targeting individuals with depression specifically are limited.

Objective: This study aimed to evaluate whether the level of user engagement, assessed on several dimensions, in an online peer support community for individuals with depression promotes empowerment and the use of self-management strategies and reduces symptom severity and disability.

Methods: In a longitudinal survey conducted from June 2019 to September 2020, we analyzed the data of the users of Depression Connect (DC), an online peer support community hosted by the Dutch Patient Association for Depression and the Pro Persona Mental Health Care institute, on measures of empowerment, self-management, depression, and disability. Of the 301 respondents, 49 (16.3%) respondents completed the survey again after 3 months and 74 (24.6%) respondents, after 6 months. Analysis of 3 parameters (ie, total time spent on the platform, number of page views, and number of posts) derived from their data logs yielded 4 engagement profiles. Linear mixed models were fitted to determine whether the outcomes had significantly changed over time and differed for the various profiles.

Results: Baseline engagement with the online peer support community was “very low” (177/301, 58.8%) or “low” (87/301, 28.9%) for most of the participants, with few showing “medium” (30/301, 9.9%) or “high” engagement patterns (7/301, 2.3%), while user profiles did not differ in demographic and clinical characteristics. Empowerment, self-management, depressive symptoms, and disability improved over time, but none were associated with the intensity or nature of user engagement.

Conclusions: With most DC members showing very low to low engagement and only a few being identified as high-engaged users, it is likely that this flexibility in use frequency is what provides value to online PSI users. In other more formal supportive environments for depression, a certain level of engagement is predetermined either by their organizational or by their societal context; at DC, users can adapt the intensity and nature of their engagement to their current needs on their personal road to recovery. This study added to the current knowledge base on user engagement for PSIs because previous studies targeting depression with an online format focused on active users, precluding passive and flexible engagement. Future studies should

explore the content and quality of the interactions in online PSIs to identify optimal user engagement as a function of current, self-reported clinical parameters and reasons to engage in the PSI.

(JMIR Form Res 2022;6(11):e39912) doi:10.2196/39912

KEYWORDS

depression; online peer support community; internet support group; experiential knowledge; self-management; empowerment; user engagement; longitudinal user survey

Introduction

Peer Support for Recovery in Depression

It is estimated that approximately 280 to 320 million people worldwide are coping with depression [1-3]. However, the availability of evidence-based mental health care, such as psychotherapy and psychopharmacology [4], is insufficient owing to high costs and a lack of skilled clinicians [5]. Moreover, persistent symptoms and high recurrence rates reported underline the chronic nature of the illness [6]. The fact that many individuals live long-term with (recurrent and persistent) depression emphasizes the need for recovery-oriented services that focus on emotional support and resilience rather than on symptom reduction [7]. Peer support interventions (PSIs) could provide such a source of support on the longer road to recovery [8], complementing professional treatment [9,10] for depression [11-13]. In particular, online PSIs meet the need for accessible and low-cost interventions [8], and offering the possibility of anonymous engagement helps circumvent the stigmatization associated with depression [14].

Principally, peer support entails giving and receiving help by exchanging personal experiences [15], where the central themes are “respect, shared responsibility, and mutual agreement of what is helpful” [16]. However, owing to the great variety of intervention types, deployment across different user groups and service delivery settings, there are multiple definitions of peer support [17,18]. Considering this heterogeneity, it is difficult to systematically disentangle the principal benefits of these systems.

The Effectiveness of Peer Support

We recently conducted a comprehensive meta-analysis of 28 randomized controlled trials (RCTs) to assess the efficacy of PSIs across a wide range of mental disorders and intervention types. Compared with control conditions, the PSIs we reviewed were associated with modest but significant positive effects on clinical symptoms and personal recovery (eg, promoting hope [7]) in individuals with mental illness [19]. Specifically, for individuals with serious mental illnesses, including major depressive disorders, peer support was associated with superior outcomes across clinical, personal, and functional recovery variables (eg, quality of life and social support) relative to control conditions.

It needs to be noted that only a limited number of trials included in our meta-analysis focused on *online* PSIs for *depression*. Nevertheless, the results of few trials were promising. Specifically, findings of the RCT conducted by Griffiths et al [20] suggested that engaging in a moderated depression internet support group may be clinically effective (ie, reducing depressive symptoms) in the long term, with potential short-term improvements for empowerment as presented in a companion paper of Crisp et al [20,21] reporting on the same trial. In addition to this quantitative evidence from a single trial for a depression PSI, descriptive systematic reviews (ie, a narrative synthesis for the efficacy of PSIs, not including a meta-analysis that systematically assesses the results of previously conducted studies) emphasize the potential of online health-related PSIs in general [22-25] and that of those specifically addressing depression [11,26].

The results of a broad systematic review [27] may help us better understand how these positive outcomes in PSIs may develop. Winsper et al [27] identified four common processes fostering change in recovery across 309 studies on recovery-oriented interventions for mental illness: (1) providing information and skills, (2) promoting a working alliance, (3) role modeling for individual recovery, and (4) increasing choice and opportunities. These processes may best be initiated within nonstigmatized, recovery-focused contexts, such as peer support where psychosocial processes of sharing lived experiences, emotional honesty, strengths-focused social and practical support, and the helper-role are important processes for mental health recovery [28]. The results of our qualitative evaluation study for users of the online peer support community Depression Connect (DC) fit these processes (eg, sense of belonging, self-efficacy, and empowerment) [29].

User Engagement Within Online PSIs

It remains unclear which PSI engagement processes are associated with these changes. In particular, for online PSIs, a high level of user engagement is considered a crucial factor for recovery [24,30,31]. A systematic review of online health communities showed that several multidimensional factors are relevant when defining user engagement, such as metrics characterizing user networks (eg, the number of people a user has interacted with), content (eg, the nature of posts), and activity (eg, the number of posts and log-in times) [32]. Use of online PSIs is mainly operationalized in terms of frequency of use [32], where the dichotomy between “lurkers” (ie, passive users, generally a substantial group, whose use is mainly restricted to reading others’ posts) and “posters” (ie, active users, generally only 1% of users [33-37]) is widely used. To accurately reflect the larger group of passive users, we need a more nuanced characterization of their engagement [32,33]. Such parameters of inactive engagement are particularly relevant for PSIs for depression, as passive behavior is associated with the condition [38]. Therefore, we conducted a *qualitative* evaluation of DC users [29], our self-developed online peer support community for individuals struggling with depression [39]. Our qualitative analysis of user experiences of DC revealed

3 successive participation styles (ie, reading, posting, and responding) that individually and together coincided with an increased sense of belonging, emotional growth, self-efficacy, and empowerment [29]. In this *quantitative* evaluation of DC, we studied engagement patterns as a possible mechanism for recovery more closely by assessing multiple metrics to define engagement as comprehensively as possible [32]. For this study, based on user data logs for DC, we included 3 parameters to operationalize the intensity level of user engagement: the number of posts, the number of page views, and the total time spent on DC. Including both active and passive user modes, it is important to acknowledge that high user engagement was not limited to active users but included users who *posted* (a substantial number of) messages on the platform. In user data logs for DC, it was not possible to distinguish the 2 active participation styles, posting and responding, that is analyzed in our qualitative evaluation. To include passively engaged users in our sample, we assessed *the number of page views* and *the total time spent on DC* per user. However, because users may have been active through sharing posts when viewing various pages on the platform during their time spent on DC, these parameters included, but were not limited to, the passive user mode of reading. Together, our operationalization of user engagement implied that both active users who posted and passive users who spent considerable time on DC and viewed many pages could be categorized as highly engaged users.

Recovery-Oriented Outcomes

In recent years, peer support studies have frequently reported on personal recovery to complement clinical recovery outcomes, with a particular focus on the benefits of online PSIs for *empowerment* as an important feature in the process of personal recovery that individuals can develop to enable them to live a meaningful life [7,40,41]. Although inconclusive, these findings were promising [21,42-45]. Within online communities, empowerment refers to enabling processes including "becoming better informed, receiving and giving emotional support by sharing relatable experiences of living with the diagnosis, helping others, and networking" [46]. Developing and exploiting self-management strategies can be seen as an active component of empowerment [47,48], and many strategies comprise individual skills "to monitor one's condition and to affect the cognitive, behavioral, and emotional responses necessary to maintain a satisfactory quality of life" [49]. However, to date, self-management has not been systematically examined as an individual outcome in peer support studies [50]. The same holds true for general well-being (ie, functional recovery, including social functioning and quality of life) [12,42-45,51-53], although both are important parameters for determining the usefulness of recovery-oriented PSIs. Our meta-analysis [19] showed that PSIs may also be effective in terms of clinical recovery (ie, symptom reduction), particularly for individuals with serious mental illness, including major depressive disorder. We also examined whether our online peer support platform helps improve depressive symptoms.

Objectives

In this longitudinal user survey, we attempted to add to the current literature on online peer support in several ways. First, we examined personal (ie, empowerment and self-management), functional (ie, well-being), and clinical (depressive symptoms) recovery parameters among DC users. Furthermore, we comprehensively explored patterns of user engagement, including parameters reflecting both the intensity and nature of DC use. Following the results of a systematic review of user engagement [32], we clustered user engagement based on the number of posts and page views and total time spent on DC. We aimed to determine whether these user engagement profiles during a 6-month interval were related to changes in empowerment. We expected that user engagement with high frequency—including active (ie, number of posts) and passive (ie, number of page views and total time spent)—would contribute to more improvement in empowerment from baseline to 3 and 6 months compared with user engagement with low frequency. In addition, we explored the use of self-management strategies and changes in depressive symptoms, levels of functioning, and disability over time.

Methods

Design

For this longitudinal study, users of our online peer support community for depression, DC, completed an online survey at 3 time points between June 19, 2019, and September 24, 2020.

Depression Connect

Launched on June 19, 2019, DC was cocreated with experiential experts, caregivers, and health professionals (therapists, psychiatrists, and psychology researchers) affiliated with the Dutch patient association for depression (The Depression Association), the Centre of Expertise for Depression as part of the Pro Persona Institute for mental health care, and the Radboud University Medical Centre. DC was developed to facilitate the exchange of personal experiences in coping with depression among peers. The online platform was easily and (if preferred) anonymously accessible to anyone dealing with depression. Potential users were not screened for depressive symptoms or other clinical characteristics before they could enter the community. Although no professionals were involved in DC, its moderators, who were all experiential experts, were able to consult a psychiatrist and psychology researchers of our team when feedback was needed. To ensure a constructive and supportive online atmosphere, DC moderators screened all new posts daily. They also generated new content or boosted user activity on the platform, for instance, by posting news items or different viewpoints on coping strategies. The DC team welcomed an average of 90 new members each month. DC members could start a new discussion topic or join an existing topic created by other users or provided by the research team. At the DC's launch, we created 8 forum topics that referred to the main themes of experiential knowledge in depression, which we identified in our qualitative interview study [29]. Widely used topics included coping with symptoms of depression (eg, concentration problems) and treatment options for depression (eg, medication and mental health care). In addition to reading and posting messages on the forum section of the community, users could read news items (posted by the DC team) and read or post blogs. There was also a function for sending private

messages to other DC users. More details about the DC's development, functionalities, and monitoring procedures are presented in our parallel qualitative evaluation of DC [29].

Participants and Procedure

All individuals who registered with DC were invited to participate in our study when the website was launched. There were no strict conditions to participate in the study regarding demographic and clinical characteristics, either for the minimum or maximum level of engagement with DC. All new DC members, and thus potential study participants, received an email to welcome them to the community, including information about our quantitative evaluation study and a link to the survey. An email address of the research team was also displayed to provide users the opportunity to ask questions about study participation. Participation was on a voluntary basis without any financial or other compensation. Interested users were invited to complete the online survey 1 or 2 days after registering and at 3 and 6 months after joining DC. Of the 1374 new members who joined DC during the recruitment period, 317 (23.07%) users completed the baseline survey. Subsequently, 5% (16/317) of participants deleted their accounts, including their user data logs, leaving 21.91% (301/1374) of participants for the final sample. The data sets of participants who completed only the baseline assessment (179/301, 59.5%) were not included in the outcome analyses.

Measurements

Participants completed the following measures at baseline and at 3 and 6 months after joining DC.

Demographic and Clinical Characteristics

At the baseline assessment, the participants were asked to list their age, sex, and level of education. At all 3 time points, we asked the participants if they received current treatment (referring to any form of mental health care) and used antidepressant medications and whether they were experiencing a depressive episode at the time of the assessment. These variables were assessed by self-report; we did not use a validated symptom-screening measure.

User Engagement

Participants' engagement in DC was determined by analyzing user data logs, which were encrypted and provided by the website host [54]. In line with the most widely used metrics to categorize user engagement in online health communities [32,37], we computed the following three parameters after 3 and 6 months of DC use: (1) total time spent on DC, (2) number of page views, and (3) number of posts entered on DC. We did not consider online activities related to survey completion.

Outcomes

Empowerment

To gauge the changes in empowerment, we used the Netherlands Empowerment List (NEL) [55], that consists of 40 questions covering the following 6 subscales: *social support, professional help, connectedness, confidence and purpose, self-management,* and *caring community*. Items were to be answered on a 5-point Likert scale ranging from 0 (*strongly disagree*) to 4 (*strongly agree*), with a *not applicable* answer option for the *professional help* subscale. We calculated the total empowerment score by summing and averaging all the completed items (range 0-4). Items on the *professional help* subscale that were scored as *not applicable* were not included in this calculation. Higher scores reflected higher levels of empowerment. Both previous research [56] and this study achieved high reliability for the total score (α=.93).

Self-management

The use of self-management strategies was evaluated using the Dutch Assessment of Self-management in Anxiety and Depression questionnaire (ASAD) [57,58]. The ASAD considers 45 self-management strategies that are presented in an equal number of statements. Respondents were asked whether and to what extent they used the strategy referred to (eg, *keep focused on the present and stop myself from looking too far ahead*). Each item is rated on a 5-point Likert scale, ranging from 0 (*not at all*) to 4 (*very much*). We used the total score (range 0-180) in our analyses. The higher the score, the higher was the use of self-management strategies. The reliability in this study was high (α=.92). Previous research only examined psychometric properties for the ASAD-Short Form, showing high levels of internal consistency (Cronbach α>.75) for the total questionnaire as well as its subscales (intraclass correlation coefficient >0.75) [46].

Depressive Symptoms

Depression severity was assessed using the Dutch version of the Beck Depression Inventory-II (BDI-II) [59], which consists of 21 questions, with each answer scored on a scale from 0 to 3. The total score ranged from 0 to 63, with higher scores reflecting more severe depressive symptoms. Specifically, scores between 0 and 13 indicate minimal symptoms, between 14 and 19 mild depression, between 20 and 28 moderate to severe depression, and the highest category with scores between 29 and 63 indicate severe depression [60]. The BDI-II has good psychometric properties [61]. In this study, the reliability of the total score was high (α=.91).

Functioning and Disability

We assessed individual functioning and disability with the Dutch version of the World Health Organization Disability Assessment Schedule (WHODAS) 2.0 [62]. A total of 6 domains (ie *cognition, mobility, self-care, getting along, life activities,* and *participation)* were evaluated with 36 items rated on a 5-point Likert scale, ranging from 0 (*no effort at all*) to 4 (*much effort*), where higher scores indicate more disability. The WHODAS 36 2.0 is a valid and reliable self-report instrument, with good psychometric properties irrespective of population type [63], which was reflected by the high reliability of the total score in our study (α=.92).

Statistical Analysis

Outcomes

Data analyses were conducted using SPSS (version 28; IBM Corp) and R (version 4.1.1) [64] using R Studio (2021.09.0+351). Longitudinal modeling was performed using the R lme4 package [65]. To determine whether the outcomes

had significantly changed over time and significantly differed between user engagement profiles, linear mixed models were fitted with the respective outcomes as dependent variables. In the model, the Restricted Maximum Likelihood Estimation calculated parameter estimates. As multiple imputation is not deemed necessary, we did not conduct a missing data analysis a priori [66]. We specified a linear mixed model regression with fixed effects: the actual day of assessment (day), engagement profiles, and the interaction effect between the engagement profile and day of assessment. The baseline value (day=0) of the outcome variable was included as a covariate, and random slopes for the repeated measures design day effect were included.

The estimated marginal means and within-group effect sizes were calculated using the emmeans package [67]. We calculated the magnitude of change between the baseline assessment (day=0) and assessment 3 (day=186), reported as the effect size, Cohen *d* [68]. To calculate the effect size, we needed an estimate of the SD of the intercept. In the model that included the baseline value of the outcome variable as a covariate, the estimate of the intercept's SD was almost 0. Therefore, we used a model without this covariate to estimate the SD of the intercept.

Engagement Profiles

We used cluster analysis to identify subgroups of participants who shared similarities in their forum use patterns. Next, we performed a K-medoids cluster analysis with the cluster package [69], using the partitioning around medoids algorithm, a more robust version of the K-means algorithm, which, instead of averages of distances between points in the sample, uses actual data as the center of a cluster. For each subject, (1) session duration, (2) number of page views, and (3) number of posts were computed for the first 3 months and the last 3 months, excluding the sessions in which the questionnaires were filled out. Because of the extreme skewness in these 6 indices, we took their square roots and transformed them into *z* scores for the cluster algorithm. Although the Tibs2001SEmax gap criterion [69] found an optimum of 7 clusters, the number of participants was very small in the high-engagement clusters (n=3 and n=4), which is why we opted for a 4-cluster solution in which the high-engagement cluster contained 7 (%) participants of the total population (N).

Ethics Approval

After evaluation, the local ethics committee (Commissie Mensgebonden Onderzoek Arnhem-Nijmegen) determined that no ethics approval was required, given the minimal burden to the study participants. The users provided passive consent to log and analyze their user data.

Results

Data Preparation

Of the 301 DC users who had provided their consent and completed the baseline measurement, 179 (59.5%) individuals did not complete the survey at the 3- and 6-month time points. In total, 15.9% (48/301) of DC users completed the 3-month and 24.5% (74/301) of users the 6-month survey. There were no missing data for the 4 main outcome measures at any of the 3 time points. For age and current depression, we noted 2 and 6 missing variables, respectively. A total of 496 observations from 301 participants were included in the mixed modeling analyses.

Baseline Characteristics of Participants

The participants' demographic and clinical characteristics as well as the means and SDs for the outcome variables at baseline are shown in Table 1. Our sample of 301 DC users included individuals with self-reported depression and a mean age of 50.2 (SD 13.12) years, 66.1% (199/301) of them were female and most of the respondents (216/301, 71.8%) had completed some form of secondary education or training. More than half of the respondents (166/301, 55.1%) reported having severe depressive symptoms (mean BDI score of 38.7, SD 6.57) and almost one-quarter of the population (72/301, 23.9%) had moderate to severe symptoms (mean BDI score of 23.8, SD 2.53). Of the remaining respondents (63/301, 20.9%), 13% (8/63) reported mild symptoms, and 8% (3/63) had minimal symptoms. The overall mean baseline BDI score for the entire sample was 29.84 (SD 11.85). Most DC users (241/301, 80.1%) received current treatment or some form of support or care from a mental health service and 69.8% (210/301) reported current use of antidepressants.

Completers, that is those respondents that had completed the baseline and at least one second assessment, were on average 4.95 years older (SD 11.4; 2-tailed t_{297}=3.25; *P*=.001) and reported significantly higher levels of empowerment, self-management, and less severe depressive symptoms and disability in major life domains compared with DC users who had only completed the baseline assessment.

Table 1. Demographic and clinical characteristics of survey respondents at baseline and of the participants having completed at least one subsequent assessment (N=301).

Characteristic	Respondents			Value		
	Total group (N=301)	Baseline only (n=179)	Completers 2 or 3 assessments (n=122)	Test statistic		*P* value
				t test (*df*)	χ^2 (*df*)	
Age (years; range 18-99), mean (SD)[a]	50.2 (13.1)	48.22 (13.9)	53.16 (11.4)	3.25 (297)	N/A[b]	<.001
Female, n (%)	199 (66.1)	119 (66.5)	80 (65.6)	N/A	0.03 (1)	.87
Educational level, n (%)				N/A	11.5 (3)	.01
None, elementary school, or vocational education	44 (14.6)	36 (12.7)	8 (2.7)			
Secondary education (middle or high school)	167 (55.5)	90 (29.9)	77 (25.6)			
Secondary vocational education and training	49 (16.3)	30 (10)	19 (6.3)			
Advanced vocational education and training and academic education	41 (13.6)	23 (7.6)	18 (6)			
Current depression (self-reported), n (%)[c]	216 (73.2)	136 (63)	80 (37)	N/A	3.7 (1)	.06
Depressive symptoms (BDI-II[d]), mean (SD)	29.84 (11.9)	31.8 (11.4)	26.97 (12)	−3.53 (299)	N/A	<.001
Severity of depressive symptoms (BDI-II)[e], n (%)				N/A	11.9 (3)	.008
Severe depressive symptoms	166 (55.1)	112 (62.6)	54 (44.3)			
Moderate to severe depressive symptoms	72 (23.9)	39 (21.8)	33 (27)			
Mild depressive symptoms	39 (13)	19 (10.6)	20 (16.4)			
Minimal depressive symptoms	24 (8)	9 (5)	15 (12.3)			
Current treatment, n (%)[f]	203 (67.4)	127 (62.6)	76 (37.4)	N/A	2.5 (1)	.12
Current antidepressant medication, n (%)	210 (69.8)	127 (60.5)	83 (39.5)	N/A	2.9 (1)	.59
Empowerment (NEL[g]), mean (SD)	2.06 (0.5)	1.99 (0.5)	2.15 (0.5)	2.78 (299)	N/A	.01
Self-management (ASAD[h]), mean (SD)	78.11 (25.1)	75.15 (26.6)	82.45 (22)	2.5 (299)	N/A	.01
Functioning and disability (WHODAS 2.0[i]), mean (SD)	35.7 (15.3)	38.11 (15.1)	32.17 (15)	−3.36 (299)	N/A	.001

[a]Owing to 2 missing variables, n=299 for the total group, n=178 for baseline only, and n=121 for completers.

[b]N/A: not applicable.

[c]Owing to 6 missing variables, n=295 for the total group, n=176 for baseline only, and n=119 for completers.

[d]BDI-II: Beck Depression Inventory-II.

[e]On the basis of the following BDI cutoff scores: 0-13, minimal depression; 14-19, mild depression; 20-28, moderate to severe depression; 29-63, severe depression.

[f]Includes any type of mental health care (eg, general or specialized mental health care and alternative support).

[g]NEL: Netherlands Empowerment List.

[h]ASAD: Dutch Assessment of Self-management in Anxiety and Depression questionnaire.

[i]WHODAS 2.0: World Health Organization Disability Assessment Schedule 2.0.

User Engagement

Our cluster analysis of the forum uses parameters as described in the Methods section and resulted in 4 user engagement profiles: *very low* (profile 1), *low* (profile 2), *medium* (profile 3), and *high* (profile 4). The user parameters for the total study period (6 months) are listed in Table 2. Baseline engagement profiles did not significantly differ for age; sex; current depression; current treatment or medication; or baseline scores on empowerment, self-management, depressive symptoms, and disability. However, results did show significant differences between participants completing the baseline assessment only and participants that completed one or 2 assessments for the "very low" engagement profile (177/301, 58.8%), of which 33.9% (60/177) participated in a second assessment and 66.1% (117/177) completed baseline only (χ^2_3=27.1; *P*<.001; Multimedia Appendix 1).

Figure 1 depicts the (changes in) outcomes and session duration for the participants who completed only the baseline assessment

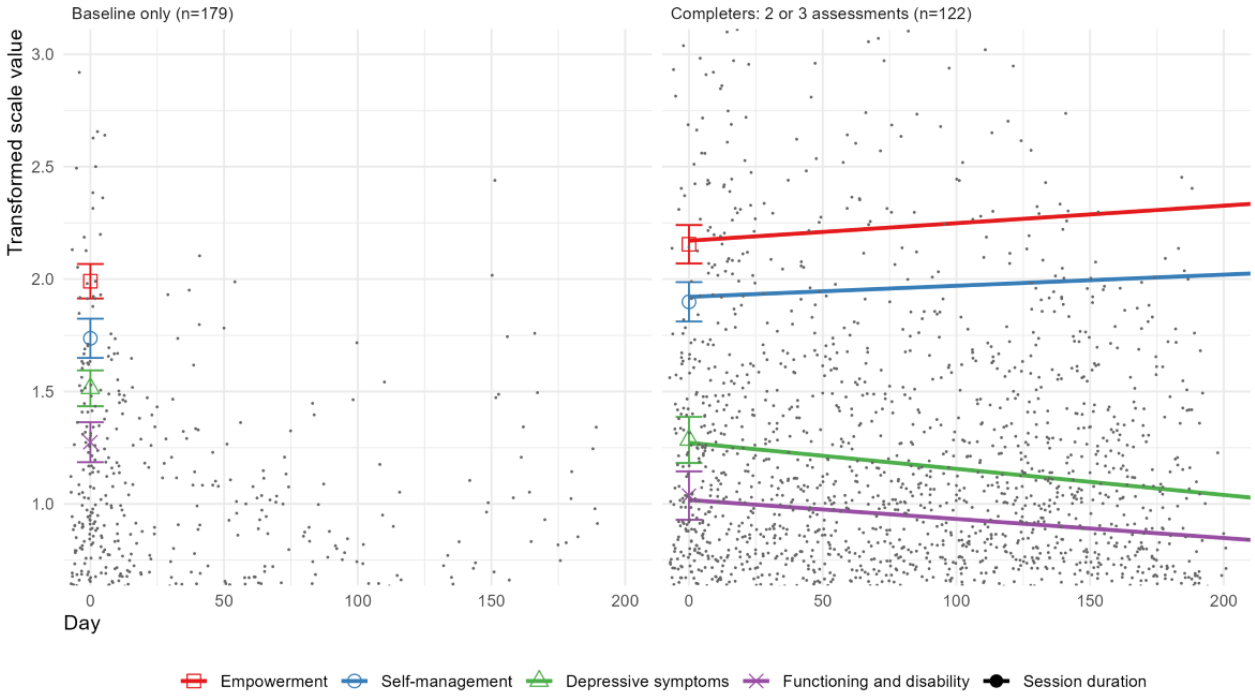
(panel 1, n=179) and for those who completed at least 2 or all 3 assessments (panel 2, n=122). The graphs present data modeled using a longitudinal mixed model regression analysis for session duration for each individual session (black dots), empowerment (NEL), self-management (ASAD), depressive symptoms (BDI-II), and functioning and disability (WHODAS 2.0) over time, including the baseline means for each outcome. They show an increase in empowerment and self-management and a decrease in depressive symptoms and disability over time (days of engagement with the online peer support community—DC).

In the figure, the scale values were transformed to a range from 0 to 4. To avoid overlap, self-management was raised by 0.4 and functioning and disability was lowered by 0.25. The session duration is transformed to the square root of the duration in minutes and divided by 5, so that most values are in the figure. These lines are based on the values predicted by the model. Values at day 0 are actual means with the confidence levels.

Table 2. Depression Connect user engagement for the total study group (N=301).

Total study group (engagement parameter)	Very low engagement (profile 1; n=177), mean (SD)	Low engagement (profile 2; n=87), mean (SD)	Medium engagement (profile 3; n=30), mean (SD)	High engagement (profile 4; n=7), mean (SD)	Value, mean (SD)	*F* test (*df*)	*P* value
Total session duration in hours	5.48 (22.37)	0.37 (1.1)	2.35 (2.43)	18.84 (17.04)	116.01 (86.20)	177.51 (3,297)	<.001
Number of page views	322.19 (1117.14)	25.85 (41.71)	161.18 (103.12)	1315.33 (1015.74)	5560 (4159.86)	175.63 (3,297)	<.001
Number of posts	14.21 (58.45)	0.33 (1.33)	5.79 (5.81)	44.8 (47.46)	338.7 (158.63)	363.26 (3,297)	<.001

Figure 1. Session duration and changes in empowerment, self-management, depressive symptoms, and functioning and disability over 6 months of using Depression Connect.



Outcomes

Table 3 lists the results of the models that investigated changes in outcomes over time (days of DC use). We computed a significant increase in empowerment (NEL) over time in days (β=.00078; SE 0.00022; *P*=.001) with a small effect size (Cohen *d*=0.36, 95% CI 0.15-0.57). Self-management also increased over time (β=.0222; SE 0.011; *P*=.046), again with a small effect size (Cohen *d*=0.22, 95% CI 0-0.43). Depressive symptoms (BDI) significantly decreased over time with a small effect size (β=−0.0244, SE 0.00612; *P*<.001; Cohen *d*=0.44, 95% CI 0.21-0.66). In addition, disability (WHODAS 2.0) significantly decreased over time with a small effect size (β=−0.0212, SE 0.00693; *P*=.001; Cohen *d*=0.29, 95% CI 0.10-0.47).

Engagement profiles were not significantly associated with changes in any of the outcomes at 3 or 6 months as indicated by the nonsignificant effects of the dependent variable by engagement profile: empowerment ($F_{3,176}$=0.07; *P*=.98); self-management ($F_{3,169}$=0.1; *P*=.96); depressive symptoms ($F_{3,184}$=0.14; *P*=.94); and functioning and disability ($F_{3,181}$=0.2; *P*=.90); and nonsignificant time by profile interactions: (empowerment ($F_{3,131}$=0.55; *P*=.65); self-management

($F_{3,126}$=.93; P=.43); depressive symptoms ($F_{3,140}$=0.09; P=.97) and functioning and disability ($F_{3,158}$=0.09; P=.96).

Table 3. Linear mixed model analysis outcomes; estimated marginal means and effect sizes[a].

Outcomes	Baseline, EMM[b] (SE)	3 months (day 95), EMM (SE)	6 months (day 186), EMM (SE)	Fixed effect of time			Cohen *d* effect size[c] day 0 to day 186 (95% CI)
				Change, per 100 days	*F* test (*df*)	*P* value	
NEL[d] (empowerment)[e]	2.17 (0.01)	2.24 (0.02)	2.32 (0.04)	.078	12.5 (1,136)	.001	0.36 (0.15-0.57)
ASAD[f], self-management[g]	83.44 (0.7)	85.56 (0.98)	87.58 (1.89)	2.22	4.05 (1,132)	.046	0.22 (0.0-0.43)
BDI-II[h] (depressive symptoms)[i]	26.71 (0.36)	24.4 (0.55)	22.18 (1.04)	−2.44	15.9 (1,147)	<.001	0.44 (0.21-0.66)
WHODAS-36 items 2.0[j] (disability)[k]	31.69 (0.53)	29.68 (0.60)	27.76 (1.11)	−2.12	9.33 (1,164)	.003	0.29 (0.10-0.47)

[a]We used a linear mixed model with time (days) and baseline value of the dependent variable as a fixed factor and subject within time (days) as random effects.

[b]EMM: estimated marginal mean.

[c]To calculate effect sizes, we used a model in which the baseline value of the outcome variable as a covariate was not included because in the model including this covariate, the estimate of the intercept's SD naturally almost 0.

[d]NEL: Netherlands Empowerment List.

[e]Scores range from 0 to 4, with higher scores indicating greater empowerment.

[f]ASAD: Dutch Assessment of Self-management in Anxiety and Depression questionnaire.

[g]Scores range from 0 to 180, with higher scores indicating higher use of self-management strategies.

[h]BDI-II: Beck Depression Inventory-II.

[i]Scores range from 0 to 63, with higher scores indicating more (severe) symptoms.

[j]WHODAS 2.0: World Health Organization Disability Assessment Schedule 2.0.

[k]Scores range from 0 to 144, with higher scores indicating greater disability.

Discussion

Principal Findings

Although the potential benefits of engaging peer support for people with severe mental illness (SMI) are widely acknowledged [25,26], peer support studies are limited for online intervention types targeting depression [20,21] and the processes for user engagement remain unclear [32,70]. In this longitudinal user survey of the online peer support community—DC, we explored patterns of user engagement and examined whether user profiles were associated with recovery-oriented outcomes. To quantify baseline to 6-month changes in empowerment, self-management, depressive symptoms, and functioning and disability in the users of DC and considering the complex interplay of relevant aspects of user engagement in PSIs, we entered the user data logs of 3 parameters (ie, total session duration, page views, and number of posts) into a cluster analysis, resulting in 4 engagement profiles. Most of the survey respondents (177/301, 58.8%) had very low or low engagement levels (87/301, 28.9%), with 9.9% (30/301) having medium and 2.3% (7/301) high user profiles. However, none of the profiles showed significant differences for age; sex; having current depression; receiving treatment at the time of assessment; or with regard to the baseline scores for empowerment, self-management, depression, and functioning and disability. All recovery-oriented outcomes improved over time; however, contrary to our hypothesis, the nature and intensity of DC user engagement were not significantly associated with any of these improvements.

Findings in Context

The number of user surveys and RCTs for online PSIs for depression is limited; however, the results are promising. Although our results did not show a significant relationship between the level of user engagement and recovery, Griffiths et al [20,21] reported positive results for engaging a online PSI for depression in their trial. They found that depressive symptoms reduced in the long-term period (6 and 12 months) and empowerment may improve in the short-term (after the intervention or at 3 months). Furthermore, reviews with and without meta-analyses of PSIs that include a heterogeneous population, which includes primarily individuals with SMI, report positive changes in psychosocial outcomes [24,25,43,45], more specifically for self-efficacy and hope [12,42,44,51-53]. We confirmed this in our new and updated meta-analysis, which included PSIs for mental illness [19]. However, research on peer support is associated with methodological issues (eg, establishing model fidelity is not possible at this point [25]). Therefore, the results of this longitudinal user survey as well as the results from the other PSI studies should be interpreted with caution.

Considering the level of user engagement, it is generally known that online communities [37,70] are associated with low engagement rates. This is often referred to as the 1% rule [36,37]

and is in line with our results. However, our study adds to the current literature, as the results improve insight into the intensity and nature of user engagement for online PSIs for depression. Further research is needed to better understand the relationship between levels of user engagement and positive changes in recovery.

Flexible User Engagement

In Search of a Valid Proxy Measure for the Nature of Forum Use

To create as true a proxy as possible for the way the participants in our study used the DC platform, we included multiple indicators that we thought would best reflect the nature of their forum use. However, our results surprisingly showed that the frequency and nature (passive vs active) of user engagement did not appear to be associated with recovery. We focused on presence and participation rates on DC, whereas the CAPE model states that a broad range of factors should be incorporated when operationalizing user engagement. CAPE is an acronym including metrics on the following factors: Connect (how many people are interested), Attend (eg, presence or how many log-ins), Participate (eg, active engagement), and Enact (making use of online learned skills in daily life) [70,71]. Adding to the current knowledge base on user engagement, our results suggested that it might be too simplistic to assume that there is an optimal or specific engagement pattern or style that is directly related to positive outcomes associated with the use of PSIs [32]. As self-determination is a crucial aspect of the recovery-oriented approach, which is reflected in our PSI, voluntary use of the program seems important [72]. Arguably, the need for support from peers or the intention to support peers depends on the stage of depression or coping levels, which affects the intensity (ie, frequency or duration) and nature of a person's forum engagement (eg, posting to ask for help or responding to help others) [73]. In line with our qualitative evaluation of DC user experiences, the data presented here might indicate that user modes are indeed used interchangeably over time, developing and deploying different engagement styles (ie, reading, posting, or responding) according to their personal needs [29]. Therefore, these shifts in forum use make it difficult to capture the effects of DC use in quantitative terms, such as engagement profiles. Nevertheless, online PSIs appear to provide users with an accessible digital realm in which they are free to choose individual modes of engagement that match their current needs in their search for recovery.

Quantity Versus Quality of User Engagement

In addition, the perceived quality of forum posts might be a relevant factor to be included when defining user engagement in terms of nature and intensity. It is possible that low engagement with the DC community suffices to benefit from peer support if a recently published post answers a specific question or explores a relevant topic effectively, satisfying the current needs of individual users [74]. DC users with queries about treatment options, for instance, may not have needed to spend much time on the platform to find pertinent information or check whether they had received a fitting response. In turn, if a user is looking for (online) friendship (to create a sense of belonging), they are likely to spend more time on the forum and engage more actively to connect with peers. Taken together, it may well be the personal needs and goals of the users and the perceived quality of the forum content that ultimately determine whether and how users engage in and benefit from online peer support communities, such as DC.

Potential Disadvantages of Active User Engagement

From a different perspective, 2 potential disadvantages of active forum engagement might have defeated the hypothesized positive association between high user engagement and the experienced benefits (recovery indices) of DC. First, the data showed that high-frequency users (high engagement) posted significantly more messages than the users with the other 3 profiles (very low, low, and medium engagement). This might imply that frequent users predominantly posted messages for (ie, responded to) peers seeking support, focusing less on their own needs and recovery. According to the helper-therapy principle [75], high-frequency users may experience positive feelings because they perceive helping peers as meaningful. In line with the central drawback that DC users emphasized in our qualitative study, this active style may also have increased distress levels by their feeling responsible for their peers' well-being or by their identification with the problems of fellow users too much [17]. In addition, as observed in clinical practice [76] and our qualitative study [29], high engagement in supportive interactions may encourage self-reflection, uncovering problems that users were not, or partly, aware of before, which might be both distressing and healing. Thus, compared with passive users, active engagers run a greater risk of being exposed to the disadvantages of peer support, possibly increasing their disease burden owing to a heightened sense of responsibility for others and an increased awareness of their personal issues.

Assessing Recovery in Online Peer Support

Finally, other recovery-oriented outcomes may be more relevant for evaluating an unstructured online peer support community such as DC. Empowerment and self-management may serve as those attributes that would characterize more advanced stages of recovery from mental illness, such as depression, as they take time to develop and generally require guidance from a nonpeer (ie, a paraprofessional) [18,46], face-to-face PSI format [42,43,77], or a wider supportive context involving family or friends [78,79]. Moreover, considering the informal nature and flexibility, free use of our platform, and the fact that our sample mostly consisted of individuals with moderate to severe depressive symptoms, smaller goals such as an increased sense of being (emotionally) supported or finding new hope are probably more feasible [80].

Limitations

This study has several limitations. The first one lies in the operationalization of user engagement. Rather than opting for (more frequently used) self-report measures, we tried to objectively quantify forum engagement using logged user data (total time spent on the platform and the number of page views and posts); however, there are other potentially relevant indicators of engagement, such as the number of posts the user reads [34], the length of threads [74], and the number of replies

received [74,81]. Particularly for individuals with depression, these activities and interactions that reflect recognition and support may reduce stress and negative emotions [82]. Unfortunately, we were unable to extract these parameters from the user data logs. Second, the results of our previous qualitative exploration of DC user experiences were primarily related to users with an active engagement style. In this quantitative study, the number of highly frequent and actively engaged users—those posting significantly more than their peers with other engagement profiles—was too small to detect any reliable effects on empowerment. Third, the lack of a comparison group in this longitudinal user survey precluded exploration of causal relationships between DC use and recovery; however, the effect sizes of RCTs comparing PSIs for mental illness with a control group that we pooled in our meta-analysis were significant both for clinical and personal recovery indices.

Finally, the generalizability of our findings is limited as we evaluated self-selected samples, where the decision to participate may contain some inherent positive bias toward engaging in (online) peer support. It is possible that users with a low engagement profile were not motivated to complete the follow-up assessments in our evaluation study because they lacked commitment to DC or may not have experienced any benefits of engagement. However, our study has an explorative character, with a naturalistic sample that informed us of the general and heterogeneous population of individuals with depression who engage in peer support. Given the observational character of our study, its internal validity is limited. We do not know whether the improved outcomes are related to DC use and to what extent other types of support or the many other variables that are part of the real-world setting (eg, the level of offline social support, self-stigma, and societal participation) influence these results. Regardless, considering the free and informal nature of our online peer-to-peer support environment that allows users to tune their use of the forum to their personal needs and the improvements observed, our survey expands the current literature by focusing on an online PSI for depression. The results underscore that this type of peer support appears to be beneficial and promotes recovery among individuals with self-reported depressive symptoms. These promising results are not only reflected in our survey but also in previously conducted user surveys in PSIs, underscoring the benefits of peer support for clinical [14] and personal recovery [33].

Future Research

As the various engagement profiles that we identified indicate that DC users appear to prefer a flexible use of the platform, insight into the content of their posts would foster the interpretation of our findings. Therefore, we recommend assessing the perceived quality of interactions (eg, “Is the content helping you to cope with your depression?”) in future research on online PSIs. The quantitative variables such as thread length and the number of posts and responses or comments might indicate how effectively a topic was explored [74]. Synthesizing qualitative data (content analysis) and quantitative data (metrics of use) of peer support user engagement would enhance our understanding of its implications for recovery. In addition to the clinical characteristics and (treatment) history of depression, it may be informative to describe the societal context of individual users. Possibly, the availability and quality of social support from family or friends may predict users’ need for online peer support and explain their low or high engagement. As peer support is considered adjunctive to formal mental health care [9] and it has been suggested that peer support encourages users to engage more actively in their professional treatment [14], it is worthwhile to investigate the usefulness and benefits of (online) peer support for concurrent professional therapy. Finally, recovery is a multidimensional concept; however, the various factors and processes involved are difficult to disentangle. Including comprehensive measurements in which the umbrella concepts of clinical, personal, and functional recovery-related indices are assessed separately and in depth, such as in the case of the Recovery Assessment Scale, might improve the validity of the findings.

Conclusions

This longitudinal user survey provides insights into the characteristics of user engagement in DC, an online peer support community for depression. Active engagement was limited to a small group of DC users and was not significantly associated with superior improvements in empowerment and secondary recovery-oriented outcomes. Users appear to attune the intensity and nature of their forum use to their personal recovery pathway and current needs, where their engagement levels may shift from low to high and from passive to active. Corresponding to the self-determination theory, the autonomy to choose the level of engagement might be one of the most valued and effective features of intervention types, such as DC, whereas in other more formal supportive environments for depression, a certain level of engagement is predetermined. Future online PSI studies should explore the content and quality of user interactions to determine what constitutes optimal user engagement, where flexibility and usefulness match users’ clinical needs and motives to seek and offer online peer support.

Acknowledgments

The authors would like to thank the Radboudumc REshape Centre of the Radboud University Medical Centre for their close collaboration and valuable contributions in developing Depression Connect (DC). In addition, the authors thank De Digitale Poli for building and hosting the DC website and the online platform. Most importantly, the authors thank all the users of DC for their commitment, with special thanks to the members who participated in this study. The “The Power of Depression” research project and this study are funded by ZonMw, the Netherlands Organization for Health Research and Development. The funder was not involved in the design of the study; or in the collection, analysis, and interpretation of the data; or the writing of the manuscript.

Data Availability and Management

All data and data management were handled in accordance with the national guidelines for university faculties of social and behavioral sciences (formulated by the DSW Committee: Deans of social sciences faculties, March 2022 [83]). Accordingly, to enhance open science and transparent research practices, we followed the FAIR principles (findable, accessible, interoperable, reusable) and stored (descriptions of) our study material and approaches as open as possible and as closed as necessary at the Radboud Data Repository.

Authors' Contributions

JNV, BG, and JS conceptualized the study; its further design was undertaken by DS and TB, who acquired, analyzed, and interpreted the data and drafted the manuscript. DS had full access to all the study data and was responsible for submitting this manuscript for publication. JNV, BG, JS, and TB participated in the critical review and revision of the manuscript. All authors provided intellectual content, reviewed, edited, and amended the manuscript. All authors gave their final approval for this version to be published and agreed to be accountable for all aspects of this work.

Conflicts of Interest

The online peer support community “Depressie Connect” was developed in collaboration with the Pro Persona Institute for mental health care, the Dutch Depression Association, and Radboud University Medical Centre, Nijmegen. The “Depressie Connect” platform, the subject of ongoing research, is run by a coordinator and moderators of the Dutch Depression Association. Decisions about the long-term development of the Depression Connect are made jointly by these 3 parties, whose affiliates have not received any financial compensation for their work in connection with the platform.

Multimedia Appendix 1
Baseline demographics and clinical characteristics for the total study group.
[DOCX File , 15 KB - formative_v6i11e39912_app1.docx]

References

1. Depression and other common mental disorders: global health estimates. No. WHO/MSD/MER/2017.2. World Health Organization. Geneva, Switzerland: World Health Organization; 2017. URL: https://apps.who.int/iris/handle/10665/254610 [accessed 2020-10-10]
2. Ustün TB, Ayuso-Mateos JL, Chatterji S, Mathers C, Murray CJ. Global burden of depressive disorders in the year 2000. Br J Psychiatry 2004 May;184:386-392. [doi: 10.1192/bjp.184.5.386] [Medline: 15123501]
3. GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct 17;396(10258):1204-1222 [FREE Full text] [doi: 10.1016/S0140-6736(20)30925-9] [Medline: 33069326]
4. Cuijpers P, Noma H, Karyotaki E, Vinkers CH, Cipriani A, Furukawa TA. A network meta-analysis of the effects of psychotherapies, pharmacotherapies and their combination in the treatment of adult depression. World Psychiatry 2020 Feb;19(1):92-107 [FREE Full text] [doi: 10.1002/wps.20701] [Medline: 31922679]
5. Patel V, Saxena S, Lund C, Thornicroft G, Baingana F, Bolton P, et al. The Lancet Commission on global mental health and sustainable development. Lancet 2018 Oct 27;392(10157):1553-1598. [doi: 10.1016/S0140-6736(18)31612-X] [Medline: 30314863]
6. Mueller TI, Leon AC, Keller MB, Solomon DA, Endicott J, Coryell W, et al. Recurrence after recovery from major depressive disorder during 15 years of observational follow-up. Am J Psychiatry 1999 Jul;156(7):1000-1006. [doi: 10.1176/ajp.156.7.1000] [Medline: 10401442]
7. Leamy M, Bird V, Le Boutillier C, Williams J, Slade M. Conceptual framework for personal recovery in mental health: systematic review and narrative synthesis. Br J Psychiatry 2011 Dec;199(6):445-452. [doi: 10.1192/bjp.bp.110.083733] [Medline: 22130746]
8. Peer support mental health services: promoting person-centred and rights-based approaches. Licence: CC BY-NC-SA 3.0 IGO. World Health Organization. Geneva, Switzerland: World Health Organization; 2021 Jun 9. URL: https://www.who.int/publications/i/item/9789240025783 [accessed 2022-01-03]
9. Jacob KS. Recovery model of mental illness: a complementary approach to psychiatric care. Indian J Psychol Med 2015;37(2):117-119 [FREE Full text] [doi: 10.4103/0253-7176.155605] [Medline: 25969592]
10. Davidson L. Recovery, self management and the expert patient – changing the culture of mental health from a UK perspective. J Mental Health 2005;14(1):25-35. [doi: 10.1080/09638230500047968]
11. Griffiths KM, Calear AL, Banfield M, Tam A. Systematic review on Internet Support Groups (ISGs) and depression (2): what is known about depression ISGs? J Med Internet Res 2009 Sep 30;11(3):e41 [FREE Full text] [doi: 10.2196/jmir.1303] [Medline: 19793718]

12. Bryan AE, Arkowitz H. Meta-analysis of the effects of peer-administered psychosocial interventions on symptoms of depression. Am J Community Psychol 2015 Jun;55(3-4):455-471. [doi: 10.1007/s10464-015-9718-y] [Medline: 25861883]
13. Pfeiffer PN, Heisler M, Piette JD, Rogers MA, Valenstein M. Efficacy of peer support interventions for depression: a meta-analysis. Gen Hosp Psychiatry 2011;33(1):29-36 [FREE Full text] [doi: 10.1016/j.genhosppsych.2010.10.002] [Medline: 21353125]
14. Houston TK, Cooper LA, Ford DE. Internet support groups for depression: a 1-year prospective cohort study. Am J Psychiatry 2002 Dec;159(12):2062-2068. [doi: 10.1176/appi.ajp.159.12.2062] [Medline: 12450957]
15. Mead S, Hilton D, Curtis L. Peer support: a theoretical perspective. Psychiatr Rehabil J 2001;25(2):134-141. [doi: 10.1037/h0095032] [Medline: 11769979]
16. Mead S. Defining peer support. Intentional peer support: an alternative approach. 2003. URL: https://www.intentionalpeersupport.org/?v=b8a74b2fbcbb [accessed 2022-04-21]
17. Shalaby RA, Agyapong VI. Peer support in mental health: literature review. JMIR Ment Health 2020 Jun 09;7(6):e15572 [FREE Full text] [doi: 10.2196/15572] [Medline: 32357127]
18. Chinman M, George P, Dougherty RH, Daniels AS, Ghose SS, Swift A, et al. Peer support services for individuals with serious mental illnesses: assessing the evidence. Psychiatr Serv 2014 Apr 01;65(4):429-441. [doi: 10.1176/appi.ps.201300244] [Medline: 24549400]
19. Smit D, Miguel C, Vrijsen JN, Groeneweg B, Spijker J, Cuijpers P. The effectiveness of peer support for individuals with mental illness: systematic review and meta-analysis. Psychol Med (forthcoming) 2022 Sep 06:1-10. [doi: 10.1017/S0033291722002422] [Medline: 36066104]
20. Griffiths KM, Mackinnon AJ, Crisp DA, Christensen H, Bennett K, Farrer L. The effectiveness of an online support group for members of the community with depression: a randomised controlled trial. PLoS One 2012;7(12):e53244 [FREE Full text] [doi: 10.1371/journal.pone.0053244] [Medline: 23285271]
21. Crisp D, Griffiths K, Mackinnon A, Bennett K, Christensen H. An online intervention for reducing depressive symptoms: secondary benefits for self-esteem, empowerment and quality of life. Psychiatry Res 2014 Apr 30;216(1):60-66. [doi: 10.1016/j.psychres.2014.01.041] [Medline: 24534125]
22. Eysenbach G, Powell J, Englesakis M, Rizo C, Stern A. Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. BMJ 2004 May 15;328(7449):1166 [FREE Full text] [doi: 10.1136/bmj.328.7449.1166] [Medline: 15142921]
23. Kingod N, Cleal B, Wahlberg A, Husted GR. Online peer-to-peer communities in the daily lives of people with chronic illness: a qualitative systematic review. Qual Health Res 2017 Jan;27(1):89-99. [doi: 10.1177/1049732316680203] [Medline: 27956659]
24. Fortuna KL, Naslund JA, LaCroix JM, Bianco CL, Brooks JM, Zisman-Ilani Y, et al. Digital peer support mental health interventions for people with a lived experience of a serious mental illness: systematic review. JMIR Ment Health 2020 Apr 03;7(4):e16460. [doi: 10.2196/16460] [Medline: 32243256]
25. Fortuna KL, Solomon P, Rivera J. An update of peer support/peer provided services underlying processes, benefits, and critical ingredients. Psychiatr Q 2022 Jun;93(2):571-586 [FREE Full text] [doi: 10.1007/s11126-022-09971-w] [Medline: 35179660]
26. Griffiths KM, Calear AL, Banfield M. Systematic review on Internet Support Groups (ISGs) and depression (1): do ISGs reduce depressive symptoms? J Med Internet Res 2009 Sep 30;11(3):e40 [FREE Full text] [doi: 10.2196/jmir.1270] [Medline: 19793719]
27. Winsper C, Crawford-Docherty A, Weich S, Fenton S, Singh SP. How do recovery-oriented interventions contribute to personal mental health recovery? A systematic review and logic model. Clin Psychol Rev 2020 Mar;76:101815. [doi: 10.1016/j.cpr.2020.101815] [Medline: 32062302]
28. Watson E. The mechanisms underpinning peer support: a literature review. J Ment Health 2019 Dec;28(6):677-688. [doi: 10.1080/09638237.2017.1417559] [Medline: 29260930]
29. Smit D, Vrijsen JN, Groeneweg B, Vellinga-Dings A, Peelen J, Spijker J. A newly developed online peer support community for depression (depression connect): qualitative study. J Med Internet Res 2021 Jul 12;23(7):e25917 [FREE Full text] [doi: 10.2196/25917] [Medline: 34255659]
30. Geramita EM, Herbeck Belnap B, Abebe KZ, Rothenberger SD, Rotondi AJ, Rollman BL. The association between increased levels of patient engagement with an internet support group and improved mental health outcomes at 6-month follow-up: post-hoc analyses from a randomized controlled trial. J Med Internet Res 2018 Jul 17;20(7):e10402 [FREE Full text] [doi: 10.2196/10402] [Medline: 30021711]
31. Hensel JM, Shaw J, Ivers NM, Desveaux L, Vigod SN, Cohen A, et al. A web-based mental health platform for individuals seeking specialized mental health care services: multicenter pragmatic randomized controlled trial. J Med Internet Res 2019 Jun 04;21(6):e10838 [FREE Full text] [doi: 10.2196/10838] [Medline: 31165710]
32. Carron-Arthur B, Ali K, Cunningham JA, Griffiths KM. From help-seekers to influential users: a systematic review of participation styles in online health communities. J Med Internet Res 2015 Dec 01;17(12):e271 [FREE Full text] [doi: 10.2196/jmir.4705] [Medline: 26627369]

33. Fullwood C, Chadwick D, Keep M, Attrill-Smith A, Asbury T, Kirwan G. Lurking towards empowerment: explaining propensity to engage with online health support groups and its association with positive outcomes. Comput Human Behav 2019 Jan;90:131-140. [doi: 10.1016/j.chb.2018.08.037]
34. van Uden-Kraan CF, Drossaert CH, Taal E, Seydel ER, van de Laar MA. Self-reported differences in empowerment between lurkers and posters in online patient support groups. J Med Internet Res 2008 Jun 30;10(2):e18 [FREE Full text] [doi: 10.2196/jmir.992] [Medline: 18653442]
35. Petrovčič A, Petrič G. Differences in intrapersonal and interactional empowerment between lurkers and posters in health-related online support communities. Comput Human Behav 2014 May;34:39-48. [doi: 10.1016/j.chb.2014.01.008]
36. van Mierlo T. The 1% rule in four digital health social networks: an observational study. J Med Internet Res 2014 Feb 04;16(2):e33 [FREE Full text] [doi: 10.2196/jmir.2966] [Medline: 24496109]
37. Carron-Arthur B, Cunningham JA, Griffiths KM. Describing the distribution of engagement in an Internet support group by post frequency: a comparison of the 90-9-1 Principle and Zipf's Law. Internet Interv 2014 Oct;1(4):165-168. [doi: 10.1016/j.invent.2014.09.003]
38. Patten SB. Depressive symptoms and disorders, levels of functioning and psychosocial stress: an integrative hypothesis. Med Hypotheses 1999 Sep;53(3):210-216. [doi: 10.1054/mehy.1998.0747] [Medline: 10580525]
39. Depressie Connect. Depressie Vereniging. URL: http://www.depressieconnect.nl [accessed 2022-06-10]
40. Slade M, Leamy M, Bacon F, Janosik M, Le Boutillier C, Williams J, et al. International differences in understanding recovery: systematic review. Epidemiol Psychiatr Sci 2012 Dec;21(4):353-364 [FREE Full text] [doi: 10.1017/S2045796012000133] [Medline: 22794507]
41. Health 2020: a European policy framework and strategy for the 21st century. World Health Organization, Regional Office for Europe. 2013. URL: https://apps.who.int/iris/handle/10665/326386?locale-attribute=es& [accessed 2021-10-01]
42. Burke E, Pyle M, Machin K, Varese F, Morrison AP. The effects of peer support on empowerment, self-efficacy, and internalized stigma: a narrative synthesis and meta-analysis. Stigma Health 2019 Aug;4(3):337-356. [doi: 10.1037/sah0000148]
43. White S, Foster R, Marks J, Morshead R, Goldsmith L, Barlow S, et al. The effectiveness of one-to-one peer support in mental health services: a systematic review and meta-analysis. BMC Psychiatry 2020 Nov 11;20(1):534 [FREE Full text] [doi: 10.1186/s12888-020-02923-3] [Medline: 33176729]
44. Lloyd-Evans B, Mayo-Wilson E, Harrison B, Istead H, Brown E, Pilling S, et al. A systematic review and meta-analysis of randomised controlled trials of peer support for people with severe mental illness. BMC Psychiatry 2014 Feb 14;14:39 [FREE Full text] [doi: 10.1186/1471-244X-14-39] [Medline: 24528545]
45. Lyons N, Cooper C, Lloyd-Evans B. A systematic review and meta-analysis of group peer support interventions for people experiencing mental health conditions. BMC Psychiatry 2021 Jun 23;21(1):315 [FREE Full text] [doi: 10.1186/s12888-021-03321-z] [Medline: 34162340]
46. Johansson V, Islind AS, Lindroth T, Angenete E, Gellerstedt M. Online communities as a driver for patient empowerment: systematic review. J Med Internet Res 2021 Feb 09;23(2):e19910 [FREE Full text] [doi: 10.2196/19910] [Medline: 33560233]
47. Risling T, Martinez J, Young J, Thorp-Froslie N. Evaluating patient empowerment in association with eHealth technology: scoping review. J Med Internet Res 2017 Sep 29;19(9):e329 [FREE Full text] [doi: 10.2196/jmir.7809] [Medline: 28963090]
48. Cerezo PG, Juvé-Udina ME, Delgado-Hito P. Concepts and measures of patient empowerment: a comprehensive review. Rev Esc Enferm USP 2016;50(4):667-674 [FREE Full text] [doi: 10.1590/S0080-623420160000500018] [Medline: 27680054]
49. Barlow J, Wright C, Sheasby J, Turner A, Hainsworth J. Self-management approaches for people with chronic conditions: a review. Patient Educ Couns 2002;48(2):177-187. [doi: 10.1016/s0738-3991(02)00032-0] [Medline: 12401421]
50. Houle J, Gascon-Depatie M, Bélanger-Dumontier G, Cardinal C. Depression self-management support: a systematic review. Patient Educ Couns 2013 Jun;91(3):271-279. [doi: 10.1016/j.pec.2013.01.012] [Medline: 23414831]
51. Chien WT, Clifton AV, Zhao S, Lui S. Peer support for people with schizophrenia or other serious mental illness. Cochrane Database Syst Rev 2019 Apr 04;4:CD010880 [FREE Full text] [doi: 10.1002/14651858.CD010880.pub2] [Medline: 30946482]
52. Fuhr DC, Salisbury TT, De Silva MJ, Atif N, van Ginneken N, Rahman A, et al. Effectiveness of peer-delivered interventions for severe mental illness and depression on clinical and psychosocial outcomes: a systematic review and meta-analysis. Soc Psychiatry Psychiatr Epidemiol 2014 Nov;49(11):1691-1702 [FREE Full text] [doi: 10.1007/s00127-014-0857-5] [Medline: 24632847]
53. Huang R, Yan C, Tian Y, Lei B, Yang D, Liu D, et al. Effectiveness of peer support intervention on perinatal depression: a systematic review and meta-analysis. J Affect Disord 2020 Nov 01;276:788-796. [doi: 10.1016/j.jad.2020.06.048] [Medline: 32738663]
54. De Digitale Poli. URL: https://digitalepoli.nl/ [accessed 2022-10-23]
55. Boevink W, Kroon H, Giesen F. De Nederlandse empowerment vragenlijst. Trimbos-instituut. Utrecht, The Netherlands: Trimbos-instituut; 2008. URL: https://www.trimbos.nl/docs/555fbd0c-384d-4ed8-b597-b3fe35cb03eb.pdf [accessed 2019-01-10]

56. Boevink W, Kroon H, Delespaul P, Van Os J. Empowerment according to persons with severe mental illness: development of the Netherlands empowerment list and its psychometric properties. OJPsych 2017 Jan;7(1):18-30. [doi: 10.4236/ojpsych.2017.71002]
57. Krijnen-de Bruin E, Draisma S, Muntingh AD, Evers A, van Straten A, Sinnema H, et al. Self-management in anxiety and depression: a psychometric evaluation of a questionnaire. Front Psychol 2021 Nov 15;12:694583 [FREE Full text] [doi: 10.3389/fpsyg.2021.694583] [Medline: 34867583]
58. Zoun MH, Koekkoek B, Sinnema H, Muntingh AD, van Balkom AJ, Schene AH, et al. Effectiveness and cost-effectiveness of a self-management training for patients with chronic and treatment resistant anxiety or depressive disorders: design of a multicenter randomized controlled trial. BMC Psychiatry 2016 Jul 07;16:216 [FREE Full text] [doi: 10.1186/s12888-016-0927-1] [Medline: 27388878]
59. Beck AT, Steer RA, Brown GK. Manual for the Beck Depression Inventory-II. San Antonio, TX, USA: Psychological Corporation; 1996.
60. von Glischinski M, von Brachel R, Hirschfeld G. How depressed is "depressed"? A systematic review and diagnostic meta-analysis of optimal cut points for the Beck Depression Inventory revised (BDI-II). Qual Life Res 2019 May;28(5):1111-1118. [doi: 10.1007/s11136-018-2050-x] [Medline: 30456716]
61. Wang YP, Gorenstein C. Psychometric properties of the Beck Depression Inventory-II: a comprehensive review. Braz J Psychiatry 2013;35(4):416-431. [doi: 10.1590/1516-4446-2012-1048] [Medline: 24402217]
62. Üstün T, Kostanjsek N, Chatterji S, Rehm J, World Health Organization. Measuring health and disability : manual for WHO Disability Assessment Schedule (WHODAS 2.0). World Health Organization. 2010. URL: https://apps.who.int/iris/handle/10665/43974 [accessed 2021-09-10]
63. Federici S, Bracalenti M, Meloni F, Luciano JV. World Health Organization disability assessment schedule 2.0: an international systematic review. Disabil Rehabil 2017 Nov;39(23):2347-2380. [doi: 10.1080/09638288.2016.1223177] [Medline: 27820966]
64. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. 2021. URL: https://www.r-project.org/ [accessed 2021-12-12]
65. Bates D, Maechler M, Bolker B, Walker S. lme4: Linear Mixed-Effects Models using 'Eigen' and S4. The Comprehensive R Archive Network. 2022 Jul 8. URL: https://cran.r-project.org/web/packages/lme4/index.html [accessed 2022-08-08]
66. Twisk JW, Rijnhart JJ, Hoekstra T, Schuster NA, Ter Wee MM, Heymans MW. Intention-to-treat analysis when only a baseline value is available. Contemp Clin Trials Commun 2020 Nov 26;20:100684 [FREE Full text] [doi: 10.1016/j.conctc.2020.100684] [Medline: 33319119]
67. Lenth RV. emmeans: Estimated Marginal Means, aka Least-Squares Means. The Comprehensive R Archive Network. 2022 Sep 8. URL: https://tinyurl.com/brkc4uvs [accessed 2022-09-25]
68. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd edition. New York, NY, USA: Routledge; 2013.
69. Maechler M, Rousseeuw P, Struyf A, Hubert M. Finding Groups in Data: Cluster Analysis Extended Rousseeuw et al. The Comprehensive R Archive Network. 2022 Aug 22. URL: https://cran.r-project.org/web/packages/cluster/cluster.pdf [accessed 2022-09-30]
70. Davis JA, Ohan JL, Gibson LY, Prescott SL, Finlay-Jones AL. Understanding engagement in digital mental health and well-being programs for women in the perinatal period: systematic review without meta-analysis. J Med Internet Res 2022 Aug 09;24(8):e36620 [FREE Full text] [doi: 10.2196/36620] [Medline: 35943773]
71. Piotrowska PJ, Tully LA, Lenroot R, Kimonis E, Hawes D, Moul C, et al. Mothers, fathers, and parental systems: a conceptual model of parental engagement in programmes for child mental health-connect, attend, participate, enact (CAPE). Clin Child Fam Psychol Rev 2017 Jun;20(2):146-161 [FREE Full text] [doi: 10.1007/s10567-016-0219-9] [Medline: 27914017]
72. Solomon P. Peer support/peer provided services underlying processes, benefits, and critical ingredients. Psychiatr Rehabil J 2004;27(4):392-401. [doi: 10.2975/27.2004.392.401] [Medline: 15222150]
73. Lueck JA. Respecting the 'stages' of depression: considering depression severity and readiness to seek help. Patient Educ Couns 2018 Jul;101(7):1276-1282. [doi: 10.1016/j.pec.2018.02.007] [Medline: 29475726]
74. Lee JT, Yang MC, Rim HC. Discovering high-quality threaded discussions in online forums. J Comput Sci Technol 2014 May 17;29(3):519-531. [doi: 10.1007/s11390-014-1446-5]
75. Riessman F. The "Helper" therapy principle. Soc Work 1965 Apr;10(2):27-32 [FREE Full text] [doi: 10.1093/sw/10.2.27]
76. Gmeinwieser S, Kuhlencord M, Ruhl U, Hagmayer Y, Probst T. Early developments in general change mechanisms predict reliable improvement in addition to early symptom trajectories in cognitive behavioral therapy. Psychother Res 2020 Apr;30(4):462-473. [doi: 10.1080/10503307.2019.1609709] [Medline: 31038052]
77. van Gestel-Timmermans H, Brouwers EP, van Assen MA, van Nieuwenhuizen C. Effects of a peer-run course on recovery from serious mental illness: a randomized controlled trial. Psychiatr Serv 2012 Jan;63(1):54-60. [doi: 10.1176/appi.ps.201000450] [Medline: 22227760]
78. Sibitz I, Amering M, Unger A, Seyringer ME, Bachmann A, Schrank B, et al. The impact of the social network, stigma and empowerment on the quality of life in patients with schizophrenia. Eur Psychiatry 2011 Jan;26(1):28-33. [doi: 10.1016/j.eurpsy.2010.08.010] [Medline: 21036554]

79. Tjaden C, Mulder CL, den Hollander W, Castelein S, Delespaul P, Keet R, et al. Effectiveness of resource groups for improving empowerment, quality of life, and functioning of people with severe mental illness: a randomized clinical trial. JAMA Psychiatry 2021 Dec 01;78(12):1309-1318 [FREE Full text] [doi: 10.1001/jamapsychiatry.2021.2880] [Medline: 34643679]
80. Borghouts J, Eikey E, Mark G, De Leon C, Schueller SM, Schneider M, et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. J Med Internet Res 2021 Mar 24;23(3):e24387 [FREE Full text] [doi: 10.2196/24387] [Medline: 33759801]
81. Pan W, Feng B, Shen C. Examining social capital, social support, and language use in an online depression forum: social network and content analysis. J Med Internet Res 2020 Jun 24;22(6):e17365 [FREE Full text] [doi: 10.2196/17365] [Medline: 32579125]
82. Wright K. Perceptions of on‐line support providers: an examination of perceived homophily, source credibility, communication and social support within on‐line support groups. Commun Q 2000 Jan;48(1):44-59. [doi: 10.1080/01463370009385579]
83. Committee Scientific Integrity, Data Storage and Reproducibility, Deans of Social Sciences in the Netherlands. Guidelines for the archiving of academic research for faculties of Behavioural and Social Sciences in the Netherlands Version 2.1. Radboud University. 2018 Apr. URL: https://www.ru.nl/publish/pages/826378/engels_-_richtlijn_archivering_wetenschappelijk_onderzoek_voor_nederlandse_faculteiten_maatschappij-.pdf [accessed 2019-09-30]

Abbreviations

ASAD: Dutch Assessment of Self-management in Anxiety and Depression questionnaire
BDI-II: Beck Depression Inventory-II
DC: Depression Connect
NEL: Netherlands Empowerment List
PSI: peer support intervention
RCT: randomized controlled trial
WHODAS: World Health Organization Disability Assessment Schedule 2.0

Edited by A Mavragani; submitted 27.05.22; peer-reviewed by E Banwell, T Boemo, T Xiong; comments to author 09.07.22; revised version received 01.09.22; accepted 06.09.22; published 02.11.22.

Please cite as:
Smit D, Vrijsen JN, Broekman T, Groeneweg B, Spijker J
User Engagement Within an Online Peer Support Community (Depression Connect) and Recovery-Related Changes in Empowerment: Longitudinal User Survey
JMIR Form Res 2022;6(11):e39912
URL: https://formative.jmir.org/2022/11/e39912
doi:10.2196/39912
PMID:36322110

Original Paper

The Use of 2 e-Learning Modalities for Diabetes Education Using Facebook in 2 Cities of Argentina During the COVID-19 Pandemic: Qualitative Study

Daniela Luz Moyano[1], BSc, MSc; María Victoria Lopez[1], BSc; Ana Cavallo[1], BSc; Julia Patricia Candia[2], MD; Aaron Kaen[3], MD; Vilma Irazola[1], MD, MSc; Andrea Beratarrechea[1], MD, MSc

[1]Department of Research on Chronic Diseases, Institute for Clinical Effectiveness and Health Policy, Buenos Aires, Argentina

[2]Ministry of Public Health of the Province of Chaco, Resistencia, Argentina

[3]Ministry of Health of the Province of La Rioja, La Rioja, Argentina

Corresponding Author:
Daniela Luz Moyano, BSc, MSc
Department of Research on Chronic Diseases
Institute for Clinical Effectiveness and Health Policy
Dr Emilio Ravignani 2024
Buenos Aires, C1414CPV
Argentina
Phone: 54 114777 8767
Email: dmoyano@iecs.org.ar

Abstract

Background: The COVID-19 pandemic and the confinement that was implemented in Argentina generated a need to implement innovative tools for the strengthening of diabetes care. Diabetes self-management education (DSME) is a core element of diabetes care; however, because of COVID-19 restrictions, in-person diabetes educational activities were suspended. Social networks have played an instrumental role in this context to provide DSME in 2 cities of Argentina and help persons with diabetes in their daily self-management.

Objective: The aim of this study is to evaluate 2 diabetes education modalities (synchronous and asynchronous) using the social media platform Facebook through the content of posts on diabetes educational sessions in 2 cities of Argentina during the COVID-19 pandemic.

Methods: In this qualitative study, we explored 2 modalities of e-learning (synchronous and asynchronous) for diabetes education that used the Facebook pages of public health institutions in Chaco and La Rioja, Argentina, in the context of confinement. Social media metrics and the content of the messages posted by users were analyzed.

Results: A total of 332 messages were analyzed. We found that in the asynchronous modality, there was a higher number of visualizations, while in the synchronous modality, there were more posts and interactions between educators and users. We also observed that the number of views increased when primary care clinics were incorporated as disseminators, sharing educational videos from the sessions via social media. Positive aspects were observed in the posts, consisting of messages of thanks and, to a lesser extent, reaffirmations, reflections or personal experiences, and consultations related to the subject treated. Another relevant finding was that the educator/moderator role had a greater presence in the synchronous modality, where posts were based on motivation for participation, help to resolve connectivity problems, and answers to specific user queries.

Conclusions: Our findings show positive contributions of an educational intervention for diabetes care using the social media platform Facebook in the context of the COVID-19 pandemic. Although each modality (synchronous vs asynchronous) could have differential and particular advantages, we believe that these strategies have potential to be replicated and adapted to other contexts. However, more documented experiences are needed to explore their sustainability and long-term impact from the users' perspective.

(JMIR Form Res 2022;6(11):e38862) doi:10.2196/38862

KEYWORDS

COVID-19; social media; diabetes mellitus; public health; qualitative research; COVID-19 pandemic; teaching and learning settings; online learning; eHealth literacy

Introduction

According to data from the International Diabetes Federation, there were 1 in 11 (32 million) adults with diabetes [1] in South and Central America in 2021. In Argentina, the prevalence of diabetes significantly increased from 8.4% to 12.7%, corresponding to a 51% rise during the 2005-2018 period [2].

Since the beginning of the COVID-19 pandemic, operational guidelines for maintaining essential health services focused on increasing awareness of persons with diabetes about their susceptibility to COVID-19 and ways to reduce transmission, increasing home supplies of medication, and modifying routine clinical visits to videocalls or phone calls, telehealth, or email [3,4]. As part of the health care system reorganization, the Ministry of Health implemented a telehealth network in the country [5] but it was mainly focused on the diagnosis and follow-up of persons with COVID-19 and was not widely implemented in public primary care clinics to provide care for underserved persons with diabetes or other chronic care conditions.

Persons with diabetes in Argentina faced different barriers to access to chronic care, which included total confinement, health care system reorganization, and the sense among patients that seeking care implied an increased risk of COVID-19 infection. In addition, the adoption of healthy habits was difficult to achieve due to confinement [6].

Diabetes self-management education (DSME) is a critical element of diabetes care as it improves patient outcomes and quality of life and prevents complications [7]. Prepandemic, only 2 of 10 patients had ever participated in either a DSME session or a workshop, although it is included as a core care element by the diabetes clinical practice guidelines (DCPG) [8].

Social media use increased during the pandemic as a new dynamic way to share information in a rapid and accessible way. In Argentina, in 2020, nearly 85.5% of people aged 4 years or more had access to and used the internet [9], and 9 of 10 internet users were on social media, Facebook being 1 of the most frequently used platforms (86%) [10]. Social media posts and reliable information can promote health and well-being, health awareness, and preventive behaviors [11,12], helping acknowledge chronic conditions and guiding the public on how to manage and contain them.

The use of social media for diabetes self-care is still incipient [13]. Although some studies have shown that social media has beneficial effects in diabetes knowledge, attitudes, and self-care activities, results regarding patient outcomes are still mixed [14-16], especially for type 2 diabetes. Because of the lockdown and restrictions due to the COVID-19 pandemic in the country [17], we implemented social media interventions to provide DSME in 2 cities of Argentina to help persons with diabetes in their daily self-management.

The aim of this study was to evaluate 2 diabetes education modalities (synchronous and asynchronous) using the social media platform Facebook through the content of posts on diabetes educational sessions in 2 cities of Argentina during the COVID-19 pandemic.

Methods

Design and Study Setting

A qualitative approach was used to explore an intervention to provide diabetes education adapted to the context of confinement during the first year of the COVID-19 pandemic in the provinces of La Rioja and Chaco, Argentina [18,19], via the social network Facebook.

The Facebook public figure pages of 2 health institutions, the Diabetes and Nutrition Service of the Julio C. Perrando Hospital in Chaco and the Directorate of Chronic Non-Communicable Diseases in La Rioja, were used to disseminate educational sessions. This initiative was carried out jointly by the Institute for Clinical Effectiveness and Health Policy and the Diabetes Programs of the provinces of Chaco and La Rioja.

Structure of the Diabetes Education Intervention

Diabetes educational sessions were conducted using 2 modalities, synchronous and asynchronous, using the Facebook social media platform. Diabetes educator–led sessions included the topics listed in Table 1.

For the transmission and dissemination of the sessions, we used the Facebook pages of 2 health institutions in Chaco and La Rioja. In the latter case, educational sessions were also disseminated through the Facebook pages of 7 primary health care clinics.

Diabetes educational sessions were directly aimed at users of both pages and open to anyone interested in diabetes.

Table 1. Topics and content of the sessions (N=9) according to the modality implemented.

Topics	Content examples	Asynchronous modality sessions (n=4, 44%), n (%)	Synchronous modality sessions (n=5, 56%), n (%)
Diabetes care during confinement	Diabetes and obesity, prevention during the pandemic, hand hygiene practices, cleaning and disinfection to reduce transmission of COVID-19, and general recommendations for healthy eating, mental health, and physical activity	1 (25)	1 (20)
Diabetic foot	Warning signs, foot examination, foot care, red flags for consultation, treatment, and complications	1 (25)	1 (20)
Insulin management	Facts and myths about insulin therapy, barriers to insulin use, beliefs, safe and effective use of insulin, and psychological aspects in insulin therapy during confinement	1 (25)	1 (20)
Healthy eating and diabetes	Recommendations of the Dietary Guidelines for the Argentine Population, homemade healthy food and portion control, meal preparation, and facts and myths about food in diabetes management	1 (25)	N/A[a]
Physical activity and diabetes	Physical activity goals during confinement, physical activity recommendations in type 1 and type 2 diabetes, physical activity prescription, nutrition in physical activity, and management of hypoglycemia	N/A	1 (20)
Mental health, diet, and diabetes	How emotions can affect us, food management at home, role of hunger and satiety in weight management, food craving, dealing with emotions, importance of sleep, psychological support, proper nutrition in the pandemic, and dietary recommendations	N/A	1 (20)

[a]N/A: not applicable.

Synchronous Modality

This modality was implemented in the province of Chaco. For this strategy, in addition to Facebook, we used Streamyard [20], a live streaming platform that allowed multiple diabetes educators to participate in the streaming. The team that conducted the workshop comprised health educators and a moderator. To market the activity, we designed an informative flyer and posted it on the Facebook page. We also shared this post through WhatsApp groups with persons with diabetes and health care providers from primary care clinics a week prior to the event. The educational sessions were held between July 30 and October 29, 2020 (Table 1).

Sessions took between 45 and 60 minutes and were structured in 3 parts. In the first part, the moderator introduced the diabetes educators and explained the activity. In the second part, the diabetes educators presented a topic included in the educational diabetes curriculum of the DCPG (Table 1), which was adapted to the context [8]. This was followed by a question-and-answer segment, during which the diabetes educators responded to the attendees' questions collected by the moderator. Finally, the workshop live-streaming recording was posted on the Diabetes and Nutrition Service of the Julio C. Perrando Hospital Facebook pages, and diabetes educators got an opportunity to review and answer further questions and comments that were posted by the audience afterward.

Asynchronous Modality

The asynchronous modality was developed in the province of La Rioja, and educational sessions were implemented from August 31 to November 30, 2020. Some of the sessions were developed using the educators' computers and others using a mobile phone and were recorded using the Zoom platform [21], a cloud-based video communications app. The diabetes educator, a health care professional, recorded a 30-minute session on a topic of interest (Table 1). The marketing of these sessions was carried out through informative flyers that were posted on Facebook and shared on WhatsApp groups a week prior to the event. On the date of the educational session, the video was released on the host's Facebook page. Afterward, the diabetes educator also functioned as a moderator, reviewing and responding to the participants' comments, questions, or doubts posted in the video. In addition to using this modality, we shared the link with 7 public primary care clinics involved in the diabetes program that had Facebook pages. Primary care clinics, through their social networks, became amplifiers.

Units of Analysis

The units of analysis were the messages posted during the synchronous and asynchronous sessions implemented.

In addition, we analyzed social media metrics, such as impression (number of views) and engagement (eg likes, comments, and shares) in both modalities [22]. These data were used to contextualize the qualitative results.

Dimensions and Categories of Analysis

Dimension: Social Media Metrics

This dimension had the following categories [22]:

- Category: Impression is the number of times the content was on a person's screen (indicator: number of views).
- Category: Engagement relates to the number of interactions with a post, and it is important when you want an engaged audience to take some form of action after reading your post (indicators: number of likes, number of comments, and number of shares).

Dimension: Role of the Moderator

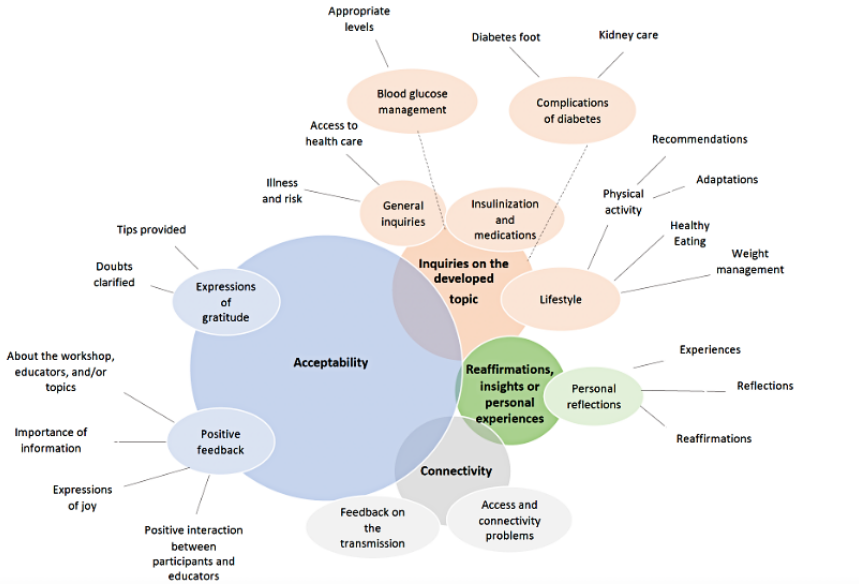
From a pedagogical point of view, the role of the moderator in information and communication technology (ICT)–mediated education acquired a crucial role; sometimes, the moderator was also the expert/educator, although these could also be separate roles [23].

Dimension: Role of the Participant

Participants were the Facebook account audience of the health institutions. Based on the analysis of the content of the posts made in each educational session, an analytical framework was developed. Four subdimensions were explored. Some similar dimensions were considered from a previous study [24].

From the analysis of the messages posted, we developed a qualitative framework (Figure 1). We included all the dimensions that emerged from the collected data regardless of the number of times they appeared. The framework reflects the preponderance that each principal subdimension had according to the frequency at which each subdimension appeared.

Figure 1. Emerging framework for the role of the participant.



Data Collection

Data collection was carried out by members of the research team through a systematic search for information on the Facebook pages of the participating institutions. It was based on the posted messages, likes, and views.

Quantitative data on social media metrics were collected weekly, including information from September 29 to December 29, 2020. Both qualitative and quantitative information fed the online matrix developed for this study.

Data Analysis

Quantitative data from social metrics indicators were analyzed using means (SDs) per educational session. Analyses were conducted using Stata 14.0 [25].

We conducted an active analysis on Facebook [18]. All wall-posted content during the sessions was analyzed through content analysis [26]. In the first stage, a process of familiarization with the information was carried out. The research team was part of the audience (without active participation) to understand the context. The second stage consisted of a first analysis of the data set before coding by a qualitative researcher when a first approximation to the dimensions and categories was elaborated.

In the third stage, data were extracted manually, a content review of the input was generated, and a new, in-depth review of all units of analysis was conducted. Subsequently, texts, visual elements, phrases, words, and emoticons were transcribed, classified, and coded. From the selected dimensions, interpretive and emerging categories were generated. To increase the internal validity of the study, 2 researchers carried out triangulation [27].

The written transcripts were entered into ATLAS.ti version 6.2 [28] and combined with the manual technique of information coding. As part of the analysis, representative direct quotations were selected and included in this paper to illustrate our findings. Messages and social metrics indicator data were reviewed at least twice by the same reviewer to improve the data quality of this study.

Ethical Considerations

Ethical approval was obtained in Chaco through the Hospital Pediátrico Avelino Castelán (Avelino Castelán Paediatric Hospital) and in La Rioja through the Consejo Provincial de Ética de las Investigaciones en Salud/CoPeis (Provincial Council of Health Research Ethics/CoPeis) (No 001). All data extracted were anonymized before being analyzed.

Results

In total, 9 educational sessions were conducted, 5 (56%) through the synchronous modality and 4 (44%) through the asynchronous modality.

Social Metrics in the Synchronous Modality

In the impression category, the mean number of views per educational session was 936.4 (SD 327.0). The total number of views of all educational sessions was 4682.

In the engagement category, the mean number of likes, comments, and shares per educational session were 51.2 (SD 10.9), 49.4 (SD 18.1), and 18.0 (SD 9.7), respectively.

Social Metrics in the Asynchronous Modality

In the impression category, the overall mean number of views was 1618.3 (SD 728.8) and the total number of views of all educational sessions was 6473.

In the engagement category, the mean number of likes, comments, and shares per educational session were 37.8 (SD 6.7), 13.8 (SD 19.7), and 36.8 (SD 14.9), respectively.

In addition to this modality, a dissemination/amplification process was also carried out through the Facebook pages of 7 primary care clinics. There was a mean increase of 109.3% (1618.2 to 3386.7) in the number of views, 238.4% (37.7 to 127. 7) in the number of likes, 52.7% (13.7 to 21) in the number of comments, and 132.0% (36.7 to 85.2) in the number of shares.

Post Content Analysis

A total of 332 wall posts corresponding to the 9 educational sessions were analyzed. Participant comments accounted for 90.1% (299/332), while educators/moderators' comments corresponded to 9.9% of all comments over the study period.

The results obtained from the analysis of the content of the comments posted are presented next and were organized into 2 main dimensions: the moderator role and the participant role. In the first case, since differences were found in emerging dimensions, the results are presented separately for the synchronous and asynchronous modalities, while we grouped the results for the participant role as no significant differences were observed during this process.

Moderator Role

In total, 4 subdimensions of post analysis emerged from this study: motivation for participation, resolution of connectivity problems, answers to specific user queries, and feedback to the expert/educator.

In the synchronous modality, there was a greater presence of the moderator's role and most of the messages were based on the "motivation for participation" subdimension. There were also messages with guidance for attendees to solve "connectivity problems," although these were few. Some of the messages were as follows:

> *Share myths and beliefs you know about Insulin and its use.*
>
> *Yes, of course. Leave your doubts and comments and we will share them.*
>
> *We look forward to your questions and concerns!!!*
>
> *Hello (attendee's name), try logging out and in again.*
>
> *Good afternoon, it's probably your internet connection because it looks and sounds fine.*

Messages from the moderator included in the "answers to specific user queries" subdimension (eg, facilitate appointment scheduling to the hospital, respond to queries on the content, market upcoming workshops) were observed in both synchronous and asynchronous modalities, although less frequently in the latter.

In addition, during the educational sessions, the expert/educator provided feedback to address participants' concerns. Some messages related to this dimension were as follows:

> *That should be answered on an individualized basis; you should schedule an in-person visit.*
>
> *(Attendee's name) definitely not. You have to look for other strategies to cope with stress, never food.*
>
> *Yes, diabetes can produce that dryness in your feet, and it is important that you moisturize them and control them every day, because if a lesion appears, you have to make the medical consultation immediately...*

Participant Role

Messages posted were short texts (256/299, 85.6%) with or without the inclusion of emoticons. In both modalities, 30 (10.0%) of the 299 messages were based only on emoticons/GIFs, such as clapping hands or smiley faces expressing joy. Some of the messages posted were participants' medical concerns that were later addressed during the educational sessions.

We did not find posts of images, videos, or links to other web pages. Although they were infrequent, tags to other Facebook users were observed that included either other people or primary care clinics (13/299, 4.3%). Table 2 presents the results of the analysis of the 4 subdimensions and their respective emerging categories from the participant role. The following new emerging subdimensions were identified: (1) inquiries on the topic developed, (2) connectivity, (3) acceptability, and (4) reaffirmations, insights, or personal experiences.

Table 2. Subdimensions, categories, descriptions, and illustrative messages on the message profiles in the synchronous and asynchronous modalities.

Subdimension and category	Description	Illustrative messages
Inquiries on the developed topic		
General inquiries	Messages about diabetes and COVID-19 risk or access to health care (inquiry about appointments)	• "What is the reason that we diabetes sufferers are a risk group?" • "...Very good!!! I am diabetic. Which are the consulting days?"
Insulinization and medications	Messages about medication, especially insulin (focused on the initiation of insulin use)	• "I would like to know why correction insulin sometimes doesn't work for me; does it make a difference if the patient starts insulin treatment directly?"
Blood glucose management	Messages about blood glucose and its appropriate levels	• "Good afternoon. What can I do? I have a blood glucose value of 110 and I would like to lower it. I am not yet diagnosed as diabetic"
Complications of diabetes	Messages related to chronic complications, kidney care, uremia, and diabetes foot	• "I would like to make a consultation. I have type 1 diabetes and chronic kidney disease on continuous ambulatory peritoneal dialysis. My feet are too dry and are have brown stains. Can it be because of increasing urea because my A1c is below 6" • "Hello, I am a type 2 diabetic. My feet are dry and itchy at night. They don't have sores or lesions"
Lifestyle	Messages about healthy eating (eg, the use of diet products, foods linked to anxiety, nuts, whole foods) and weight management, physical activity (recommended time, exercise that they can do at home, routines adapted to specific groups)	• "My mother asks: to satiate the sugar craving, are nuts or cereals with sweetener useful? • "What foods do you recommend for sugar craving when you are anxious?" • "An exercise routine to do at home" • "Regarding physical activity, e.g., walking. How long is recommended?"
Connectivity		
Access and connectivity problems	Comments on problems to access, video, connectivity, or sound	• "I can't see the whole picture" • "I can't hear well" • "Many people do not know how to connect to the diabetes talk"
Feedback on transmission	Messages related to watching educational sessions at a later time and indicating that it is heard well, whether they can post queries, and the frequency in which that the workshops will be held	• "Hi, we are watching the recorded educational session…" • "I welcome this initiative and also to have more channels of communication. The Whastapp group we have is daily and helps a lot. I wanted to know how often we are going to see you and get in touch through Facebook and through live broadcasts"
Acceptability		
Expressions of gratitude	Messages with thank-you comments for the tips provided and doubts clarified	• "Thank you for everything and especially for offering online care. Priceless advice" • "Very interesting!!! Thank you for educating us"
Positive feedback	Messages with positive feedback about the workshop, educators, or topics addressed	• "I liked it very much, very useful for those of us who live far from the Capital city and do not have a diabetologist...thank you" • "Very interesting, exposing what happens to us daily helps us to understand it better and to be able to focus in another way" • "Very good explanation. You help us a lot. I'm looking forward to the next session! • "Congratulations to the two speakers, very clear, and using simple language without loosing scientific rigor"
Reaffirmations, insights, or personal experiences		
Personal reflections	Comments of users on their own experiences, reflections, or reaffirmations about what was presented in the session (eg, importance of physical activity, diet, and emotions)	• "It was hard for me to get out of this, but I am achieving it little by little and I started walking, cycling and dancing zumba and when the pandemic started, I felt like eating all the time, but I started with my diet. Thanks to the group and the doctor and (says the name) for helping me" • "It is important that we can learn for ourselves and teach our children to express our emotions", ... we can always improve" • "… exposing what happens to us daily helps us to understand it better and be able to approach the disease in a different way"

Subdimension: Inquiries on the Developed Topic

The majority of the messages posted in this subdimension were not directly related to the topic presented during the educational session but questions and concerns participants had regarding diabetes care (Table 2).

Subdimension: Connectivity

A few messages were posted under the "connectivity problems and feedback on transmission" category, which expressed connectivity difficulties related to video or audio problems mainly in the synchronous modality (Table 2).

Subdimension: Acceptability

Under the "expressions of gratitude" category, we observed that a large portion of the messages posted during the sessions (in both synchronous and asynchronous modalities) were thank-you messages and positive feedback. Most of these messages were accompanied by emoticons/GIFs (Table 2).

In some cases, the messages consisted of joy or satisfaction or both with the educational session and expressions of gratitude toward the educators.

Under the "positive feedback" category, some of the messages expressed the importance of and the help provided by the educational sessions and asked about the possibility of continuing with these activities. In some cases, positive interaction was observed among participants and between participants and educators.

There were no negative messages or expressions of discomfort in either of the 2 modalities under analysis.

Subdimension: Reaffirmations, Insights, or Personal Experiences

To a lesser extent and only in the synchronous modality, comments related to short stories about the participants' own individual experiences, reflections, or reaffirmations related to the content of the educational sessions were observed. The statements were based on sharing the positive changes made to manage diabetes (Table 2).

Discussion

Principal Findings

Our findings showed that both modalities (synchronous and asynchronous) have their own advantages and limitations. During the implementation of the synchronous modality, we observed a greater interaction between educators and participants; the educators played a proactive role by answering questions in real time, motivating the interaction, and helping participants solve technical problems that occurred during livestream educational sessions. These particular advantages of synchronous e-learning sessions were highlighted in other studies conducted with students during the COVID-19 pandemic [29,30]. A limitation found in implementing this modality was the need to have access to good internet connectivity during live streaming. To ensure equitable access using this modality, it is important to provide asynchronous options by posting the video after live streaming in the social media platform. Connectivity barriers were also reported in other studies when using the synchronous modality [29,31].

Asynchronous e-learning provides an opportunity for learning at a time and place that are convenient and not competing with personal time or work demands [32]. This could be a possible explanation for a higher number of visualizations of the videos posted in the asynchronous modality compared to the synchronous modality, which implied an increased social reach.

In both modalities, although more frequently in the synchronous modality, participants interacted mainly with educators by posting questions and messages regarding diabetes care and COVID-19 that in another context would have been addressed during a clinic consultation with health care professionals.

Although social media platforms foster interaction between users, peer interaction was infrequent during synchronous and asynchronous educational sessions. The Association of Diabetes Care & Education Specialists pointed out the clinical, behavioral, psychological, and educational benefits of online peer support [33]. In this sense, strategies to foster peer exchange and connection through social media need to be included as part of the educational intervention.

An interesting finding was that sharing the link to primary care clinics' social media pages increased the social media reach of the proposed educational intervention as well as users' engagement.

We found that participants posted messages related to the emotional effects of living with diabetes in the context of the pandemic and the impact it had on their self-care activities. These findings are in line with other studies that highlighted the emotional impact the pandemic had on people with chronic conditions [34,35].

In both modalities, a large proportion of the messages posted expressed positive perceptions and experienced acceptability of the social media intervention to educate and support persons with diabetes and their families in a context of confinement.

To the best of our knowledge, this is the first study to analyze the content of the posts of 2 diabetes education modalities using the Facebook social media platform during the COVID-19 pandemic. In agreement with our findings, Thomas et al [36] found that the use of social media platforms (Twitter and YouTube) constituted a reliable method of education for people with diabetes during the COVID-19 pandemic. A distinctive fact that arises from our work is that we added content analysis of the posts to the social media metrics analysis.

A previous local experience before the COVID-19 pandemic showed that educational strategies based on information technologies aimed at people with diabetes in a context of vulnerability had positive contributions in different dimensions of integral health [37].

However, a greater body of knowledge is still needed regarding the use of social media to provide diabetes education in low- and middle-income countries (LMICs). A scoping review published before the pandemic by Dol et al [38] identified 414 papers that used social media for research purposes, with 1.2% (5/414) of the first authors based in LMICs and only 3 studies

focusing on patient education and care. In the same line, Hagg et al [39] explored the state of literature on the use of social media for health purposes in LMICs and found that 28 of 40 papers focused on health education. However, education campaigns using social media included themes such as tobacco use, support for HIV patients under treatment, sexual health behaviours, Middle East respiratory syndrome (MERS), malaria, and dengue but not chronic diseases.

Some authors stated that telehealth and other digital media were used safely and effectively for diabetes training and education for persons with diabetes during the pandemic [36,40,41].

The use of social networks in Latin America and the Caribbean has a great reach, with Facebook being 1 of the most popular social networks [42] and granting unique communication capabilities that can serve different educational purposes. It appears to be a promising technique for promoting health and wellness [43] and helping propagate health information [44].

However, studies based on social networks, and even more so for chronic diseases and diabetes education, are still currently limited and robust evidence is not yet available [45]. More work is needed in this field postpandemic.

Strengths and Limitations

One of the main strengths of our study was the qualitative analysis of a significant number of posts. From the analysis of the content of the posts, we developed a framework (emerging dimensions) to theorize about the contributions the use of social media for diabetes education had for the users in the context of the COVID-19 pandemic and the confinement. It would be a helpful next step to validate this framework in other contexts. Another strength to highlight is that we explored a topic with little precedent in the current literature in an LMIC that could be useful for present and future virtual educational interventions.

There are some limitations to be considered in our study. The study was conducted using public Facebook accounts, so we did not have information regarding the users who participated in educational sessions. In Argentina, there are 33.8 million people who access Facebook, most are between 25 and 34 years old, and, therefore, the older population would be excluded from diabetes education provision [46]. We are also aware that there may be a selection bias of the population under analysis as messages posted might correspond to adherent persons or people who can be considered more aware of the health topics covered. Another limitation is derived from the qualitative research itself, such as the lack of generalizability of the results to a nonpandemic context and the use of other social networks not reached in the study that might be more widely used in the future.

Implications for Practice

Based on the results and the analysis carried out, the following general recommendations are proposed for health teams and decision makers when designing and implementing educational strategies for diabetes using social media:

- When designing e-learning strategies using social media, it is necessary to analyze what type of modality best suits the needs of the target population and consider access to social networks, the type of social network, and connectivity in the target population.
- It is important to address the level of support needed by diabetes educators to plan educational sessions and develop educational materials for social networks. A synchronous e-learning modality needs more interaction and exchange between educators and users. Educators have to provide timely support to users so they can follow the educational sessions. If there are barriers to technologies or to high-speed internet, resulting in difficulties to participate, an asynchronous flexible modality should be proposed.
- Regardless of the modality chosen, a monitoring process of the posts is necessary to provide answers to users and avoid possible situations of misinformation.
- Strategies to motivate peer-to-peer interaction should be included.
- Social media reach can be enhanced by the incorporation of the social networks of health centers as disseminating/amplifying channels.
- Both modalities may have differential advantages, so a combination of synchronous and asynchronous strategies may be optimal in the framework of e-learning based on an analysis of the local context, needs, and barriers of the communities.

Conclusion

The study findings show the positive contribution of implementing e-learning strategies using both synchronous and asynchronous modalities through a social media platform in the context of the COVID-19 pandemic for diabetes care support and education.

The differential potential of each modality (synchronous or asynchronous) needs, barriers, and particularities of the users should be considered.

Further studies are needed to explore their sustainability and impact in the medium and long term and within other implementation contexts, including barriers and facilitators.

Acknowledgments

This study was funded by a grant provided by the World Diabetes Foundation (WDF 17-538).

The authors wish to acknowledge the Ministry of Health of the province of La Rioja and the Ministry of Public Health of the province of Chaco for their collaboration and support, as well as the teams of the Provincial Diabetes Program of Chaco, the Diabetes Program of La Roja, the Directorate of Chronic Non-Communicable Diseases, the primary health care centers involved (Benjamin Rincón, Faldeo del Velazco, San Vicente, La Cañada, El Libertador, Antártida Argentina, and Islas Malvinas) in La Rioja, and the Diabetes and Nutrition Service of the Julio C Perrando Hospital in Chaco.

The authors acknowledge the kind participation of the people and patients who participated in the diabetes educational sessions, including María Olivera Mores for proofreading and for the English language review.

Additional nonauthor members of the working group who contributed to the design and implementation of the diabetes educational sessions are Marta Fontela, Mariana Ojeda, Itati Roldan, Yasmine Mina, Cecilia Chandia, Mariana Gasko, Rocío Zabala, Susana Gouric, and Juliana Figueroa all from the Diabetes and Nutrition Service of the Julio C. Perrando Hospital, Chaco, Argentina, and Josefina Arnaudo, Bibiana Escudero, Yolanda Cerreño, and Marcela Nallim all from the Directorate of Chronic Non-Communicable Diseases, Ministry of Health of the Province of La Rioja.

Conflicts of Interest

None declared.

References

1. IDF. Diabetes around the World in 2021. URL: https://diabetesatlas.org/ [accessed 2022-01-02]
2. Ministry of Health and Social Development of Argentina. 4th National Survey of Risk Factors. URL: https://www.indec.gob.ar/indec/web/Nivel4-Tema-4-32-68 [accessed 2022-01-02]
3. National Ministry of Health. COVID-19 Orientaciones para el equipo de salud en el manejo de las personas con diabetes mellitus durante la pandemia. URL: https://bancos.salud.gob.ar/sites/default/files/2020-07/covid-19-orientaciones-equipo-salud-manejo-de-personas-con-diabetes-mellitus-V2.pdf [accessed 2022-01-04]
4. Word Health Organization. Maintaining Essential Health Services: Operational Guidance for the COVID-19 Context. Technical Documents. URL: https://apps.who.int/iris/handle/10665/332240 [accessed 2022-01-04]
5. Lopez E, Berlin M, Stein R, Cozzi E, Bermudez A, Mandirola Brieux H, et al. Results of the use of the teleconsultation platform after 2 months of implementation. Stud Health Technol Inform 2020 Jun 16;270:1377-1378. [doi: 10.3233/SHTI200450] [Medline: 32570667]
6. Ruiz-Roso MB, Knott-Torcal C, Matilla-Escalante DC, Garcimartín A, Sampedro-Nuñez MA, Dávalos A, et al. COVID-19 lockdown and changes of the dietary pattern and physical activity habits in a cohort of patients with type 2 diabetes mellitus. Nutrients 2020 Aug 04;12(8):2327 [FREE Full text] [doi: 10.3390/nu12082327] [Medline: 32759636]
7. Brunisholz K, Briot P, Hamilton S, Joy E, Lomax M, Barton N, et al. Diabetes self-management education improves quality of care and clinical outcomes determined by a diabetes bundle measure. J Multidiscip Healthc 2014 Nov;7:533-542. [doi: 10.2147/jmdh.s69000]
8. National Ministry of Health. National Clinical Practice Guideline on Prevention, Diagnosis and Treatment of Diabetes Mellitus Type 2 (DM2). URL: https://bancos.salud.gob.ar/sites/default/files/2020-09/guia-nacional-practica-clinica-diabetes-mellitius-tipo2_2019.pdf [accessed 2022-01-04]
9. National Institute of Statistics and Census. Access to and Use of Information and Communication Technologies. EPH. Fourth Quarter of 2020. URL: https://www.indec.gob.ar/uploads/informesdeprensa/mautic_05_213B13B3593A.pdf [accessed 2022-01-04]
10. Statista. Percentage of Users of Some Social Networks in Argentina in 2022. URL: https://es.statista.com/estadisticas/1218938/argentina-porcentaje-de-usuarios-por-red-social/ [accessed 2022-02-22]
11. Mheidly N, Fares J. Leveraging media and health communication strategies to overcome the COVID-19 infodemic. J Public Health Policy 2020 Dec;41(4):410-420 [FREE Full text] [doi: 10.1057/s41271-020-00247-w] [Medline: 32826935]
12. Li X, Liu Q. Social media use, eHealth literacy, disease knowledge, and preventive behaviors in the COVID-19 pandemic: cross-sectional study on Chinese netizens. J Med Internet Res 2020 Oct 09;22(10):e19684 [FREE Full text] [doi: 10.2196/19684] [Medline: 33006940]
13. Elnaggar A, Ta Park V, Lee SJ, Bender M, Siegmund LA, Park LG. Patients' use of social media for diabetes self-care: systematic review. J Med Internet Res 2020 Apr 24;22(4):e14209 [FREE Full text] [doi: 10.2196/14209] [Medline: 32329745]
14. Petrovski G, Zivkovic M. Are we ready to treat our diabetes patients using social media? Yes, we are. J Diabetes Sci Technol 2019 Mar;13(2):171-175 [FREE Full text] [doi: 10.1177/1932296818795441] [Medline: 30132685]
15. Kim SH, Utz S. Effectiveness of a social media-based, health literacy-sensitive diabetes self-management intervention: a randomized controlled trial. J Nurs Scholarsh 2019 Nov 17;51(6):661-669. [doi: 10.1111/jnu.12521] [Medline: 31622033]
16. Gabarron E, Årsand E, Wynn R. Social media use in interventions for diabetes: rapid evidence-based review. J Med Internet Res 2018 Aug 10;20(8):e10303 [FREE Full text] [doi: 10.2196/10303] [Medline: 30097421]
17. Ministry of Health. What Measures Is the Government Taking?. URL: https://www.argentina.gob.ar/coronavirus/medidas-gobierno [accessed 2022-01-04]
18. Franz D, Marsh HE, Chen JI, Teo AR. Using Facebook for qualitative research: a brief primer. J Med Internet Res 2019 Aug 13;21(8):e13544 [FREE Full text] [doi: 10.2196/13544] [Medline: 31411143]
19. Hernán-García M, Lineros-González C, Ruiz-Azarola A. [How to adapt qualitative research to confinement contexts]. Gac Sanit 2021 May;35(3):298-301 [FREE Full text] [doi: 10.1016/j.gaceta.2020.06.007] [Medline: 32680659]
20. StreamyYard. The Easiest Way to Create Professional Live Streams. URL: https://streamyard.com/ [accessed 2022-09-28]
21. Zoom. URL: https://zoom.us/ [accessed 2022-09-28]

22. Teoh JY, Mackenzie G, Tortolero L, Rivas JG. Social media analytics: what you need to know as a urologist. Eur Urol Focus 2020 May 15;6(3):434-436. [doi: 10.1016/j.euf.2019.08.005] [Medline: 31471218]
23. Filius R. De e-moderator: voorwaarde voor succesvol online onderwijs. (Succesful online education). Opleiding Ontwikkeling 2012;25(2):38-43.
24. Merolli M, Gray K, Martin-Sanchez F. Therapeutic affordances of social media: emergent themes from a global online survey of people with chronic pain. J Med Internet Res 2014 Dec 22;16(12):e284 [FREE Full text] [doi: 10.2196/jmir.3494] [Medline: 25533453]
25. Stata. URL: https://www.stata.com [accessed 2022-09-28]
26. Macnamara J. Media content analysis: its uses, benefits and best practice methodology. Asia-Pacific Public Relat J 2005;6(1):1-34.
27. Carter N, Bryant-Lukosius D, DiCenso A, Blythe J, Neville AJ. The use of triangulation in qualitative research. Oncol Nurs Forum 2014 Sep;41(5):545-547. [doi: 10.1188/14.ONF.545-547] [Medline: 25158659]
28. ALTAS.ti. LEARN MORE ABOUT OUR LATEST LAUNCH yesYompheandmandnyouand ganand mandjeitherrands ceitherneithercYomYoandnyoueithers dand sors dayoueithers. URL: https://atlasti.com/es/ [accessed 2022-09-28]
29. Yadav SK, Para S, Singh G, Gupta R, Sarin N, Singh S. Comparison of asynchronous and synchronous methods of online teaching for students of medical laboratory technology course: a cross-sectional analysis. J Educ Health Promot 2021;10:232 [FREE Full text] [doi: 10.4103/jehp.jehp_1022_20] [Medline: 34395669]
30. Alamer A, Alharbi F. Synchronous distance teaching of radiology clerkship promotes medical students' learning and engagement. Insights Imaging 2021 Mar 25;12(1):41 [FREE Full text] [doi: 10.1186/s13244-021-00984-w] [Medline: 33765254]
31. Sweetman DS. Making virtual learning engaging and interactive. FASEB Bioadv 2021 Jan;3(1):11-19 [FREE Full text] [doi: 10.1096/fba.2020-00084] [Medline: 33490881]
32. Sinclair P, Kable A, Levett-Jones T. The effectiveness of internet-based e-learning on clinician behavior and patient outcomes: a systematic review protocol. JBI Database Syst Rev Implement Rep 2015 Jan;13(1):52-64. [doi: 10.11124/jbisrir-2015-1919] [Medline: 26447007]
33. The Importance of Peer Support: Online and In-Person Support for People Living with Diabetes. URL: https://www.diabeteseducator.org/living-with-diabetes/Tools-and-Resources/peer-support [accessed 2022-09-28]
34. Munekawa C, Hosomi Y, Hashimoto Y, Okamura T, Takahashi F, Kawano R, et al. Effect of coronavirus disease 2019 pandemic on the lifestyle and glycemic control in patients with type 2 diabetes: a cross-section and retrospective cohort study. Endocr J 2021 Feb 28;68(2):201-210 [FREE Full text] [doi: 10.1507/endocrj.EJ20-0426] [Medline: 32999133]
35. Ammar A, Mueller P, Trabelsi K, Chtourou H, Boukhris O, Masmoudi L, ECLB-COVID19 Consortium. Psychological consequences of COVID-19 home confinement: the ECLB-COVID19 multicenter study. PLoS One 2020;15(11):e0240204 [FREE Full text] [doi: 10.1371/journal.pone.0240204] [Medline: 33152030]
36. Thomas RL, Alabraba V, Barnard S, Beba H, Brake J, Cox A, et al. Use of social media as a platform for education and support for people with diabetes during a global pandemic. J Diabetes Sci Technol 2021 Oct 31:19322968211054862 [FREE Full text] [doi: 10.1177/19322968211054862] [Medline: 34719972]
37. Moyano D, Morelli D, Santero M, Belizan M, Irazola V, Beratarrechea A. Perceptions and acceptability of text messaging for diabetes care in primary care in Argentina: exploratory study. JMIR Diabetes 2019 Mar 18;4(1):e10350 [FREE Full text] [doi: 10.2196/10350] [Medline: 30882362]
38. Dol J, Tutelman PR, Chambers CT, Barwick M, Drake EK, Parker JA, et al. Health researchers' use of social media: scoping review. J Med Internet Res 2019 Nov 13;21(11):e13687 [FREE Full text] [doi: 10.2196/13687] [Medline: 31719028]
39. Hagg E, Dahinten VS, Currie LM. The emerging use of social media for health-related purposes in low and middle-income countries: a scoping review. Int J Med Inform 2018 Jul;115:92-105 [FREE Full text] [doi: 10.1016/j.ijmedinf.2018.04.010] [Medline: 29779724]
40. Garg SK, Rodbard D, Hirsch IB, Forlenza GP. Managing new-onset type 1 diabetes during the covid-19 pandemic: challenges and opportunities. Diabetes Technol Ther 2020 Jun 17;22(6):431-439. [doi: 10.1089/dia.2020.0161] [Medline: 32302499]
41. Lian X, Dalan R, Seow CJ, Liew H, Jong M, Chew D, et al. Diabetes care during covid-19 pandemic in Singapore using a telehealth strategy. Horm Metab Res 2021 Mar;53(3):191-196. [doi: 10.1055/a-1352-5023] [Medline: 33530117]
42. Statista. Penetration Rate of Social Networks in Latin America and the Caribbean by Country in January 2021. URL: https://es.statista.com/estadisticas/1073796/alcance-redes-sociales-america-latina/ [accessed 2022-02-22]
43. Dixit S, Nandakumar G. Promoting healthy lifestyles using information technology during the COVID-19 pandemic. Rev Cardiovasc Med 2021 Mar 30;22(1):115-125 [FREE Full text] [doi: 10.31083/j.rcm.2021.01.187] [Medline: 33792253]
44. Katz M, Nandi N. Social media and medical education in the context of the covid-19 pandemic: scoping review. JMIR Med Educ 2021 Apr 12;7(2):e25892 [FREE Full text] [doi: 10.2196/25892] [Medline: 33755578]
45. Partridge SR, Gallagher P, Freeman B, Gallagher R. Facebook groups for the management of chronic diseases. J Med Internet Res 2018 Jan 17;20(1):e21 [FREE Full text] [doi: 10.2196/jmir.7558] [Medline: 29343460]
46. Statista. Distribution of Facebook Users in Argentina as of February 2022, by Age Group. URL: https://www.statista.com/statistics/282333/number-of-facebook-users-in-argentina/ [accessed 2022-09-22]

Abbreviations

DCPG: diabetes clinical practice guidelines
DSME: diabetes self-management education
LMIC: low- and middle-income country

Edited by A Mavragani; submitted 19.04.22; peer-reviewed by C Memering, E Petrosyan, L Zang; comments to author 07.09.22; revised version received 28.09.22; accepted 21.10.22; published 16.11.22.

Please cite as:
Moyano DL, Lopez MV, Cavallo A, Candia JP, Kaen A, Irazola V, Beratarrechea A
The Use of 2 e-Learning Modalities for Diabetes Education Using Facebook in 2 Cities of Argentina During the COVID-19 Pandemic: Qualitative Study
JMIR Form Res 2022;6(11):e38862
URL: https://formative.jmir.org/2022/11/e38862
doi:10.2196/38862
PMID:36322794

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org