JMIR Formative Research

Impact Factor (2022): 2.2
Volume 6 (2022), Issue 2 ISSN 2561-326X Editor in Chief: Gunther Eysenbach, MD, MPH, FACMI

Contents

Original Papers

Viewpoint

Original Paper

The Impact of Artificial Intelligence on Waiting Time for Medical Care in an Urgent Care Service for COVID-19: Single-Center Prospective Study

Kaio Jia Bin[1], MD, MBA; Adler Araujo Ribeiro Melo[2], MD; José Guilherme Moraes Franco da Rocha[3]; Renata Pivi de Almeida[1], BA; Vilson Cobello Junior[1], BEng; Fernando Liebhart Maia[1], MBA; Elizabeth de Faria[1], MD; Antonio José Pereira[1], BEng; Linamara Rizzo Battistella[4], MD, PhD; Suzane Kioko Ono[5], MD, PhD

[1]Hospital das Clinicas, Faculdade de Medicina, Universidade de Sao Paulo, Sao Paulo, Brazil

[2]Department of Internal Medicine, Federal University of Sao Paulo, Sao Paulo, Brazil

[3]Escola de Engenharia de Sao Carlos, Universidade de Sao Paulo, Sao Carlos, Brazil

[4]Faculdade de Medicina, Universidade de Sao Paulo, Sao Paulo, Brazil

[5]Department of Gastroenterology, Faculdade de Medicina, Universidade de Sao Paulo, Sao Paulo, Brazil

Corresponding Author:
Kaio Jia Bin, MD, MBA
Hospital das Clinicas
Faculdade de Medicina
Universidade de Sao Paulo
Rua Dr. Ovidio Pires de Campos 225
Sao Paulo, 05403-110
Brazil
Phone: 55 1126616208
Fax: 55 1126616200
Email: kaiobin@gmail.com

Abstract

Background: To demonstrate the value of implementation of an artificial intelligence solution in health care service, a winning project of the Massachusetts Institute of Technology Hacking Medicine Brazil competition was implemented in an urgent care service for health care professionals at Hospital das Clínicas of the Faculdade de Medicina da Universidade de São Paulo during the COVID-19 pandemic.

Objective: The aim of this study was to determine the impact of implementation of the digital solution in the urgent care service, assessing the reduction of nonvalue-added activities and its effect on the nurses' time required for screening and the waiting time for patients to receive medical care.

Methods: This was a single-center, comparative, prospective study designed according to the Public Health England guide "Evaluating Digital Products for Health." A total of 38,042 visits were analyzed over 18 months to determine the impact of implementing the digital solution. Medical care registration, health screening, and waiting time for medical care were compared before and after implementation of the digital solution.

Results: The digital solution automated 92% of medical care registrations. The time for health screening increased by approximately 16% during the implementation and in the first 3 months after the implementation. The waiting time for medical care after automation with the digital solution was reduced by approximately 12 minutes compared with that required for visits without automation. The total time savings in the 12 months after implementation was estimated to be 2508 hours.

Conclusions: The digital solution was able to reduce nonvalue-added activities, without a substantial impact on health screening, and further saved waiting time for medical care in an urgent care service in Brazil during the COVID-19 pandemic.

(JMIR Form Res 2022;6(2):e29012) doi:10.2196/29012

KEYWORDS

COVID-19; artificial intelligence; robotic process automation; digital health; health care management; pandemic; waiting time; queue; nonvalue-added activities

Introduction

Background

Artificial intelligence (AI) has arrived in the field of health care as an aurora, gradually bringing changes and innovations in medical practices. AI-based medical apps and platforms can assist physicians to make better clinical decisions such as in radiology imaging by replacing potentially subjective judgments made by the human eye [1,2]. However, pressures of cost, high expectations, uncertain benefits, a large variety of stakeholders involved, data sharing, and patient safety remain challenging obstacles in the implementation of AI in health care [3-5].

A digital solution was recently developed for the Hospital das Clínicas of the Faculdade de Medicina da Universidade de São Paulo (HCFMUSP) to assist with the reception of patients at the urgent care service. To demonstrate the value of such AI implementation in a health care service, we performed a comparative before-and-after study to understand the impact of the digital solution on the waiting time for medical care [6].

Electronic Health Records and Waiting Time

Since their initial development in the 1970s [7], the use of electronic health records (EHRs) has become increasingly common in most hospital centers worldwide with advances in the area of information technology [8]. However, use of an EHR is associated with an increase in the time needed to fill out the information [9]. Studies in this area have indicated that doctors spend more time with an EHR system than with direct patient care [10,11].

Despite general satisfaction of doctors with EHR systems [12], the increase in the time spent to use an EHR can contribute to burnout and reduce the quality of the doctor-patient relationship, resulting in a worse interaction [13,14]. Accordingly, several medical universities, especially in the United States, have been seeking strategies to reduce the time spent on EHRs [15], as a prolonged waiting time is one of the most common reasons for a patient to give up on being seen by a doctor in an emergency department [16,17]. The assessment of dissatisfaction is related to a waiting time that is twice as long as that of a fully satisfied patient [18]. Thus, in addition to increasing the patient's level of satisfaction [19,20], a reduction of waiting time also reduces patient evasion due to fatigue and frustration in the emergency room [17].

The Pandemic

With arrival of the COVID-19 pandemic in Brazil [21], the HCFMUSP established a special operation for their health care professionals with respiratory symptoms at Centro Especializado em Atendimento ao Colaborador (CEAC). Under this system, health care workers with respiratory symptoms are screened quickly to rapidly confirm or dismiss COVID-19 infection. In the first months of the pandemic, the service volume exceeded 3000 visits.

To reduce the length of stay and the face-to-face interaction between symptomatic health care professionals and the CEAC's administrative team, a digital solution was proposed to be implemented with the primary goal of saving the waiting time to receive medical care.

Motivation and Aim

To demonstrate the benefits of an AI solution in health care and to reduce the face-to-face interaction during the pandemic, the digital solution was proposed to be implemented. The aim of this study was to determine the impact caused by implementation of the digital solution in the urgent care service for health care professionals of HCFMUSP by assessing the reduction of nonvalue-added activities and its effect on the time required for nurses to screen patients and the waiting time for patients to receive medical care.

Main Research Questions

We addressed the following questions: Was the digital solution able to automate the medical care registration after nurses perform health screening in the CEAC's urgent care center? Did use of the digital solution by the health screening team increase the time of screening in CEAC's emergency department [22]? Did implementation of the AI-based digital solution reduce the waiting time to receive medical care in comparison with attendance without automation at the medical care registration?

Methods

Study Design and Setting

This was a single-center, comparative, prospective study that followed the Public Health England guide "Evaluating Digital Products for Health" [6], published in January 2020 by the British government to guide and evaluate the development or implementation of digital products in the field of health. According to the guide, a comparative "before-and-after" study model [23] was applied to evaluate the digital solution using data collected from the Queue Management System module of the CEAC's EHR system.

Population and Sample

Administrative data from the EHR between January 2020 and June 2021 at the emergency department of the CEAC were extracted and anonymized, in which each attendance was individualized according to the number of services, which is a unique entry in the EHR system.

A total of 38,042 visits were identified, 98 of which were eliminated due to inconsistent data, resulting in a database of 37,944 visits, which corresponds to 99.74% of all visits received during the study period (Table 1).

For each visit, we checked the EHR for the time record of the following events: medical care registration, nurse health screening, and medical care.

Table 1. Exclusion of data per year.

Reason for exclusion	2020 (n=25,578)	2021 (n=12,464)	Total (N=38,042)
System error on the health screening's end time record, n	0	3	3
Patient leaving before receiving medical care, n	29	2	31
Absence of health screening time record, n	36	4	40
Medical record initiated another day	7	17	24
All reasons, n (%)	72 (0.28)	26 (0.21)	98 (0.26)
Data used for analysis in this work, n (%)	25,506 (99.72)	12,438 (99.76)	37,944 (99.74)

Study Period

The digital solution was implemented over the month of June in 2020, with repeated tests and follow-up performed by the research team and the information technology team, and the solution has been considered to be 100% functional and operational since July 2020. For the before-and-after study, we considered the preimplementation period from January to May of 2020 and the postimplementation period from July 2020 to June 2021.

Variables

All variables were obtained through the EHR's Queue Management System module of the CEAC. Importantly, only released data were collected, meaning data that were released to the EHR system that cannot be changed, edited, or deleted by a user through the EHR itself.

To assess the impact of the digital solution, we analyzed the events, medical care registration, nurse's health screening, and medical care received for each visit during the study period, grouped by month.

For medical care registration automation, we checked the time record in the EHR: if the time marker was "null," this indicated that the digital solution filled in the information needed for the registration by robotic process automation (RPA).

A coefficient was created to measure the medical care registration automation, calculated based on the total medical care registration automated divided by the total number of visits per month. The time record for medical care registration in the EHR was checked, which was considered to be the runtime for calculating the time difference between the beginning and end of the registration (see Step 6 of Figure 1 and Figure 2).

Figure 1. Flowchart before digital solution implementation. Step 1: A patient arrives at the CEAC and takes a queue system number (QSN) for the electronic health record (EHR). Step 2: The CEAC health screening team (nurse) calls the next person in the screening queue. Step 3: The nurse starts the health screening in the EHR. Step 4: At end of the health screening, the nurse transfers the QSN to the registration queue. Step 5: The CEAC administrative team (Admin. Officer) calls the next person in the registration queue. Step 6: The Admin. Officer fills in the medical care registration in the EHR. Step 7: The Admin. Officer transfers the QSN to the medical care queue. Step 8: The CEAC doctor calls the next person in the medical care queue. Step 9: The doctor starts providing medical care. CEAC: Centro Especializado em Atendimento ao Colaborador.

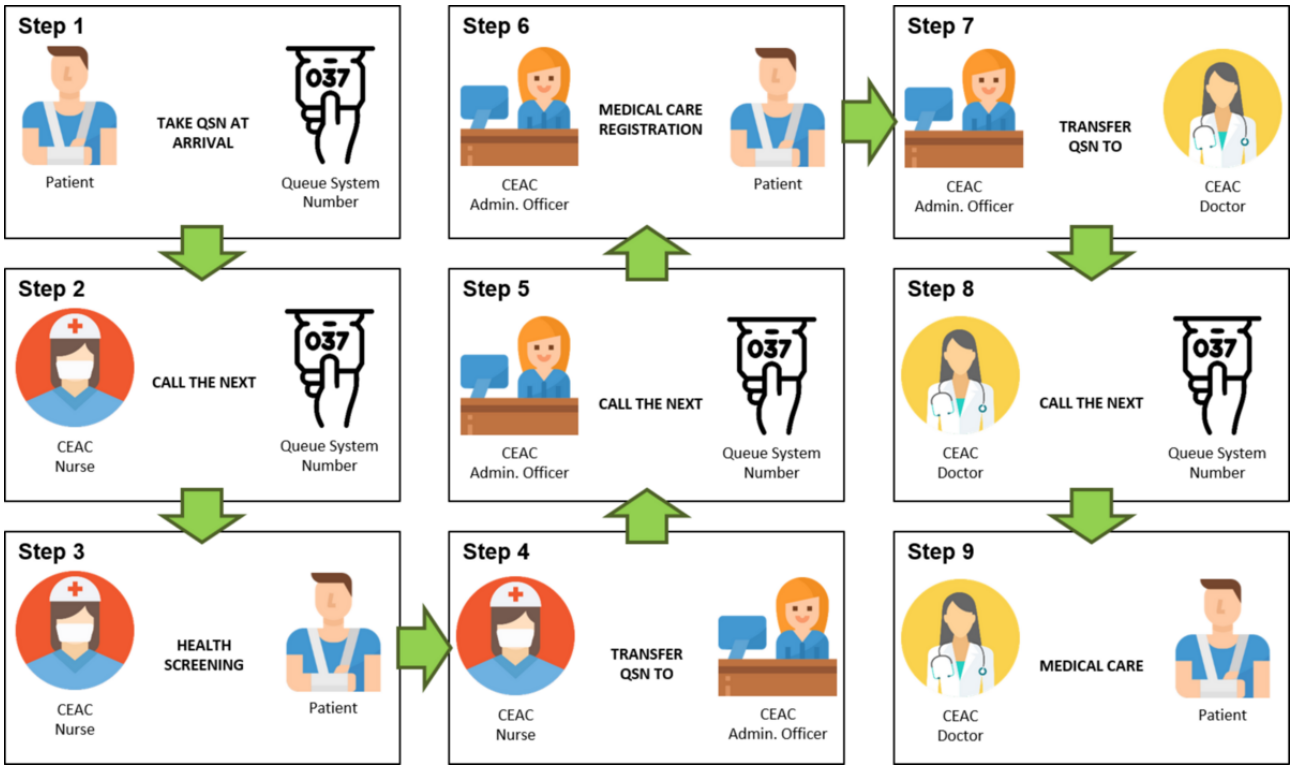
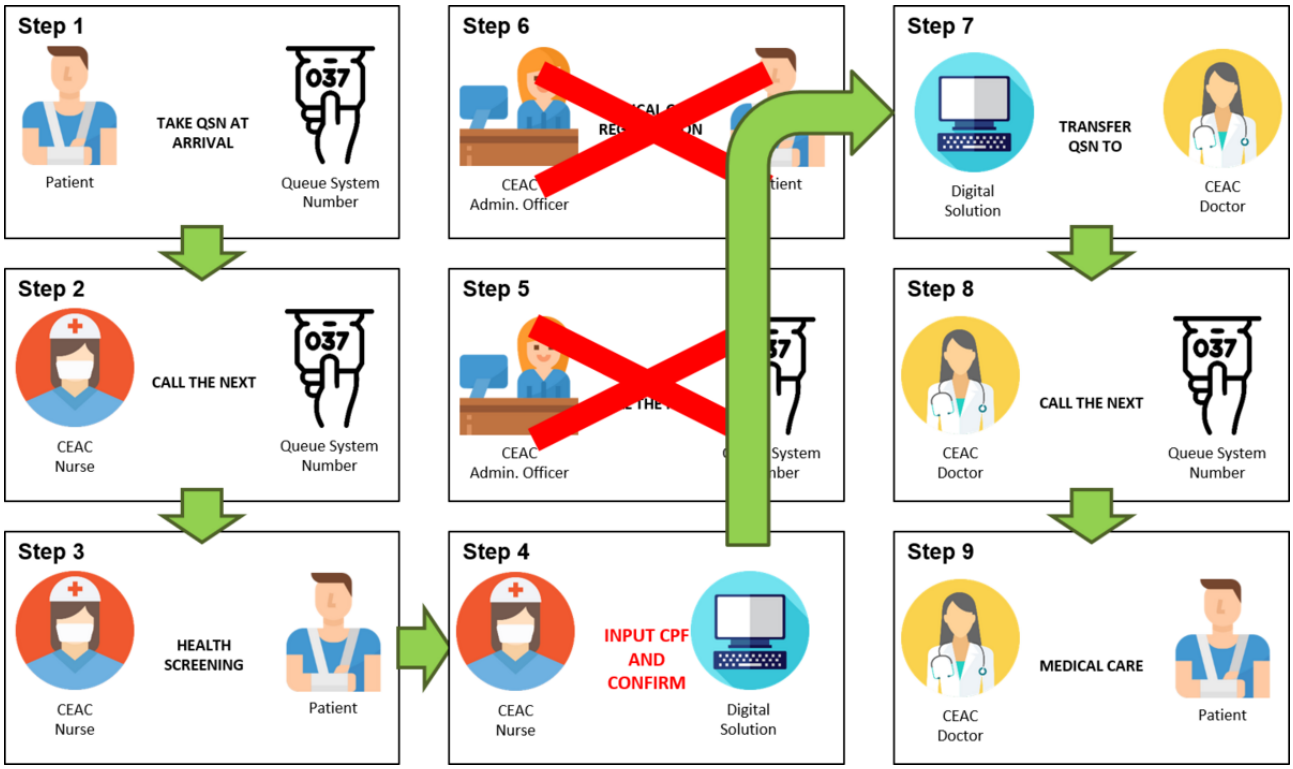


Figure 2. Flowchart after digital solution implementation. Step 1: A patient arrives at the CEAC and takes a queue system number (QSN) for the electronic health record (EHR). Step 2: The CEAC health screening team (nurse) calls the next person in the screening queue. Step 3: The nurse starts the health screening in the EHR. Step 4: At end of the health screening, the nurse inputs the CPF and QSN to the digital solution. Steps 5 to 7: The digital solution fills in the medical care registration in the EHR and transfers the QSN to the medical care queue. Step 8: The CEAC doctor calls the next person in the medical care queue. Step 9: The doctor starts providing medical care. CEAC: Centro Especializado em Atendimento ao Colaborador; CPF: Cadastro de Pessoas Físicas (natural persons number).



The waiting time for registration was used to calculate the time difference between the beginning and end of the registration for the nurse health screening stage (see Steps 3 to 6 of Figure 1 and Figure 2).

The time of screening variable was calculated as the difference in the time for nurses to perform the health screening from the beginning to the end of the screening stage (see Step 3 of Figure 1 and Figure 2).

Waiting time to medical care was based on the time record in the EHR, which was calculated as the time difference between the beginning of the medical care and the end of the nurse's health screening (Steps 3 to 9 of Figure 1 and Figure 2).

Data Collection and Analysis

The data were extracted from the EHR and compiled in Microsoft Excel 2010. Each visit was counted per month with Excel's Pivot Table function. The arithmetic average was used as a measure to assess the data per month.

The Digital Solution

All activities after the first health screening were identified, and the RPA technology was used to build a software application to replace the tasks of the CEAC's administrative team before providing medical care. After 8 weeks of development, the digital solution was finally delivered in June 2020 to the CEAC's health screening team.

Developed for a web platform, the application uses modern Node-JS, Restful, and SOAP [24] technologies to process data entries generated by the nursing team, and automates the administrative processes at the EHR, interpreting data entry without human interaction.

Two key pieces of information were considered essential for this application: (1) the queue system number, which is a key number issued by the CEAC's EHR for each patient to organize the call order that is printed on paper and displayed on the panel; and (2) the Cadastro de Pessoas Físicas (CPF; translation: Natural Persons Number), which is an 11-digit federal registration number assigned to individual Brazilian taxpayers [25].

At the end of the health screening, the nurse inputs the patient's CPF to the digital solution, and through the RPA, the digital solution searches for data linked to the entered CPF and fills in the medical care registration form with the necessary information.

As the health screening is completed before the medical care registration, all screening information is linked only to the queue system number. Therefore, with the RPA, the digital solution searches for the screening information linked to the queue system number and makes a definitive link to the patient's health record after the medical care registration. That is, if the digital solution successfully finds the patient's profile data in the EHR through the CPF, the subsequent medical care registration process will be automated and the patient will go straight to the next step, which is to be called by the doctor (Step 8 of Figure 2).

The digital solution was installed at the workstation of the nursing team responsible for the health screening, and the entire nursing team of the CEAC was trained to use the system during the month of June.

Results

Digital Solution for Medical Care Registration Automation

A total of 37,944 visits over the 18 months between January 2020 and June 2021 were analyzed. We calculated the registration automation coefficient to demonstrate the degree to which the digital solution automated the medical care registration process (Table 2).

With the digital solution, medical care registration has been automated since the month of its implementation (June 2020) in 69% of visits, ranging from a rate of 87% to 95% by month, with an RAC of 92% from July 2020 to June 2021, its first year of implementation (Table 2).

Table 2. Distribution of total health screening, total medical care registration (MCR) in person and with automation, and the corresponding registration automation coefficient (RAC).

Period	Total visits, N	MCR with automation, n	MCR with administrative team, n	RAC, %
Preimplementation				
January 2020	1322	0	1322	0
February 2020	992	0	992	0
March 2020	3361	0	3361	0
April 2020	3338	0	3338	0
May 2020	2916	0	2916	0
Sum	11,929	0	11,929	0
Implementation: June 2020	2199	1527	672	69
Postimplementation				
July 2020	2143	2043	100	95
August 2020	1915	1711	204	89
September 2020	1562	1466	96	94
October 2020	1512	1396	116	92
November 2020	2261	2031	230	90
December 2020	1985	1871	114	94
January 2021	1743	1640	103	94
February 2021	1872	1780	92	95
March 2021	2518	2276	242	90
April 2021	1780	1645	135	92
May 2021	2265	2072	193	91
June 2021	2260	1962	298	87
Sum	23,816	21,893	1923	92
Total	37,944	23,420	14,524	—[a]

[a]Not applicable.

Nurse Health Screening

In the period prior to implementation of the digital solution, the monthly average time of screening from January to May 2020 ranged from 2 minutes and 37 seconds to 3 minutes and 2 seconds, with a mean in this 5-month period of 2 minutes and 54 seconds (Table 3).

During the month of implementation of the digital solution, in June 2020, the mean time of screening was 3 minutes and 21 seconds, representing an increase of 16% compared with the average of the preimplementation period (Table 3). In the postimplementation period, the mean time of screening per month ranged from 2 minutes and 41 seconds to 3 minutes and 23 seconds, with an overall average of 3 minutes for this period (Table 3).

Table 3. Distribution by month of total health screenings and average time of screening divided by period and analysis groups.

Period	Health screenings, n	Mean time of screening (minutes:seconds)	Change from the average for the preimplementation period, %
Before automation with the digital solution			
January 2020	1322	2:58	2
February 2020	992	2:54	0
March 2020	3361	2:37	–9
April 2020	3338	3:01	4
May 2020	2916	3:02	5
Total	11,929	2:54	N/A[a]
Implementation: June 2020	2199	3:21	16
After automation with the digital solution			
July 2020	2143	3:21	16
August 2020	1915	3:23	17
September 2020	1562	3:14	12
October 2020	1512	2:55	1
November 2020	2261	2:44	–6
December 2020	1985	2:41	–7
January 2021	1743	3:03	6
February 2021	1872	2:45	–5
March 2021	2518	3:01	4
April 2021	1780	2:59	3
May 2021	2265	2:55	1
June 2021	2260	3:02	5
Total	23,819	3:00	N/A
Overall	37,944	2:59	N/A

[a]N/A: not applicable.

The variation in the arithmetic average of the time of screening from pre- to postimplementation showed an increase ranging from 12% to 17% from June to September 2020, whereas the change for the other months of the postimplementation period ranged from –7% to 5%; thus the change for the 5 months prior to the implementation ranged from –9% to 5% in relation to the average for the period from January to May 2020 (Table 3).

Waiting Time for Medical Care

With the digital solution, the medical care registration was automated for 92% of visits. This meant that the HCFMUSP patients did not need to wait for the administrative officer to call the queue service number or to wait for the registration procedure. Thus, two nonvalue-added activities were reduced before the final product of receiving medical care (Figure 2).

Table 4 compares the waiting time for medical care with automation of medical care registration through the digital solution and with that performed by an administrative officer.

Compared with the group without automation, the group with automation by the digital solution had a reduction in waiting time ranging from 5 minutes to 12 minutes and 45 seconds, with an average in the postimplementation period (July 2020 to June 2021) of 11 minutes and 51 seconds in waiting for medical care after the nurse's health screening (Table 4).

Table 4. Waiting time for medical care with and without automation from January 2020 to June 2021.

Period	With automation		Without automation		Percent change	Time reduction
	Visits, n	Mean waiting time (hours:minutes:seconds)	Visits, n	Mean waiting time (hours:minutes:seconds)		
Before implementation						
January 2020	N/A[a]	N/A	1322	0:45:46	N/A	N/A
February 2020	N/A	N/A	992	0:35:01	N/A	N/A
March 2020	N/A	N/A	3361	1:15:28	N/A	N/A
April 2020	N/A	N/A	3338	0:30:25	N/A	N/A
May 2020	N/A	N/A	2916	0:36:10	N/A	N/A
Total	N/A	N/A	2386	0:46:36	N/A	N/A
Implementation: June 2020	1527	0:13:45	672	0:24:24	N/A	N/A
After implementation						
July 2020	2043	0:20:12	100	0:31:18	–35	0:11:06
August 2020	1711	0:26:58	204	0:39:05	–31	0:12:08
September 2020	1466	0:17:33	96	0:26:51	–35	0:09:17
October 2020	1396	0:29:12	116	0:38:20	–24	0:09:08
November 2020	2031	0:41:18	230	0:52:10	–21	0:10:52
December 2020	1871	0:30:48	114	0:37:13	–17	0:06:25
January 2021	1640	0:19:28	103	0:24:28	–20	0:05:00
February 2021	1780	0:25:14	92	0:31:55	–21	0:06:42
March 2021	2276	0:32:32	242	0:39:34	–18	0:07:01
April 2021	1645	0:30:17	135	0:37:49	–20%	0:07:32
May 2021	2072	0:53:12	193	1:05:57	–19	0:12:45
June 2021	1962	0:40:40	298	0:53:07	–23	0:12:27
Total	1824	0:31:20	160	0:43:12	–27	0:11:51

[a]N/A: not applicable.

Medical Care Registration Before and After Implementing the Digital Solution

Prior to implementation of the digital solution, 100% of the medical care registrations were performed by the administrative team after they were transferred the queue system number by the nursing team (see Steps 4 to 7 in Figure 1).

In this period, the average waiting time for the medical care registration was 4 minutes and 48 seconds and the average runtime for registration was 2 minutes and 4 seconds, for an overall average 6 minutes and 52 seconds spent for the activity (Table 5).

Table 5. Arithmetic averages of waiting time and runtime of the medical care registration performed by the administrative team before and after implementation of the digital solution.

Period	Waiting time (minutes:seconds)	Runtime (minutes:seconds)	Total activity time (minutes:seconds)
Before implementation			
January 2020	3:34	1:48	5:22
February 2020	3:56	1:42	5:38
March 2020	8:21	2:12	10:32
April 2020	3:02	2:07	5:09
May 2020	3:37	2:06	5:43
Overall mean	4:48	2:04	6:52
After implementation			
July 2020	8:38	3:04	11:42
August 2020	10:55	2:41	13:36
September 2020	10:14	3:06	13:20
October 2020	8:40	3:46	12:26
November 2020	19:14	2:57	22:11
December 2020	9:40	3:18	12:58
January 2021	9:29	3:33	13:02
February 2021	9:45	3:00	12:45
March 2021	13:18	3:16	16:33
April 2021	8:51	3:35	12:26
May 2021	13:15	3:04	16:19
June 2021	13:17	2:51	16:08
Overall mean	12:11	3:07	15:18

With automation by the digital solution, the time taken to perform the medical care registration became null, as the AI by RPA could fill in the data needed for the registration instantly, whereas humans need to manually enter each letter or number by keyboard, in addition to the waiting time for the activity to be executed.

The average waiting time for the start of the registration by the administrative officer (Step 5 of Figure 1) in the postimplementation period of the digital solution was 12 minutes and 11 seconds, whereas the average time for the administrative officer to manually complete the registration was 3 minutes and 7 seconds, representing a total activity time of 15 minutes and 18 seconds (Table 5).

Overall, there was a 154% increase in the waiting time and a 51% increase in the runtime of the medical care registration performed by the administrative team after the digital solution was implemented.

The monthly averages before and after automation with the digital solution are summarized in Table 6. According to these values, the monthly time savings (monthly mean visits×12 months×total activity time for medical registration) realized by adopting the digital solution is estimated at 2507 hours, 42 minutes, and 24 seconds.

Table 6. Summary of monthly arithmetic average values before and after implementation of the digital solution with robotic process automation.

Item	Before implementation (January to May 2020)	After implementation (July 2020 to June 2021)	
		Without digital solution	With digital solution
Monthly mean number of visits	2386	160	1826
Mean time of screening (minutes:seconds)	2:54	N/A[a]	3:00
Mean waiting time for MCR[b]	4:48	12:11	N/A
Mean runtime of MCR	2:04	3:07	N/A
Total activity of MCR (waiting+runtime)	6:52	15:18	N/A
Mean waiting time for medical care	46:36	43:12	31:20

[a]N/A: not applicable.

[b]MCR: medical care registration.

Discussion

Principal Findings

In this study, we analyzed 99.7% of the visits performed at the CEAC in a period of 18 months (January 2020 to June 2021) spanning the period before, during, and after a digital solution with RPA was implemented (June 2020), which was considered 100% functional as of July 2021.

Before the solution was implemented, all (100%) health care professionals screened by the nursing team had to pass through the administrative team for medical care registration, which is an administrative task aimed at ensuring the appropriate entry of information into the EHR; although this is very important, it is an extra step in the process (Figure 2).

After implementation of the digital solution and over its first 12 months of operation, there was a 92% success rate in medical care registration automation at the end of the health screening by entering the patient's CPF number in the computer (Table 2). Therefore, this digital solution is effective for the great majority of visits.

As a result, the direct contact of a health care professional presenting to the HCFMUSP with predominantly respiratory symptoms with the CEAC administrative team was reduced, which contributed to maintaining social distancing during the pandemic. This also allowed for relocation of a portion of the administrative team responsible for the medical care registration to other activities, thus increasing the mean waiting time for the medical care registration performed by an administrative officer (Table 5).

The 8% of HCFMUSP health care professionals who needed to be registered by an administrative officer were identified by the local CEAC managers as those with an error in the CPF number in the EHR data profile. As a result, the mean runtime of medical care registration for this group was 51% higher than that of the period prior to implementation of the digital solution (Table 5), based on the need for data correction in the patient's profile.

This activity of medical care registration performed by an administrative officer is considered to be a nonvalue-added activity according to the lean manufacturing concept [26], because even without this activity, the customer reaches the final product (medical care in this case) and in a shorter time [27,28]. This digital solution with automated medical care registration will therefore help administrative officers to dedicate their time to more important activities with more human contact, compassion, and empathy.

We also found an increase in the overall time taken for the nurse's health screening with introduction of the digital solution in the first 3 months after implementation, which is consistent with results found in previous studies related to the implementation of electronic record systems in emergency care [22,29]. In the following months, as of October 2020, the mean time of screening demonstrated similar variations to the period prior to implementation, indicating a possible learning curve for the digital solution in the first 3 months (Table 3).

Moreover, with automation of the medical care registration process, the patients had a shorter wait to be seen by the doctor (Table 4). This saving in waiting time is consistent with the total time required for medical care registration performed by the administrative team (Table 5). By eliminating the nonvalue-added activity, the patient gained the time they would otherwise waste with the medical care registration process [26,27].

Considering that the activity involving medical care registration prior to implementation of the solution took an average of 6 minutes and 52 seconds (Table 6) and a monthly average of 1826 visits involved automated registration, the savings time was approximately 209 hours per month or 2508 hours in this 12-month period.

Limitations

The total time in the patient's journey was not evaluated in this study, because this could be affected by other influencing factors such as the number of nurses, administrative officers, and doctors on duty. We also did not administer a satisfaction survey to verify whether the savings in waiting time reflected a positive perception for the patient's experience.

Cost

Developing this digital solution at the CEAC took 160 hours of labor by a senior programmer, at a total final cost of Brazil real (R$) 16,000 (US $2915.18 based on an exchange rate of

US $1.00=R $5.4885 at the time of the study). The application was initially tested on a tablet and then installed on the desktop computer of the health screening room. There were no other expenses, which means that in the 12 months of operation, for each hour saved in waiting time, US $1.16 was spent.

Conclusions

The year 2020 will always be remembered as the first year of the COVID-19 pandemic. On February 3, 2020, the Ministry of Health of Brazil declared a health emergency of national importance due to human infection by the new coronavirus through Ordinance No. 188 [21], and the following month, the World Health Organization declared a global pandemic [30,31].

In the midst of the pandemic, implementation of the digital solution eliminated the need for an administrative office to register medical care at the CEAC in 92% of cases, thus reducing contact between potentially symptomatic patients and administrative staff.

The introduction of the digital solution in the nursing routine increased the health screening runtime in the first 3 months of use by around 16%, which then returned to the standard of preimplementation after the fourth month of full use.

The waiting time for medical care with automated medical care registration by the digital solution was, on average, almost 12 minutes shorter than that required when automation was not possible. This time saving of around 2500 hours fully justifies the cost and time invested in developing the solution, bringing the prospect of investments in new functionalities for the digital solution.

The same RPA could be applied to other medical activities such as helping the search for exam results or information in the EHR. Such automation can help to reduce the time a doctor must spend in front of the computer, thus providing more time available for human contact between the doctor and patient.

Acknowledgments

The authors would like to thank the entire Centro Especializado em Atendimento ao Colaborador team for welcoming and trusting in the implementation of the digital solution and Fundação Faculdade de Medicina for equipment costs and general support. This study was funded by a grant from the Brazilian Council for Development of Science and Technology (PQ308609/2018-2) to SO.

Conflicts of Interest

None declared.

References

1. Dilsizian SE, Siegel EL. Artificial intelligence in medicine and cardiac imaging: harnessing big data and advanced computing to provide personalized medical diagnosis and treatment. Curr Cardiol Rep 2014 Jan 13;16(1):441-449. [doi: 10.1007/s11886-013-0441-8] [Medline: 24338557]
2. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. Nat Rev Cancer 2018 Aug;18(8):500-510 [FREE Full text] [doi: 10.1038/s41568-018-0016-5] [Medline: 29777175]
3. Strohm L, Hehakaya C, Ranschaert ER, Boon WPC, Moors EHM. Implementation of artificial intelligence (AI) applications in radiology: hindering and facilitating factors. Eur Radiol 2020 Oct 26;30(10):5525-5532 [FREE Full text] [doi: 10.1007/s00330-020-06946-y] [Medline: 32458173]
4. Zuo KJ, Saun TJ, Forrest CR. Facial recognition technology: a primer for plastic surgeons. Plast Reconstr Surg 2019 Jun;143(6):1298e-1306e. [doi: 10.1097/PRS.0000000000005673] [Medline: 31136498]
5. Shaw J, Rudzicz F, Jamieson T, Goldfarb A. Artificial intelligence and the implementation challenge. J Med Internet Res 2019 Jul 10;21(7):e13659 [FREE Full text] [doi: 10.2196/13659] [Medline: 31293245]
6. Evaluating digital health products. UK Health Security Agency. 2020 Jan 30. URL: https://www.gov.uk/government/collections/evaluating-digital-health-products [accessed 2020-05-04]
7. Evans RS. Electronic health records: then, now, and in the future. Yearb Med Inform 2018 Mar 06;25(S01):S48-S61. [doi: 10.15265/iys-2016-s006]
8. Adler-Milstein J, Holmgren AJ, Kralovec P, Worzala C, Searcy T, Patel V. Electronic health record adoption in US hospitals: the emergence of a digital "advanced use" divide. J Am Med Inform Assoc 2017 Nov 01;24(6):1142-1148 [FREE Full text] [doi: 10.1093/jamia/ocx080] [Medline: 29016973]
9. Joukes E, Abu-Hanna A, Cornet R, de Keizer N. Time spent on dedicated patient care and documentation tasks before and after the introduction of a structured and standardized electronic health record. Appl Clin Inform 2018 Jan 17;9(1):46-53 [FREE Full text] [doi: 10.1055/s-0037-1615747] [Medline: 29342479]
10. Arndt BG, Beasley JW, Watkinson MD, Temte JL, Tuan W, Sinsky CA, et al. Tethered to the EHR: primary care physician workload assessment using EHR event log data and time-motion observations. Ann Fam Med 2017 Sep;15(5):419-426 [FREE Full text] [doi: 10.1370/afm.2121] [Medline: 28893811]
11. Sinsky C, Colligan L, Li L, Prgomet M, Reynolds S, Goeders L, et al. Allocation of physician time in ambulatory practice: a time and motion study in 4 specialties. Ann Intern Med 2016 Dec 06;165(11):753-760. [doi: 10.7326/M16-0961] [Medline: 27595430]

12. Williams DC, Warren RW, Ebeling M, Andrews AL, Teufel Ii RJ. Physician use of electronic health records: survey study assessing factors associated with provider reported satisfaction and perceived patient impact. JMIR Med Inform 2019 Apr 04;7(2):e10949 [FREE Full text] [doi: 10.2196/10949] [Medline: 30946023]
13. Street RL, Liu L, Farber NJ, Chen Y, Calvitti A, Zuest D, et al. Provider interaction with the electronic health record: the effects on patient-centered communication in medical encounters. Patient Educ Couns 2014 Sep;96(3):315-319 [FREE Full text] [doi: 10.1016/j.pec.2014.05.004] [Medline: 24882086]
14. Marmor RA, Clay B, Millen M, Savides TJ, Longhurst CA. The impact of physician EHR usage on patient satisfaction. Appl Clin Inform 2018 Jan;9(1):11-14 [FREE Full text] [doi: 10.1055/s-0037-1620263] [Medline: 29298451]
15. Chen L, Guo U, Illipparambil LC, Netherton MD, Sheshadri B, Karu E, et al. Racing against the clock: internal medicine residents' time spent on electronic health records. J Grad Med Educ 2016 Feb;8(1):39-44 [FREE Full text] [doi: 10.4300/JGME-D-15-00240.1] [Medline: 26913101]
16. Mohsin M, Forero R, Ieraci S, Bauman AE, Young L, Santiano N. A population follow-up study of patients who left an emergency department without being seen by a medical officer. Emerg Med J 2007 Mar;24(3):175-179 [FREE Full text] [doi: 10.1136/emj.2006.038679] [Medline: 17351221]
17. Spaite DW, Bartholomeaux F, Guisto J, Lindberg E, Hull B, Eyherabide A, et al. Rapid process redesign in a university-based emergency department: decreasing waiting time intervals and improving patient satisfaction. Ann Emerg Med 2002 Feb;39(2):168-177. [doi: 10.1067/mem.2002.121215] [Medline: 11823772]
18. McMullen M, Netland PA. Wait time as a driver of overall patient satisfaction in an ophthalmology clinic. Clin Ophthalmol 2013;7:1655-1660. [doi: 10.2147/OPTH.S49382] [Medline: 23986630]
19. Sun J, Lin Q, Zhao P, Zhang Q, Xu K, Chen H, et al. Reducing waiting time and raising outpatient satisfaction in a Chinese public tertiary general hospital-an interrupted time series study. BMC Public Health 2017 Aug 22;17(1):668 [FREE Full text] [doi: 10.1186/s12889-017-4667-z] [Medline: 28830400]
20. Viotti S, Cortese CG, Garlasco J, Rainero E, Emelurumonye IN, Passi S, et al. The buffering effect of humanity of care in the relationship between patient satisfaction and waiting time: a cross-sectional study in an emergency department. Int J Environ Res Public Health 2020 Apr 24;17(8):2939 [FREE Full text] [doi: 10.3390/ijerph17082939] [Medline: 32344530]
21. Portaria nº 188, de 3 de fevereiro de 2020. Brasil Ministério da Saúde. 2020 Feb 3. URL: https://www.in.gov.br/en/web/dou/-/portaria-n-188-de-3-de-fevereiro-de-2020-241408388 [accessed 2020-02-10]
22. Vartak S, Crandall DK, Brokel JM, Wakefield DS, Ward MM. Transformation of emergency department processes of care with EHR, CPOE, and ER event tracking systems. Health Inf Manag 2009;38(2):27-32. [doi: 10.1177/183335830903800204] [Medline: 19546485]
23. Before-and-after study: comparative studies. UK Health Security Agency. 2020 Jan 30. URL: https://www.gov.uk/guidance/before-and-after-study-comparative-studies [accessed 2020-05-04]
24. Soni A, Ranga V. API Features Individualizing of Web Services: REST and SOAP. Int J Innov Technol Explor Eng 2019 Aug 23;8(9S):664-671. [doi: 10.35940/ijitee.i1107.0789s19]
25. Perguntas e Respostas. Brasil Receita Federal. URL: https://receita.economia.gov.br/orientacao/tributaria/cadastros/cadastro-de-pessoas-fisicas-cpf/assuntos-relacionados/perguntas-e-respostas [accessed 2020-05-04]
26. Ikeziri LM, Melo JC, Campos RT, Okimura LI, Gobbo Junior JA. A perspectiva da indústria 4.0 sobre a filosofia de gestão Lean Manufacturing. Braz J Dev 2020;6(1):1274-1289. [doi: 10.34117/bjdv6n1-089]
27. Sanders A, Elangeswaran C, Wulfsberg J. Industry 4.0 implies lean manufacturing: Research activities in industry 4.0 function as enablers for lean manufacturing. J Ind Eng Manage 2016 Sep 21;9(3):811. [doi: 10.3926/jiem.1940]
28. Sundar R, Balaji A, Kumar RS. A review on lean manufacturing implementation techniques. Procedia Engin 2014;97:1875-1885. [doi: 10.1016/j.proeng.2014.12.341]
29. Baumann LA, Baker J, Elshaug AG. The impact of electronic health record systems on clinical documentation times: A systematic review. Health Policy 2018 Aug;122(8):827-836. [doi: 10.1016/j.healthpol.2018.05.014] [Medline: 29895467]
30. WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020. World Health Organization. URL: https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020 [accessed 2020-03-12]
31. Rolling updates on coronavirus disease (COVID-19). World Health Organization. URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen [accessed 2020-05-04]

Abbreviations

AI: artificial intelligence
CEAC: Centro Especializado em Atendimento ao Colaborador
CPF: Cadastro de Pessoas Físicas (Natural Persons Number)
EHR: electronic health record
HCFMUSP: Hospital das Clínicas of The Faculdade de Medicina da Universidade de São Paulo
RPA: robotic process automation

Edited by G Eysenbach; submitted 22.03.21; peer-reviewed by M von Wagner, YC Kato-Lin; comments to author 13.04.21; revised version received 14.05.21; accepted 22.11.21; published 01.02.22.

Please cite as:
Bin KJ, Melo AAR, da Rocha JGMF, de Almeida RP, Cobello Junior V, Maia FL, de Faria E, Pereira AJ, Battistella LR, Ono SK
The Impact of Artificial Intelligence on Waiting Time for Medical Care in an Urgent Care Service for COVID-19: Single-Center Prospective Study
JMIR Form Res 2022;6(2):e29012
URL: https://formative.jmir.org/2022/2/e29012
doi:10.2196/29012
PMID:35103611

Original Paper

Peer Support Specialists' Perspectives of a Standard Online Research Ethics Training: Qualitative Study

Karen L Fortuna[1*], LICSW, PhD; Skyla R Marceau[2*], MSW; Arya Kadakia[3*]; Sarah I Pratt[2*], PhD; Joy Varney[4*]; Robert Walker[5*], MSc; Amanda L Myers[6*], MPH; Shavon Thompson[4*]; Katina Carter[4*]; Kaycie Greene[7*]; Willie Pringle[7*]

[1]Department of Psychiatry, Geisel School of Medicine, Dartmouth College, Lebanon, NH, United States

[2]Geisel School of Medicine, Dartmouth College, Concord, NH, United States

[3]Dartmouth College, Hanover, NH, United States

[4]Centerstone Kentucky, Louisville, KY, United States

[5]Office of Recovery and Empowerment, Massachusetts Department of Mental Health, Boston, MA, United States

[6]Institute for Behavioral Health, Heller School, Brandeis University, Waltham, MA, United States

[7]Centerstone Tennessee, Clarksville, TN, United States

[*]all authors contributed equally

Corresponding Author:
Karen L Fortuna, LICSW, PhD
Department of Psychiatry
Geisel School of Medicine
Dartmouth College
46 Centerra Pkwy
Lebanon, NH, 03766
United States
Phone: 1 6037225727
Email: karen.l.fortuna@dartmouth.edu

Abstract

Background: Certified peer support specialists (CPS) have a mental health condition and are trained and certified by their respective state to offer Medicaid reimbursable peer support services. CPS are increasingly involved as partners in research studies. However, most research ethics training in the protection of human subjects is designed for people who, unlike CPS, have had exposure to prior formal research training.

Objective: The aim of this study is to explore the perspectives of CPS in completing the Collaborative Institutional Training Initiative Social and Behavioral Responsible Conduct of Research online training.

Methods: A total of 5 CPS were recruited using a convenience sample framework through the parent study, a patient-centered outcomes research study that examined the comparative effectiveness of two chronic health disease management programs for people with serious mental illness. Participants independently completed the Collaborative Institutional Training Initiative Social and Behavioral Responsible Conduct of Research online training. All participants completed 15 online modules in approximately 7-9 hours and also filled out a self-report measure of executive functioning (the Adult Executive Functioning Inventory [ADEXI]). Qualitative data were collected from a 1-hour focus group and qualitative analysis was informed by the grounded theory approach. The codebook consisted of codes inductively derived from the data. Codes were independently assigned to text, grouped, and checked for themes. Thematic analysis was used to organize themes.

Results: Passing scores for each module ranged from 81%-89%, with an average of 85.4% and a median of 86%. The two themes that emerged from the focus group were the following: comprehension (barrier) and opportunity (facilitator). Participants had a mean score of 27.4 on the ADEXI.

Conclusions: The CPS perceived the research ethics online training as an opportunity to share their lived experience expertise to enhance current research efforts by nonpeer scientists. Although the CPS completed the online research ethics training, the findings indicate CPS experienced difficulty with comprehension of the research ethics online training materials. Adaptations may be needed to facilitate uptake of research ethics online training by CPS and create a workforce of CPS to offer their lived experience expertise alongside peer and nonpeer researchers.

(JMIR Form Res 2022;6(2):e29073) doi:10.2196/29073

KEYWORDS

peer support specialists; community engagement; research ethics; mental health; peer support; codebook; online health; online training; education; ethics

Introduction

It is widely recognized that the inclusion of the insights of people with mental health conditions in psychiatric research is needed to advance mental health care [1]. Certified peer support specialists (CPS) are people with a lived experience of a mental health condition who have been trained and accredited by their state to provide Medicaid reimbursable support services [2], including peer support [3]. Peer support is a nonmanualized, nonstandardized form of social support that a person provides to others experiencing a mental health and/or chronic health condition to bring about change [2]. Peer support has been shown to increase individuals' hope, their sense of personal control, and their ability to make positive changes, and decrease psychiatric symptoms [4] through listening, sharing one's lived experiences, and role modeling [2]. The inclusion of CPS into the workforce has transformed the mental health system globally in that peer support services have the capacity to support people in between clinical encounters. As such, CPS is one of the fastest-growing sectors of the mental health workforce providing community-based services [5].

CPS are also increasingly involved in research-related activities (eg, recruitment, retention efforts, obtaining informed consent, collecting sensitive data, publishing findings in peer-reviewed journals [6]). For example, the Quality of Patient-Centered Outcomes Research Partnerships instrument was developed by researchers and CPS to offer quality improvement opportunities related to developing equitable research partnerships [7]. Moreover, community-engaged research that includes CPS in all stages of research may also lead to the dissemination of study results to a wider audience by incorporating CPS in conferences and publications as well as by using social media to showcase results in a more digestible manner for service users of the mental health system and other CPS [6].

As the influence of CPS in research grows, adequately training CPS in research ethics human subjects protection is an emerging concern [8]. Since regulations that required organizations to establish institutional review boards have been instituted, multiple online training programs in research ethics have been developed, such as the Collaborative Institutional Training Initiative (CITI). The CITI Program covers various domains, such as human subject research, and includes courses on various topics (eg, research ethics, protocol development, and safety practices). Each course includes webinars, text, and quizzes that overall take approximately 30-45 minutes to complete [9]. Such courses are designed for university researchers and students, as well as health care providers and governmental agencies, for basic accreditation [10]. Human research ethics training currently exists for special populations, including Aboriginal peoples [11], faith-based communities [12], community health workers [13], and *Promotores* (ie, Latino community mental health workers) [13]. Nevertheless, while these trainings target lay populations with limited research backgrounds, no human research ethics trainings (to our knowledge) have been designed specifically to address the cognitive challenges these trainings pose to individuals new to research who may also experience a mental health condition.

Hence, the purpose of this study was to examine the perspectives of CPS in completing the CITI Social and Behavioral Responsible Conduct of Research online training. Findings can be used to determine the feasibility of current community-engaged research training and identify opportunities for improvement.

Methods

Overview

A total of 5 CPS were hired as part of a Patient-Centered Outcomes Research Institute study. The parent study is in the process of conducting a randomized control trial (N=600 people with a lived experience of a mental health condition) to examine the impact of two chronic health disease management programs for people with serious mental illness. The outcomes of interest include change in knowledge and skills related to illness self-management, patient activation, and acute hospital events. As part of this study, researchers employed the Peer and Academic Partnership model, which has been described in detail elsewhere [6]. Briefly, CPS were hired and financially reimbursed to work as partners in the implementation of the randomized control trial detailed above. To be eligible, individuals had to meet the following criteria: (1) have lived experience of any mental health condition, (2) have a peer support certification, (3) reside in the United States, and (4) be aged ≥18 years. Prior to conducting that study, the CPS did not have prior research training or research experience. As part of the parent study, the CPS met monthly through videoconference with a trained PhD researcher (first author) to provide input on randomized control trial implementation challenges (eg, recruitment, retention). Improvements included inclusion of innovative recruitment strategies and development of new COVID-19–related intervention materials. For this study, a convenience sample framework of those CPS hired to provide patient-centered insights for the implementation of the parent study was used. The participants were hired to be part of this study if they had time available and had an interest in taking part in the research. All participants were already employees of the organization. Prior to completing the online research ethics training, the voluntary and confidential nature of this study was explained, and verbal informed consent was obtained. This study was approved by the Dartmouth College Institutional Review Board.

Participants then completed the CITI Social and Behavioral Responsible Conduct of Research online training that consisted of 15 online modules: (1) Belmont Report and Its Principles,

(2) Conflicts of Interest in Human Subjects Research, (3) Cultural Competence in Research, (4) History and Ethical Principles, (5) Defining Research with Human Subjects, (6) The Federal Regulations, (7) Assessing Risk, (8) Informed Consent, (9) Course Privacy and Confidentiality, (10) Research with Prisoners, (11) Research with Children, (12) Research in Public Elementary and Secondary Schools, (13) International Research, (14) Course Internet-Based Research, and (15) Unanticipated Problems and Reporting Requirements in Social and Behavioral Research. Participants completed the modules online independently and had their scores emailed to the first author (KLF) along with their passing certificate. The first author was present to help while they completed the training in a single session. They were financially reimbursed regardless of whether they passed the training or not.

Following the completion of the online research ethics training, a 1-hour focus group was cofacilitated by a trained PhD-level facilitator (KLF) and a masters-level CPS (JV) to collect qualitative data. An interview guide aided the cofacilitators (see Textbox 1). The PhD-level facilitator and the masters-level CPS, both of whom have extensive research experience and training, developed the interview guide. The participants were encouraged to express their views of the online research ethics training and their role as a CPS research partner. No follow-up interviews were conducted. The focus group was audio recorded and the data were transcribed. The focus group was conducted in October 2019 and lasted approximately 1 hour.

Textbox 1. Focus group interview guide.

Perspectives on the Social and Behavioral Responsible Conduct of Research training
1. What are your views on clinical research in general?
2. What are your thoughts on the content of the Social and Behavioral Responsible Conduct of Research training?
3. What are your thoughts on the quizzes in the Social and Behavioral Responsible Conduct of Research training?
4. How would you feel as a peer support specialist being part of a research team?
5. How did you feel going through the training?

Study Sample

The sample included 5 CPS from 2 states. The majority of participants were female (n=4, 80%) and African American (n=4, 80%). Ages ranged from 40-52 years, with a mean age of 45.8 years. All participants had a high school education/General Educational Development certificate and were providing services in their respective communities. Tenure in CPS positions ranged from 3 months to 4 years. None of the CPS had completed research ethics training in the past or had any research experience. As CPS are protected by the Americans with Disabilities Act, data on their diagnosis remained confidential.

Instruments

Each participant also completed the Adult Executive Functioning Inventory (ADEXI), a self-report measure that assessed their executive functioning. It is a reliable 14-item self-report scale (Table 1) assessing working memory and inhibition in adulthood using a 5-point Likert scale ranging from 1 (definitely not true) to 5 (definitely true). The summed score (ranging from 14-70) is used to assess executive functioning. Higher scores indicate greater difficulty with executive functioning.

Table 1. Adult Executive Functioning Inventory.

Number	Statement[a]
1	I have difficulty remembering lengthy instructions.
2	I sometimes have difficulty remembering what I am doing in the middle of an activity.
3	I have a tendency to do things without first thinking about what could happen.
4	I sometimes have difficulty stopping myself from doing something that I like even though someone tells me that it is not allowed.
5	When someone asks me to do several things, I sometimes remember only the first or last.
6	I sometimes have difficulty refraining from smiling or laughing in situations where it is inappropriate.
7	I have difficulty coming up with a different way of solving a problem when I get stuck.
8	When someone asks me to fetch something, I sometimes forget what I am supposed to fetch.
9	I have difficulty planning for an activity (eg, remembering to bring everything necessary when going on a trip/to work/to school).
10	I sometimes have difficulty stopping an activity that I like (eg, I watch TV or sit in front of the computer in the evening even though it is time to go to bed).
11	I sometimes have difficulty understanding verbal instructions unless I am also shown how to do something.
12	I have difficulties with tasks or activities that involve several steps.
13	I have difficulty thinking ahead or learning from experience.
14	People that I meet sometimes seem to think that I am more lively/wilder compared to other people my age.

[a]Participants are asked to circle a number to indicate how well that statement describes how they are as a person (1=definitely not true, 2=not true, 3=partially true, 4=true, and 5=definitely true).

Analytic Plan

Qualitative analysis was informed by the "grounded theory" approach [14]. The codebook included codes inductively derived from qualitative data [14]. KLF and SRM read transcribed qualitative data. Codes were independently assigned to text, grouped, and thematically analyzed to check for emerging themes. Analyses included within-group consensus or disagreement. Member checking with CPS was used to verify and/or resolve any dissimilar findings. Member checking is a method used to validate findings that involves discussing findings with respondents and examining the accuracy of findings [15].

Results

Social and Behavioral Responsible Conduct of Research Training

A total of 5 CPS independently completed the CITI Social and Behavioral Responsible Conduct of Research online training. All participants completed 15 online modules. Completion of the research ethics training took approximately 7-9 hours. Scores were calculated and collected through the CITI platform and were then emailed to KLF. Passing scores for each module ranged from 81%-89%, with an average of 85.4% and a median of 86%.

We identified a final set of 9 codes (ie, sentence structure, grade level, length, new knowledge, acronyms, cognitive load and emotional response retention, retention recommendations, lived experience expertise, and new opportunity) relating to 2 themes from the focus group. Themes related to peer support specialists' perspective of the barriers and facilitators to completing the research ethics training included the following: comprehension (barrier) and opportunity (facilitator).

Comprehension

The first theme, comprehension, included two subcategories: (1) cognitive complexity of content and quizzes and (2) learning and retention.

Cognitive Complexity of Content and Quizzes

All CPS respondents referred to the cognitive complexity of the content and quizzes as a barrier to completing the research ethics training (eg, "give me simple sentences. That was too much"). Sentence structure was multisyllabic and written at a 12th grade level (as assessed by the Flesch-Kincaid Grade Level test in Microsoft Word), which required extensive cognitive effort (eg, "The big words, the paragraph questions, I like that it was only five questions but they were just long, they were real big").

Learning and Retention

All CPS respondents referred to learning and retention as a barrier to completing the research ethics training. Most CPS respondents (n=4, 80%) indicated the need to comprehend information to retain knowledge (eg, "I couldn't even say oh that's interesting because my brain was so overwhelmed by the way the content was written?"). Unfamiliarity with terms and acronyms such as "NIH" or "PHS" led to confusion, and produced feelings of frustration and exhaustion in completing the training. CPS respondents recommended hyperlinks for unfamiliar terms to assist CPS test takers in understanding definitions of words.

Opportunity

All CPS respondents perceived research ethics education as an opportunity to share their lived experience expertise to enhance current research efforts by nonpeer scientists (eg, "If the world is open to it I think peer support specialists could have a great role in research because we have a different perspective... we come from a perspective of experience" and "this is a once-in-a-lifetime opportunity").

ADEXI

Participants had a mean executive functioning score of 27.4 (SD 7.83) on the ADEXI. Participants had a mean working memory score of 17.2 (SD 5.50) and inhibition score of 10.2 (SD 2.59).

Discussion

Principal Findings

The purpose of this study was to explore the perspectives of CPS regarding an online research ethics training. All participants completed the CITI Social and Behavioral Responsible Conduct of Research online training. The findings indicate difficulty with comprehension of research ethics online training materials. The ADEXI scale indicated that participants had little self-reported difficulty with executive functioning. This suggests that perhaps CPS involved in this study were not hindered by poor working memory and inhibition, and that difficulties emerged from research-specific complexity such as the use of jargon and the need to absorb new and unfamiliar information. Nevertheless, CPS perceived the research ethics online training as an opportunity to share their lived experience expertise to enhance current research efforts by nonpeer scientists.

All participants completed the CITI research ethics training in 7-9 hours—approximately 2.5-4.5 hours longer than the average completion time [16]. As CPS are increasingly involved in research as partners, adaptations may be needed to facilitate the uptake of online research ethics training. Adaptations can include providing information on the time commitment needed to complete these tests, multiple testing sessions on different days, technological capacity to sign in and out of training portals and begin where one had left off, and opportunities to complete testing at home.

The cognitive complexity of content and quizzes can be addressed by reducing the number of multisyllabic words and sentences that included compound subjects, which required extensive cognitive effort to comprehend. Although programs that provide research ethics training specifically for community research partners exist—for example, CIRTification: Community Involvement in Research Training—these community research partner trainings are designed for people with a high school education [17]. Although such trainings may be more digestible for CPS (as compared to the CITI training, which requires prior research experience), many CPS may not meet the educational requirement, as a high school education is not required to become a CPS in many states [18]. Existing research suggests that plain-language summaries and the use of language written at a fourth-grade level [19] could improve the readability and comprehensibility of trainings—thus, existing research ethics trainings could be improved by following plain-language summary and reading-level recommendations.

The use of acronyms such as "NIH" or "PHS" in research ethics training impacted comprehension. Making meanings explicit is an essential web-based design feature necessary to facilitate learning among people with mental health conditions [20]. Including definitions can improve comprehension by making meanings explicit. CPS respondents recommended hyperlinks in web-based applications to link test takers to definitions of words [20]. Evidence-based guidelines for this population support this finding [20]. Additional evidence-based, web-based design guidelines that may be useful in online research ethics training include a singular focus in website content, simple architecture, prominent content, and explicit navigation [20].

As they did not have prior knowledge of research processes, terminology, and the history of research ethics, CPS were learning new information. Determining the factors that influence how CPS learn may aid researchers in understanding better ways to help CPS grasp new content. Research indicates that factors that may impact the learning experience of CPS include the following: (1) life experiences, (2) work experience, (3) previous adult learning experiences [21,22], and (4) potential cognitive deficits related to mental health conditions [23]. As such, real-world examples in the context of research ethics training may facilitate learning and retention of new knowledge. In addition, people with mental health conditions may have potential cognitive impairments and need repetition to reinforce new knowledge [20]. Exploring ongoing learning methods such as "audit and feedback" or "learning collaboratives" may offer opportunities to facilitate reinforcement of research ethics training and lead to greater retention of new knowledge [20].

All CPS respondents perceived research ethics education as an opportunity to share their lived experience expertise to enhance current research efforts by nonpeer scientists.

This study is not without limitations. First, this study included a nonprobability convenience sample. As such, the study sample may not be representative of the greater population of CPS. In addition, the severity of mental health conditions among CPS is not known. Although the severity of a mental health condition could impact comprehension of CITI modules, CPS are employees of organizations, and therefore the Americans with Disabilities Act protects them from reporting on a diagnosis they have been given. As such, we did not request information related to a participant's mental health diagnosis. Second, due to the small sample size, it is not known if saturation was met. However, this study is the first to explore peer support specialists' perceptions of human research ethics training and delineates recommendations to develop an online training for this group. Third, it is not known if participants needed to take a CITI module more than once before receiving a passing score. Consequently, it would have been informative to collect the number of times a participant needed to retake a module to pass. In addition, our sample was largely from the same demographic, and it is possible that stratifying by different demographics may produce alternate outcomes. Finally, we examined only one standard online research ethics training. Other trainings for community partners may be better suited for CPS. However, to

our knowledge, standard community research partner trainings are designed for people with a high school education [17]. A research ethics training designed for community health workers [13] may potentially have higher levels of acceptability among CPS. Nevertheless, this is the first study, to our knowledge, to examine perspectives of CPS in completing an online research ethics training.

Next steps could potentially include the development of an ethics training that is tailored to research-naive individuals and those with limited education using universal design principles. Universal design is the process of creating products that are accessible to people with a wide range of abilities [24]. Based on our findings, we suggest reducing verbosity and the use of jargon. A more fundamental explanation of the importance of research ethics and the inclusion of more definitions for certain terms would reduce the cognitive complexity of such training for many CPS. Concurrently, repetition—along with the use of visual aids and demo videos—would aid in the learning and retention of the content. Moreover, ease of navigation can play a critical role in how peer specialists interact with the content they need to learn [25]. Improving usability would consequently help CPS maximize opportunities to share their lived experience with researchers. Future studies should use a larger sample size and include CPS from multiple demographics to better examine how CPS respond to research ethics trainings.

Conclusion

As the inclusion of CPS as research partners with shared decision-making authority becomes more commonplace, this may result in complications in research if research ethics training is not designed for this population. Although research ethics training programs for community research partners exist, they may need further adaptation for CPS.

Acknowledgments

This research was supported in part by the National Institutes of Mental Health (K01MH117496) and Patient-Centered Outcomes Research Institute (NCT03966872).

Conflicts of Interest

KLF offers consulting services through Social Wellness, LLC and has received funding from K Health.

References

1. DuBois JM, Bailey-Burch B, Bustillos D, Campbell J, Cottler L, Fisher CB, et al. Ethical issues in mental health research: the case for community engagement. Current Opinion in Psychiatry 2011;24(3):208-214. [doi: 10.1097/yco.0b013e3283459422]
2. 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]
3. Whiteman KL, Naslund JA, DiNapoli EA, Bruce ML, Bartels SJ. Systematic Review of Integrated General Medical and Psychiatric Self-Management Interventions for Adults With Serious Mental Illness. Psychiatr Serv 2016 Nov 01;67(11):1213-1225 [FREE Full text] [doi: 10.1176/appi.ps.201500521] [Medline: 27301767]
4. 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]
5. Fortuna KL, Brooks JM, Umucu E, Walker R, Chow PI. Peer Support: a Human Factor to Enhance Engagement in Digital Health Behavior Change Interventions. J Technol Behav Sci 2019 Jun 15;4(2):152-161 [FREE Full text] [doi: 10.1007/s41347-019-00105-x] [Medline: 34337145]
6. Fortuna KL, Barr P, Goldstein C, Walker R, Brewer L, Zagaria A, et al. Application of Community-Engaged Research to Inform the Development and Implementation of a Peer-Delivered Mobile Health Intervention for Adults With Serious Mental Illness. J Particip Med 2019 Mar 19;11(1):e12380 [FREE Full text] [doi: 10.2196/12380] [Medline: 32095314]
7. Fortuna K, Myers A, Brooks J, Collins-Pisano C, Marceau S, Pratt S, et al. Co-production of the quality of patient-centered outcomes research partnerships instrument for people with mental health conditions. Patient Experience Journal 2021 Apr 28;8(1):148-156 [FREE Full text] [doi: 10.35680/2372-0247.1533]
8. National Institutes of Health. Required education in the protection of human research participants (OD-00-039). 2000. URL: https://grants.nih.gov/grants/guide/notice-files/NOT-OD-00-039.html [accessed 2021-07-21]
9. CITI Program. The Trusted Standard in Research, Ethics, and Compliance Training. URL: https://about.citiprogram.org/en/homepage/ [accessed 2021-08-06]
10. Frequently Asked Questions. CITI Program. URL: https://about.citiprogram.org/en/faqs/ [accessed 2021-08-06]
11. Shannon C, Shibasaki S, National Health and Medical Research Council Committee, Australian Health Ethics Committee. Keeping research on track: A guide for Aboriginal and Torres Strait Islander peoples about health research ethics. Canberra, Australia: Australian Government; 2005.
12. Hatcher J, Schoenberg NE. Human subjects protection training for community workers: an example from "Faith Moves Mountains". Prog Community Health Partnersh 2007;1(3):257-265. [doi: 10.1353/cpr.2007.0022] [Medline: 20208288]

13. Terpstra J, Coleman KJ, Simon G, Nebeker C. The role of community health workers (CHWs) in health promotion research: ethical challenges and practical solutions. Health Promot Pract 2011 Jan 03;12(1):86-93 [FREE Full text] [doi: 10.1177/1524839908330809] [Medline: 19346410]
14. Martin PY, Turner BA. Grounded Theory and Organizational Research. The Journal of Applied Behavioral Science 2016 Jul 26;22(2):141-157. [doi: 10.1177/002188638602200207]
15. Birt L, Scott S, Cavers D, Campbell C, Walter F. Member Checking: A Tool to Enhance Trustworthiness or Merely a Nod to Validation? Qual Health Res 2016 Nov 10;26(13):1802-1811. [doi: 10.1177/1049732316654870] [Medline: 27340178]
16. Protecting Human Subjects in Research. CITI Program. URL: https://www.citiprogram.org/citidocuments/doe/doe.htm [accessed 2020-02-14]
17. Anderson EE. CIRTification: Training in Human Research Protections for Community-Engaged Research Partners. Prog Community Health Partnersh 2015;9(2):283-288 [FREE Full text] [doi: 10.1353/cpr.2015.0044] [Medline: 26412769]
18. Fortuna K, 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 [FREE Full text] [doi: 10.2196/16460] [Medline: 32243256]
19. Fortuna KL, Lohman MC, Gill LE, Bruce ML, Bartels SJ. Adapting a Psychosocial Intervention for Smartphone Delivery to Middle-Aged and Older Adults with Serious Mental Illness. Am J Geriatr Psychiatry 2017 Aug;25(8):819-828 [FREE Full text] [doi: 10.1016/j.jagp.2016.12.007] [Medline: 28169129]
20. Rotondi AJ, Eack SM, Hanusa BH, Spring MB, Haas GL. Critical design elements of e-health applications for users with severe mental illness: singular focus, simple architecture, prominent contents, explicit navigation, and inclusive hyperlinks. Schizophr Bull 2015 Mar 27;41(2):440-448 [FREE Full text] [doi: 10.1093/schbul/sbt194] [Medline: 24375458]
21. Knowles M. The Modern Practice of Adult Education: From Pedagogy to Andragogy. Cambridge: Cambridge Book Co; 1988.
22. Kolb D, Fry R. Toward an applied theory of experiential learning. In: Cooper C, editor. Theories of group processes. London: John Wiley & Sons; 1975:33-57.
23. Trivedi J. Cognitive deficits in psychiatric disorders: Current status. Indian J Psychiatry 2006 Jan;48(1):10-20 [FREE Full text] [doi: 10.4103/0019-5545.31613] [Medline: 20703409]
24. Edyburn D. Universal Usability and Universal Design for Learning. Intervention in School and Clinic 2020 Oct 12;56(5):310-315 [FREE Full text] [doi: 10.1177/1053451220963082]
25. Rotondi A, Grady J, Hanusa BH, Haas GL, Spring MR, Abebe KZ, et al. Key Variables for Effective eHealth Designs for Individuals With and Without Mental Health Disorders: 2^12-4 Fractional Factorial Experiment. J Med Internet Res 2021 Mar 24;23(3):e23137 [FREE Full text] [doi: 10.2196/23137] [Medline: 33759796]

Abbreviations

ADEXI: Adult Executive Functioning Inventory
CITI: Collaborative Institutional Training Initiative
CPS: certified peer support specialist

Edited by G Eysenbach; submitted 24.03.21; peer-reviewed by A Gold, M Lewis; comments to author 11.05.21; revised version received 09.08.21; accepted 13.09.21; published 01.02.22.

Please cite as:
Fortuna KL, Marceau SR, Kadakia A, Pratt SI, Varney J, Walker R, Myers AL, Thompson S, Carter K, Greene K, Pringle W
Peer Support Specialists' Perspectives of a Standard Online Research Ethics Training: Qualitative Study
JMIR Form Res 2022;6(2):e29073
URL: https://formative.jmir.org/2022/2/e29073
doi:10.2196/29073
PMID:35103606

Original Paper

An Internet-Based Cognitive Behavioral Therapy Program for Anxiety and Depression (Tranquility): Adaptation Co-design and Fidelity Evaluation Study

Victoria C Patterson[1], BA; Meghan A Rossi[1], BSc; Alissa Pencer[1], PhD; Lori Wozney[2], PhD

[1]Department of Psychology and Neuroscience, Dalhousie University, Halifax, NS, Canada

[2]Nova Scotia Health, Halifax, NS, Canada

Corresponding Author:
Alissa Pencer, PhD
Department of Psychology and Neuroscience
Dalhousie University
P.O. Box 15000
Halifax, NS, B3H 4R2
Canada
Phone: 1 9024944466
Email: alissa.pencer@dal.ca

Abstract

Background: Internet-based cognitive behavioral therapy (iCBT) is a necessary step toward increasing the accessibility of mental health services. Yet, few iCBT programs have been evaluated for their fidelity to the therapeutic principles of cognitive behavioral therapy (CBT) or usability standards. In addition, many existing iCBT programs do not include treatments targeting both anxiety and depression, which are commonly co-occurring conditions.

Objective: This study aims to evaluate the usability of Tranquility—a novel iCBT program for anxiety—and its fidelity to CBT principles. This study also aims to engage in a co-design process to adapt Tranquility to include treatment elements for depression.

Methods: CBT experts (n=6) and mental health–informed peers (n=6) reviewed the iCBT program Tranquility. CBT experts assessed Tranquility's fidelity to CBT principles and were asked to identify necessary interventions for depression by using 2 simulated client case examples. Mental health–informed peers engaged in 2 co-design focus groups to discuss adaptations to the existing anxiety program and the integration of interventions for depression. Both groups completed web-based surveys assessing the usability of Tranquility and the likelihood that they would recommend the program.

Results: The CBT experts' mean rating of Tranquility's fidelity to CBT principles was 91%, indicating a high fidelity to CBT. Further, 5 out of 6 CBT experts and all mental health–informed peers (all participants: 11/12, 88%) rated Tranquility as satisfactory, indicating that they may recommend Tranquility to others, and they rated its usability highly (mean 76.56, SD 14.07). Mental health–informed peers provided suggestions on how to leverage engagement with Tranquility (eg, adding incentives and notification control).

Conclusions: This preliminary study demonstrated the strong fidelity of Tranquility to CBT and usability standards. The results highlight the importance of involving stakeholders in the co-design process and future opportunities to increase engagement.

(JMIR Form Res 2022;6(2):e33374) doi:10.2196/33374

KEYWORDS

cognitive behavioral therapy; anxiety; depression; fidelity; usability

Introduction

Background

The in-person delivery of cognitive behavioral therapy (CBT) is an effective, evidence-based approach for treating a wide range of mental health conditions with a substantial amount of research indicating its efficacy for the treatment of anxiety [1] and depression [2]. In Canada, mild to moderate depression and anxiety disorders affect up to 11.7% and 10.8% of the adult population, respectively [3]. Although these mental health challenges affect many Canadians, there are numerous barriers to accessing evidence-based treatment. The availability of treatment is hindered by lengthy wait lists, therapists' caseload

limitations, and the moderate to low availability of clinicians trained in the delivery of evidence-based CBT [4]. The accessibility of treatment is further hindered by the monetary costs of treatment outside the public system and geographic access for those living in rural areas [5]. During the COVID-19 pandemic, additional challenges to accessing in-person treatment included physical distancing and self-isolation measures related to COVID-19, in addition to the sharp increase in stress, anxiety, and depression associated with the pandemic [6,7].

Previous research has found that internet-based CBT (iCBT) interventions provide increased access to evidence-based treatment, which can improve the mental health of the general population by effectively reducing symptoms associated with anxiety and depression [8,9]. However, few web-based interventions have been reviewed by therapy experts and have demonstrated treatment fidelity to core evidence-based CBT treatment protocols. Treatment fidelity is defined as "the extent to which an intervention is delivered as intended by the protocol" [10,11]. For an iCBT program targeting anxiety and depression, treatment fidelity indicates how closely the program follows the standard CBT approach for these mental health challenges and whether the necessary therapeutic interventions (eg, exposure and behavioral activation) are included. Bowie-DaBreo et al [12] described how there is significant uncertainty about the effectiveness of many mental health apps and programs, and much of that uncertainty appears to stem from a lack of information about treatment fidelity. In addition, few programs appeared to follow the guidelines for evidence-based care for specific challenges (eg, anxiety). Furthermore, a recent study examining mobile apps to treat depression [13] found that few programs followed a CBT approach, even though the literature supports CBT as the gold standard for the treatment of anxiety and depression [14]. In 2018, an international group of researchers developed a list of recommendations for the development of iCBT programs [15]—one of their core recommendations was the use of a co-design process (ie, involvement of intended users in the development of the iCBT intervention). Co-design has many benefits (eg, improved program tailoring) [16], yet this approach appears underused in the literature. In sum, there appears to be a gap in the literature pertaining to evidence-based iCBT programs that also follow the iCBT recommendations for development, especially for programs targeting co-occurring depression and anxiety [17].

Tranquility Program

Tranquility was designed to be an iCBT intervention that increases cognitive and behavioral skills (eg, thought challenging, exposure, and behavioral experiments) of individuals with mild to moderate anxiety. Users learn ways to manage their symptoms in addition to gaining access to personalized support through video, phone, and in-app messaging with a web-based coach. Results from a pilot program evaluation of Tranquility illustrated decreases in anxiety and stress levels in individuals who completed 3 or more modules, meaning that people who engaged with the program were noted to benefit, and most program users found the program to be helpful [18,19]. Though designed by a team including licensed psychologists with advanced skills in CBT techniques, researchers, and people with lived experience, Tranquility has not yet undergone a rigorous evaluation of its fidelity to core CBT components linked to improved client outcomes by a broader team of experts in this field. Furthermore, engaging mental health–informed peers (eg, people with lived experience of mental health challenges or peer counselors) in program development and evaluation could further illuminate usability characteristics (ie, aspects that contribute to the technical effectiveness and efficiency of a program) [20,21] that enable and increase engagement. Finally, Tranquility has the potential to be adapted for dual treatment streams related to anxiety and depression but requires both clinical and first voice input to map development.

Study Aims

This study aims to address these gaps by (1) evaluating Tranquility's fidelity to CBT principles; (2) assessing CBT experts' and mental health–informed peers' perceptions of the usability of the Tranquility anxiety program and the likelihood they would recommend it to others; and (3) co-designing an adaptation of the Tranquility program to include treatment for depression with a group of mental health–informed peers.

Methods

Recruitment

To recruit focus group participants, relevant networks (eg, local university programs and peer counseling programs) were contacted and asked to disseminate study information to all local individuals on their newsletter mailing list. The focus group participants had to identify as a first voice advocate or have experience with mental health conditions in a near-peer role where they would have advanced knowledge of needs and experiences of the population for which Tranquility was designed (eg, peer mentors on university campuses). CBT experts were recruited through the authors' professional networks of CBT clinicians and North American CBT organizations (eg, Canadian Association of Cognitive and Behavioural Therapies and Association for Behavioral and Cognitive Therapies), and individuals who met the study criteria were contacted directly via email. CBT experts had to have 5 years of experience delivering CBT for depression and anxiety in adults and must be licensed by a professional body. All participants were required to be ≥18 years of age.

We recruited 6 mental health–informed peers, referred to as focus group participants throughout the remainder of the paper, and 6 CBT experts to participate in the study. Sociodemographic information of all participants is presented in Table 1. Across both groups, participants were predominately White, heterosexual women. The mean age of the focus group was 39.25 (SD 8.88) years, and the mean age of experts was 47.6 (SD 14.72) years. CBT experts were entered into a draw to win a web-based gift card, while the focus group participants received an honorarium following each focus group. The study procedures were approved by the Research Ethics Board of the Nova Scotia Health Authority (file number 1025561) and conformed to the ethical standards of research set out by the Canadian Tri-Council Policy Statement 2 [22].

Table 1. Participant demographics.

Demographics	Focus group participants (n=6)	CBT[a] experts (n=6)
Age (years), mean (SD)	39.25 (8.88)	47.6 (14.72)
Gender, n (%)[b]		
Man	—[c]	1 (17)
Woman	5 (83)	5 (83)
Preferred not to say	1 (17)	—
Sexual orientation, n (%)[b]		
Heterosexual	4 (66)	6 (100)
Bisexual	1 (17)	—
Preferred not to say	1 (17)	—
Race, n (%)[b]		
White	5 (83)	6 (100)
Preferred not to say	1 (17)	—

[a]CBT: cognitive behavioral therapy.

[b]Participants were provided with a comprehensive list of gender identities, sexual orientations, and racial identities from which to choose when reporting their demographic information—only those selected by at least 1 participant are listed in this table.

[c]Not available. These categories were not used by the focus group or cognitive behavioral therapy experts.

Measures

Focus Groups

Overview

Mental health–informed peers were involved in 2 focus groups to co-design adaptations to Tranquility for both anxiety and depression and the benefit of these changes. A script with a combination of question types (eg, open-ended and close-ended questions) for the focus group was developed by the research team and can be found in Multimedia Appendix 1. The ECOUTER (Employing Conceptual Schema for Policy and Translation Engagement in Research) [23] framework was used in the facilitation of the focus groups. This framework positions participants in a prominent role in the research process by encouraging knowledge exchange, developing a conceptual schema, analyzing discussion contributions, and refining recommendations. Following each focus group, participants completed a web-based survey, which included demographic questions (eg, age and gender), the usability of the web-based platform, and their likelihood of recommending Tranquility to others.

The System Usability Scale

The System Usability Scale (SUS) [24] is a validated 10-item scale that can be adapted to assess facets of usability in different programs (eg, "I thought that the Tranquility program was easy to use"). Items are rated on a 5-point scale ranging from *strongly disagree* (score=1) to *strongly agree* (score=5). To obtain an overall usability score, a value of 1 is subtracted from the score of odd-numbered items, and the respondent's score is subtracted from 5 for each even-numbered question. These new values for each item are summed, and the total is multiplied by 2.5 to create scores ranging from 0 to 100. Scores >69 reflect appropriate ratings of usability, with higher scores indicating greater usability [25]. Scores in the high 70s to 80s indicate good usability, while scores ≥90 indicate high overall usability. The SUS is the most widely used measure that is validated for testing usability across various iCBT programs [26,27], and it is designed to be tailored to the program being tested.

The Likelihood to Recommend Scale

Participants rated how likely they were to recommend Tranquility via a 1-item measure: "How likely are you to recommend Tranquility to your friends, family, or associates?" [28]. Responses were rated on an 11-point scale ranging from *not at all likely* (score=0) to *extremely likely* (score=10); higher scores reflect a greater likelihood to recommend the program. Ratings between 0 and 6 indicate dissatisfaction and a low likelihood to recommend; ratings of 7 to 8 indicate satisfaction and a moderate likelihood to recommend; and ratings of 9 or 10 indicate high satisfaction and a strong likelihood to recommend to others.

CBT Experts

CBT experts completed questionnaires assessing demographics, the SUS, the Likelihood to Recommend scale, and Tranquility's level of fidelity to CBT protocols for anxiety. They also followed clinical case vignettes through the program to detail the therapeutic components of CBT that are necessary to treat the depicted cases of depression.

The Component Analysis of Tranquility

To assess fidelity to CBT components for the treatment of anxiety, we adapted the evaluation criteria for web-based apps treating depression by Huguet et al [13]. We adapted that measure to be specific to the Tranquility program and the treatment of anxiety. The measure includes 10 items, such as "In your opinion, does the Tranquility program provide an

explanation of the CBT model?" Responses were rated on a 3-point scale, *none* (score=0), *some* (score=1), and a tailored third option indicating correct and specific use of a therapeutic component (eg, *clear explanation* [score=2]) for each question. All items were summed and divided by the maximum possible score (ie, 20). This value was then multiplied by 100 to create a total score percentage ranging from 0 to 100, with higher scores reflecting greater fidelity to CBT components for treating anxiety.

Adaptations to Tranquility

We adapted 2 vignettes of individuals with depression (Multimedia Appendix 1) from the depressive disorders section of the Diagnostic and Statistical Manual of Mental Disorders–5 Clinical Case Handbook [29]—Diane, who has some social anxiety and has been depressed for the last 2 years and experiences little to no interest or pleasure (case 4.6) [30], and Helen, who has been feeling depressed for the last 2 months, drinks 4-5 alcoholic drinks per day, has significant insomnia, had childhood anxiety, and has recently attempted suicide (case 4.10) [31].

CBT experts were asked which CBT components Tranquility should include to offer treatment for depression for the individual in the vignette. Each vignette had the same list of 19 components (eg, behavioral activation and mood tracking; Multimedia Appendix 1). CBT experts were also asked whether there were necessary special considerations when treating individuals seeking treatment for both anxiety and depression.

Data Collection Procedures

Upon completion of informed consent, all participants were provided with a video link outlining how to access and use Tranquility, and participants were then asked to review Tranquility in detail. To fulfill the 3 aims of the study, data collection occurred in 3 phases. In phase 1, following their review of Tranquility for anxiety, the CBT experts and focus group participants completed measures of usability (ie, SUS) and the likelihood to recommend (ie, Likelihood to Recommend scale). In addition, CBT experts evaluated Tranquility's treatment fidelity to CBT. All measures were completed on the REDCap (Research Electronic Data Capture; Vanderbilt University) platform, a web-based data collection platform [32,33]. In phase 2, focus group participants participated in a co-design meeting with the researchers using the Zoom Professional platform (Zoom Video Communications Inc); discussions focused on existing anxiety components and adaptations for depression. In phase 3, focus group participants were presented with the survey results from phase 1 to provide a foundation for their feedback. The themes that emerged from the focus group meeting in phase 2 were shown to focus group participants to refine this information and correct any discrepancies. Next, focus group participants were presented with the adaptations made to Tranquility following the previous co-design meeting, and final feedback regarding the adaptations and the integration of anxiety and depression interventions in Tranquility was obtained.

Data Analysis

As the purpose of the study was not to test a specific hypothesis or compare groups but rather to explore usability, assess treatment fidelity, and engage in the co-design process, data from both groups were pooled for all measures except for CBT fidelity measures only completed by CBT experts. Descriptive statistics were used to summarize the data. Focus group transcripts were analyzed using the principles of thematic content analysis, and we used a data-driven approach to inductively establish themes [34]. The themes were created using the guiding principle of selecting feedback that highlighted the strengths and weaknesses of the Tranquility program. Units of analysis included discrete words, sentences, and paragraphs. In line with the ECOUTER framework, mind maps were used to develop the conceptual schema of themes and subthemes; 2 experienced raters reviewed the anonymized interview transcripts and independently coded the data using Microsoft Excel [35]. Any discrepancies were resolved through discussion.

Results

Usability and Likelihood to Recommend

The participants' ratings of the usability of Tranquility are presented in Table 2. Data were missing from several focus group members (n=4) on measures of usability, resulting in a final sample of 8 participants for the SUS measure. Most participants agreed that Tranquility was easy to use, and all participants agreed with the statement that they felt confident in their ability to use the Tranquility program and that the functions of Tranquility were well-integrated into the program. In contrast, there was less agreement about whether participants would use the program frequently in the future or whether most people would learn to use Tranquility quickly. When looking at reverse-coded items, it was apparent that most participants believed that program users would not need the support of a technical person nor would they need to learn much before being able to use the Tranquility program. Nearly all participants strongly agreed that inconsistency was not a problem across the Tranquility program. The SUS questionnaire results indicated that the average overall usability score was 76.56 (SD 14.07; range 52.5-97.5), which was above the evidence-based cut-off for program usability. Finally, the mean Likelihood to Recommend scale score was 7 (SD 1.07; range 5-8), indicating user satisfaction; 88% (11/12) of participants rated Tranquility as being within the *satisfactory* category, while 1 participant's rating of Tranquility indicated dissatisfaction and a low likelihood to recommend.

Table 2. Usability ratings of the Tranquility program.

Usability component[a]	Rating, mean (SD; range)	Agreement with this statement, n (%)	Disagreement with this statement[b], n (%)
I think that I would like to use the Tranquility program frequently	3.75 (0.71; 3-5)	5 (63)	N/A[c]
I thought that the Tranquility program was easy to use	4.29 (0.76; 3-5)	6 (86)[d]	N/A
I found the various functions in the Tranquility program were well integrated	4.25 (0.46; 4-5)	8 (100)	N/A
I would imagine that most people would learn to use the Tranquility program very quickly	3.88 (0.99; 2-5)	6 (75)	N/A
I felt very confident using the Tranquility program	4.13 (0.35; 4-5)	8 (100)	N/A
I needed to learn a lot of things before I could get going with the Tranquility program [e]	1.75 (1.04; 1-4)	N/A	7 (86)
I found the Tranquility program unnecessarily complex [e]	2.13 (0.99; 1-4)	N/A	6 (75)
I think that I would need the support of a technical person to be able to use the Tranquility program [e]	1.5 (0.53; 1-2)	N/A	8 (100)
I thought there was too much inconsistency in the Tranquility program [e]	1.75 (0.46; 1-2)	N/A	8 (100)
I found the Tranquility program cumbersome to use [e]	2.13 (0.83; 1-4)	N/A	7 (86)

[a]Ratings ranged from *strongly disagree* (score=1) to *strongly agree* (score=5).

[b]Reverse-coded items.

[c]N/A: not applicable.

[d]There was missing data for 1 participant (n=7).

[e]Items that are reverse-scored (ie, low scores mean higher usability) are italicized.

Focus Group Content Analysis and Adaptations to Tranquility

Content Analysis

Focus group feedback was organized into 8 themes and 8 subthemes (Multimedia Appendix 2). Broadly, feedback from the focus groups focused on suggestions to increase engagement (eg, personalization and reduced psychoeducation) and increase the accessibility of the written content. For the newly added depression material, participants suggested using the word *depression* rather than alternative terms (eg, low mood), ways to tailor the material to a variety of users, the addition of several treatment targets (eg, functioning and self-care), and how to increase the clarity of advertisements about Tranquility.

Program Changes

As a part of the co-design process, changes were made to Tranquility across the existing anxiety components and the new depression components (Multimedia Appendix 2). CBT experts' strong agreement about the necessary components to treat comorbid anxiety and depression resulted in the adaptation of the existing cognitive elements for the treatment of anxiety (eg, thought records and cognitive distortions) for the treatment of depression, and the inclusion of behavioral activation strategies.

The focus group participants suggested providing the program user with more control over their experience and more support for engagement, resulting in increased control over aspects of notifications (eg, notification type), and ongoing developments for Tranquility include gamification and new incentives to foster engagement. Focus group participants asked that coaches initiate coaching appointments, help personalize the messaging experience with clients, and aid users in selecting content that is best suited to their needs. All requested changes applicable to coaching were made to foster connections with the coach and enrich the experience with Tranquility.

The focus group also encouraged the addition of information about program fit before beginning the program; Tranquility now begins with a screening assessment to ensure a good fit between the user's identified needs and the Tranquility program before payment, and feedback is provided to the user during onboarding. The focus group participants said that there was too much psychoeducation, and the language level was too high for most laypeople; all language was adjusted by increasing layperson terminology and reducing jargon. Finally, the group discussed program tailoring, the addition of treatment targets, and more accurate advertising. As a result, Tranquility can be tailored to begin with either anxiety or depression content, depending on user needs; now includes quality of life and well-being tracking; and is more specifically advertised as a daily use program targeting mild to moderate anxiety and depression.

CBT Components

CBT experts evaluated the fidelity of Tranquility to CBT for the treatment of anxiety. The average rating for fidelity to core CBT components was 91% (SD 11.14; range 75-100), and half of the CBT experts indicated that Tranquility included 100% of all required CBT components.

In terms of specific components, all 6 experts reported that Tranquility included clear explanations of both anxiety and the CBT model, including cognitive and behavioral techniques, and Tranquility provided formal ratings of anxiety (eg, 0 to 10 scale) to program users. In total, 4 of 5 (80%) CBT experts agreed that Tranquility included specific emotion monitoring, 1 of 5 (20%) therapists agreed that Tranquility included only some emotion monitoring, and 1 expert did not respond. Similarly, 5 of 6 (83%) experts agreed that Tranquility provided specific monitoring of cognitions, and 1 of 6 (17%) therapists indicated that Tranquility included only some monitoring of cognitions. This exact pattern of results was also seen when evaluating whether Tranquility provided a method to monitor behaviors. Only 4 of 6 (67%) CBT experts reported that Tranquility offered a way to monitor specific physical sensations and allowed for adequate case conceptualization of anxiety. For both physical sensation monitoring and case conceptualization, one CBT expert believed there was some tracking of physical sensation monitoring or case conceptualization, while another expert believed there was none.

Adaptations to Depression Content

Common Findings Across Vignettes

All CBT experts indicated that the following components should be included to treat both individuals depicted in the vignettes: psychoeducation about depression, behavioral activation, mood tracking, and case conceptualization. In addition, 5 of 6 (83%) experts agreed that mood ratings (eg, rating mood from 0 to 10), behavioral experiments, and problem-solving skills would also be important to include in Tranquility to treat both cases, while 4 of 6 (67%) experts agreed that coping strategies and sleep hygiene information should be included.

Diane Vignette

All 6 (100%) experts agreed that the Diane case would also require pleasant activity scheduling and symptom or outcome tracking. In total, 5 of 6 (83%) experts indicated that it would be helpful to include thought records, while 4 of 6 (67%) experts believed that the identification of cognitive distortions was important to include. Only 2 of 6 (33%) experts indicated that it was important to include motivational interviewing, physical symptom monitoring, substance use tracking, exposure stepladders, or psychoeducation about safety behaviors. Experts also suggested that the following components should also be added: mindfulness, meditation, open journaling, and CBT for insomnia.

Helen Vignette

There was unanimous agreement that Helen would benefit from the inclusion of thought records within Tranquility (6/6, 100%). In total, 5 of 6 (83%) experts indicated their belief that pleasant activity scheduling and the identification of cognitive distortions should also be used within Tranquility to treat depression affecting Helen. In total, 4 of 6 (67%) experts believed that substance use tracking, psychoeducation about safety behaviors, and symptom or outcome tracking should also be added. Half of the experts (3/6, 50%) thought motivational interviewing could be beneficial to include, while one-third (2/6, 33%) thought that physical symptom tracking should be included. No CBT expert endorsed the addition of exposure stepladders. Experts also suggested that the following components be added: safety planning, list of crisis resources, mindfulness, meditation, open journaling, and CBT for insomnia.

Discussion

Principal Findings

This study represents an assessment of both usability and treatment fidelity to CBT and a co-design adaptation of an iCBT program. Overall, participants highly rated the usability of Tranquility and indicated satisfaction with the program, and CBT experts provided high ratings for Tranquility's treatment fidelity to CBT. The co-design adaptation process resulted in several improvements to Tranquility.

Experts agreed that Tranquility had high fidelity to the CBT model and included the most necessary components for the treatment of anxiety. There was high agreement across most components (eg, inclusion of behavioral techniques), although to obtain perfect fidelity, Tranquility would need to make emotion, behavior, physical sensation, and cognition monitoring more explicit. These elevated fidelity ratings were not unexpected, as the Tranquility program development team included a licensed clinical psychologist with extensive expertise in CBT. Notably, a recent functionality analysis of apps for depression [36] found that of a possible 8 CBT components, 22% of apps only included ≥3 CBT components, 68% included only 1 or 2, and 10% of apps did not include any CBT components. Furthermore, only 45% of these apps had expert involvement (eg, health professionals) during the app development process. These findings highlight the gap that Tranquility fills with regard to fidelity to CBT components and the importance of fidelity evaluations by objective experts.

Across vignettes, CBT experts typically agreed on which CBT interventions were necessary to treat the depicted cases. Moreover, both cases required the same set of interventions, except for the addition of safety behavior psychoeducation and substance use tracking for the Helen vignette, given the depicted substance use. It was an expected finding that experts would strongly suggest using interventions such as behavioral activation, given its robust efficacy in depression treatment, including within an iCBT format [37]. The integration of therapeutic strategies for depression within an iCBT program for anxiety is likely to be beneficial, given the overlap between the symptoms and the core therapeutic strategies used to treat each disorder [38]. However, the addition of unique therapeutic elements also used to treat either anxiety or depression can overwhelm program users; individuals may benefit from external support to navigate and select the most helpful strategies. A recent literature review [39] suggested that guided (eg, coaches or therapist support) programs had similar rates of adherence as in-person CBT treatment, which is higher than self-guided programs—guidance appears to positively impact treatment program use [39,40].

Comparison With Prior Work

Current guidelines for digital mental health interventions strongly suggest using a co-design process [41,42], and the

results of this study illustrate the importance of that process—the mental health peers suggested crucial improvements to program customizability, delivery, accessibility, and available interventions. Previous research has found that collaborative study designs, such as the co-design approach (see Hill et al [15] for a review), can have a positive effect on aspects such as user adherence, usability, and uptake by involved stakeholders (eg, program users and clinicians).

Mental health programs targeting depression and anxiety need to have both high fidelity to the CBT treatment model and be engaging, flexible, and allow user personalization. Similarly, usability testing is critical, and Kushniruk [43] found that usability testing with small groups (ie, 8-10 users) can drastically reduce the number of usability issues experienced by users in the future. Usability testing within this study revealed areas upon which Tranquility program developers can improve, such as increasing user engagement to encourage users to continue to use this program frequently and reducing the initial learning curve of how to use Tranquility.

The findings of this study are in line with much of the iCBT literature—users want engagement with and control over their web-based therapeutic experiences. Stawarz et al [36] noted that although approximately 58% of CBT apps included at least 1 engagement strategy, they used a less-varied complement of engagement strategies compared with other kinds of mental health apps. Most CBT apps included strategies such as visual aids (eg, graphs and charts) but few included more complex strategies, for example, gamification and coaching or therapist chat functions. Stawarz et al [36] also conducted a qualitative analysis of app reviews, which highlighted the value of leveraging engagement strategies (eg, personalization) to increase user satisfaction with mental health programs [44]. Similarly, a qualitative study of engagement with an iCBT program found that participants wanted to choose what information to learn, accessible content (eg, audiovisual content and appropriate reading level), and a more tailored coaching experience [16,45]. In sum, users want personalized content and more complex engagement strategies.

Usability of Methodology

The ECOUTER framework provided an accessible approach that centered potential users in the feedback process and provided a procedure guide, which was beneficial when recruiting and working collaboratively with focus group participants as well as during data analysis. Engaging in an adaptation co-design process significantly improved program design and delivery for this iCBT program because it permitted the integration of researchers' expertise pertaining to the design and delivery of iCBT and participants' expertise pertaining to mental health both as a clinician and as a client. Moreover, seeking usability feedback from all participants aided in the rigor of this mixed methods evaluation and co-design adaptation of Tranquility because it allowed for multiple perspectives to be considered as CBT experts and focus group members may have different experiences and expectations of iCBT programs and may place value on differing program elements measured within the usability questionnaire we used, the SUS. However, although this approach had many benefits, it did include challenges such as recruiting CBT experts. Many of the approached individuals declined to participate, likely due to demands in excess of their available time, and scheduling focus groups was especially challenging. The usability of the design used in this study, although challenging at times, provides a greater *real-world* approach to mental health program design and both fidelity and usability evaluation.

Future Directions and Limitations

These findings have therapeutic implications for the iCBT literature. Taken together, future research and iCBT program development should consider the involvement of program users and clinical experts to ensure that fidelity to CBT and engagement strategies are implemented, given its association with adherence and program effectiveness [46]. The effectiveness of gamification and incentives offered within eHealth tools, including within iCBT programs, remains an understudied element of the literature [15], and future work should aim to assess the effectiveness of these user engagement strategies in addition to measuring client satisfaction. In addition, future work is needed to examine the effectiveness of iCBT interventions for comorbid mental health challenges, such as anxiety and depression, and the addition of iCBT protocols to address other challenges (insomnia, trauma, etc) may also help meet treatment demands for other comorbid presentations (eg, anxiety and insomnia).

Similarly, many studies, including this one, included predominantly White, heterosexual women. As a result, there are additional considerations related to race, gender, and sexual orientation that were not captured in the participant feedback or in the resultant changes to Tranquility (eg, using acceptance or values-based strategies vs cognitive restructuring for microaggressions; greater role for family or community members).

Understanding the perspectives and unique needs of marginalized groups is necessary to increase iCBT treatment accessibility and effectiveness—efforts should be made to include participants who are racialized and 2SLGBTQ+ (ie, two-spirit, lesbian, gay, bisexual, transgender, queer, and all other members of this community) and individuals who live in rural areas or are older adults. Efforts to recruit a larger and more diverse sample will afford a greater range of perspectives. Furthermore, this study included a small sample of CBT experts and focus group participants. This smaller sample was a significant strength of the study in that it afforded the focus group participants an opportunity to provide rich and in-depth perspectives on Tranquility and suggestions for adaptation. However, we recognize that the sample size does impact the generalizability of these findings to individuals who were underrepresented in the participant pool (eg, men, nonbinary individuals, or Black people) and who may or may not have been represented at all (eg, people with disabilities). Finally, it is of note that technological confidence and knowledge were not measured; therefore, it is possible that the ratings of this program were influenced by this variable.

Conclusions

CBT experts and mental health peers agreed that Tranquility, a web-based program treating anxiety and depression, had high usability, and both groups would be likely to recommend this program to others. CBT experts scored Tranquility as having high fidelity to CBT, and nearly all intervention components needed to treat depression were included as a part of Tranquility. Finally, the co-design process was key to refining the existing anxiety content and for the creation and integration of the new depression content. These results provide a preliminary evaluation of the Tranquility program, and they may provide user-centered engagement strategies that may help increase adherence and effectiveness for the iCBT treatment of anxiety and depression.

Acknowledgments

The authors are grateful to the individuals who participated in the focus groups and to the cognitive behavioral therapy experts who committed time to review this program and provide valuable feedback. They would also like to thank Stephanie Sellars for helping with the focus groups.

Authors' Contributions

AP, LW, and VCP conceptualized the study. VCP created project materials and recruited and interviewed participants with support from LW and MAR. VCP and MAR wrote the manuscript with support from AP and LW.

Conflicts of Interest

AP is a cofounder and senior scientific director of Tranquility, along with being a paid consultant. AP had no direct contact with the study participants during this study and did not participate in the data analysis process.

Multimedia Appendix 1
Cognitive behavioral therapy expert and focus group question list and case vignettes.
[DOC File , 39 KB - formative_v6i2e33374_app1.doc]

Multimedia Appendix 2
Focus group feedback and changes to the Tranquility program.
[DOCX File , 34 KB - formative_v6i2e33374_app2.docx]

References

1. Hofmann SG, Asnaani A, Vonk IJJ, Sawyer AT, Fang A. The efficacy of cognitive behavioral therapy: a review of meta-analyses. Cognit Ther Res 2012 Oct 1;36(5):427-440 [FREE Full text] [doi: 10.1007/s10608-012-9476-1] [Medline: 23459093]
2. Weitz ES, Hollon SD, Twisk J, van Straten A, Huibers MJ, David D, et al. Baseline depression severity as moderator of depression outcomes between cognitive behavioral therapy vs pharmacotherapy: an individual patient data meta-analysis. JAMA Psychiatry 2015 Nov;72(11):1102-1109. [doi: 10.1001/jamapsychiatry.2015.1516] [Medline: 26397232]
3. Trends in the prevalence of depression and anxiety disorders among working-age Canadian adults between 2000 and 2016. Statistics Canada. 2020. URL: https://www150.statcan.gc.ca/n1/pub/82-003-x/2020012/article/00002-eng.htm [accessed 2021-08-15]
4. 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]
5. Wright JH, Mishkind M, Eells TD, Chan SR. Computer-assisted cognitive-behavior therapy and mobile apps for depression and anxiety. Curr Psychiatry Rep 2019 Jun 27;21(7):62. [doi: 10.1007/s11920-019-1031-2] [Medline: 31250242]
6. Dozois DJ. Anxiety and depression in Canada during the COVID-19 pandemic: a national survey. Can Psychol / Psychologie canadienne 2021 Feb 27;62(1):136-142. [doi: 10.1037/cap0000251]
7. 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]
8. Kumar V, Sattar Y, Bseiso A, Khan S, Rutkofsky IH. The effectiveness of internet-based cognitive behavioral therapy in treatment of psychiatric disorders. Cureus 2017 Aug 29;9(8):e1626 [FREE Full text] [doi: 10.7759/cureus.1626] [Medline: 29098136]
9. Spek V, Cuijpers P, Nyklícek I, Riper H, Keyzer J, Pop V. Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: a meta-analysis. Psychol Med 2006 Nov 20;37(03):319. [doi: 10.1017/s0033291706008944]
10. Bassett SS, Stein L, Rossi JS, Martin RA. Evaluating measures of fidelity for substance abuse group treatment with incarcerated adolescents. J Subst Abuse Treat 2016 Jul;66:9-15 [FREE Full text] [doi: 10.1016/j.jsat.2016.02.011] [Medline: 27211991]

11. Perepletchikova F, Treat TA, Kazdin AE. Treatment integrity in psychotherapy research: analysis of the studies and examination of the associated factors. J Consult Clin Psychol 2007;75(6):829-841. [doi: 10.1037/0022-006x.75.6.829]
12. Bowie-DaBreo D, Sünram-Lea SI, Sas C, Iles-Smith H. Evaluation of treatment descriptions and alignment with clinical guidance of apps for depression on app stores: systematic search and content analysis. JMIR Form Res 2020 Dec 13;4(11):e14988 [FREE Full text] [doi: 10.2196/14988] [Medline: 33185566]
13. Huguet A, Rao S, McGrath PJ, Wozney L, Wheaton M, Conrod J, et al. A systematic review of cognitive behavioral therapy and behavioral activation apps for depression. PLoS One 2016;11(5):e0154248 [FREE Full text] [doi: 10.1371/journal.pone.0154248] [Medline: 27135410]
14. David D, Cristea I, Hofmann SG. Why cognitive behavioral therapy is the current gold standard of psychotherapy. Front Psychiatry 2018 Jan;9:4 [FREE Full text] [doi: 10.3389/fpsyt.2018.00004] [Medline: 29434552]
15. Hill C, Creswell C, Vigerland S, Nauta MH, March S, Donovan C, et al. Navigating the development and dissemination of internet cognitive behavioral therapy (iCBT) for anxiety disorders in children and young people: a consensus statement with recommendations from the #iCBTLorentz Workshop Group. Internet Interv 2018 Jul;12:1-10 [FREE Full text] [doi: 10.1016/j.invent.2018.02.002] [Medline: 30135763]
16. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and promoting effective engagement with digital behavior change interventions. Am J Prev Med 2016 Nov;51(5):833-842. [doi: 10.1016/j.amepre.2016.06.015] [Medline: 27745683]
17. Sartorius N, Üstün TB, Lecrubier Y, Wittchen H. Depression comorbid with anxiety: results from the WHO study on psychological disorders in primary health care. Br J Psychiatry 2018 Aug 06;168(S30):38-43. [doi: 10.1192/s0007125000298395]
18. Pencer A, Tucker R, Muise J. Breaking down barriers: evaluating an internet-based CBT program for adults experiencing anxiety. In: Proceedings of the 9th World Congress of Behavioural and Cognitive Therapies. 2019 Jul 17 Presented at: 9th World Congress of Behavioural and Cognitive Therapies; July 17-20, 2019; Berlin, Germany URL: https://wcbct2019.org/
19. Pencer A, Tucker R, Muise J, Wozney L. Breaking down barriers and optimizing retention: a novel approach to delivering iCBT. In: Proceedings of the Poster Presentation at the 9th Annual E-mental Health Conference. 2020 Presented at: 9th Annual E-mental Health Conference; March 2020; Toronto, ON URL: https://www.mentalhealthcommission.ca/wp-content/uploads/drupal/2020-09/EMental_Health_Conference_2020_Summary_Report_eng.pdf
20. Thüring M, Mahlke S. Usability, aesthetics and emotions in human–technology interaction. Int J Psychol 2007 Aug;42(4):253-264. [doi: 10.1080/00207590701396674]
21. Bevan N. Measuring usability as quality of use. Software Qual J 1995 Jun;4(2):115-130. [doi: 10.1007/BF00402715]
22. Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada, Social Sciences and Humanities Research Council. Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans – TCPS 2 (2018). Ottawa: Government of Canada; 2018 Dec. URL: https://ethics.gc.ca/eng/policy-politique_tcps2-eptc2_2018.html [accessed 2020-04-20]
23. Murtagh MJ, Minion JT, Turner A, Wilson RC, Blell M, Ochieng C, et al. The ECOUTER methodology for stakeholder engagement in translational research. BMC Med Ethics 2017 Apr 04;18(1):24 [FREE Full text] [doi: 10.1186/s12910-017-0167-z] [Medline: 28376776]
24. Brooke J. SUS: A 'Quick and Dirty' usability scale. In: Jordan P, Thomas B, Weerdmeester B, editors. Usability Evaluation In Industry. Oxfordshire, United Kingdom: Taylor & Francis; 1996.
25. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Hum-Comput Interact 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]
26. Lewis JR. The system usability scale: past, present, and future. Int J Hum–Comput Interact 2018 Mar 30;34(7):577-590. [doi: 10.1080/10447318.2018.1455307]
27. Mol M, van Schaik A, Dozeman E, Ruwaard J, Vis C, Ebert DD, et al. Dimensionality of the system usability scale among professionals using internet-based interventions for depression: a confirmatory factor analysis. BMC Psychiatry 2020 May 12;20(1):218 [FREE Full text] [doi: 10.1186/s12888-020-02627-8] [Medline: 32398111]
28. Reichheld FF. The one number you need to grow. Harv Bus Rev 2003 Dec;81(12):46-54, 124. [Medline: 14712543]
29. Barnhill J, Wood W, Yonkers K. Depressive disorders. In: DSM-5 Clinical Cases. Washington (DC): American Psychiatric Association Publishing; 2013.
30. Brody B. Depressive disorders case 4.6. In: DSM-5 Clinical Cases: Depressive Disorders. Washington (DC): American Psychiatric Association Publishing; 2013.
31. Goldberg J. Depressive disorders case 4.10. In: DSM-5 Clinical Cases. Washington (DC): American Psychiatric Association Publishing; 2013.
32. 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]
33. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]

34. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
35. Bree RT, Gallagher G. Using Microsoft Excel to code and thematically analyse qualitative data: a simple, cost-effective approach. All Ireland J High Edu 2016;8(2):2811-2824 [FREE Full text]
36. Stawarz K, Preist C, Tallon D, Wiles N, Coyle D. User experience of cognitive behavioral therapy apps for depression: an analysis of app functionality and user reviews. J Med Internet Res 2018 Jun 06;20(6):e10120 [FREE Full text] [doi: 10.2196/10120] [Medline: 29875087]
37. Furukawa TA, Suganuma A, Ostinelli EG, Andersson G, Beevers CG, Shumake J, et al. Dismantling, optimising, and personalising internet cognitive behavioural therapy for depression: a systematic review and component network meta-analysis using individual participant data. Lancet Psychiatry 2021 Jun 22;8(6):500-511 [FREE Full text] [doi: 10.1016/S2215-0366(21)00077-8] [Medline: 30798295]
38. Păsărelu CR, Andersson G, Bergman NL, Dobrean A. Internet-delivered transdiagnostic and tailored cognitive behavioral therapy for anxiety and depression: a systematic review and meta-analysis of randomized controlled trials. Cogn Behav Ther 2017 Jan;46(1):1-28. [doi: 10.1080/16506073.2016.1231219] [Medline: 27712544]
39. Radomski AD, Wozney L, McGrath P, Huguet A, Hartling L, Dyson MP, et al. Design and delivery features that may improve the use of internet-based cognitive behavioral therapy for children and adolescents with anxiety: a realist literature synthesis with a persuasive systems design perspective. J Med Internet Res 2019 Feb 05;21(2):e11128 [FREE Full text] [doi: 10.2196/11128] [Medline: 30720436]
40. Wozney L, Huguet A, Bennett K, Radomski AD, Hartling L, Dyson M, et al. How do eHealth programs for adolescents with depression work? A realist review of persuasive system design components in internet-based psychological therapies. J Med Internet Res 2017 Aug 09;19(8):e266 [FREE Full text] [doi: 10.2196/jmir.7573] [Medline: 28793983]
41. Achilles MR, Anderson M, Li SH, Subotic-Kerry M, Parker B, O'Dea B. Adherence to e-mental health among youth: considerations for intervention development and research design. Digit Health 2020 May 21;6:2055207620926064 [FREE Full text] [doi: 10.1177/2055207620926064] [Medline: 32547775]
42. Scariot CA, Heemann A, Padovani S. Understanding the collaborative-participatory design. Work 2012;41 Suppl 1:2701-2705. [doi: 10.3233/WOR-2012-0656-2701] [Medline: 22317129]
43. Kushniruk A. Evaluation in the design of health information systems: application of approaches emerging from usability engineering. Comput Biol Med 2002 May;32(3):141-149. [doi: 10.1016/s0010-4825(02)00011-2]
44. Batterham PJ, Calear AL, Farrer L, McCallum SM, Cheng VW. FitMindKit: Randomised controlled trial of an automatically tailored online program for mood, anxiety, substance use and suicidality. Internet Interv 2018 Jul;12:91-99 [FREE Full text] [doi: 10.1016/j.invent.2017.08.002] [Medline: 30135773]
45. Walsh A, Richards D. Experiences and engagement with the design features and strategies of an internet-delivered treatment programme for generalised anxiety disorder: a service-based evaluation. Br J Guid Counsell 2016 Feb 28;45(1):16-31. [doi: 10.1080/03069885.2016.1153039]
46. Hilvert-Bruce Z, Rossouw PJ, Wong N, Sunderland M, Andrews G. Adherence as a determinant of effectiveness of internet cognitive behavioural therapy for anxiety and depressive disorders. Behav Res Ther 2012 Aug;50(7-8):463-468. [doi: 10.1016/j.brat.2012.04.001] [Medline: 22659155]

Abbreviations

CBT: cognitive behavioral therapy
ECOUTER: Employing Conceptual Schema for Policy and Translation Engagement in Research
iCBT: internet cognitive behavioral therapy
REDCap: Research Electronic Data Capture
SUS: System Usability Scale
2SLGBTQ+: two-spirit, lesbian, gay, bisexual, transgender, queer, and all other members of this community

Edited by J Torous; submitted 04.09.21; peer-reviewed by K Stearns-Yoder, J Bierbooms; comments to author 04.10.21; revised version received 29.11.21; accepted 05.12.21; published 02.02.22.

Please cite as:
Patterson VC, Rossi MA, Pencer A, Wozney L
An Internet-Based Cognitive Behavioral Therapy Program for Anxiety and Depression (Tranquility): Adaptation Co-design and Fidelity Evaluation Study
JMIR Form Res 2022;6(2):e33374
URL: https://formative.jmir.org/2022/2/e33374
doi:10.2196/33374
PMID:34910660

Original Paper

Implementation of an mHealth App to Promote Engagement During HIV Care and Viral Load Suppression in Johannesburg, South Africa (iThemba Life): Pilot Technical Feasibility and Acceptability Study

Samanta Tresha Lalla-Edward[1], PhD; Nonkululeko Mashabane[1], MBL; Lynsey Stewart-Isherwood[2], MSc; Lesley Scott[2], PhD; Kyle Fyvie[2], MSc; Dana Duncan[3], MD; Betiel Haile[3], MPH; Kamal Chugh[3], MS; Yiyong Zhou[3], PhD; Jacob Reimers[4], MS; Matteus Pan[3], BS; Maya Venkatraman[3], PhD; Wendy Stevens[2,5], PhD

[1]Ezintsha, Faculty of Health Sciences, University of the Witwatersrand, Parktown, South Africa

[2]Molecular Medicine & Haematology, University of the Witwatersrand, Parktown, South Africa

[3]Roche Molecular Systems, Inc, Pleasanton, CA, United States

[4]Roche Molecular Systems, Rotkreuz, Switzerland

[5]National Priorities Programme, National Health Laboratory Systems, Parktown, South Africa

Corresponding Author:
Samanta Tresha Lalla-Edward, PhD
Ezintsha
Faculty of Health Sciences
University of the Witwatersrand
Sunnyside Office Park
32 Princess of Wales Terrace
Parktown, 2193
South Africa
Phone: 27 0826172490
Email: slallaedward@ezintsha.org

Abstract

Background: South Africa has the largest HIV treatment program worldwide. Retention in care and medication adherence remain problematic necessitating innovative solutions for improving HIV care. The increasing availability and use of mobile technology can support positive clinical outcomes for persons living with HIV. iThemba Life is a mobile health app designed with input from South African health professionals and patients, promoting engagement with HIV care through access to medical results.

Objective: This study aimed to test the feasibility and acceptability of receiving HIV viral load (VL) results through the app and compare the time to HIV VL result return for study participants before and after app use.

Methods: Using convenience sampling, adults having routine VL phlebotomy were recruited from 2 Johannesburg health facilities. After signed consent, the app was downloaded on their Android smartphones, phlebotomy was performed, and the sample barcode was scanned through their phone to link the sample and app. Participants received a notification of the result availability and logged into the app to view results, their explanation and recommended action.

Results: Overall, 750 people were screened to enroll 500 participants. Of 750, 113 (15.1%) failed eligibility screening. 21.5% (137/637) had smartphone technical limitations preventing enrollment. Results were released to 92.2% (461/500) of participants' phones. App technical issues and laboratory operational issues limited the number of released results. Approximately 78.1% (360/461) results were viewed in the app. Median time from notification of availability to result viewed being 15.5 hours (0.6; range 0-150 days). Turnaround time from phlebotomy to the result being received was 6 (range 1-167) days for users versus 56 days (range 10-430 days; $P<.001$) before app use. Overall, 4% (20/500) of participants received unsuppressed results (VL>1000 copies/mL). Turnaround time for unsuppressed results was 7 days for participants versus 37.5 days before app use ($P<.001$). The difference before and after app use in the suppressed and unsuppressed users for time from sample collection to result delivery was statistically significant. Of 20 participants, 12 (60%) returned for a confirmatory VL during the study period. The time from an unsuppressed VL to a confirmatory VL was 106 days for app users versus 203 days before app use ($P<.001$). Overall, 52.4%

(262/500) of participants completed an exit survey; 23.2% (58/250) reported challenges in viewing their VL results. Moreover, 58% (35/60) reported that they overcame challenges with technical assistance from others, and 97.3% (255/262) wanted to continue using the app for VL results.

Conclusions: Using iThemba Life for VL results was well-received despite limited smartphone access for some participants. App users received results 10 times sooner than before the app and 5 times sooner if their VL >1000 copies/mL. This increased notification speed led to participants wanting to continue using iThemba Life.

(JMIR Form Res 2022;6(2):e26033) doi:10.2196/26033

KEYWORDS

HIV; virological suppression; mHealth; digital health; South Africa; patient-centric; disease management; mobile phone

Introduction

As of 2018, there are 37.9 million people living with HIV (PLWH) worldwide, of which South Africa accounts for 7.5 million (19%) [1]. South Africa has the largest HIV treatment program [2], providing antiretroviral therapy (ART) to 4.4 million people [3]. Despite the high number of treatments, retention of care and medication adherence remain problematic, as only 86% of PLWH successfully achieve viral suppression [4], which is short of the 95% goal outlined by the 95-95-95 Joint United Nations Programme on HIV/AIDS initiative to end the AIDS epidemic by 2030 [5]. Despite increasing viral load (VL) testing, the results do not often lead to changes in the clinical management needed to achieve viral suppression because of clinic attendance and health system factors that delay clinical decision-making [6-8]. Improvements in health care delivery are required to achieve improved clinical outcomes expected with routine VL testing.

The increasing availability and use of mobile technology provide an opportunity for improving health outcomes in low- and middle-income countries [9], and a variety of mobile health (mHealth) interventions have been developed to support the HIV care cascade [10-13]. Traditionally, mHealth platforms have concentrated on text-messaging or phone call interventions [10,11], and SMS which has been shown to improve HIV retention and VL suppression [14]. However, current trends are moving toward app-based interventions that allow the tailoring of content to match the needs and preferences of users and can provide multimedia content to enhance participation and further motivate behavior change. Apps have been shown to be acceptable options in South Africa along the HIV care cascade. In a randomized controlled trial, the SmartLink app was shown to improve linkage to care for those newly diagnosed as HIV-positive by 20% [12]. The Aspect HIVST app has recently undergone a feasibility study in Johannesburg, where the acceptability of the app was high [13].

iThemba Life complements existing conventional ART adherence and viral suppression strategies through the electronic delivery of health information and HIV VL results directly from the laboratory to the user. Our hypothesis is that the iThemba app can provide simplified logistics for timely delivery of HIV VL results and support linkage to care. The objective of this pilot study was to assess the technical feasibility and acceptability of the use of iThemba Life to return HIV VL results and compare the time to HIV VL result return for study participants before and after app use. This is the first report of an evaluation of the iThemba Life app.

Methods

Study Design and Participants

In this mixed methods study, 750 HIV-positive adults receiving ART at the Hillbrow Community Health Center and Yeoville Clinic in Johannesburg, South Africa, were screened to enroll a convenience sample of 500 participants. Participants were recruited while waiting to have their blood collected for HIV VL testing and were included if they were aged ≥18 years, comfortable reading English, reported regular access to an Android phone and internet, had blood collected for an HIV VL test on the day of enrollment, and signed the consent form. Participants were excluded if they were not currently on ART, were receiving antenatal care services, had iThemba Life previously downloaded on their phone, or refused to participate. After enrollment, the app was downloaded on their Android smartphones, phlebotomy was performed, and the sample barcode was scanned using their phone camera to link the sample and app.

Recruitment and Data Collection

A trained, good clinical practice–certified study team, comprising a nurse, recruiter, and counselor, implemented the study at each site. The study team members approached clients, informed them about the study, and invited them to be screened for participation. Potential participants who met the eligibility criteria were provided with an overview of the study, asked to read the informed consent, and then asked to sign it if they agreed to participate.

A study team member assisted with the installation of the app on the participant's phone (Wi-Fi was provided by study-supported mobile hot spots) and then assisted the participant with setting up an account and provided an overview of app use. Any user unable to download or register on the app was withdrawn from the study, and the consent form was updated to reflect that no medical information would be collected. The recruiter recorded the number of approached clients and the reason for nonparticipation in a screening log for any patient found to be ineligible. If a participant was taken out of the queue for the recruitment process, they had their routine HIV VL collected by the study nurse, and if, instead, they returned to the blood collection queue, their place was maintained. Recruitment took place in April 2019, and

participants were followed up until October 2019, when a study exit interview was conducted.

Several logs were maintained during the recruitment and data collection processes (ie, screening log [mentioned above], enrollment log, and VL result log). At enrollment, data collected included ART regimen, ART initiation date, standard time for receiving an HIV VL result, HIV status disclosure, type of phone, and perceived level of HIV knowledge. The VL result log contained the results received through iThemba Life and previous results extracted from the user's medical records. In addition, there were in-app user experience questions. All users were invited to participate in an exit interview. The interviews were conducted in English using a structured interview guide. Data were collected on demographics, satisfaction with the app, experiences with viewing results, willingness to recommend the app, likes, dislikes, and recommendations for subsequent iterations. Finally, 2 focus group discussions (FGDs) were conducted with a total of 9 (Yeoville Clinic: n=5, 56%, including n=4, 80% male and n=1, 20% female; Hillbrow Community Health Center: n=4, 44%, who were all female) consenting nurses. The discussions were facilitated in English. Trained staff used a semistructured interview guide to gather information on the nurses' opinions and perceptions of iThemba Life. FGDs were audio recorded and transcribed verbatim.

Throughout the study, app users received all HIV care from the clinic staff per standard of care and not from the study team. HIV VL testing frequency followed national guidelines, and no additional specimens were collected as part of this evaluation. The samples were transported to a central laboratory for testing following standard sample transport procedures for the health facility. The study team did not provide HIV clinical care to standardize clinical care for app users and nonusers.

Delivery of Results

Once the central laboratory performed the testing, participants received a notification of a message waiting to be viewed on the app. Users could log into iThemba Life to view their results. The laboratory results were presented clearly, with information explaining the results and recommended actions for the user to take (Figure 1). Participants who did not achieve viral suppression were told to return to the clinic for adherence counseling and a repeat HIV VL test in 2 to 3 months according to national guidelines. If participants returned for another HIV VL test during the study period, they scanned their sample barcode with their phone to link the specimen and the app and followed the same procedure described above. Participants who did not view their results were sent a reminder notification 3 days and 7 days after the result was available.

The study flow describing the recruitment, data collection, and delivery of results is shown in Figure 2.

Figure 1. Screenshot of in-app laboratory result explanations.

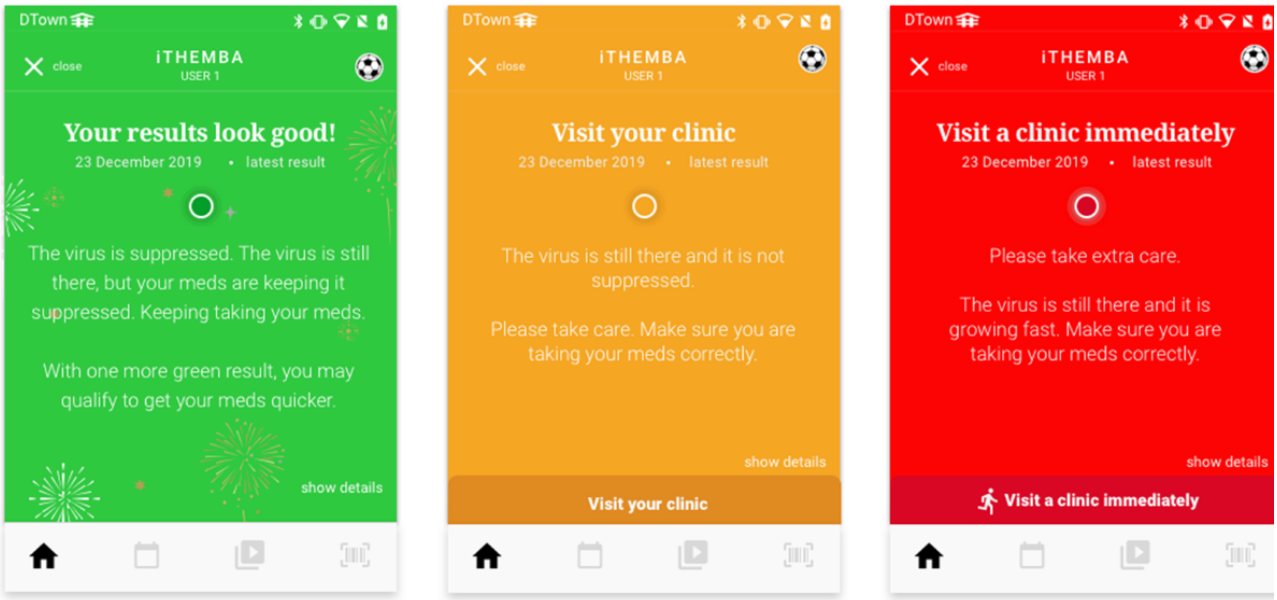
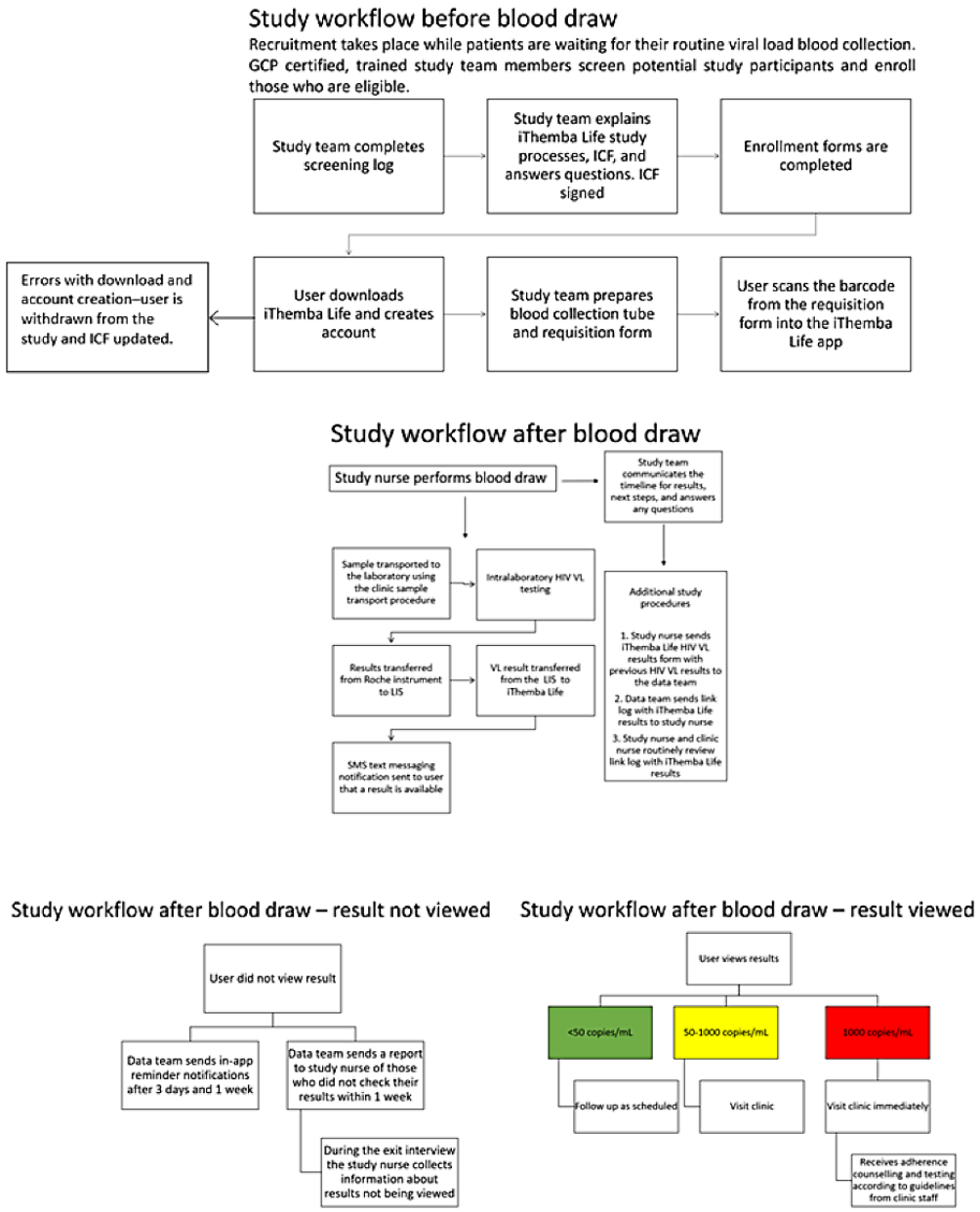


Figure 2. Study workflow. GCP: good clinical practice; ICF: informed consent form; LIS: laboratory information system; VL: viral load; SMS: short messaging service.



Outcome Measures

The primary objective was to assess the technical feasibility and acceptability of the use of iThemba Life to return HIV VL results. The secondary objective was to compare the time to HIV VL result return for study participants before and after app use.

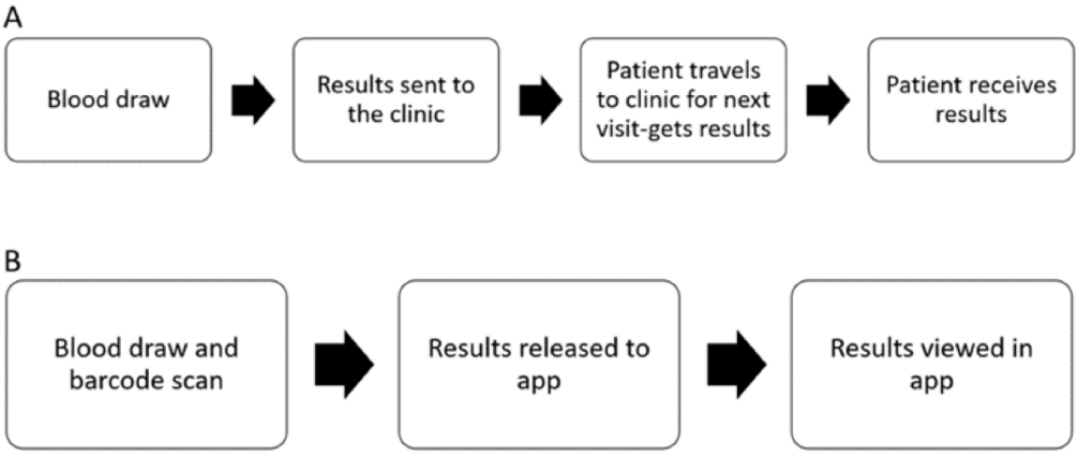
The technical feasibility of receiving HIV VL results through the app was assessed by process indicators, including the number of individuals screened, HIV VL specimens collected, HIV VL results returned, HIV VL results viewed, time to viewing HIV VL result, and the time to next clinic visit for those with unsuppressed HIV VL requiring follow-up according to national HIV care and treatment guidelines. The acceptability of using the app to return HIV VL results was assessed through in-app user experience surveys using a 5-point Likert scale and through a structured study exit interview. The interviews were conducted in English using a structured interview guide. The collected data included demographics, experience with viewing the results, willingness to recommend the app, likes, dislikes, and recommendations for subsequent iterations. FGDs were conducted with health care workers from the study sites to understand their opinions and perceptions of the app.

To assess the impact of the app on service delivery, study staff extracted data from health facility registers used for routine patient management and from enrollment questionnaires, including the most recent HIV VL results, date of the results being communicated to the individual, and the date of subsequent clinic visits for those with an unsuppressed HIV VL. The median time between sample collection and HIV VL result return was compared for study participants' before app use when receiving routine service delivery and after app use

when receiving app-supported service delivery. As an exploratory end point, the time from the result being received to repeat, confirmatory VL for individuals who were unsuppressed was analyzed. Both the World Health Organization and South African HIV Care and Treatment Guidelines recommend a confirmatory VL before ART regimen change [15,16].

The logistics required for HIV VL result return in routine service delivery is complex and often delays result delivery [17]. The use of the app allows for direct communication between the laboratory and the app user, reducing the time from sample collection to the results being returned to the individual and potentially reducing the time to a confirmatory VL and clinical management when needed (Figure 3).

Figure 3. Routine HIV viral load result return (A) before app use and (B) app-supported HIV viral load result return workflow.



Data Analysis

Quantitative data were analyzed with descriptive statistics using SAS (version 9.4; SAS Institute Inc). A chi-square test was performed to evaluate the relationship between demographic factors and time to results being viewed after notification of result availability. The Fisher exact test was used to compare the acceptability of the app by age, and a median 2-sample test was used to compare time from sample collection to result delivery before and after app use. $P<.001$ was considered significant. Qualitative data involved thematic analysis using SAS (version 9.4).

App Development

iThemba Life was developed by Roche Molecular Systems and designed with the input and feedback of >40 health care professionals and >60 PLWH in South Africa. Demonstrations of prototypes provided qualitative insights used to improve and refine the app in preparation for the pilot.

iThemba Life integrates with the local laboratory information system or results database and allows the delivery of any laboratory result from any manufacturer.

The iThemba Life mHealth platform safeguards and manages sensitive patient data. iThemba Life is designed to track sensitive data, both in transit and at rest in the cloud, with prevailing data privacy and security regulations in mind. The security mechanisms of iThemba Life are built on the principles of segregation, minimization, encryption, and monitoring.

Under the principle of minimization of protected health information, users' names or ID numbers were not collected. The collected data included the phone number, gender, year of birth and name of clinic, and the HIV VL result from the laboratory results database. Data privacy was considered as a design principle throughout the system development. As a result, data collection and representation sought to minimize the risk of disclosure and maximize user privacy. The app icon and name were not suggestive of a health app, and all app-based push notifications provided no detail about the nature of the information, ensuring that all health-related details were only displayed within the password-protected section of the app. Users could delete their profile information from the iThemba Life platform whenever they so chose to comply with local data rights and privacy laws. This would not delete any data that the laboratory or clinic had in their medical files.

Ethical Considerations

The human research ethics committee of the University of the Witwatersrand provided approval for the study (M181025), and permission was granted by the Department of Health Johannesburg Health District office (National Health Research Database reference GP_201810_009). Trained study staff obtained written informed consent from all the study participants. Participants completing the exit interview received a ZAR 150 (US $10.70) reimbursement for their time.

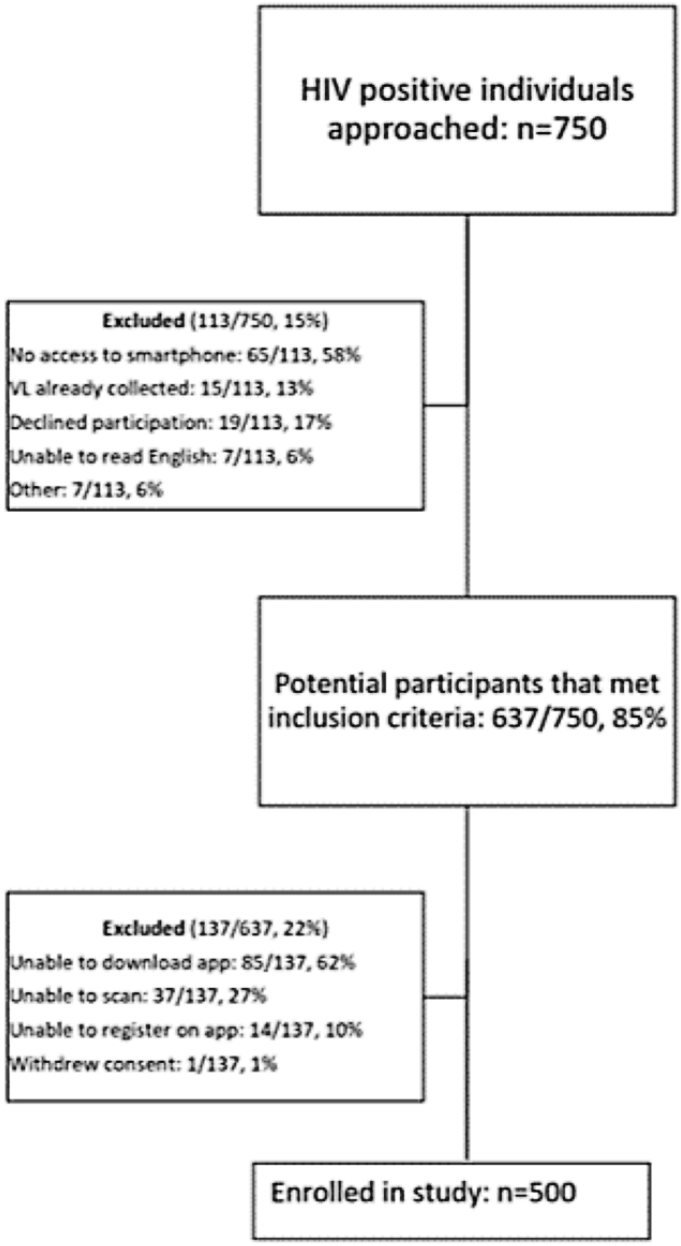
Results

Screening and Enrollment

Over the course of the study, 750 (Hillbrow Community Health Center n=383, 51.1% and Yeoville Clinic n=367, 48.9%) people were screened to enroll 500 (66.7%; Hillbrow Community Health Center n=238, 47.6% and Yeoville Clinic n=262, 52.4%) participants (Figure 4). Of the 750 people, 113 (15.1%) did not meet inclusion criteria, as 65 (57.5%) people did not have access to a smartphone, 15 (13.3%) had a VL sample already collected, 19 (16.8%) declined to participate, 7 (6.2%) could not read English, and 7 (6.2%) were excluded for other reasons. Of the remaining 637 people, 137 (21.5%) of those eligible were not

enrolled because of their smartphone's technical limitations, as 85 (62%) individuals were unable to download the app, 37 (27%) were unable to scan the sample barcode, 14 (10.25%) were unable to register on the app, and 1 (0.7%) withdrew consent.

Figure 4. Recruitment flow. VL: viral load.



Participant Characteristics

Of the 500 participants, 325 (65%) were female. The median age of the patients was 37 years. Of the 500 participants, age disaggregation comprised 3 (0.6%) participants being aged <20 years, 79 (15.8%) being aged 21 to 30 years, 246 (49.2%) being aged 31 to 40 years, 139 (27.8%) being aged 41 to 50 years, 27 (5.4%) being aged 51 to 60 years, and 1 (0.2%) being aged >60 years. Approximately 91.8% (459/500) of users were on first-line ART (tenofovir disoproxil fumarate/lamivudine [or emtricitabine]/efavirenz), whereas 8% (40/500) were on second-line treatment (zidovudine/lamivudine/ritonavir-boosted atazanavir [or ritonavir-boosted lopinavir]), and 0.2% (1/500) had an unknown treatment regimen. Of the 500 participants, 249 (49.8%) could not recall how long it usually took to receive their VL result after a blood draw, and 234 (46.8%) stated that it took >1 month to receive their VL result.

Regarding disclosure of HIV status, although most (481/500, 96.2%) had disclosed their HIV status to someone in their household/family, there was concern about disclosure of their status to others (303/500, 60.6%). Just under half (247/500, 49.4%) of the participants reported that they were very informed about HIV, and some (54/500, 10.8%) participants reported that they regularly used a shared phone. Table 1 shows the baseline characteristics of the study participants.

The exit interview was completed by 52.2% (261/500) of participants: 52.1% (136/261) from Hillbrow Community Health

Center and 47.9% (125/261) from Yeoville Clinic. Most (209/261 80.1%) had completed high/secondary school, and some (133/261, 50.9%) were employed full-time. Approximately 70.1% (183/261) of participants earned <ZAR 5000 (US $357) per month. Most respondents (218/261, 83.5%) reported having consistent network coverage throughout the day, and some (197/261, 75.5%) stated that they would have downloaded the iThemba Life app without access to free Wi-Fi at the health facility. Users also reported that the median time to travel to the clinic was 30 minutes, at the cost of <ZAR 20 (US $1.40) for 51.3% (134/261) of participants. A median of 2 hours was spent at the clinic, and 96.2% (251/261) reported that they received the services they required at their last clinic visit.

Table 1. Demographic and other characteristics of participants (N=500).

Characteristics	Overall (N=500)	Hillbrow Community Health Center (n=238)	Yeoville Clinic (n=262)
Sex, n (%)			
Male	175 (35)	79 (33.2)	96 (36.6)
Female	325 (65)	159 (66.8)	166 (63.4)
Age (years), n (%)			
<20	3 (0.6)	2 (0.8)	1 (0.4)
21-30	79 (15.8)	30 (12.6)	49 (18.7)
31-40	246 (49.2)	111 (46.6)	135 (51.5)
41-50	139 (27.8)	75 (31.5)	64 (24.4)
51-60	27 (5.4)	16 (6.7)	11 (4.2)
>60	1 (0.2)	1 (0.4)	0 (0)
Unknown	5 (1)	3 (1.3)	2 (0.8)
Age (years)			
Values, mean (SD; range)	38 (7.8; 19-61)	38 (8; 19-61)	37 (7.5; 19-60)
Values, median (IQR)	37 (32-42)	39 (33-44)	37 (32-42)
ARV[a] treatment, n (%)			
First-line: TDF[b]/3TC[c] (or FTC[d])/EFV[e]	459 (91.8)	203 (85.3)	256 (97.7)
Second-line: AZT[f]/3TC/ATV(r)[g] or LPV(r)[h]	40 (8)	34 (14.3)	6 (2.3)
Unknown	1 (0.2)	1 (0.4)	0 (0)
Have users disclosed status to anyone in household or family?, n (%)			
Yes	481 (96.2)	222 (93.3)	259 (98.9)
No	19 (3.8)	16 (6.7)	3 (1.1)
User concerned about disclosing condition to others?, n (%)			
Very concerned	91 (18.2)	47 (19.7)	44 (16.8)
Concerned	78 (15.6)	63 (26.5)	15 (5.7)
Somewhat concerned	134 (26.8)	69 (29)	65 (24.8)
Not concerned	60 (12)	31 (13)	29 (11.1)
Not at all Concerned	137 (27.4)	28 (11.8)	109 (41.6)
Informed about HIV, n (%)			
Very informed	247 (49.9)	119 (50)	128 (48.9)
Somewhat informed	231 (46.2)	112 (47.1)	119 (45.4)
Not informed	22 (4.4)	7 (2.9)	15 (5.7)
Shared phone, n (%)			
Yes	54 (10.8)	23 (9.7)	31 (11.8)
No	446 (89.2)	215 (90.3)	231 (88.2)

[a]ARV: antiretroviral.

[b]TDF: tenofovir disoproxil fumarate.

[c]3TC: lamivudine.

[d]FTC: emtricitabine.

[e]EFV: efavirenz.

[f]AZT: zidovudine.

[g]ATV(r): atazanavir/ritonavir.

[h]LPV(r): lopinavir/ritonavir.

Technical Feasibility

All participants were able to download and register for the iThemba Life app and successfully scan the sample barcode (Table 2). Of the 500 sample bar codes scanned, 461 (92.2%) HIV VL results were obtained from the app database, and of these 461 results, 360 (78.1%) were viewed by users.

All participants managed to scan in their barcoded samples; however, only 92.2% (461/500) results were linked to the iThemba Life database. Of those, most of the results (360/461, 78.1%) were viewed by the participants.

Table 2. Feasibility of receiving viral load results through the app (N=500).

	Overall (N=500)	Hillbrow Community Health Center (n=238)	Yeoville Clinic (n=262)
Successful initial viral load samples scanned	500 (100)	238 (100)	262 (100)
Viral load results received by app database	461 (92.2)	218 (91.6)	243 (92.7)
Viral load viewed by user	360 (78.1)	178 (81.7)	182 (74.9)

Acceptability

Of the 500 users, 480 (96%) responded on the ease of scanning barcode questions through in-app surveys, and of these 480 users, 347 (72.3%) found it easy to perform. When asked about the ease of logging into the app, 52.2% (261/500) of users responded, and of these, 78.2% (204/261) found it easy to perform. For the question on the ease of receiving a VL result through iThemba Life, 19.2% (96/500) of users responded, and of these, 82.3% (79/96) found it easy.

During the exit survey, 79.1% (204/258) of respondents reported that they were very happy with the app, 79.6% (191/240) completely trusted the information from the app, 75.7% (196/259) were likely to recommend the app to others, with 78.4% (203/259) reporting that they already told someone else about the app.

Most participants (227/244, 93%) reported that the app was very helpful, 85.5% (207/242) reported it was easy or very easy to understand the result, and 97.3% (255/262) reported that they wanted to continue to use the app to receive their HIV VL results. Most participants noted the benefit of the app in assisting with the self-management of their disease. Some of their statements included the following:

> *Most of the time when I get results no one explains them to me but with iThemba Life I managed to get an explanation.* [Yeoville Clinic, female]
>
> *It helped me because I didn't have to spend money to come to fetch my results* [Yeoville Clinic, male]
>
> *Because it was easy to open it and get results immediately* [Yeoville Clinic, female]

In the exit survey, of the 500 people, 55 (11%) noted experiencing challenges while viewing their VL results. These included opening the result (30/55, 55%) and data and network access (16/55, 27%) issues. Considering the challenges experienced with iThemba Life (60/500, 12%), these were primarily because of limited technical skills, and many (35/60, 58%) reported that they overcame the challenge with technical assistance from others. Challenges that could not be resolved were because of factors beyond the user's control (eg, lost phone or broken phone). Of the 500 participants, <3% (n=7, 1.4%; n=4, 57% from Hill Brow Community Health Center and n=3, 43% from Yeoville Clinic) indicated that they did not want to continue using the app. Similar to the challenges reported above, these were not for app-related reasons but rather for other reasons such as no longer having a phone.

The age distribution of the participants who responded to the acceptability questions is reported in Multimedia Appendix 1. There was no statistically significant difference in the acceptability of the app for any age category.

When asked about potential future features of the app, >248 respondents wanted health information and reminders for clinic appointments and medication pickups sent by the app, whereas 75.1% (196/261) participants wanted a daily reminder to take their medication. Some of their requests included the following:

> *Give information on types of food we can eat or not.* [Hillbrow Community Health Center, female]
>
> *To get updates about HIV every month.* [Hillbrow Community Health Center, female]

During the FGDs, nurses noted the value of iThemba Life for the users and for themselves. Some of their statements included the following:

> *I have seen patients showing the result with excitement. Before iThemba Life, they never discussed their viral load.* [Hillbrow Community Health Center, nurse]
>
> *It will be useful, instead of sending the patient to obtain their result they will have access to it.* [Hillbrow Community Health Center, nurse]
>
> *It cuts the time when I need to contact the patient. There will be time saved.* [Yeoville Clinic, nurse]
>
> *Two patients came with high viral load and they were referred to the clinic and switched their treatment.* [Hillbrow Community Health Center, nurse]

Service Delivery Outcome

Of the 360 results viewed by app users, 245 (68.1%) were <50 copies/mL, 92 (25.6%) were 50 to 1000 copies/mL, and 20 (5.6%) were >1000 copies/mL. Approximately 0.6% (2/360) of results were reported as invalid and 0.3% (1/360) that was reported as an inadequate specimen by the laboratory.

The overall median number of days from sample collection to the result being received for the app users was 6 (range 1-167) days, whereas the median number of days before app use was 56 days (range 10-430 days; $P<.001$; Table 3). The median time

from notification of result availability to results being viewed was 0.6 (range 0-150) days. For unsuppressed users, the overall median days from sample collection to results being received was 7 days for app users and 37.5 days before app use (*P*<.001). For suppressed users, the overall median days from sample collection to the results being received was 6 days for app users and 56 days before app use (*P*<.001). An exploratory analysis of the time to a confirmatory HIV VL found that 60% (12/20) of app users with unsuppressed results returned for a confirmatory VL within the study period. The time from an unsuppressed VL to a confirmatory VL was 106 days for app users and 203 days before app use (*P*<.001).

Before app use, participants would have a VL sample collected and receive their results at their next scheduled appointment, which could be 1 to 6 months later. When there were unsuppressed results, the facility may call the patient to return to the facility sooner. This telephonic follow-up is not consistent and contingent on the availability of both human resources and a working phone and whether the results were reviewed by clinic staff when the facility received them from the laboratory.

Table 3. Time from sample collection to result delivery before and after app use (N=500)[a].

Results delivery	Overall, median	Participants, n (%)	P value[b]
Number of days from specimen collection to result receipt: all users (days)			
Before app use[c]	56.0	251 (50.2)	<.001
App use	6.0	360 (72)	—[d]
Time from notification of result availability to results being viewed (hours)			
App use	15.5	360 (72)	—
Before app use	—	—	—
Number of days from specimen collection to result receipt: unsuppressed users (days)			
Before app use[e]	37.5	14 (2.8)	<.001
App use	7.0	20 (4)	—
Number of days from specimen collection to result receipt: suppressed users (days)			
Before app use	56.0	234 (46.8)	<.001
App use	6.0	337 (67.4)	—
Number of days from first to second sample collection: unsuppressed users (days)			
Before app use	203.0	27 (5.4)	<.001
App use	106.0	12 (2.4)	—

[a]Users included in this table have either preapp or app use turnaround data.

[b]Median 2-sample test was used to estimate the P values.

[c]The standard-of-care viral load result delivery times for nonapp users were determined by the date of the next visit after viral load sample collection for users before using the app.

[d]Not available.

[e]The standard-of-care viral load result delivery times for unsuppressed nonapp users were determined by the result delivery times for viral load results before using the app users.

Association Analysis

The chi-square tests of independence were performed to examine the relationships between the results viewed and age, sex, and time since ART initiation and results were not significant (Table 4).

Table 4. Sociodemographic characteristics and feasibility of receiving HIV viral load (VL) result through a mobile app (N=500)[a].

Variable	HIV VL result received through mobile app	HIV VL result received through mobile app (viewed)	HIV VL result received through mobile app (not viewed)	P value[b]
Age (years), n (%)				.12
18-25	26 (5.2)	21 (4.2)	5 (1)	
26-45	359 (71.8)	287 (57.4)	72 (14.4)	
46-65	71 (14.2)	49 (9.8)	22 (4.4)	
Unknown	5 (1)	3 (0.6)	2 (0.4)	
Sex, n (%)				.25
Male	165 (33)	124 (24.8)	41 (8.2)	
Female	296 (59.2)	236 (47.2)	60 (12)	
Unknown	0 (0)	0 (0)	0 (0)	
Time since ART[c] initiation (years), n (%)				.64
<1	31 (6.2)	9 (1.8)	22 (4.4)	
1-2	103 (20.6)	22 (4.4)	81 (16.2)	
3-5	147 (29.4)	29 (5.8)	118 (23.6)	
>5	150 (30)	36 (7.2)	114 (22.8)	
Unknown	30 (6)	25 (5)	5 (1)	

[a]Only participants who received HIV viral load through the mobile app are included in this table.

[b]Chi-square test of independence was used to estimate P values. Data from unknown categories were not included in the P value estimates.

[c]ART: antiretroviral therapy.

Discussion

Principal Findings

Similar to a few other digital health efforts aimed at supporting HIV care [12,13,18], this intervention demonstrated notable outcomes and was highly accepted by users. Although the SmartLink adherence app increased linkage to care in youth by 20% [12], iThemba Life users received VL results 10 times sooner than before app use (6 days vs 56 days) and 5 times sooner (7 days vs 37.5 days) than before app use if their HIV VL was >1000 copies/mL. There was a significant difference before and after app use in both suppressed and unsuppressed users with respect to time from sample collection to result delivery. Faster delivery of HIV VL results may significantly improve clinical outcomes.

As HIV/AIDS programs have matured, the fast, reliable delivery of laboratory results has become essential to achieve disease elimination. Digital health can be a powerful tool for increasing the delivery of, and access to, high-quality health services [19]. Front-line health workers are often considered as first-level recipients of digital health [19]; however, the increasing availability of smartphones facilitates increased patient engagement by promoting a shift from text-based to interactive app-based solutions [20].

Access to smartphone technology is rapidly increasing, leapfrogging traditional service delivery constraints, and overcoming infrastructure barriers. The penetration of smartphones is not yet universal throughout Sub-Saharan Africa but has accelerated significantly in recent years with the availability of lower-priced devices, and it is expected to continue increasing to 67% in 2025 [21]. iThemba Life leverages the increasing availability and use of mobile technology to simplify the manual and complex logistics currently used for HIV VL result delivery and ensures that results are available for timely clinical decision-making. The app was designed to complement existing conventional ART adherence and viral suppression strategies through the electronic delivery of health information and HIV VL results directly from the laboratory to the user. This is the first report of an evaluation of the iThemba Life app.

iThemba Life followed an iterative development strategy with continual input and collaboration from patients, health care workers, and other key stakeholders to address sociocultural and workflow aspects related to HIV program delivery. During the development of iThemba Life, focusing on understanding user needs and experience with early prototypes, limiting the installation size, and ensuring that the app could be used on older and less expensive smartphones was essential. This early focus ensured that operational and technical challenges experienced by similar interventions in this setting [22] were minimized, and most of those screened in this study were able to access the intervention and had a positive experience with the app.

Data costs are considered high by many South Africans and typically hinder the uptake of digital health interventions [23]; however, most exit survey respondents reported that they would still have downloaded the app without free Wi-Fi access during this pilot. This not only qualifies satisfaction with the app but also highlights health care recipients' urgent need for improved

service delivery in that costs were outweighed by the ability to receive HIV VL results sooner. iThemba Life users recognized the value of quickly receiving their VL result as a motivator of adherence to treatment, a way to make health management more convenient, and as a source of health information. Many felt that the app was supportive in managing their disease and equipping them to assume control of their health. Health care workers echoed this value of the app for users and noted the benefits of having more informed clients and reliable access to VL results.

Innovative app-based technologies are accompanied by challenges concerning the phone capability and technical proficiency of users, and these areas have been prioritized for improvement following this pilot assessment. The inability to use this app on certain phones or by some users created barriers to entry for a few individuals and has been addressed, as iThemba Life is positioned to scale up with updates to the software to decrease technical challenges.

However, despite these challenges and recent studies advocating for non–app-based mHealth interventions [18,24], the high level of user acceptability coupled with the marked decrease in VL result delivery times indicates that iThemba Life provides value to potential users and the national HIV program. iThemba Life users with unsuppressed results returned for a confirmatory VL in half the typical time they would have, which may decrease the time spent on an ineffective ART regimen, potentially having individual health and population benefits by lessening the time to viral resuppression. The ability to receive VL results electronically decreases visits to health care facilities, which may reduce strain on facilities while also saving patients' time and money. South Africa has a proven track record of promoting and adopting innovative mHealth interventions such as MomConnect [25], which has >2 million registered users who receive prenatal support messages through the platform [26]. If South Africa was able to similarly scale up iThemba Life, PLWH could have significant financial and time savings on transportation and clinic waiting [27] while also improving their health outcomes and reducing the risk of transmission because of earlier resuppression times.

Using smartphones, iThemba Life is at the beginning of a new frontier of communication between patients and health care workers. This evolving platform will continually incorporate feedback from active users to improve its functionality and features. In the future, iThemba Life can continue to deliver information and services to patients and also create bidirectional feedback from patients to further assess the quality of health care delivery and foster a richer, more meaningful interaction with the health system. In addition, through smartphone technologies, iThemba Life offers the capability to create password-protected multiple user profiles on a single device to allow a device to be shared by families or communities, further expanding access to these services.

Limitations

Because of the mobile requirements of the app, connectivity issues and incompatibility with some phone models prevented some prospective users from participating. Furthermore, some of the phone cameras were not able to scan the barcode because of the fixed focal length of the camera, which did not permit close-up scanning of the barcode. A dry run of the complete app workflow was not conducted before the start of the study. Consequently, some app technical issues were not identified until after the study enrollment began. Technical issues with the app database itself and operational issues because of sample transport and testing limited the number of results released to the app. The placement of the user survey within the app flow resulted in many users not viewing and, consequently, not completing some or all of these in-app questions. The study staff faced challenges with users returning for the exit survey, which was due in part to the 2019 Johannesburg riots, limiting travel to the sites.

Johannesburg is a large urban center with more resources and services than many other cities in the country and on the African continent; therefore, the current data are not generalizable outside of Johannesburg. In addition, most participants in this study had been receiving ART for many years and were virally suppressed. Future studies would need to target users who have recently initiated ART or who are facing challenges in achieving or maintaining viral suppression.

Conclusions

This evaluation demonstrated the feasibility and acceptability of receiving HIV VL results directly to users' smartphones through the iThemba Life app while also decreasing the time for PLWH to receive their VL results. iThemba Life has the potential to increase treatment adherence and literacy, reduce barriers to entry for special populations, and reduce the burden of care for PLWH. Future studies are needed to further our understanding of how the support and information provided by iThemba Life affect the health outcomes of PLWH and how iThemba Life can support other disease areas.

Acknowledgments

The study team would like to thank the health care personnel at the Hillbrow Community Health Centre and Yeoville Clinic in Johannesburg and all the patients who participated in the screening and enrollment.

Authors' Contributions

All authors contributed to the study design, data analysis, and manuscript preparation. MP, MV, and JR developed the app software with input from all the authors. SLE and NM led the data collection. SLE wrote the first draft. All the authors critically reviewed and approved the final draft.

Conflicts of Interest

DD, BH, KC, YZ, JR, MP, and MV are employed by Roche Molecular Systems, Inc.

Multimedia Appendix 1
Acceptability of iThemba app by age (years; N=500).
[XLSX File (Microsoft Excel File), 11 KB - formative_v6i2e26033_app1.xlsx]

References

1. UNAIDS data 2019. UNAIDS. 2019. URL: https://www.unaids.org/en/resources/documents/2019/2019-UNAIDS-data [accessed 2022-01-02]
2. Mid-year population estimates 2018. Statistics South Africa. 2018. URL: http://www.statssa.gov.za/?p=11341 [accessed 2022-01-02]
3. UNAIDS data 2018. UNAIDS. 2018. URL: https://www.unaids.org/en/resources/documents/2018/unaids-data-2018 [accessed 2022-01-02]
4. The Fifth South African National HIV Prevalence, Incidence, Behaviour and Communication Survey, 2017. HIV Impact Assessment Summary. 2018. URL: http://www.hsrc.ac.za/uploads/pageContent/9234/SABSSMV_Impact_Assessment_Summary_ZA_ADS_cleared_PDFA4.pdf [accessed 2022-01-02]
5. Understanding fast-track: accelerating action to end the aids epidemic by 2030. UNAIDS. 2015. URL: https://jliflc.com/resources/understanding-fast-track-accelerating-action-to-end-the-aids-epidemic-by-2030/ [accessed 2022-01-02]
6. Girdwood SJ, Crompton T, Sharma M, Dorward J, Garrett N, Drain PK, et al. Cost-effectiveness of adoption strategies for point of care HIV viral load monitoring in South Africa. EClinicalMedicine 2020 Dec;28:100607 [FREE Full text] [doi: 10.1016/j.eclinm.2020.100607] [Medline: 33294817]
7. Apollo T, Takarinda KC, Phillips A, Ndhlovu C, Cowan FM. Provision of HIV viral load testing services in Zimbabwe: secondary data analyses using data from health facilities using the electronic Patient Monitoring System. PLoS One 2021 Jan 22;16(1):e0245720 [FREE Full text] [doi: 10.1371/journal.pone.0245720] [Medline: 33481931]
8. Glass TR, Motaboli L, Nsakala B, Lerotholi M, Vanobberghen F, Amstutz A, et al. The viral load monitoring cascade in a resource-limited setting: a prospective multicentre cohort study after introduction of routine viral load monitoring in rural Lesotho. PLoS One 2019 Aug 28;14(8):e0220337 [FREE Full text] [doi: 10.1371/journal.pone.0220337] [Medline: 31461455]
9. Stephani V, Opoku D, Quentin W. A systematic review of randomized controlled trials of mHealth interventions against non-communicable diseases in developing countries. BMC Public Health 2016 Jul 15;16(1):572 [FREE Full text] [doi: 10.1186/s12889-016-3226-3] [Medline: 27417513]
10. Georgette N, Siedner MJ, Zanoni B, Sibaya T, Petty CR, Carpenter S, et al. The acceptability and perceived usefulness of a weekly clinical SMS program to promote HIV antiretroviral medication adherence in KwaZulu-Natal, South Africa. AIDS Behav 2016 Nov 18;20(11):2629-2638 [FREE Full text] [doi: 10.1007/s10461-016-1287-z] [Medline: 26781866]
11. Rana Y, Haberer J, Huang H, Kambugu A, Mukasa B, Thirumurthy H, et al. Short message service (SMS)-based intervention to improve treatment adherence among HIV-positive youth in Uganda: focus group findings. PLoS One 2015 Apr 16;10(4):e0125187 [FREE Full text] [doi: 10.1371/journal.pone.0125187] [Medline: 25881059]
12. Venter WD, Fischer A, Lalla-Edward ST, Coleman J, Lau Chan V, Shubber Z, et al. Improving linkage to and retention in care in newly diagnosed hiv-positive patients using smartphones in South Africa: randomized controlled trial. JMIR Mhealth Uhealth 2019 Apr 02;7(4):e12652 [FREE Full text] [doi: 10.2196/12652] [Medline: 30938681]
13. Gous N, Fischer AE, Rhagnath N, Phatsoane M, Majam M, Lalla-Edward ST. Evaluation of a mobile application to support HIV self-testing in Johannesburg, South Africa. South Afr J HIV Med 2020 Jun 30;21(1):1088 [FREE Full text] [doi: 10.4102/sajhivmed.v21i1.1088] [Medline: 32670629]
14. Kinuthia J, Ronen K, Unger JA, Jiang W, Matemo D, Perrier T, et al. SMS messaging to improve retention and viral suppression in prevention of mother-to-child HIV transmission (PMTCT) programs in Kenya: a 3-arm randomized clinical trial. PLoS Med 2021 May 24;18(5):e1003650 [FREE Full text] [doi: 10.1371/journal.pmed.1003650] [Medline: 34029338]
15. Cameron E. Introduction to the Southern African HIV Clinicians Society harm reduction guidelines. South Afr J HIV Med 2020 Dec 17;21(1):1179 [FREE Full text] [doi: 10.4102/sajhivmed.v21i1.1179] [Medline: 33391837]
16. World Health Organization. Updated Recommendations on HIV Prevention, Infant Diagnosis, Antiretroviral Initiation and Monitoring. Geneva: World Health Organization; 2021.
17. Peter T, Zeh C, Katz Z, Elbireer A, Alemayehu B, Vojnov L, et al. Scaling up HIV viral load - lessons from the large-scale implementation of HIV early infant diagnosis and CD4 testing. J Int AIDS Soc 2017 Nov 12;20 Suppl 7:e25008 [FREE Full text] [doi: 10.1002/jia2.25008] [Medline: 29130601]
18. DiAndreth L, Jarrett BA, Elf JL, Nishath T, Donville B, Heidari O, et al. Secure delivery of HIV-related and tuberculosis laboratory results to patient cell phones: a pilot comparative study. AIDS Behav 2020 Dec 15;24(12):3511-3521. [doi: 10.1007/s10461-020-02912-3] [Medline: 32415616]

19. Agarwal S, Perry HB, Long L, Labrique AB. Evidence on feasibility and effective use of mHealth strategies by frontline health workers in developing countries: systematic review. Trop Med Int Health 2015 Aug 14;20(8):1003-1014 [FREE Full text] [doi: 10.1111/tmi.12525] [Medline: 25881735]
20. Lee M, Lee H, Kim Y, Kim J, Cho M, Jang J, et al. Mobile app-based health promotion programs: a systematic review of the literature. Int J Environ Res Public Health 2018 Dec 13;15(12):2838 [FREE Full text] [doi: 10.3390/ijerph15122838] [Medline: 30551555]
21. The mobile economy 2020. GSM Association. 2020. URL: https://www.gsma.com/mobileeconomy/wp-content/uploads/2020/03/GSMA_MobileEconomy2020_Global.pdf [accessed 2022-01-02]
22. Venter W, Coleman J, Chan VL, Shubber Z, Phatsoane M, Gorgens M, et al. Improving linkage to HIV care through mobile phone apps: randomized controlled trial. JMIR Mhealth Uhealth 2018 Jul 17;6(7):e155 [FREE Full text] [doi: 10.2196/mhealth.8376] [Medline: 30021706]
23. Fischer AE, Sebidi J, Barron P, Lalla-Edward ST. The MomConnect nurses and midwives support platform (NurseConnect): a qualitative process evaluation. JMIR Mhealth Uhealth 2019 Feb 13;7(2):e11644 [FREE Full text] [doi: 10.2196/11644] [Medline: 30758298]
24. DiAndreth L, Krishnan N, Elf JL, Cox S, Tilchin C, Nthulana M, et al. Formative research for an mHealth program to improve the HIV care continuum in South Africa. AIDS Care 2020 Jun 12;32(6):744-748. [doi: 10.1080/09540121.2019.1640850] [Medline: 31298566]
25. Skinner D, Delobelle P, Pappin M, Pieterse D, Esterhuizen TM, Barron P, et al. User assessments and the use of information from MomConnect, a mobile phone text-based information service, by pregnant women and new mothers in South Africa. BMJ Glob Health 2018 Apr 24;3(Suppl 2):e000561 [FREE Full text] [doi: 10.1136/bmjgh-2017-000561] [Medline: 29713504]
26. Seebregts C, Dane P, Parsons AN, Fogwill T, Rogers D, Bekker M, et al. Designing for scale: optimising the health information system architecture for mobile maternal health messaging in South Africa (MomConnect). BMJ Glob Health 2018 Apr;3(Suppl 2):e000563 [FREE Full text] [doi: 10.1136/bmjgh-2017-000563] [Medline: 29713506]
27. Wang Z, Zhang L, Ma L, Liu B. Modeling medical services with mobile health applications. J Healthc Eng 2018;2018:1385034 [FREE Full text] [doi: 10.1155/2018/1385034] [Medline: 29849991]

Abbreviations

ART: antiretroviral therapy
FGD: focus group discussion
mHealth: mobile health
PLWH: people living with HIV
VL: viral load

Edited by G Eysenbach; submitted 25.11.20; peer-reviewed by J Simoni, R Mpofu, J Coleman; comments to author 05.03.21; revised version received 30.06.21; accepted 23.11.21; published 02.02.22.

Please cite as:
Lalla-Edward ST, Mashabane N, Stewart-Isherwood L, Scott L, Fyvie K, Duncan D, Haile B, Chugh K, Zhou Y, Reimers J, Pan M, Venkatraman M, Stevens W
Implementation of an mHealth App to Promote Engagement During HIV Care and Viral Load Suppression in Johannesburg, South Africa (iThemba Life): Pilot Technical Feasibility and Acceptability Study
JMIR Form Res 2022;6(2):e26033
URL: https://formative.jmir.org/2022/2/e26033
doi:10.2196/26033
PMID:35107427

Original Paper

Remote Assessment of Cardiovascular Risk Factors and Cognition in Middle-Aged and Older Adults: Proof-of-Concept Study

Jennifer A Eastman[1,2*], PhD; Allison R Kaup[1,2,3*], PhD; Amber L Bahorik[2*], PhD; Xochitl Butcher[4*], BSA; Mouna Attarha[5*], PhD; Gregory M Marcus[4], MD, MAS; Mark J Pletcher[6], MD, MPH; Jeffrey E Olgin[4*], MD; Deborah E Barnes[1,2,6*], PhD, MPH; Kristine Yaffe[1,2,6,7*], MD

[1]San Francisco VA Medical Center, San Francisco, CA, United States
[2]Department of Psychiatry and Behavioral Sciences, University of California, San Francisco, CA, United States
[3]The Neurology Center of Southern California, Carlsbad, CA, United States
[4]Department of Medicine, University of California, San Francisco, CA, United States
[5]Posit Science Corporation, San Francisco, CA, United States
[6]Department of Epidemiology and Biostatistics, University of California, San Francisco, CA, United States
[7]Department of Neurology, University of California, San Francisco, CA, United States
[*]these authors contributed equally

Corresponding Author:
Jennifer A Eastman, PhD
San Francisco VA Medical Center
4150 Clement St.
San Francisco, CA, 94121
United States
Phone: 1 951 760 6711
Email: jennifer.eastman@ucsf.edu

Abstract

Background: Adults with cardiovascular disease risk factors (CVRFs) are also at increased risk of developing cognitive decline and dementia. However, it is often difficult to study the relationships between CVRFs and cognitive function because cognitive assessment typically requires time-consuming in-person neuropsychological evaluations that may not be feasible for real-world situations.

Objective: We conducted a proof-of-concept study to determine if the association between CVRFs and cognitive function could be detected using web-based, self-administered cognitive tasks and CVRF assessment.

Methods: We recruited 239 participants aged ≥50 years (mean age 62.7 years, SD 8.8; 42.7% [n=102] female, 88.7% [n=212] White) who were enrolled in the Health eHeart Study, a web-based platform focused on cardiac disease. The participants self-reported CVRFs (hypertension, high cholesterol, diabetes, and atrial fibrillation) using web-based health surveys between August 2016 and July 2018. After an average of 3 years of follow-up, we remotely evaluated episodic memory, working memory, and executive function via the web-based Posit Science platform, BrainHQ. Raw data were normalized and averaged into 3 domain scores. We used linear regression models to examine the association between CVRFs and cognitive function.

Results: CVRF prevalence was 62.8% (n=150) for high cholesterol, 45.2% (n=108) for hypertension, 10.9% (n=26) for atrial fibrillation, and 7.5% (n=18) for diabetes. In multivariable models, atrial fibrillation was associated with worse working memory (β=-.51, 95% CI -0.91 to -0.11) and worse episodic memory (β=-.31, 95% CI -0.59 to -0.04); hypertension was associated with worse episodic memory (β=-.27, 95% CI -0.44 to -0.11). Diabetes and high cholesterol were not associated with cognitive performance.

Conclusions: Self-administered web-based tools can be used to detect both CVRFs and cognitive health. We observed that atrial fibrillation and hypertension were associated with worse cognitive function even in those in their 60s and 70s. The potential of mobile assessments to detect risk factors for cognitive aging merits further investigation.

(JMIR Form Res 2022;6(2):e30410) doi:10.2196/30410

KEYWORDS

mHealth; internet; mobile health; digital health; eHealth; cardiovascular; risk factors; cognition; cognitive impairment; remote cognitive assessment; aging

Introduction

In-person neuropsychological evaluation through conventional paper-and-pencil tests administered one-to-one in a clinic or lab by a trained psychometrist represents the gold standard for testing cognitive function [1-3]. However, there are drawbacks associated with this method of evaluation including limited availability, high cost, lengthy time commitment for both study participants and personnel, and limited accessibility of even the most commonly used cognitive testing measures in community settings with hard-to-reach or at-risk populations. These factors alone are reason enough to seek out more efficient means of collecting cognitive data for large-scale research. In the face of a global pandemic, the vital importance of developing remote cognitive assessment tools for both research productivity and improved patient care has never been more apparent.

Remote cognitive assessment as a mobile health (mHealth) technology offers the potential of increased flexibility and portability, and improved efficiency in neuropsychological testing [4]. mHealth is the use of mobile technology including phones, tablets, and wearable devices to collect data and interface with patients for treatment or research [5]. It provides the unique opportunity to study health risk factors remotely through web-based platforms without in-person visits. Web-based cognitive testing can address the weaknesses of traditional neuropsychological evaluation by showing improved accuracy in administration and scoring, adaptability to the performance of the test taker, decreased cost of administration, shorter and more precise batteries, and an increased accessibility for diverse and rural populations [1,6,7].

Traditional, in-person studies of cardiovascular risk factors (CVRFs) such as diabetes mellitus, atrial fibrillation, hypertension, and high cholesterol have reported associations with higher risk for cognitive decline and dementia in older adulthood [8-13]. With this in mind, we developed the Health eBrain study. We sought to determine whether a self-administered mobile cognitive intervention tool could detect the associations between CVRFs and cognitive performance using completely remote measures. We hypothesized that mobile measures of cardiovascular health would be associated with mobile measures of cognition in adults aged ≥50, such that those with CVRFs would demonstrate significantly worse cognitive performance.

Methods

Recruitment

Health eHeart conducts enrollment, consent, and participation entirely through the internet. Participation is open to any individual aged 18 or older with an email address. Recruitment for Health eHeart is worldwide and accomplished through news stories, email, social media, and word of mouth [14]. Health eBrain was designed as a companion study with recruitment through Health eHeart.

Statistical Analysis

To assess the association between CVRFs and cognitive performance, we sent email invitations to all Health eHeart Study participants enrolled from August 2016 to July 2018 who were aged 50 years or older, were English speaking, were not colorblind, and had internet access and sufficient computer proficiency to engage in online assessment. We used this time point as our goal was to assess cognition 2-3 years following enrollment. A total of 741 eligible individuals consented to participate in the study. Of these, 502 did not complete testing, and the final analytical cohort was 239 participants (mean age 62.74 years, SD 8.8; 42.7% [n=102] female, 88.7% [n=212] White), who completed the web-based testing between March 20, 2019, and May 31, 2019.

CVRF Measurements

Hypertension, high cholesterol, and diabetes diagnoses were established through self-report of a physician's diagnosis or by laboratory values obtained from self-report or mobile device monitoring and were defined by American Heart Association criteria for diabetes (fasting glucose >126 units), high blood pressure (BP >140/90 units), and high cholesterol (total cholesterol >240 units). Atrial fibrillation was identified through participant self-report of a physician's diagnosis, a method previously validated in the Health eHeart cohort [15].

Cognitive Assessment

Between March 2019 and June 2019, Health eBrain participants completed a series of cognitive tasks on the Posit Science BrainHQ platform [16], which was integrated with the Eureka Research Platform through an application programming interface. BrainHQ is a series of web-based, self-administered cognitive tasks engineered to run on a computer or as a mobile app on iOS devices. It provides game-like cognitive exercises, with the ability to assess the performance of individuals relative to their peers. Of the tasks available through BrainHQ, 6 were used to assess 3 cognitive domains: episodic memory, working memory, and executive function. One cognitive task (To and Fro Motor Speed) was used solely to adjust for motor speed processing.

While BrainHQ was originally designed around the concept of brain plasticity for the purpose of cognitive training and intervention, the study participants were not engaged in brain training. Each was given a single administration of the cognitive tasks linked to their unique email address. After completion, they received a summary of performance results from BrainHQ. Our data were audited to confirm the use of the primary testing session data and to prevent the analysis of repeat testing among participants who registered a second email address to improve their performance.

Data Analysis

Descriptive statistics were used to characterize the sample. BrainHQ raw data were normalized (z-score) and averaged to derive cognitive domain scores. Table 1 presents the cognitive domains used in the study and details the BrainHQ tasks that comprised each domain. We then used unadjusted and adjusted linear regression models to examine associations between the CVRFs and each cognitive domain. Adjusted models included demographic characteristics (age, race, and sex). We also conducted a sensitivity analysis, which further corrected for a measure of motor speed processing. Tests of statistical significance were two-tailed with significance set to $P<.05$. The analyses were performed using R version 4.0.2 (R Foundation for Statistical Computing).

Table 1. Cognitive domains and associated tasks.

Cognitive domain or test		Description
Episodic memory		
	Face facts	Recall for names, faces, and facts
	Auditory paired associates	Auditory recognition task
	Pathfinder	Visuospatial learning and memory
Working memory	Target tracker	Multiple-object target tracking
Executive function		
	Word conflict	Stroop task
	Rule switcher	Color or shape rule-switching for set shifting
Motor speed adjustment	To and Fro test	Test of motor speed

Results

At the time of cognitive testing, the 239 participants had a mean age of 62.74 years (SD 8.8). Moreover, 42.7% (n=102) were female, and 88.7% (n=212) had non-Hispanic White racial backgrounds. CVRF prevalence was 62.8% (n=150) for high cholesterol, 45.2% (n=108) for hypertension, 10.9% (n=26) for atrial fibrillation, and 7.5% (n=18) for diabetes (Table 2).

Table 2. Participant characteristics.

Characteristics		Values
Age (years), mean (SD)		62.7 (8.8)
	<60	101 (42.3)
	60-70	87 (36.4)
	>70	51 (21.3)
Gender, n (%)	Female	102 (42.7)
Race, n (%)		
	White	212 (88.7)
	Black	8 (3.3)
	Asian	10 (4.2)
	Other	7 (2.9)
	Not disclosed	2 (0.8)
Cardiovascular risk factors, n (%)		
	High cholesterol	150 (62.8)
	Atrial fibrillation	26 (10.9)
	Diabetes	18 (7.5)
	Hypertension	108 (45.2)
Cognitive assessment, n (%), range		
	Episodic memory	17.4 (6.3), 5.2 to 32.4
	Working memory	4.4 (0.7), 2.8 to 5.8
	Executive function	211.7 (111.2), -56.2 to 515

In unadjusted models, the participants with atrial fibrillation performed more poorly than those without atrial fibrillation on measures of episodic memory (β=-.31, 95% CI -0.59 to -0.04) and working memory (β=-.51, 95% CI -0.91 to -0.11), and a trend association was observed for executive function (β=-.33, 95% CI -0.74 to -0.07). The participants with hypertension performed more poorly than those without hypertension on measures of episodic memory (β=-.27, 95% CI -0.44 to -0.11), with a trend association for working memory (β=-.22, 95% CI -0.47 to 0.03) and no association for executive function (β=.03, 95% CI -0.22 to 0.28). There was no significant association between high cholesterol or diabetes and cognitive test performance. After adjustments for sex, age, and race, the findings remained unchanged from those of the unadjusted model (Table 3).

Table 3. Associations between cardiovascular risk factors and cognitive performance.

Cognitive performance according to CVRF[a,b]	Unadjusted β (95% CI)	*P* value	Demographics-adjusted[c] β (95% CI)	*P* value
Atrial fibrillation				
Episodic memory	-.31 (-0.59 to -0.04)	.02	-.28 (-0.54 to -0.02)	.03
Working memory	-.51 (-0.91 to -0.11)	.01	-.49 (-0.89 to -0.09)	.01
Executive function	-.33 (-0.74 to -0.07)	.10	-.31 (-0.71 to 0.09)	.12
Hypertension				
Episodic memory	-.27 (-0.44 to -0.11)	<.001	-.20 (-0.36 to -0.04)	.01
Working memory	-.22 (-0.47 to 0.03)	.09	-.15 (-0.40 to 0.10)	.24
Executive function	.03 (-0.22 to 0.28)	.80	.10 (-0.14 to 0.36)	.40
High cholesterol				
Episodic memory	-.01 (-0.18 to 0.16)	.91	.02 (-0.14 to 0.19)	.78
Working memory	-.08 (-0.34 to 0.17)	.52	-.07 (-0.33 to 0.18)	.58
Executive function	-.13 (-0.39 to 0.12)	.31	-.10 (-0.36 to 0.16)	.44
Diabetes				
Episodic memory	-.23 (-0.55 to 0.08)	.15	-.14 (-0.45 to 0.16)	.27
Working memory	.07 (-0.40 to 0.56)	.75	.18 (-0.29 to 0.66)	.45
Executive function	-.29 (-0.77 to 0.18)	.22	-.19 (-0.67 to 0.28)	.43

[a]CVRF: cardiovascular risk factors.

[b]Standardized difference in cognitive test (95% CI)

[c]Adjusted for age, race, and sex.

In a sensitivity analysis, we further adjusted for a measure of motor speed processing derived from the To and Fro cognitive task and found almost identical results.

Discussion

The results of this proof-of-concept study suggest that mHealth assessment tools can be used to effectively detect the association between CVRFs and cognitive function. We were successful at recruiting, consenting, and measuring the participants' cognitive function all by remote assessment. Furthermore, we found that those with atrial fibrillation and hypertension had worse cognitive performance in the area of memory.

Our investigation contributes to the field by examining the mHealth data and cognitive performance of middle-aged to older adults entirely through remote technology. While numerous studies have observed the utility of mHealth tool across a spectrum of neurological conditions (eg, dementia, stroke, and multiple sclerosis) [3,4,6,7,17-19], our study is unique in that we demonstrated that this technology can be used to identify subtle cognitive deficits in otherwise healthy middle-aged and older adults, without in-person patient contact. Additionally, our results contradict the perception that older adults do not or cannot participate in technology-driven mHealth assessment.

Our findings are consistent with in-person neuropsychological testing results that have demonstrated that individuals with CVRFs are at increased risk for cognitive decline and dementia as they age [9,13,20,21]. Specifically, atrial fibrillation absent stroke has been associated with cognitive decline and dementia and may accelerate cognitive decline through mechanisms of hypoperfusion, systemic inflammation, and cerebral small vessel diseases [22,23], while hypertension has been associated with increased risk for cognitive decline and microstructural white matter alterations [24,25]. The results for high cholesterol have been mixed [26-30]; however, many studies have reported an association with diabetes and worse cognition. Our lack of finding an association with diabetes may have been due to limited power as only 7% (n=18) had diabetes.

The high CVRF prevalence in the US population is concerning. The Third National Health and Nutrition Examination Survey (NHANES III) estimates that 60% of men and 50% of women have 1 to 2 CVRFs, and this increases with age [31,32]. Given

these striking numbers, mHealth studies such as Health eHeart and remote cognitive assessment tools may represent a viable strategy for closely tracking and monitoring the cognitive effects of CVRFs in at-risk patients, with the hope of improving outcomes.

The Health eBrain proof-of-concept study utilized well-established platforms to remotely evaluate the relationship between CVRF exposure and cognition. Despite the study's strengths, several limitations need to be considered. Our sample size was relatively small, particularly regarding participants with diabetes, and this may have reduced our ability to detect subtle cognitive changes. Second, our cohort was not very diverse as compared to the US population. This is in part a reflection of the Health eHeart Study cohort, which is less likely to be from Black, Hispanic, or Asian (versus White or non-Hispanic) racial backgrounds, relative to all adults in the United States [32]. This lack of diversity in Health eHeart may be due to a recruitment strategy that lacks the specific targeting of minority populations. It also may be due to the "digital divide" and associated issues with access to technology related to socioeconomic and cultural factors. While our results indicate that middle-aged and older adults can actively engage in the use of cell phone and computer technologies to participate in remote research and health tracking, our participants may represent an overall higher-functioning group relative to their peers. Therefore, our results may not be generalizable to the broader population. Finally, while remote, self-guided cognitive assessments have a great potential to provide large-scale cognitive data, they are not without their weaknesses. Control of the testing environment is an important component to standardizing and interpreting results to prevent distractions that can artificially lower the testing scores. Cognitive performance in remote testing is vulnerable to extraneous circumstances, level of task engagement, display of emotion and frustration, or the tendency to give up easily, and this study cannot account for that variability [3].

The results suggest that mHealth assessments can be used to detect the association between CVRFs and cognition function. mHealth tools demonstrated specific sensitivity for detecting memory deficits associated with atrial fibrillation and hypertension in middle-aged and older adults. Future research would benefit from a larger study of this remote assessment platform with longitudinal monitoring of CVRFs and cognitive performance. The ability to remotely track cognitive health in individuals with modifiable CVRFs could represent a unique opportunity to target high-risk individuals for early education, frequent monitoring, and interventions with the hope of preventing accelerated cognitive decline with aging.

Acknowledgments

The writing of this manuscript was supported by the Office of Academic Affiliations, Advanced Fellowship Program in Mental Illness Research and Treatment, Department of Veteran Affairs.

Conflicts of Interest

KY serves on Data and Safety Monitoring Boards for Eli Lilly and trials sponsored by the National Institutes of Health, as well as on the Board of Directors for Alector. MA is an employee and senior research scientist at Posit Science, the company that develops the cognitive assessments discussed in this study. She helped manage study activities related to assessment, development, selection, and technical integration. She is a shareholder of Posit Science stock. DEB is a cofounder of Together Senior Health Inc, which offers online group programs to help people living with cognitive decline or dementia remain independent and reduce isolation.

References

1. Morrison GE, Simone CM, Ng NF, Hardy JL. Reliability and validity of the NeuroCognitive Performance Test, a web-based neuropsychological assessment. Front Psychol 2015;6:1652 [FREE Full text] [doi: 10.3389/fpsyg.2015.01652] [Medline: 26579035]
2. Dobbin JE, Lennon R, Milholland JE, Hopkins KD, Himelstein P, Garlock JG, et al. Book Reviews: Standards for Educational and Psychological Tests and Manuals, prepared by a joint committee of the American Psychological Association, the American Educational Research Association, and the National Council on Measurement in Education. In: Educational and Psychological Measurement. Washington, DC, US: American Educational Research Association; Jul 02, 2016:751-767.
3. Bauer RM, Iverson GL, Cernich AN, Binder LM, Ruff RM, Naugle RI. Computerized neuropsychological assessment devices: joint position paper of the American Academy of Clinical Neuropsychology and the National Academy of Neuropsychology. Arch Clin Neuropsychol 2012 May;27(3):362-373 [FREE Full text] [doi: 10.1093/arclin/acs027] [Medline: 22382386]
4. Larson S. Technology in Psychological and Neuropsychological Testing. America Psychological Association. 2019 Mar 5. URL: https://www.apaservices.org/practice/good-practice/psychological-neuropsychological.pdf [accessed 2021-12-21]
5. Adibi S. Mobile Health: A Technology Road Map. In: Champringer International Publishing : Imprintpringer. New York, US: Springer; Mar 19, 2015.
6. Kane RL, Kay GG. Computerized assessment in neuropsychology: a review of tests and test batteries. Neuropsychol Rev 1992 Mar;3(1):1-117. [doi: 10.1007/BF01108787] [Medline: 1300218]

7. Kueider AM, Parisi JM, Gross AL, Rebok GW. Computerized cognitive training with older adults: a systematic review. PLoS One 2012;7(7):e40588 [FREE Full text] [doi: 10.1371/journal.pone.0040588] [Medline: 22792378]
8. Buford TW. Hypertension and aging. Ageing Res Rev 2016 Mar;26:96-111 [FREE Full text] [doi: 10.1016/j.arr.2016.01.007] [Medline: 26835847]
9. Knopman D, Boland LL, Mosley T, Howard G, Liao D, Szklo M, Atherosclerosis Risk in Communities (ARIC) Study Investigators. Cardiovascular risk factors and cognitive decline in middle-aged adults. Neurology 2001 Jan 09;56(1):42-48. [doi: 10.1212/wnl.56.1.42] [Medline: 11148234]
10. Launer LJ, Masaki K, Petrovitch H, Foley D, Havlik RJ. The association between midlife blood pressure levels and late-life cognitive function. The Honolulu-Asia Aging Study. JAMA 1995 Dec 20;274(23):1846-1851. [Medline: 7500533]
11. Mayeda ER, Whitmer RA, Yaffe K. Diabetes and cognition. Clin Geriatr Med 2015 Feb;31(1):101-15, ix [FREE Full text] [doi: 10.1016/j.cger.2014.08.021] [Medline: 25453304]
12. Whitmer RA, Sidney S, Selby J, Johnston SC, Yaffe K. Midlife cardiovascular risk factors and risk of dementia in late life. Neurology 2005 Jan 25;64(2):277-281. [doi: 10.1212/01.WNL.0000149519.47454.F2] [Medline: 15668425]
13. Yaffe K, Bahorik AL, Hoang TD, Forrester S, Jacobs DR, Lewis CE, et al. Cardiovascular risk factors and accelerated cognitive decline in midlife: The CARDIA Study. Neurology 2020 Aug 18;95(7):e839-e846 [FREE Full text] [doi: 10.1212/WNL.0000000000010078] [Medline: 32669394]
14. Study to End Heart Disease. Health eHeart. URL: https://www.health-eheartstudy.org/study [accessed 2021-01-05]
15. Dixit S, Pletcher MJ, Vittinghoff E, Imburgia K, Maguire C, Whitman IR, et al. Secondhand smoke and atrial fibrillation: Data from the Health eHeart Study. Heart Rhythm 2016 Jan;13(1):3-9 [FREE Full text] [doi: 10.1016/j.hrthm.2015.08.004] [Medline: 26340844]
16. About the BrainHQ Exercises. Brain HQ. URL: https://www.brainhq.com/why-brainhq/about-the-brainhq-exercises [accessed 2021-01-05]
17. Mahncke HW, Connor BB, Appelman J, Ahsanuddin ON, Hardy JL, Wood RA, et al. Memory enhancement in healthy older adults using a brain plasticity-based training program: a randomized, controlled study. Proc Natl Acad Sci U S A 2006 Aug 15;103(33):12523-12528 [FREE Full text] [doi: 10.1073/pnas.0605194103] [Medline: 16888038]
18. Jung H, Daneault J, Lee H, Kim K, Kim B, Park S, et al. Remote Assessment of Cognitive Impairment Level Based on Serious Mobile Game Performance: An Initial Proof of Concept. IEEE J Biomed Health Inform 2019 May;23(3):1269-1277. [doi: 10.1109/JBHI.2019.2893897] [Medline: 30668485]
19. Settle JR, Robinson SA, Kane R, Maloni HW, Wallin MT. Remote cognitive assessments for patients with multiple sclerosis: a feasibility study. Mult Scler 2015 Jul 12;21(8):1072-1079. [doi: 10.1177/1352458514559296] [Medline: 25583842]
20. Sabia S, Nabi H, Kivimaki M, Shipley MJ, Marmot MG, Singh-Manoux A. Health behaviors from early to late midlife as predictors of cognitive function: The Whitehall II study. Am J Epidemiol 2009 Aug 15;170(4):428-437 [FREE Full text] [doi: 10.1093/aje/kwp161] [Medline: 19574344]
21. Knecht S, Oelschläger C, Duning T, Lohmann H, Albers J, Stehling C, et al. Atrial fibrillation in stroke-free patients is associated with memory impairment and hippocampal atrophy. Eur Heart J 2008 Sep 09;29(17):2125-2132. [doi: 10.1093/eurheartj/ehn341] [Medline: 18667399]
22. Ding M, Qiu C. Atrial Fibrillation, Cognitive Decline, and Dementia: an Epidemiologic Review. Curr Epidemiol Rep 2018;5(3):252-261 [FREE Full text] [doi: 10.1007/s40471-018-0159-7] [Medline: 30148041]
23. Ott A, Breteler MM, de Bruyne MC, van Harskamp F, Grobbee DE, Hofman A. Atrial fibrillation and dementia in a population-based study. The Rotterdam Study. Stroke 1997 Feb;28(2):316-321. [doi: 10.1161/01.str.28.2.316] [Medline: 9040682]
24. McEvoy LK, Fennema-Notestine C, Eyler LT, Franz CE, Hagler DJ, Lyons MJ, et al. Hypertension-related alterations in white matter microstructure detectable in middle age. Hypertension 2015 Aug;66(2):317-323 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.115.05336] [Medline: 26056337]
25. Reitz C, Tang M, Manly J, Mayeux R, Luchsinger JA. Hypertension and the risk of mild cognitive impairment. Arch Neurol 2007 Dec 01;64(12):1734-1740 [FREE Full text] [doi: 10.1001/archneur.64.12.1734] [Medline: 18071036]
26. Carlsson CM, Nondahl DM, Klein BEK, McBride PE, Sager MA, Schubert CR, et al. Increased atherogenic lipoproteins are associated with cognitive impairment: effects of statins and subclinical atherosclerosis. Alzheimer Dis Assoc Disord 2009;23(1):11-17 [FREE Full text] [doi: 10.1097/wad.0b013e3181850188] [Medline: 19266697]
27. Elias PK, Elias MF, D'Agostino RB, Sullivan LM, Wolf PA. Serum cholesterol and cognitive performance in the Framingham Heart Study. Psychosom Med 2005;67(1):24-30. [doi: 10.1097/01.psy.0000151745.67285.c2] [Medline: 15673620]
28. Mielke MM, Zandi PP, Sjogren M, Gustafson D, Ostling S, Steen B, et al. High total cholesterol levels in late life associated with a reduced risk of dementia. Neurology 2005 Apr 20;64(10):1689-1695. [doi: 10.1212/01.wnl.0000161870.78572.a5]
29. Solomon A, Kåreholt I, Ngandu T, Wolozin B, Macdonald SWS, Winblad B, et al. Serum total cholesterol, statins and cognition in non-demented elderly. Neurobiol Aging 2009 Jun;30(6):1006-1009. [doi: 10.1016/j.neurobiolaging.2007.09.012] [Medline: 18022292]
30. Yaffe K, Barrett-Connor E, Lin F, Grady D. Serum lipoprotein levels, statin use, and cognitive function in older women. Arch Neurol 2002 Mar;59(3):378-384. [doi: 10.1001/archneur.59.3.378] [Medline: 11890840]

31. National Health and Nutrition Examination Survey. Center for Disease Control. URL: https://www.cdc.gov/nchs/nhanes/index.htm [accessed 2021-01-05]
32. Guo X, Vittinghoff E, Olgin JE, Marcus GM, Pletcher MJ. Volunteer Participation in the Health eHeart Study: A Comparison with the US Population. Sci Rep 2017 May 16;7(1):1956 [FREE Full text] [doi: 10.1038/s41598-017-02232-y] [Medline: 28512303]

Abbreviations

CVRF: cardiovascular risk factor
mHealth: mobile health
NHANES III: Third National Health and Nutrition Examination Survey

Edited by A Mavragani; submitted 13.05.21; peer-reviewed by C Reis, A Hidki; accepted 08.12.21; published 02.02.22.

Please cite as:
Eastman JA, Kaup AR, Bahorik AL, Butcher X, Attarha M, Marcus GM, Pletcher MJ, Olgin JE, Barnes DE, Yaffe K
Remote Assessment of Cardiovascular Risk Factors and Cognition in Middle-Aged and Older Adults: Proof-of-Concept Study
JMIR Form Res 2022;6(2):e30410
URL: https://formative.jmir.org/2022/2/e30410
doi:10.2196/30410
PMID:35107430

Original Paper

Exploring Relationships Between Tweet Numbers and Over-the-counter Drug Sales for Allergic Rhinitis: Retrospective Analysis

Shoko Wakamiya[1], PhD; Osamu Morimoto[2], BPharm; Katsuhiro Omichi[2], BPharm; Hideyuki Hara[2], MPharm; Ichiro Kawase[2], MPharm; Ryuji Koshiba[3], PhD; Eiji Aramaki[1], PhD

[1]Graduate School of Science and Technology, Nara Institute of Science and Technology, Nara, Japan

[2]Scientific Affairs, SSP Co Ltd, Tokyo, Japan

[3]Medical Affairs, Sanofi KK, Tokyo, Japan

Corresponding Author:
Shoko Wakamiya, PhD
Graduate School of Science and Technology
Nara Institute of Science and Technology
8916-5, Takayama-cho, Ikoma
Nara, 6300192
Japan
Phone: 81 743 72 5250
Email: wakamiya@is.naist.jp

Abstract

Background: Health-related social media data are increasingly being used in disease surveillance studies. In particular, surveillance of infectious diseases such as influenza has demonstrated high correlations between the number of social media posts mentioning the disease and the number of patients who went to the hospital and were diagnosed with the disease. However, the prevalence of some diseases, such as allergic rhinitis, cannot be estimated based on the number of patients alone. Specifically, individuals with allergic rhinitis typically self-medicate by taking over-the-counter (OTC) medications without going to the hospital. Although allergic rhinitis is not a life-threatening disease, it represents a major social problem because it reduces people's quality of life, making it essential to understand its prevalence and people's motives for self-medication behavior.

Objective: This study aims to explore the relationship between the number of social media posts mentioning the main symptoms of allergic rhinitis and the sales volume of OTC rhinitis medications in Japan.

Methods: We collected tweets over 4 years (from 2017 to 2020) that included keywords corresponding to the main nasal symptoms of allergic rhinitis: "sneezing," "runny nose," and "stuffy nose." We also obtained the sales volume of OTC drugs, including oral medications and nasal sprays, for the same period. We then calculated the Pearson correlation coefficient between time series data on the number of tweets per week and time series data on the sales volume of OTC drugs per week.

Results: The results showed a much higher correlation (r=0.8432) between the time series data on the number of tweets mentioning "stuffy nose" and the time series data on the sales volume of nasal sprays than for the other two symptoms. There was also a high correlation (r=0.9317) between the seasonal components of these time series data.

Conclusions: We investigated the relationships between social media data and behavioral patterns, such as OTC drug sales volume. Exploring these relationships can help us understand the prevalence of allergic rhinitis and the motives for self-care treatment using social media data, which would be useful as a marketing indicator to reduce the number of out-of-stocks in stores, provide (sell) rhinitis medicines to consumers in a stable manner, and reduce the loss of sales opportunities. In the future, in-depth investigations are required to estimate sales volume using social media data, and future research could investigate other diseases and countries.

(JMIR Form Res 2022;6(2):e33941) doi:10.2196/33941

KEYWORDS

infoveillance; social media; Twitter; over-the-counter drugs; allergic rhinitis; hay fever; drug; treatment; allergy; immunology; surveillance; monitoring; prevalence; motivation; Japan; symptom

Introduction

Social media data are a valuable source for rapidly exploring and understanding various real-world phenomena. Because many people share their health conditions on social media, a high volume of health-related social media data are available and the use of these data for large-scale quantitative analyses [1,2] and disease surveillance (referred to as “infoveillance”) is gaining much interest [3]. In particular, major advances have been made in the use of social media data to track the prevalence and spread of infectious diseases, including influenza and other conditions [4-13]. These studies have contributed to public health by demonstrating high correlations between fluctuations in the number of relevant social media posts and patients for a specific disease.

This study focuses on allergic rhinitis, also called hay fever, which is one of the most common allergic diseases worldwide [14]. In Japan, many people suffer from seasonal allergic rhinitis induced by Japanese cedar pollen between February and April each year. According to the results of the Japan National Epidemiological Studies in 1998, 2008, and 2019, the prevalence of allergic rhinitis in Japan has increased significantly over the past 20 years (49.2% increase in overall allergic rhinitis and 42.5% increase in cedar pollen–induced allergic rhinitis). It is now a national disease that affects the majority of the population [15].

Although allergic rhinitis is not a life-threatening disease, its main symptoms—sneezing, nasal discharge (watery), and nasal obstruction—significantly impair the quality of life (QOL) of patients, causing a major social problem [15-17]. According to the practical guideline for the management of allergic rhinitis in Japan 2020 [18], approximately 40% of people in Japan said they had allergic rhinitis, of which 30% had nasal allergies caused by pollen. In addition, since most self-medication for allergic rhinitis in Japan is for seasonal allergic rhinitis caused by cedar and cypress tree pollen, which are dispersed in the spring, the number of patients and the sales amount of over-the-counter (OTC) rhinitis drugs are greatly affected by the pollen conditions of the year or region. About three-fourths of OTC rhinitis drugs are oral medicines and one-fourth are nasal sprays; for oral medicines, more than 60% of annual sales are concentrated in the 3 main pollen dispersal months, from February to April. Retailers are required to assess the situation and make accurate predictions during the annual spring pollen season, as they are required to manage products without causing opportunity loss. Therefore, a real-time understanding of the prevalence of allergic rhinitis is important for providing necessary solutions, such as appropriate pharmaceutical distribution.

The surveillance methods currently available rely on the following three types of statistical data:

1. Amount of pollen. This value is not precise because of the complex mechanisms that cause allergic rhinitis. First, when antigens such as pollen or house dust enter the nose, sneezing and nasal discharge occur immediately. As antigens repeatedly enter the nose, a reaction centered on nasal congestion occurs, and the symptoms of allergic rhinitis intensify [15]. Thus, there is a time lag between exposure to pollen and the onset of allergic rhinitis. In addition, the timing of symptom onset varies from person to person. For some patients, symptoms appear as soon as pollen starts to disperse, while for others, symptoms do not appear until there is a large amount of pollen in the air. The intensity of symptoms is also not the same, with some people having mild symptoms and others having severe symptoms [15]. As a result, there is no strong association between pollen count and patient numbers, complicating disease surveillance.
2. Number of outpatients. This is the number of patients who visit the hospital for allergic rhinitis. Since many patients try to self-medicate using OTC drugs instead of visiting a medical institution, such data do not provide an overall picture of the trend.
3. Volume of OTC drug sales. This value would be more reliable than the above two types of data. The trend of self-medicating using OTC drugs is being accelerated by the introduction of many new OTC medicines that switched from prescription to OTC status [19].

So far, correlations of hay fever–related tweets with pollen counts [12] and reported incidents of hay fever [13] have been investigated. Our previous study [20] analyzed data on pollen count, the number of hay fever–related tweets, and the number of patients during the seasonal allergic rhinitis period in Japan to explore their relationships. The results showed that increased pollen counts were associated with increased numbers of tweets and patients. In addition, increases in the number of tweets were also associated with increased numbers of patients.

This study explores the relationships between social media data related to allergic rhinitis and OTC allergic rhinitis drug sales as an outcome of consumer behavior. To the best of our knowledge, this is the first study to compare tweet trends with drug sales trends. Specifically, we investigate the correlation between the weekly number of tweets related to 3 main symptoms of allergic rhinitis—paroxysmal repetitive sneezing (sneezing), watery rhinorrhea (runny nose), and nasal obstruction (stuffy nose)—and the weekly sales volume of OTC allergic rhinitis medication (oral medicine and nasal spray).

Methods

Data

Number of Tweets Related to Allergic Rhinitis Nasal Symptoms

We collected tweets that included any of the following Japanese keywords for major nasal symptoms of allergic rhinitis: kushami (くしゃみ; sneezing), hanamizu (鼻水; runny nose), and hanadumari (鼻づまり; stuffy nose). These keywords were selected by analyzing co-occurrence words in tweets concerning hay fever (kafunsho or 花粉症 in Japanese) and extracting typical notations with high frequency in our preliminary experiments. These tweets were crawled using the Twitter application programming interface. After removing retweets, we obtained 5,834,920 tweets concerning sneezing, 7,695,598 tweets concerning runny nose, and 274,119 tweets related to

stuffy nose between January 2, 2017, and January 3, 2021 (209 weeks).

OTC Allergic Rhinitis Medication Sales Volume

In addition to visiting a medical institution for the treatment of allergic rhinitis, patients self-medicate using OTC drugs based on their own judgment. OTC allergic rhinitis medications are typically oral medicines or nasal sprays. In recent years, many medicines have switched from prescription to OTC status. Allergic rhinitis is one of the health complaints for which self-medication is most common.

Since there are no comprehensive government statistics or other survey information on the number of users of OTC drugs, estimates are made using sales data for OTC drugs. However, the reporting of OTC sales information also involves some delays. In addition to information on the shipment value of OTC drugs from manufacturers and distributors, point-of-sale (POS) data from retail stores (such as supermarkets, convenience stores, home centers/discount stores, drugstores, and pharmacies) provided by private research companies are used. POS data are collected almost in real time from approximately 6000 panel retailers nationwide through in-store cash registers and systems, and include information such as which products are sold, when, where, at what price, and how many; these data are provided after aggregation and are an important source of information for understanding consumer behavior regarding self-medication [21].

For this study, we used data on the sales volume of OTC allergic rhinitis drugs; the data were obtained from INTAGE Healthcare Inc's nationwide drugstore panel research [19]. During the study period between January 2, 2017, and January 3, 2021, a total of 205 oral medicines (OTC allergic rhinitis drugs) recorded a weekly market share of 0.0001% or more (including Alesion 20 by SSP Co Ltd and Allegra FX by Hisamitsu Pharmaceutical Co Inc), as did 118 nasal spray products (such as Pabron Nasal Spray by Taisho Pharmaceutical Co Ltd and Contac Rhinitis Spray for seasonal allergies by GlaxoSmithKline plc). No new product launched after January 2017 had a weekly market share of more than 10% [21]. Therefore, we consider that new products have not had a significant impact on sales.

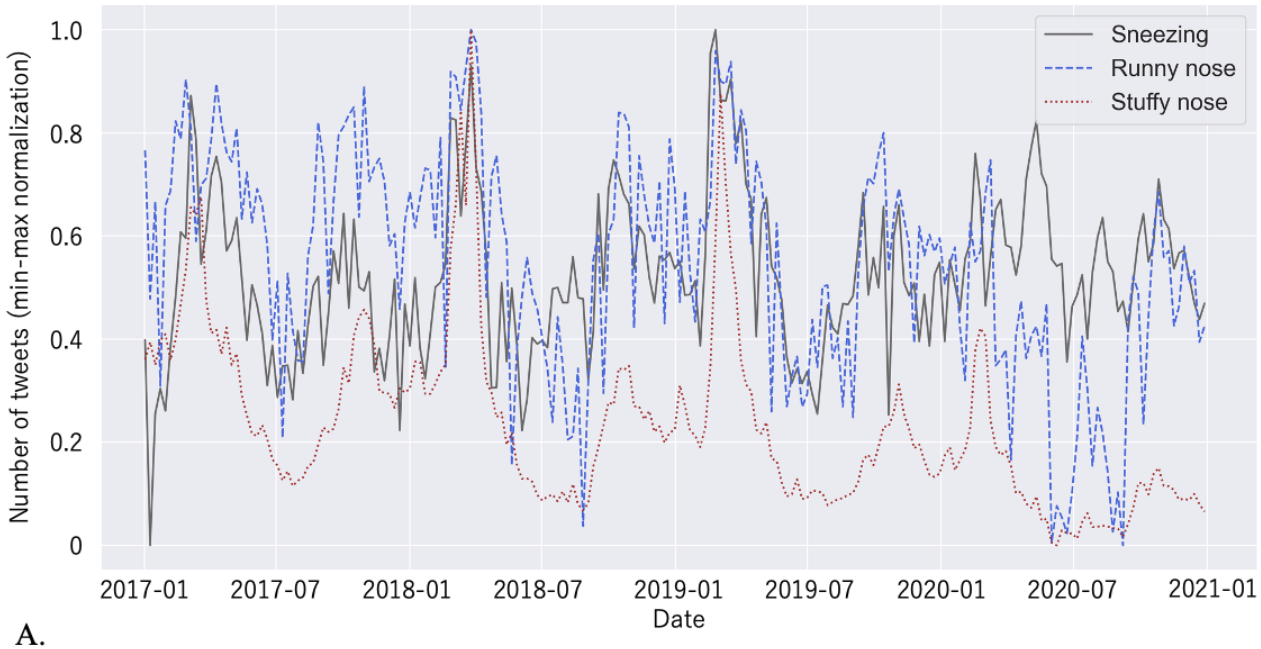
Correlation Coefficient Calculation

We aimed to examine the relationships between allergic rhinitis nasal symptom–related tweet numbers and OTC allergic rhinitis medication sales volume (oral medicine or nasal spray). To this end, we calculated the Pearson correlation coefficients between the time series data. In addition to the correlations between the observed time series data, we also investigated correlations between the trend, seasonality, and residual components of these time series data. The time series decomposition was performed using the seasonal_decompose function from the statsmodels module [22] in Python.

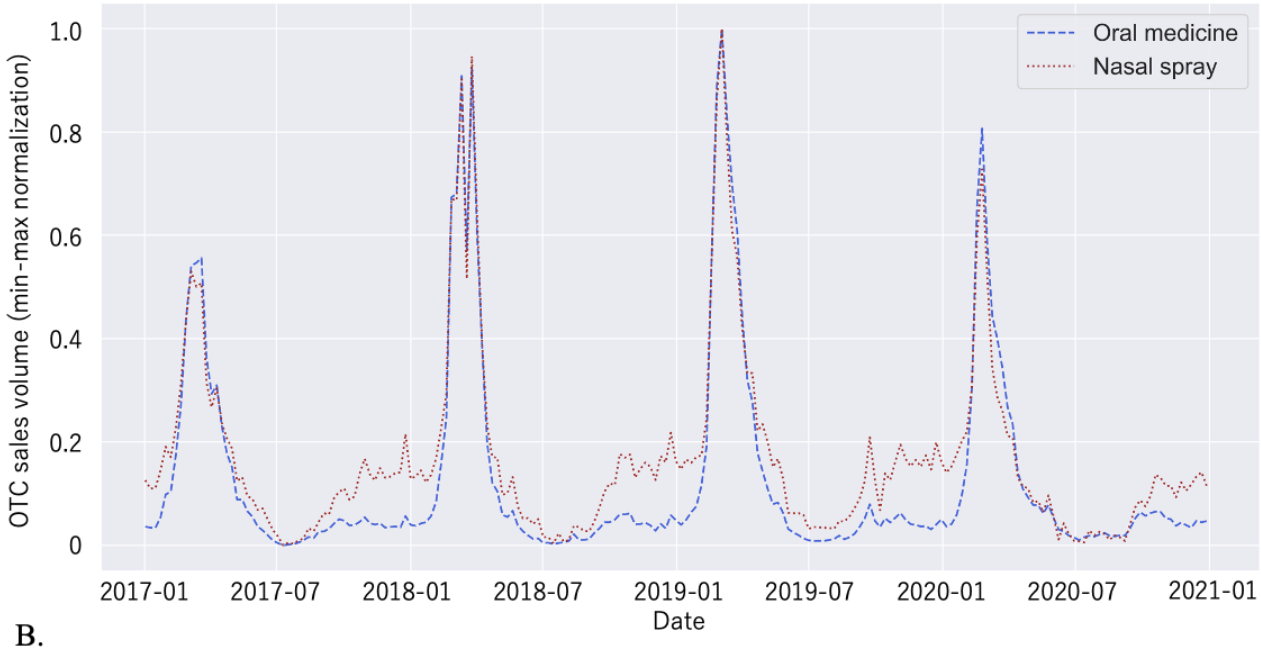
Results

Figure 1A shows the changes in the weekly number of tweets related to the 3 main symptoms of allergic rhinitis for the target period. Figure 1B shows the changes in weekly sales volume of OTC allergic rhinitis medication, including oral medication and nasal spray, from 2017 to 2020. The most common causative antigen of seasonal allergic rhinitis in Japan is cedar pollen, which disperses between February and April. In Figure 1B, there is a clear peak during this period each year.

Figure 1. Time-based changes in data from 2017 to 2020 in Japan. The x-axes represent the date and the y-axes represent data counts, to which min-max normalization is applied for the following variables: (A) weekly changes in the number of tweets for the 3 main symptoms of allergic rhinitis (sneezing, solid gray line; runny nose, blue dashed line; and stuffy nose, red dotted line), and (B) weekly changes in the sales volume of OTC allergic rhinitis medication (oral medicine, blue dashed line; nasal spray, red dotted line). OTC: over-the-counter.



A.

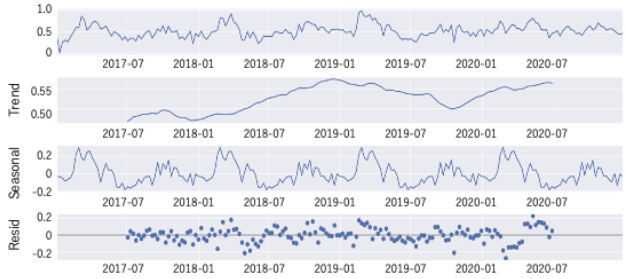


B.

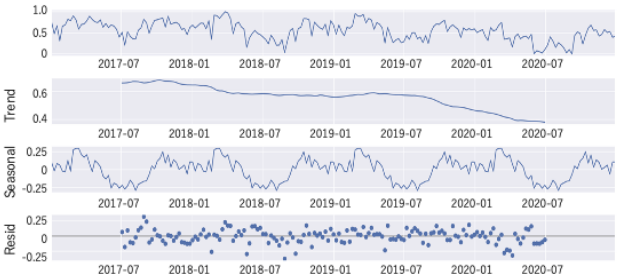
Figure 2 shows the time series of the observed data and its decomposed components: trend, seasonality, and residual. Figure 3 shows heat maps of correlations of all pairs of time series data. Figure 3A shows correlations between observed time series data. Figures 3B-D show correlations between the trend, seasonality, and residual components, respectively.

Among pairs of the observed time series of tweets, the positive correlation between tweets concerning stuffy nose and tweets concerning runny nose was the highest (r=0.7349), as shown in Figure 3A. The time series of the trend components showed the highest positive correlation (r=0.9613; Figure 3B), and the time series of the seasonal components was also highly correlated (r=0.8483; Figure 3C). On the other hand, there were positive correlations with tweets about sneezing, but they were not high (r=0.3564 for tweets concerning runny nose and r=0.4382 for tweets concerning stuffy nose; Figure 3A), due to negative correlations between the time series of the trend components (r=–0.6005 for tweets concerning runny nose and r=–0.6197 for tweets concerning stuffy nose; Figure 3B). As for OTC medication sales volume, the highest positive correlation (r>0.95) was between oral medicine and nasal spray.

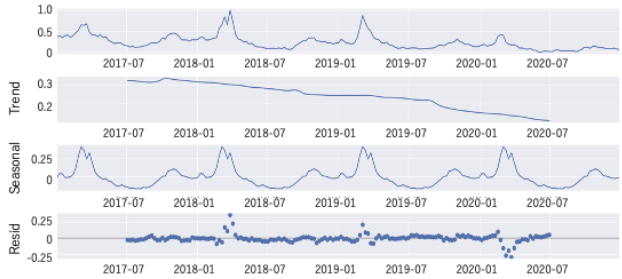
Figure 2. Time series of observed data (the top graph) and its decomposed components (the second, third, and fourth graphs represent trend, seasonality, and residual components, respectively) from 2017 to 2020. (A) Time series of the number of tweets concerning sneezing, (B) time series of the number of tweets concerning runny nose, (C) time series of the number of tweets concerning stuffy nose, (D) time series of oral medicine sales volume, and (E) time series of nasal spray sales volume. Resid: residual.



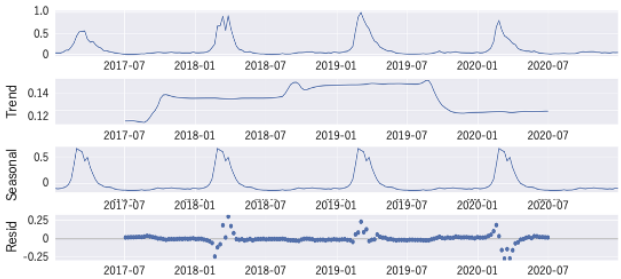
A. The number of tweets concerning sneezing



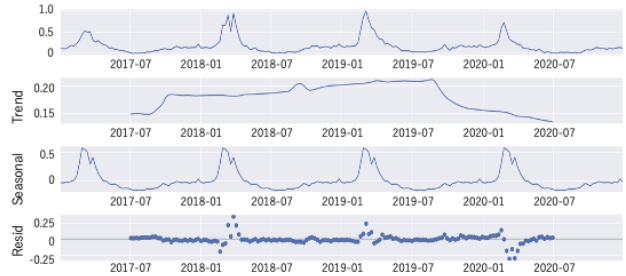
B. The number of tweets concerning runny nose



C. The number of tweets concerning stuffy nose

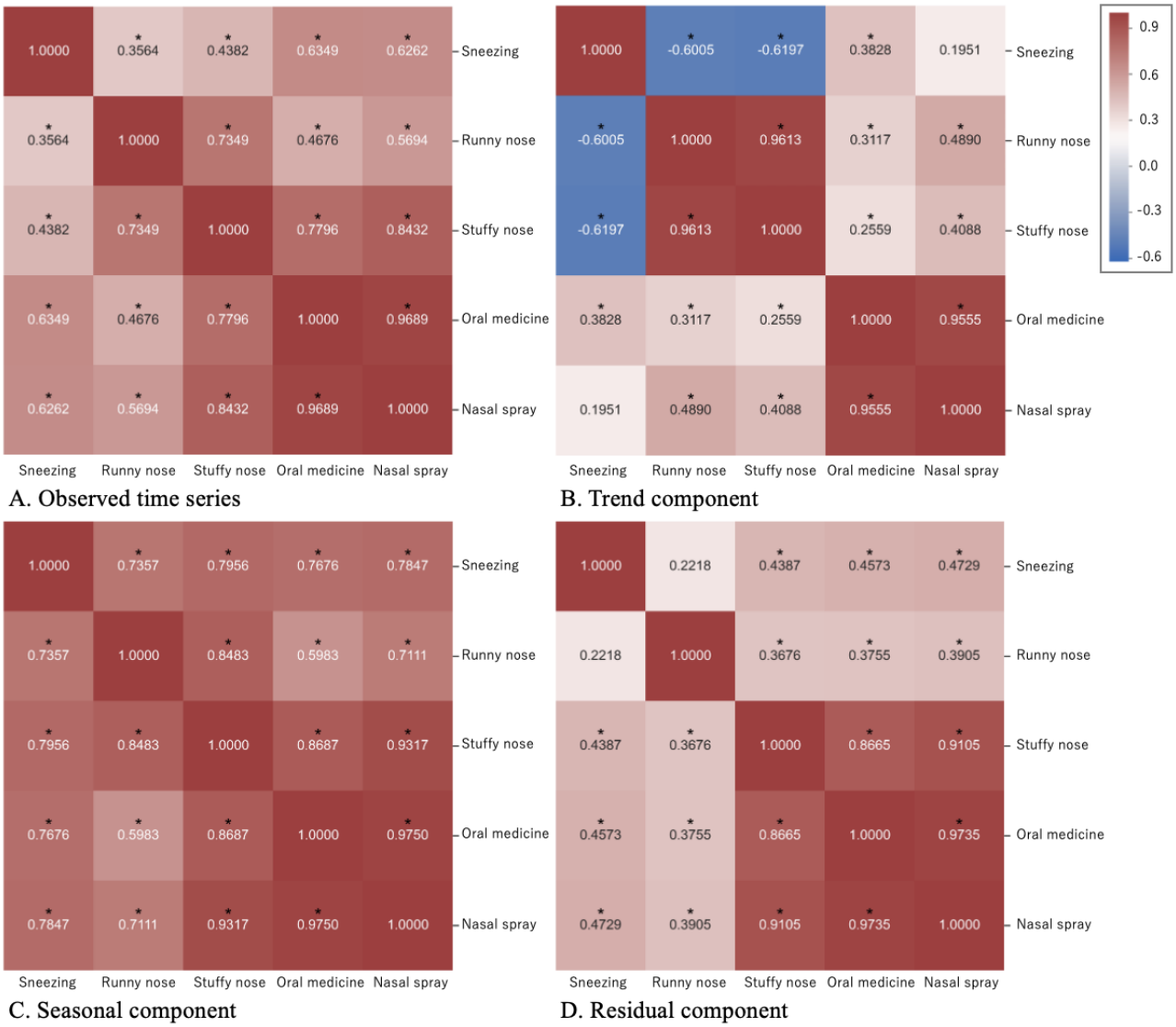


D. Oral medicine sales volume



E. Nasal spray sales volume

Figure 3. Heat map of time series correlation. (A) Correlations between observed time series data, (B) correlations between the trend components, (C) correlations between the seasonality components, and (D) correlations between the residual components. The trend, seasonality, and residual components were estimated by decomposing the observed time series data. *P<.001.



As for the correlations between tweets and OTC medication sales volume, the highest positive correlation (r=0.8432) was between the time series of tweets concerning stuffy nose and the time series of nose spray sales volume, as shown in Figure 3A. The correlation of the trend component was positive and not high (r=0.4088; Figure 3B). On the other hand, the correlation of the seasonal component was positive and high (r=0.9317; Figure 3C), indicating the tweet numbers of allergic rhinitis keywords and OTC drug sales have a seasonal pattern due to the seasonality of allergic rhinitis in Japan. Several Twitter-based surveillance studies [7-11] dealt with infectious diseases that demonstrate seasonality, such as influenza, and some of them effectively utilized such seasonal features.

Discussion

Principal Results

We found that the positive correlation between the number of tweets concerning stuffy nose and the weekly sales volume of nasal spray is the highest. In various surveys, it has been reported that among the 3 main symptoms of allergic rhinitis, nasal congestion is the most unpleasant and difficult to cure and it reduces patients' QOL, especially their mental QOL [15]. Therefore, it can be inferred that the inability to seek medical attention or the desire to deal with the symptoms immediately may appear stronger when people are experiencing nasal obstruction than sneezing or runny nose, which may lead to the purchasing of OTC allergic rhinitis medications. In particular, patients with nasal obstruction symptoms often use OTC nasal drops, which can be administered directly to the affected area (nasal cavity) in the hope that these will provide immediate relief, and a high correlation was observed in this study between tweets about nasal obstruction and sales volume of nasal sprays.

Sneezing and runny nose are symptoms that are obvious to people other than the patient and may also be tweeted about by other people, not just patients, which can create noise. Compared to these visible symptoms, symptoms such as nasal obstruction are less likely to be noticed by others and are often tweeted about by patients themselves, suggesting that tweets related to a stuffy nose create less noise. The numbers of tweets about

sneezing and runny nose are much larger than the number of tweets about stuffy nose. Thus, when dealing with tweets about symptoms that are easily observable by others, it would be reasonable to distinguish between tweets by patients and those by others using natural language processing.

Furthermore, according to a survey in Japan [21], an increasing number of people are becoming sensitive to people sneezing near them, which spreads droplets, due to the COVID-19 pandemic. This may be one of the reasons why the time series of tweets concerning sneezing has a different trend than tweets about the other two symptoms, as shown in Figures 2A-C and Figure 3B.

Limitations

This study applied simple statistical methods to explore the relationship between the number of social media posts mentioning the main symptoms of allergic rhinitis and the sales volume of OTC rhinitis medications in Japan. In the future, we need to further explore relations between the variables, including causal, temporal, and confounding relations.

Although this study focused on nasal symptoms of allergic rhinitis and investigated their relationship with OTC drug sales volume, there are other symptoms associated with allergic rhinitis, including sleep disturbance, olfactory disturbance, nasal itching, problems with learning, poor concentration, inattention, fatigue, irritability, lightheadedness, and headache. In the future, these symptoms should be considered when conducting further investigations. In addition, this study dealt with the sales amount of OTC drugs as the sales volume; in the future, we should consider the number of sales as well.

Another limitation is language bias. This paper focuses only on the Japanese language, in which several allergy symptoms are not polysemic. However, allergy symptoms in other languages may have multiple meanings. In addition to polysemic words, idioms, including allergy symptoms, could bias word frequencies. This aspect is worth studying in the future.

However, the actual number of patients with allergic rhinitis is still unknown. This study revealed only a correlation between the sales volume of OTC drugs and the number of tweets about the main symptoms of allergic rhinitis; the relationship of the sales volume with the number of patients will require further investigation.

Conclusions

This study investigated correlations between social media data related to allergic rhinitis symptoms and OTC allergic rhinitis drug sales volume as an outcome of consumer behavior. We analyzed time series data for 4 years and showed a strong positive correlation between the number of tweets regarding stuffy nose and the sales volume of nasal spray. Regardless of the temporal dependency direction between the two variables, understanding such relationships has great potential as a market indicator to reduce the number of out-of-stocks in stores, provide (sell) rhinitis medicines to consumers in a stable manner, and reduce the loss of sales opportunities. In the future, additional relationships, such as causal, temporal, and confounding relations, should be explored by employing sophisticated time series analysis methods. In-depth investigations are also required to make estimations of sales volume using social media data, and future research could investigate other diseases and countries.

Acknowledgments

This research was partly supported by SSP Co Ltd and JSPS KAKENHI grant JP19K20279.

Authors' Contributions

SW and EA collected the social media data. OM, KO, HH, IK, and RK obtained the OTC drug sales volume data. SW and OM conducted the experiments. All authors analyzed the results. All authors contributed to the study design, wrote the main manuscript, and reviewed the manuscript.

Conflicts of Interest

SSP Co Ltd pays joint research funds to the Nara Institute of Science and Technology. SSP Co Ltd is a subsidiary of Sanofi KK. OM, KO, HH, and IK are current employees of SSP Co Ltd and RK is a current employee of Sanofi KK. They may hold shares and stock options in their respective companies.

References

1. Lee JL, DeCamp M, Dredze M, Chisolm MS, Berger ZD. What are health-related users tweeting? A qualitative content analysis of health-related users and their messages on twitter. J Med Internet Res 2014;16(10):e237 [FREE Full text] [doi: 10.2196/jmir.3765] [Medline: 25591063]
2. Alnemer KA, Alhuzaim WM, Alnemer AA, Alharbi BB, Bawazir AS, Barayyan OR, et al. Are Health-Related Tweets Evidence Based? Review and Analysis of Health-Related Tweets on Twitter. J Med Internet Res 2015 Oct 29;17(10):e246 [FREE Full text] [doi: 10.2196/jmir.4898] [Medline: 26515535]
3. 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;11(1):e11 [FREE Full text] [doi: 10.2196/jmir.1157] [Medline: 19329408]

4. Charles-Smith LE, Reynolds TL, Cameron MA, Conway M, Lau EHY, Olsen JM, et al. Using Social Media for Actionable Disease Surveillance and Outbreak Management: A Systematic Literature Review. PLoS One 2015 Oct;10(10):e0139701 [FREE Full text] [doi: 10.1371/journal.pone.0139701] [Medline: 26437454]
5. Paul M, Sarker A, Brownstein J, Nikfarjam A, Scotch M, Smith K. Social media mining for public health monitoring and surveillance. : World Scientific Publishing Co. Pte Ltd; 2016 Presented at: 21st Pacific Symposium on Biocomputing, PSB; January 4, 2016; Hawaii, United States p. 468-479. [doi: 10.1142/9789814749411_0043]
6. Al-Garadi MA, Khan MS, Varathan KD, Mujtaba G, Al-Kabsi AM. Using online social networks to track a pandemic: A systematic review. J Biomed Inform 2016 Aug;62:1-11 [FREE Full text] [doi: 10.1016/j.jbi.2016.05.005] [Medline: 27224846]
7. Aramaki E, Maskawa S, Morita M. Twitter catches the flu: Detecting influenza epidemics using Twitter. In: Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing. 2011 Jul Presented at: 2011 Conference on Empirical Methods in Natural Language Processing; July 27-31, 2011; Edinburgh, Scotland, UK p. 1568-1576 URL: https://aclanthology.org/D11-1145/
8. Paul M, Dredze M. You are what you tweet: Analyzing Twitter for public health. In: Proceedings of the International AAAI Conference on Web and Social Media. 2011 Aug 03 Presented at: Fifth International AAAI Conference on Weblogs and Social Media; 17-21 July 2011; Barcelona, Spain p. 265-272 URL: https://ojs.aaai.org/index.php/ICWSM/article/view/14137
9. Broniatowski DA, Paul MJ, Dredze M. National and local influenza surveillance through Twitter: an analysis of the 2012-2013 influenza epidemic. PLoS One 2013 Dec;8(12):e83672 [FREE Full text] [doi: 10.1371/journal.pone.0083672] [Medline: 24349542]
10. Paul MJ, Dredze M, Broniatowski D. Twitter improves influenza forecasting. PLoS Curr 2014 Oct 28;6:PMC4234396 [FREE Full text] [doi: 10.1371/currents.outbreaks.90b9ed0f59bae4ccaa683a39865d9117] [Medline: 25642377]
11. Wakamiya S, Kawai Y, Aramaki E. Twitter-Based Influenza Detection After Flu Peak via Tweets With Indirect Information: Text Mining Study. JMIR Public Health Surveill 2018 Sep 25;4(3):e65 [FREE Full text] [doi: 10.2196/publichealth.8627] [Medline: 30274968]
12. Gesualdo F, Stilo G, D'Ambrosio A, Carloni E, Pandolfi E, Velardi P, et al. Can Twitter Be a Source of Information on Allergy? Correlation of Pollen Counts with Tweets Reporting Symptoms of Allergic Rhinoconjunctivitis and Names of Antihistamine Drugs. PLoS One 2015 Jul 21;10(7):e0133706 [FREE Full text] [doi: 10.1371/journal.pone.0133706] [Medline: 26197474]
13. Quincey E, Kyriacou T, Pantin T. #hayfever; A longitudinal study into hay fever related tweets in the UK. In: Proceedings of the 6th International Conference on Digital Health Conference. 2016 Apr Presented at: 6th International Conference on Digital Health Conference; 2016; Montréal, Québec, Canada p. 85-89. [doi: 10.1145/2896338.2896342]
14. Bousquet J, Anto JM, Bachert C, Baiardini I, Bosnic-Anticevich S, Walter Canonica G, et al. Allergic rhinitis. Nat Rev Dis Primers 2020 Dec 03;6:95 [FREE Full text] [doi: 10.1038/s41572-020-00227-0] [Medline: 33273461]
15. Okubo K, Kurono Y, Ichimura K, Enomoto T, Okamoto Y, Kawauchi H, Japanese Society of Allergology. Japanese guidelines for allergic rhinitis 2020. Allergol Int 2020 Jul;69(3):331-345 [FREE Full text] [doi: 10.1016/j.alit.2020.04.001] [Medline: 32473790]
16. Yamada T, Saito H, Fujieda S. Present state of Japanese cedar pollinosis: the national affliction. J Allergy Clin Immunol 2014 Mar;133(3):632-9.e5. [doi: 10.1016/j.jaci.2013.11.002] [Medline: 24361081]
17. Oseroff C, Pham J, Frazier A, Hinz D, Sidney J, Paul S, et al. Immunodominance in allergic T-cell reactivity to Japanese cedar in different geographic cohorts. Ann Allergy Asthma Immunol 2016 Dec;117(6):680-689.e1 [FREE Full text] [doi: 10.1016/j.anai.2016.10.014] [Medline: 27979027]
18. Okano M. Practical guideline for the management of allergic rhinitis in Japan 2020. Arerugi 2021;70(3):166-170. [doi: 10.15036/arerugi.70.166] [Medline: 34011770]
19. SRI+ (nationwide drugstore panel research). INTAGE Healthcare Inc. URL: https://www.intage.co.jp/english/service/platform/sdi/ [accessed 2021-08-25]
20. Wakamiya S, Matsune S, Okubo K, Aramaki E. Causal Relationships Among Pollen Counts, Tweet Numbers, and Patient Numbers for Seasonal Allergic Rhinitis Surveillance: Retrospective Analysis. J Med Internet Res 2019 Feb 20;21(2):e10450 [FREE Full text] [doi: 10.2196/10450] [Medline: 30785411]
21. Press release [in Japanese]. NOVARTIS. URL: https://www.novartis.co.jp/news/media-releases/prkk20210122 [accessed 2021-08-25]
22. statsmodels v0.13.1. URL: https://www.statsmodels.org/stable/generated/statsmodels.tsa.seasonal.seasonal_decompose.html [accessed 2021-08-25]

Abbreviations

OTC: over-the-counter
POS: point-of-sale
QOL: quality of life

Edited by A Mavragani; submitted 30.09.21; peer-reviewed by S Matsuda, E Daskalaki, A Allam; comments to author 22.10.21; revised version received 19.12.21; accepted 04.01.22; published 02.02.22.

Please cite as:
Wakamiya S, Morimoto O, Omichi K, Hara H, Kawase I, Koshiba R, Aramaki E
Exploring Relationships Between Tweet Numbers and Over-the-counter Drug Sales for Allergic Rhinitis: Retrospective Analysis
JMIR Form Res 2022;6(2):e33941
URL: https://formative.jmir.org/2022/2/e33941
doi:10.2196/33941
PMID:35107434

Original Paper

An Android-Based Mobile App (ARVPredictor) for the Detection of HIV Drug-Resistance Mutations and Treatment at the Point of Care: Development Study

Beatrice Ongadi[1], MSc; Raphael Lihana[2*], PhD; John Kiiru[3*], PhD; Musa Ngayo[3*], PhD; George Obiero[1*], PhD

[1]Centre for Biotechnology and Bioinformatics, University of Nairobi, Nairobi, Kenya

[2]Centre for Virus Research, Kenya Medical Research Institute, Nairobi, Kenya

[3]Centre for Microbiology Research, Kenya Medical Research Institute, Nairobi, Kenya

[*]these authors contributed equally

Corresponding Author:
Beatrice Ongadi, MSc
Centre for Biotechnology and Bioinformatics
University of Nairobi
Department of Biochemistry - CEBIB
P.O. Box 30197
Nairobi, 00100
Kenya
Phone: 254 728324324
Email: betongadi@gmail.com

Abstract

Background: HIV/AIDS remains one of the major global human health challenges, especially in resource-limited environments. By 2017, over 77.3 million people were infected with the disease, and approximately 35.4 million individuals had already died from AIDS-related illnesses. Approximately 21.7 million people were accessing ART with significant clinical outcomes. However, numerous challenges are experienced in the delivery and accurate interpretation of data on patients with HIV data by various health care providers at different care levels. Mobile health (mHealth) technology is progressively making inroads into the health sector as well as medical research. Different mobile devices have become common in health care settings, leading to rapid growth in the development of downloadable software specifically designed to fulfill particular health-related purposes.

Objective: We developed a mobile-based app called ARVPredictor and demonstrated that it can accurately define HIV-1 drug-resistance mutations in the HIV pol gene for use at the point of care.

Methods: ARVPredictor was designed using Android Studio with Java as the programming language and is compatible with both Android and iOS. The app system is hosted on Nginx Server, and network calls are built on PHP's Laravel framework handled by the Retrofit Library. The DigitalOcean offers a high-performance and stable cloud computing platform for ARVPredictor. This mobile app is enlisted in the Google Play Store as an "ARVPredictor" and the source code is available under MIT permissive license at a GitHub repository. To test for agreement between the ARVPredictor and Stanford HIV Database in detecting HIV subtype and NNRT and NRTI mutations, a total of 100 known HIV sequences were evaluated.

Results: The mobile-based app (ARVPredictor) takes in a set of sequences or known mutations (protease, reverse transcriptase and integrase). It then returns inferred levels of resistance to selected nucleoside, nonnucleoside protease, and integrase inhibitors for accurate HIV/AIDS management at the point of care. The ARVPredictor identified similar HIV subtypes in 98/100 sequences compared with the Stanford HIV Database (κ=0.98, indicating near perfect agreement). There were 89/100 major NNRTI and NRTI mutations identified by ARVPredictor, similar to the Stanford HIV Database (κ=0.89, indicating near perfect agreement). Eight mutations classified as major by the Stanford HIV Database were classified as others by ARVPredictor.

Conclusions: The ARVPredictor largely agrees with the Stanford HIV Database in identifying both major and minor proteases, reverse transcriptase, and integrase mutations. The app can be conveniently used robustly at the point of care by HIV/AIDS care providers to improve the management of HIV infection.

(JMIR Form Res 2022;6(2):e26891) doi:10.2196/26891

KEYWORDS

database; mobile Android app; HIV/AIDS; mutation; pol gene; protease; reverse transcriptase; integrase; ARVPredictor; mobile app; mHealth; HIV; Android; digital health

Introduction

The dynamic and exponential growth of economical mobile phones with sufficient built-in resources has been experienced worldwide today [1]. Business systems can currently leverage the strength of smartphones to affect long-distance transactions and undertake business consultations off-site robustly and accurately at low operating costs [2]. Mobile technology has progressively found vital utility in various health care platforms globally [3]. After the introduction of the first cellular phone by Motorola in the 1990s [4], faster processors, improved storage spaces, smaller longer life batteries, and superefficient mobile operating systems have been invented, paving avenues for wide useful application developments [5]. Mobile phones have become a common device in health care settings. This has in effect led to the proliferation of small self-contained pieces of downloadable software. Numerous medical software apps are now available to assist different health care professionals with many tasks, such as health record maintenance, time management, patient management and monitoring, clinical decision making, medical education and training, and online diagnosis [6]. In the recent past, mobile health (mHealth) has been used instrumentally to address various mental conditions [7], such as substance abuse disorders [8], depression [9], psychosis [10], and suicide. This and other related mHealth-based systematic reviews give an indication of the very positive future of human health care via smartphones [11,12].

Globally, by 2017, HIV/AIDS had infected over 77.3 million people, among which 35.4 million people or more have so far died from AIDS-related illness [13]. Kenya is among the worst hit countries, recording high prevalence and mortality rates alongside Mozambique and Uganda [14]. The advent of antiretroviral (ARV) therapy has led to a significant reduction in morbidity and mortality among different populations [15]. By 2017, over 21.7 million people were accessing antiretroviral therapy with significant clinical outcomes. Unfortunately, the development of HIV drug-resistance (HIVDR) mutations is currently jeopardizing these clinical benefits [16]. The United Nations Programme on HIV/AIDS (UNAIDS) 90-90-90 initiative [17] is also likely to be affected by increased HIV drug-resistant mutations, especially in developing countries. Therefore, drug-resistance testing is vital for accurate patient management [18]. Additionally, with the rolling out of "Treat all" in the HIV world, the occurrence of drug-resistant HIV is likely to become the key public health threat, thereby hindering treatment options available to people with HIV/AIDS [14]. The global prevalence of HIVDR is rising, mainly due to resistance to non-nucleoside reverse transcriptase inhibitor (NNRTI) drugs, which make up the backbone of World Health Organization's (WHO) first-line antiretroviral treatment procedures [14]. Hence, there is an increase in the need for available and easy-to-access viral load testing avenues and other patient monitoring methods to alleviate the rise in HIVDR at the point of care [19].

ARVPredictor, which was designed and developed in this study, is a mobile app that bridges the glaring gap by providing point-of-care results to clinicians, most of whom in Kenya are now trained on HIVDR result interpretation. It comes with the advantages of mobile telephony and the capability to serialize evolutionary and drug-related sequence variation in HIV reverse transcriptase and protease, which are targeted by different antiretroviral drugs. It is enlisted in the Google Play Store as "ARVPredictor" [20]. The app is designed using Android Studio with Java as the programming language and is compatible for both Android and iOS with user-friendly graphical interfaces. This app accurately predicts HIVDR mutations in the *pol* region similar to those obtained from the web-based Stanford HIV Drug Resistance Database (HIVdb) [21,22].

Methods

Design and Setting

We obtained a replicated worksheet for developing ARVPredictor from different drug-resistance databases. The list included mutations existing in the ANRS (the French National Agency for Research on AIDS), HIVdb, IAS–USA (International Antiviral Society–USA), Los Alamos, and Rega algorithm lists [23]. We summarized the final register to a total of 19 normalized core tables, 10 lookup tables, and up to 20 derived tables. This was in line with other simulated databases mostly implemented using MySQL on Linux platforms. We ended up with ordered relationships linking main entities within, including antiretroviral therapy history, isolated drug-susceptibility outcome, and patient plasma HIV-1 RNA levels. The final database allows users to retrieve and analyze different sets of sequences that meet particular criteria. Common queries envisioned included retrieval of sequences of HIV-1 isolates containing mutations at specific positions, patients receiving a specific drug regimen, and drug-susceptibility data on HIV-1 isolates containing combinations of mutations. Every designed query provided the following category of data: hyperlinks to MEDLINE and GenBank records, a list of mutations in the sequence, a classification of the sequence, drug-susceptibility results, and some technical data. We availed options for downloading or viewing raw sequence data at the back end; each predesigned table returned 8 or more columns of data.

Participants

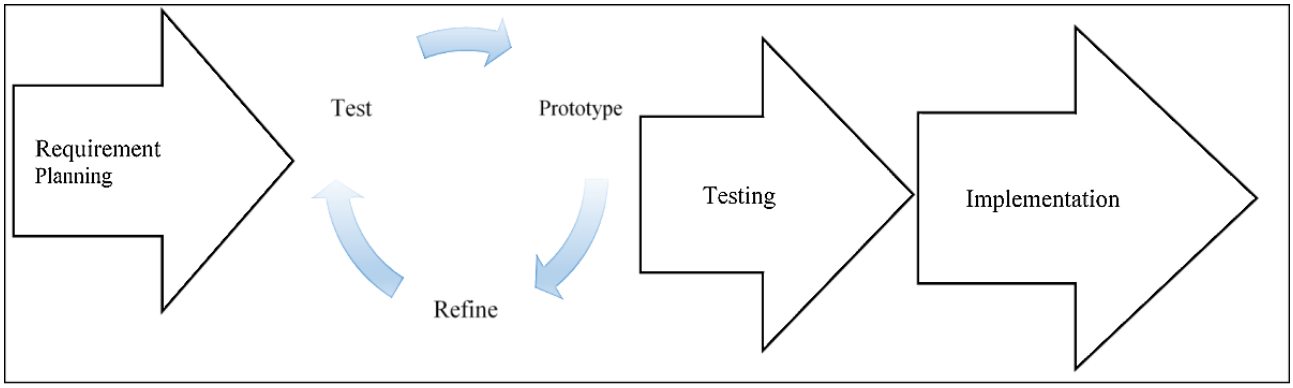
To test the app's usability, a random population comprising 100 health practitioners actively involved in HIV/AIDS management, app developers, and information and communications technology (ICT) students was recruited. The health practitioners enrolled in the usability survey included 10 HIV experts from the Kenya HIVDR Technical Working Group on drug-resistance approval and regimen guidance, 25 HIV scientists based at the Kenya Medical Research Institute and various universities in Kenya. Others included 20 medical and pharmacy students from the University of Nairobi, Kenya, 20 graduates of the National

Advanced HIV Clinical Course (NAHCC) class of 2015, 10 app developers, and 15 ICT students from Jomo Kenyatta University of Agriculture and Technology, Kenya. All the enrolled health practitioners were required to own or have access to Android-based smartphones; those who needed intensive hands-on training on downloading and using their smartphones were automatically excluded in the survey. The health practitioners were invited to participate in this survey using open invites through email to selected human health research institutions, universities (especially ICT and virology departments), and HIV comprehensive care centers. The survey also invited other participants directly via phone calls. Participants willing to be engaged recorded their interest through replying to our email by filling out a short acceptance/consenting online form and were enrolled on a first-come-first-enrolled basis. Those selected received a link on how to download and use our ARVPredictor Android app for test purposes.

App Development Process

We designed and developed ARVPredictor from models of a combination of software development life cycle methodology [24] and rapid application development (RAD) [25]. RAD is an agile strategy for developing software that has proven to be fast and helps complete a project within a shorter timeframe. It achieves this by reducing the time spent in planning and maximizing prototyping development. The RAD enables faster communication between a developer and the end user; hence, its high efficiency is achieved by following 4 main phases at a reasonably lower cost [26]. The initial working prototype is developed and is improved gradually through discussions until a satisfying output is reached (Figure 1).

Figure 1. The ARVPredictor rapid application development model.



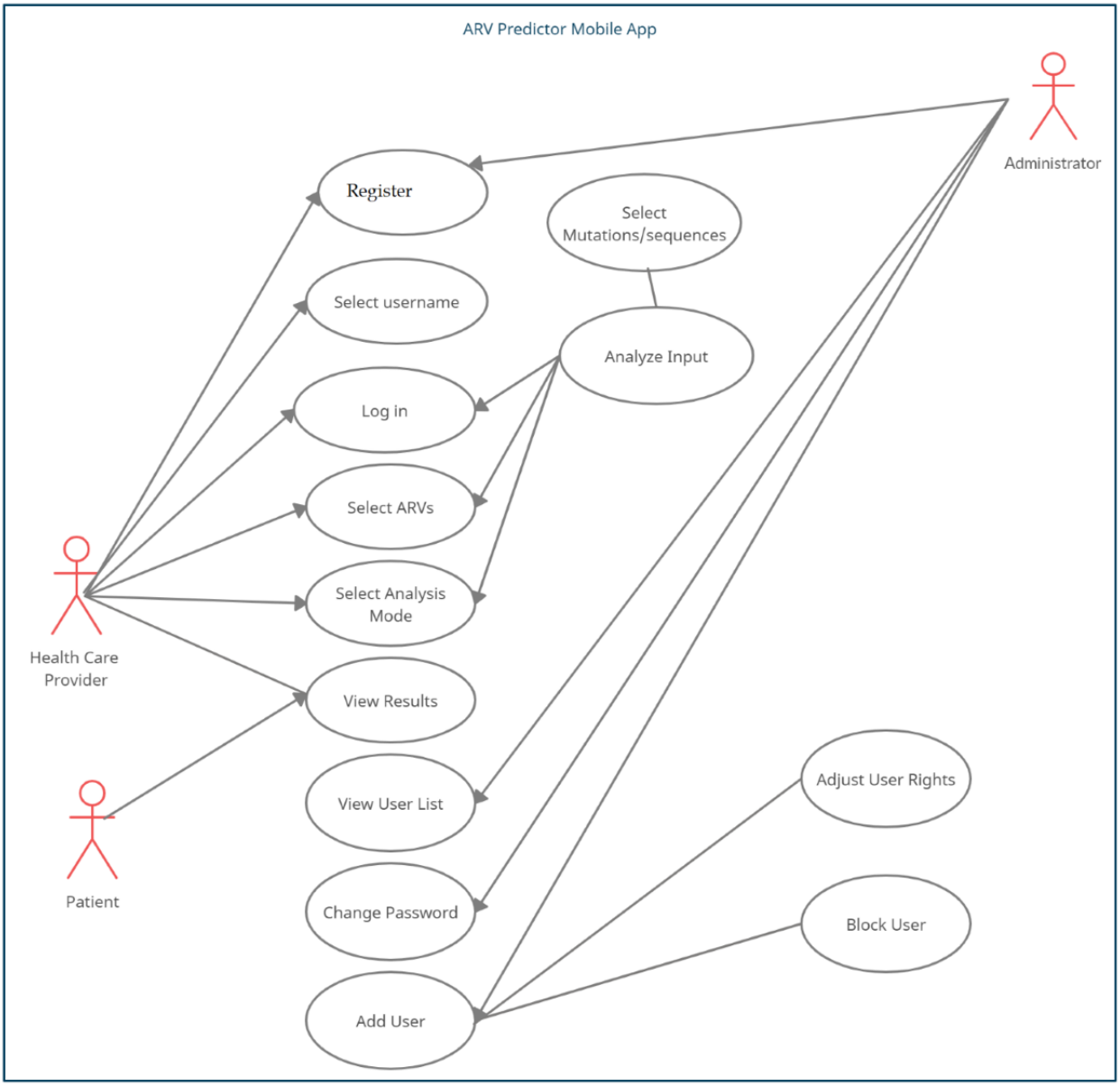
Gap Identification and Requirements Gathering

A combination of 3 main activities was carried out at the initial stages of the development to meet both functional and nonfunctional requirements of the app. A brief questionnaire was randomly administered to consented selected HIV/AIDS doctors/clinicians and health care providers to determine and understand the current treatment process. This was followed by a face-to-face chat with a few health care providers who were very keen in using their smartphones to support their daily services to their patients. By performing a quick analysis of the results and with reference to relevant literature, a glaring gap was identified in the turnaround time and availability of point-of-care resources for interpreting the genetically modified HIV strains for appropriate antiretroviral drug choices. It was also understood that with the current mobile telephony systems, different forms of data can now be shared easily among various devices [27]. To create the initial prototype of the app, the output identified during this stage was translated into modules for better understanding and ownership of the whole app [28].

Use Case Modeling

Use case diagrams show the interaction between users and the system [29]. In this app, we first identified the main actors and their respective interactions with the app. Each identified role was assigned relevant access rights and hence classified as *patient*, *health care givers*, and *administrators* of the system (Figure 2). Administrators have the capability to manage the critical functions of the app, whereas a normal user has very limited access to the back end functionality. By signing in using an existing email address and an active phone number, a normal user can query the database using mutations/sequences. A willing patient also has the capability to view the analysis output with predefined limitations.

Figure 2. ARVPredictor use case diagram illustrating graphical interaction between users and the app. ARV: antiretroviral.



Operational Process of the App

A simple plain text prototype identified several app components, including mobile clients, web application clients, real-time servers, and databases. Following that, the development environment was set up for full development of the mobile app based on the various hardware and software specifications listed in the following subsections.

Android Studio 4.1 and Java 10

The team developed this Android app using Android Studio 4.1 [30] and Java version 10 [31] as the programming language. Android Studio is the official Integrated Development Environment (IDE) for Android development and usually includes several aspects required for building various Android apps. It is based on IntelliJ IDEA [32] with powerful code editors and other developer-preferred tools. Integrated aspects of IntelliJ IDEA are essential in maximizing productivity with intelligent coding support. This aspect of the tool proved very handy in code management and provided coding hints during the development of ARVPredictor.

The Java programming language used in this app was developed in early 1910 by Sun Microsystems [33]. It is simple and efficient and can be used for various programming purposes. The combined and ultimate strength of these tools provided a suitable development platform for this app.

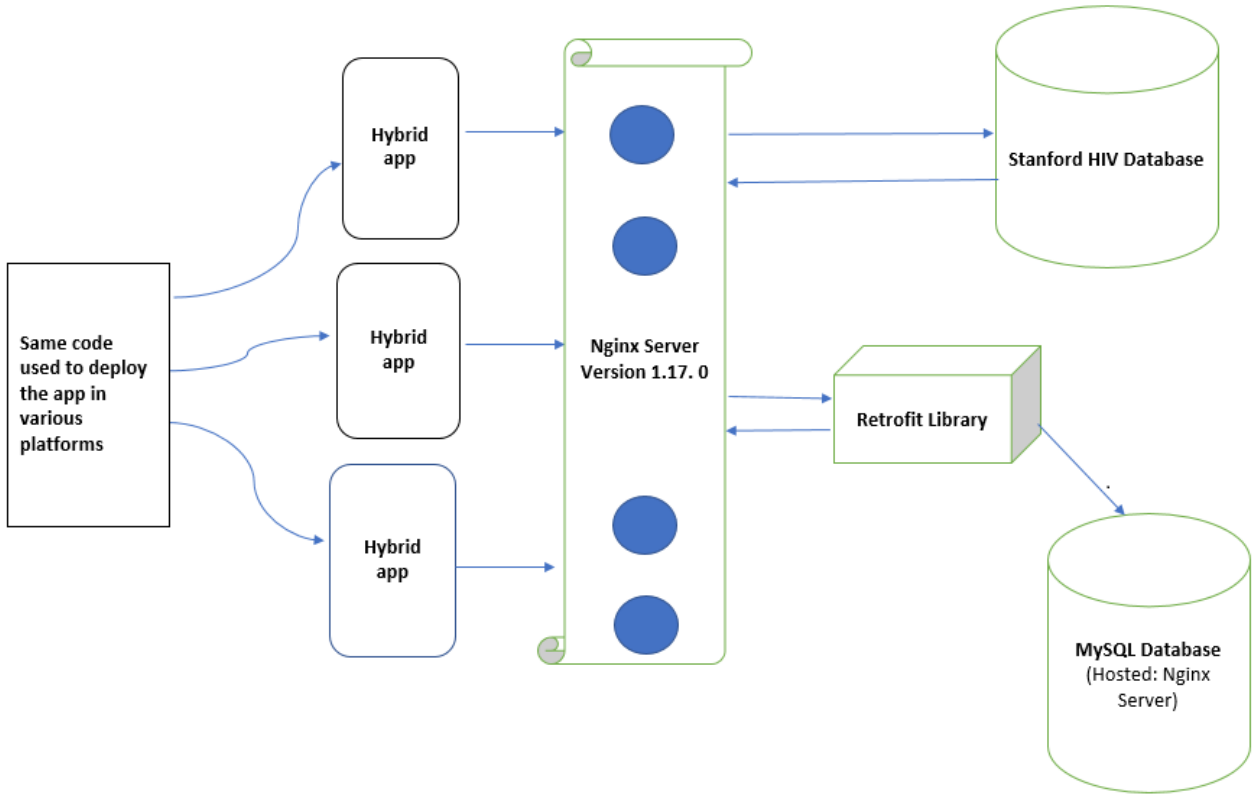
Nginx Server Version 1.17.0

The development of this app involved Nginx Server [34] written in C programming language. This is an open source and a high-performance development platform. It is characterized by reasonable resource consumption, ease of configuration, stability, and a comprehensive feature set. Nginx supports both a high-performance HTTP server and a reverse proxy together with POP3/IMAP proxy servers. The developed app (ARVPredictor) is hosted on the Nginx server with DigitalOcean Droplets [35] offering the cloud hosting platform.

MySQL

We projected massive growth of data in ARVPredictor and used MySQL [36], which is an open source document database written in C++. It is capable of creating and deploying a highly scalable database with high performance capability. It was selected due to its ability to work across platforms, high querying capability, availability, and predictable online professional support. The MySQL database management system has the capability to store data from a single record to a large amount of information. User data in ARVPredictor are stored in the MySQL database hosted on the same Nginx Server version 1.17.0 [34]. The actual arrangement of the development heap comprising mobile phones, servers, and databases is presented in Figure 3.

Figure 3. App development stack: back end arrangement of numerous connecting devices, main server, and active connection between the app (ARVPredictor) and Stanford HIVdb [21,22].



Apollo Android

The ARVPredictor app was developed to connect to the Stanford Database [21] query language (GraphQL) through the Apollo Android library [37]. This platform converts and transfers data between the HIV Stanford Database and the ARVPredictor user interface.

Retrofit Android Library

This is a rest client library [38]; in this app, retrofitting was used to handle all network calls from the app's back end built on PHP's Laravel framework.

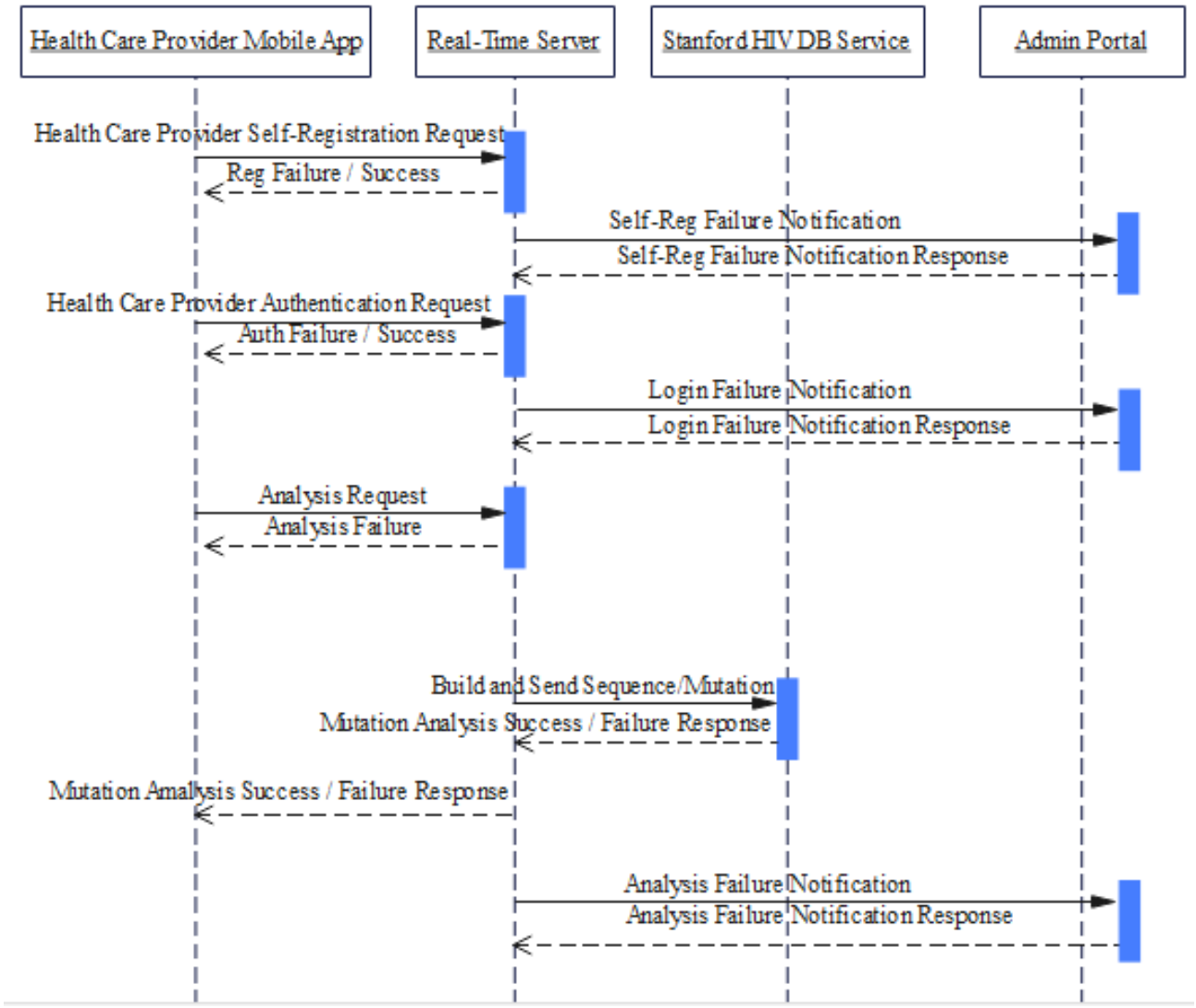
Mobile Client Environment

To ensure proper functionality and accurate returning of results, the app requires phone-free space of not less than 50 MB, touch screen display of 3.0 inches or higher, Android OS version 3.2 or later, and adequate data bundles.

App System Design

The design of the system was created and subjected to a simple objective versus output evaluation to ensure that all aspects of the app requirement were captured (Figure 4). The health care provider signs up using a valid and authenticated email address and uses a phone number and a password to access the app. A real-time server is set to build and send sequence or mutation files to the Stanford Database and return the expected analysis outcome. Different variables are captured immediately into an Excel sheet (Microsoft) and available for statistical analysis of the app.

Figure 4. Application system design and preliminary testing procedure: Model indicating all probable activities by the user of ARVPredictor and predictable respective responses by both real-time server and Stanford HIVdb.



ARVPredictor Deployment and Testing

ARVPredictor is hosted on the Nginx server with DigitalOcean Linux Droplets offering the cloud hosting platform. For security, public access rights were set to be limited for the MySQL database. The Android app is freely distributed through Google Play (Google Play Store). To use ARVPredictor, users first need to create an account with very limited personal details (eg, valid email addresses, passwords, and preferred active phone numbers; see Multimedia Appendix 1). We deemed data gathered during the sign-in process as important for future growth of the project and would be kept under controlled access. The fully developed mobile app was then subjected to 2 levels of evaluation and testing, first through random entry of data to ensure system component interactivity and proper functionality. The second is to guarantee accurate analysis and result output. Black-box testing was used where the functionality of each simple app was subjected to testing without minding its internal structures or workings. This preliminary testing helped to reduce the cost and time spent in the final testing stages of the app development.

ARVPredictor Maintenance

Regular and administrative monitoring steps were set up to assist in the continuous review of the app's system logs. We considered this information useful in understanding and maintaining the app's service health status. The output of this process informs maintenance-related needs such as any faults, downtimes, and any unauthorized activities. It helps in debugging and resolving any issues that arise as well as provides a platform for future upgrades.

Input of Actual Data for Analysis

Two types of data, namely, sequences and point mutations, are usable for analysis in ARVPredictor. The said sequences can be keyed in or pasted directly onto the screen of the phone or alternatively uploaded from a separate file. HIV point mutations are preconfigured and are selectable based on the WHO 2009 listing [23]; both can be analyzed and confirmed according to the latest available version of the Stanford Database [21].

Test Performance

The test performance of ARVPredictor against the Stanford HIV Database to determine HIV subtypes and both major and

minor proteases, reverse transcriptase, and integrase mutations was determined using kappa statistics [39]. Accuracy of the ARVPredictor was then determined. A set of 100 sequences was blasted using both platforms, and the similarity was identified (Multimedia Appendix 2). The subtypes and minor and major mutations were noted for both platforms. To test for method accuracy, each sequence was blasted 3 times, and in each case, any variation (if any) in the subtype, base pairs, and mutations was recorded.

Results

Overview

The mobile app ARVPredictor was designed to take in data on protease, reverse transcriptase, and integrase mutations or sequences. It returns inferred levels of resistance to selected proteases, nucleosides, non-nucleosides, and integrase inhibitors for accurate HIV/AIDS management at the point of care.

Sequence Analysis

We demonstrate the value and usability of ARVPredictor by analyzing different sets of reference sequences from the National Center for Biotechnology Information (NCBI) and other related studies. From numerous test results, we diagrammatically present sampled comparison results for the ARVPredictor against the gold-standard Stanford HIV Database [21]. The test sequence data set is presented in Textbox 1.

Textbox 1. Test sequence data.

Accession No: KX505361.1: HIV-1 isolate 5873 from Kenya pol protein (pol)gene, partial cds [40].

AATGGCCATTGACRGAAGAAAAATAAAGGCATTGATAGAAATTTGTACAGAGATGGAAAAGGAAGGAAA

AATTTCAAGAATTGGGCCTGAGAATCCATACAATACTCCAGTATTTGCCATAAAAARGAAAGACAGTACT

AAGTGGAGAAAATTAGTAGATTTCAGGGAACTCAATAAAAGAACCCAAGACTTTTGGGAAGTTCAATTAG

GRATACCACACCCAGCAGGGTTAAAARAGAAAAAATCAGYGACAGTACTAGATGTGGGGGATGCRTATTT

TTCAGTWCCTTTAGATGAAAGCTTCAGGAAATATACTGCATTYACCATACCRAGTRTAAACAATGAGACA

CCAGGAATCAGRTATCAGTACAATGTGCTTCCACAAGGATGGAAAGGATCACCRGCAATATTCCAAGCTA

GCATGACAAAAATYCTGGAACCTTTTAGGAAACAAAATCCAGAAATGATTATCTATCAATACATGGATGA

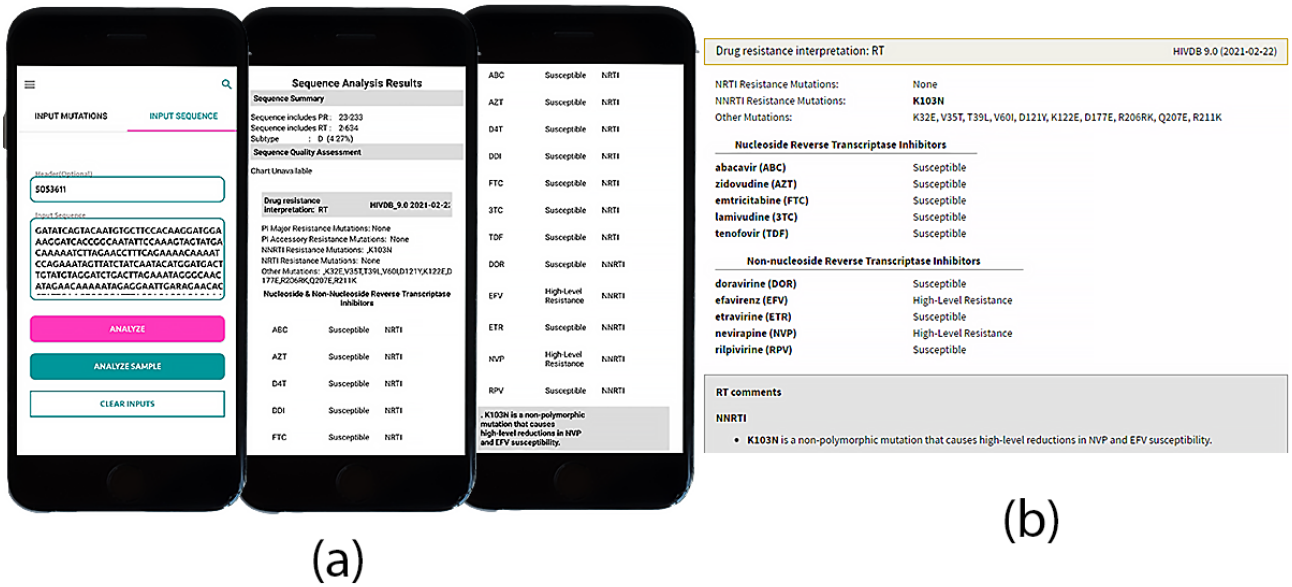
TTTGTATGTAGGATCTGACTTAGAAATAGGGCAACATAGAGCAAAAATAGAGRAATTAAGGGAACACCTG

TTAAAGTGGGGGTTTACTACACCAGACAAAAAGCATCAGAAAGAACCTCCAYTCCTTTGGATTGGTTAT

Sequence Comparative Output

Figure 5 shows comparable sequence analysis results for both the Stanford Database [21] and ARVPredictor.

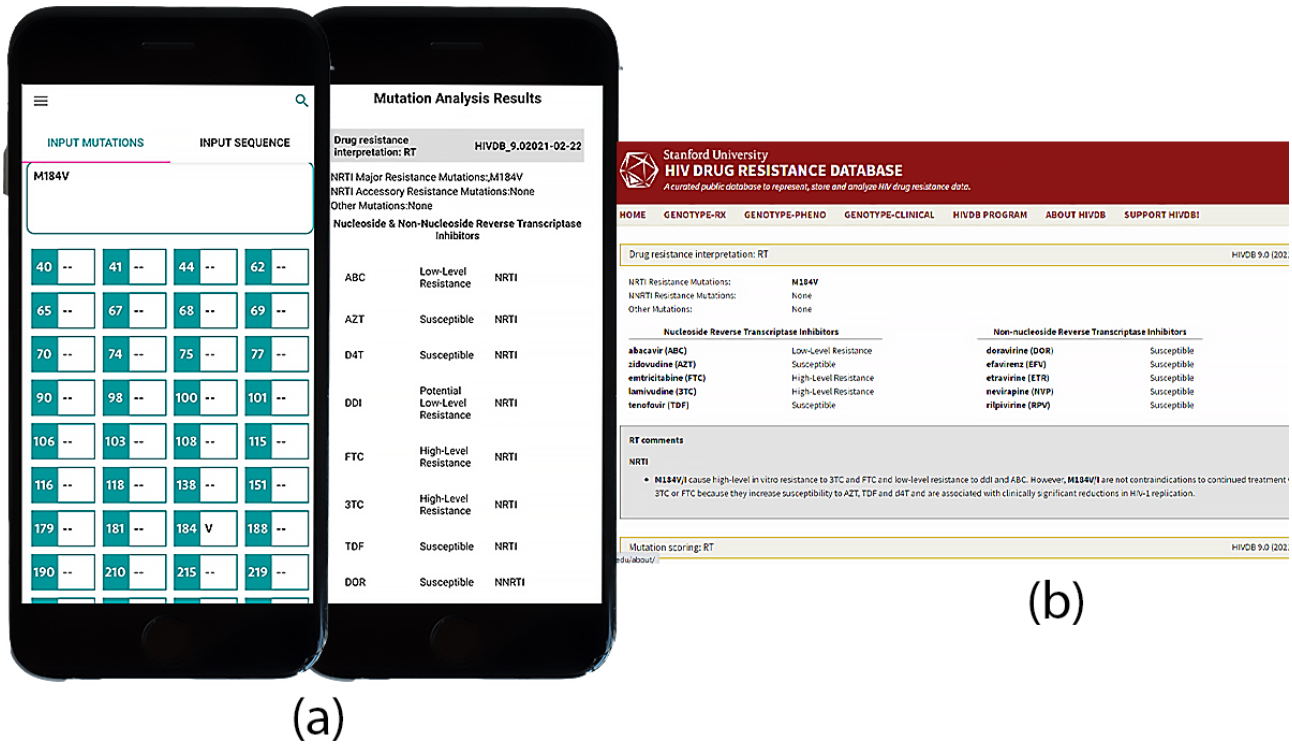
Figure 5. Sequence analysis results: Demonstrating similarity between sequence analysis results output for both (a) ARVPredictor and (b) Stanford HIV Database.



Mutation Analysis

From the WHO Major HIV-1 Drug Resistance Mutations list [23], we tabulated the analysis results of the predefined HIV mutation M184V. The most common resistance mutations occur in nucleoside reverse transcriptase inhibitors (NRTIs) in vitro for both the ARVPredictor and Stanford Database. There is a high level of resistance to emtricitabine (FTC) and lamivudine (3TC) and potentially low level of resistance to didanosine (DDI) while displaying susceptibility to both zidovudine (ZDV) and tenofovir (TDV). The results shown in Figure 6 are comparable.

Figure 6. Mutation Analysis Results. Demonstrating similarity between mutation analysis results for (a) ARVPredictor and (b) Stanford HIV Database.



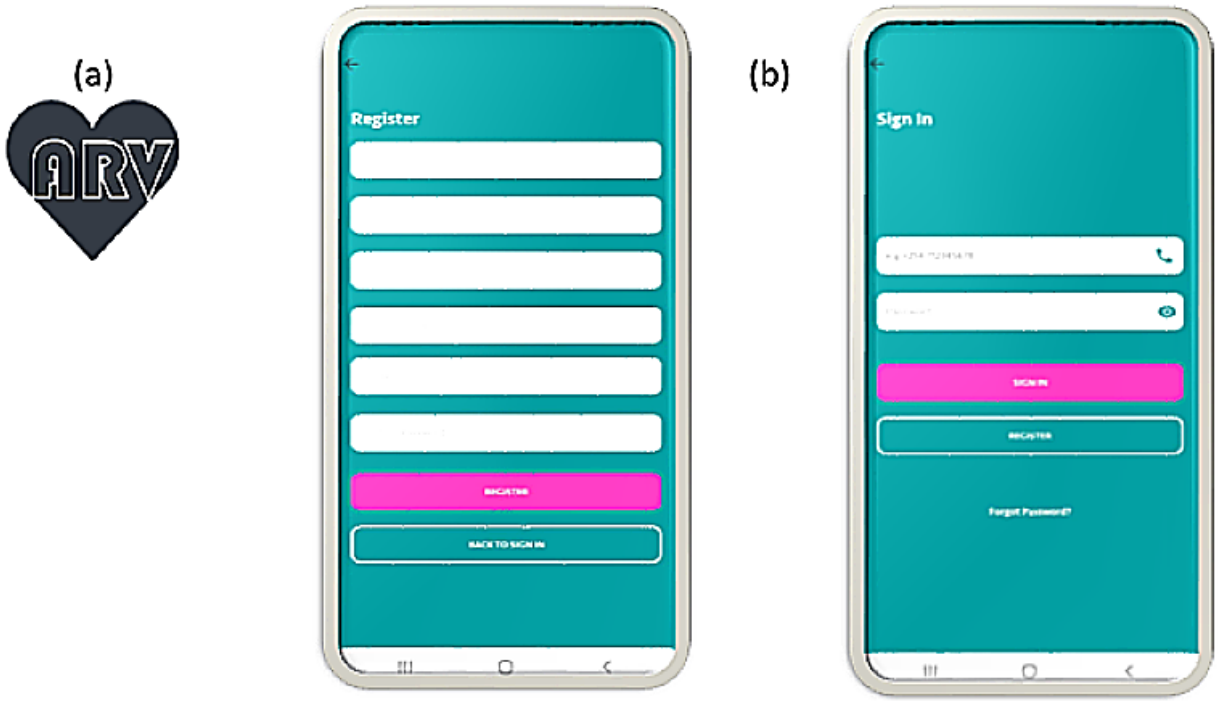
Test Performance and Agreement

Multimedia Appendix 3 summarizes sequences with variation in subtype and mutations as determined by both the ARVPredictor app and the Stanford Database. The ARVPredictor app identified similar HIV subtypes in 98/100 sequences compared with the Stanford HIV Database (κ=0.98, indicating near perfect agreement). There were 89/100 major NNRTI and NRTI mutations identified by ARVPredictor, similar to the Stanford HIV Database (κ=0.89, indicating near perfect agreement). Seven mutations classified as major mutations by the Stanford HIV Database were classified as other mutations by ARVPredictor. This further indicates that the Stanford-confirmed GraphQL web service works fairly well, and all the results are in sync with most parts of the web version. Both tools found several minor/other mutations, but depending on small phone display window, some may be hidden from view.

ARVPredictor Availability

ARVPredictor is currently distributed freely through the Google Play Store and App Store (Apple Inc.), with basic rules requiring users to create an account and provide very limited personal details such as email addresses, passwords, and preferred telephone numbers. Different user-friendly interfaces viewable through the usage of the app are shown in Figures 7-11. The source code for ARVPredictor is available under the MIT permissive license in a GitHub repository [41].

Figure 7. ARVPredictor Preliminary User Interfaces. (a) The pictorial icon shows how ARVPredictor appears ready for download from the app/play store. (b) Potential users must register in order to use the application, registered users are required to sign-in.



User Friendly Interfaces

User Registration and User login

First-time users of the app require registration into the system. This process helps prevent unauthorized access and use of the app. An SMS text message alert is then sent through the given phone number with an activation code. System registration can occur in 2 different ways: through the app itself (Figure 7) or by the administrator remotely on request. Figure 8 shows the respective back end code used in developing this registration process. Only authenticated and valid email addresses/phone numbers are allowed for registration.

Figure 8. User Registration and Sign in source code: Set of computer logical instructions aiding registration and sign in process before using the app.

```
call.enqueue(new Callback<Register>() {
    @Override
    public void onResponse(@NotNull Call<Register> call, @NotNull Response<Register> response) {
        Register respo = response.body();

        if (respo != null) {
            UserData userData = respo.getData();

            if (respo.getToken().equals("NONE")) {
                Message.makeToast(activity, activity, respo.getMessage());
            } else {

                if (userData != null) {
                    userDataBox.removeAll();
                    userDataBox.put(userData);
                    Session.sessionStoreData(respo.getToken(), userData.getFirstName(), userData.getSecondName(),
                            userData.getPhoneNumber(), userData.getEmail(), userData.getStatus().toString(), activity, activity);
                    StaticVariables.first_name = userData.getFirstName();
                    StaticVariables.email = userData.getEmail();

                }
                onLoginFormActivityListener.doLogin();
                Message.makeToast(activity, activity, respo.getMessage());
                PreferenceManager preferenceManager = new PreferenceManager(activity);
                preferenceManager.setLoginStatus(true);

            }

        } else {
            Message.makeToast(activity, activity, message: "Response is null!");
        }
```

Antiretroviral Drug Options/Input Window

Upon successful registration and verification, the next button opens the known and common antiretroviral drug display window (Figure 9). Some antiretroviral drugs will be premarked by default, but checkboxes can be used to add or remove any drug. A "save" button at the bottom of the phone screen allows one to save and use his/her final choice. Figure 10 illustrates the back end code used in the development of this app's preferred drug selection process.

Figure 9. Antiretroviral Display Options: Displaying all current ARVs and categorizes them as Nucleoside reverse transcriptase inhibitors (NRTIs), Non-nucleoside reverse transcriptase inhibitors (NNRTIs), Integrase Strand Transfer Inhibitor (INSTIs) and Protease Inhibitor (PI).

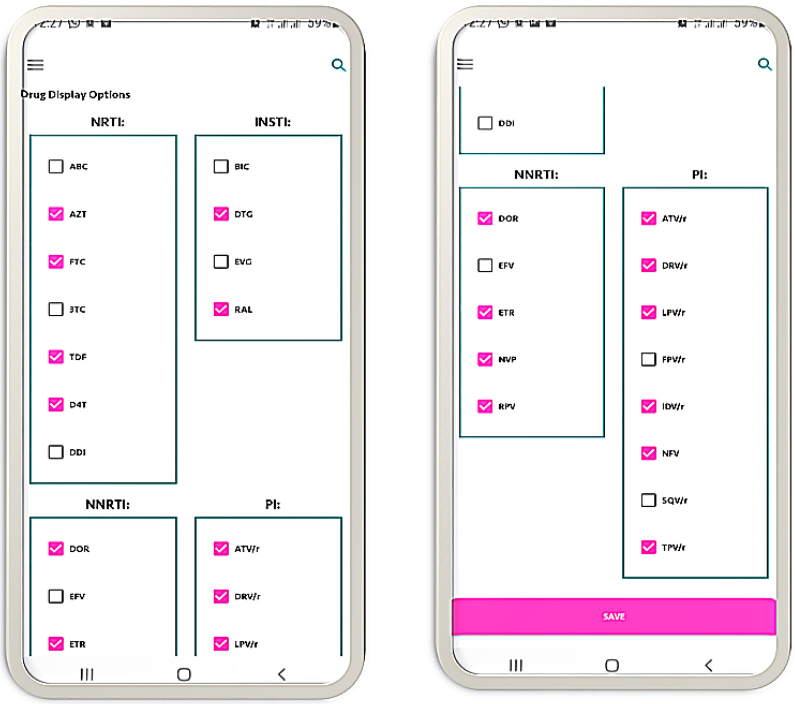


Figure 10. ARV Drug Selection code by category. Set of back-end computer related logical instructions of selection of different ARV drugs available for use by various HIV Patients.

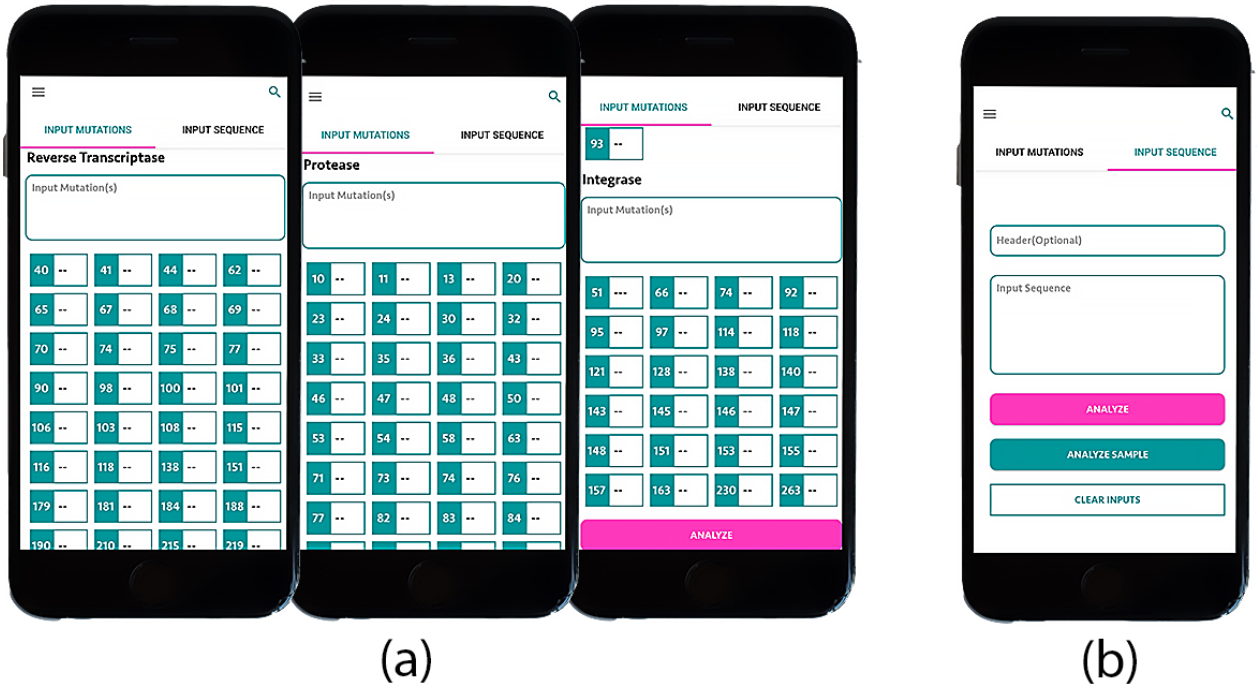
```
try{
    holder.drug.setText(data.get(position).getName());
    holder.drug.setChecked(data.get(position).getSelected());
    holder.drug.setTag(position);
    holder.drug.setOnClickListener(new View.OnClickListener(){

        @Override
        public void onClick(View view) {
            Integer pos = (Integer) holder.drug.getTag();
            String name = data.get(pos).getName();
            if (preferenceManager.getDrugStatus(name)){
                preferenceManager.setDrugSelected( selected: false,name);
            }
            else {
                preferenceManager.setDrugSelected( selected: true,name);
            }
        }
    });
}catch (Exception e){
    e.printStackTrace();
}
```

HIV Mutation Selection Window

The next screen displays 2 options of the input type for analysis: point mutations or sets of sequences (Figure 11). The first group is a list of all the registered major HIV-1 drug-resistance mutations as per WHO 2009 data [23]. Three mutation categories, namely, reverse transcriptase, protease, and integrase, can be selected from a scroll down window and analyzed at the bottom of the screen. The second part of this figure is a sequence input window. It provides an option to either upload a set of sequences from a separate source or manually key them in.

Figure 11. ARVPredictor. (a) Major HIV-1 drug resistance mutation list. (b) Sequence entry option window.



ARVPredictor Mutation Analysis Results Window

The results are analyzed and sent back in comparison to the latest version of the Stanford HIV Database. The output is packaged in terms of the region targeted, for example, protease, reverse transcriptase, and integrase, and preferred antiretrovirals are listed alongside. Output of a mutation analysis is shown in Figure 12. The development code for displaying the mutation analysis output in this app is shown in Figure 13. Figures 14 and 15 display the analyzed sequences and respective back end development codes, respectively.

Figure 12. ARVPredictor: Mutation Analysis Results.

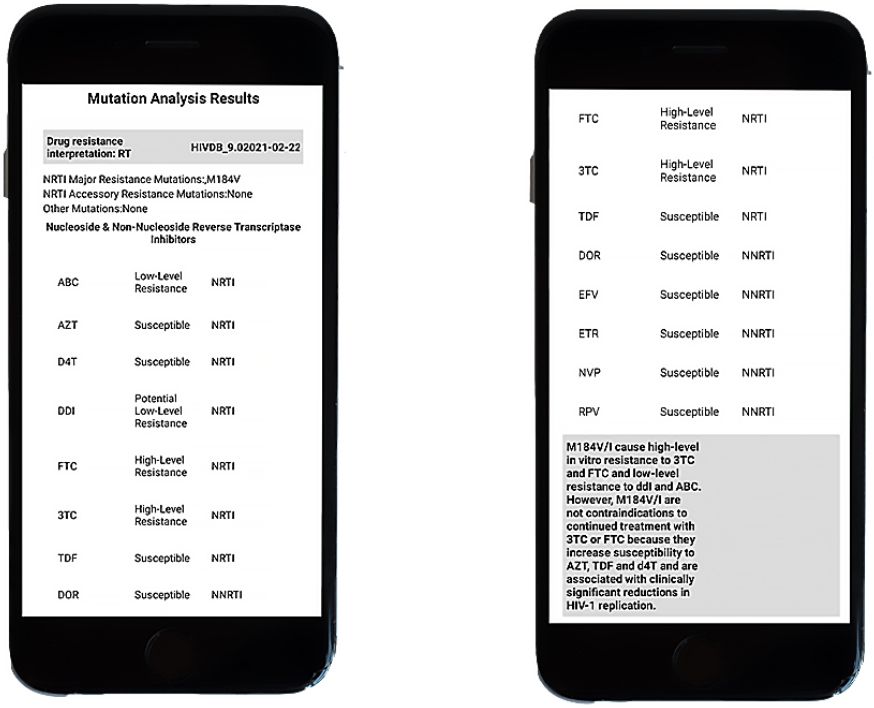


Figure 13. ARVPredictor Mutation Analysis Code.

```
myApolloClient.query(MutationAnalysisRequestQuery.builder().mutations(myMutations).build())
        .enqueue(new ApolloCall.Callback<MutationAnalysisRequestQuery.Data>() {
            @Override
            public void onResponse(@NotNull Response<MutationAnalysisRequestQuery.Data> response) {
                getActivity().runOnUiThread(
                        new Runnable() {
                            @Override
                            public void run() {
                                List<MutationAnalysisRequestQuery.DrugResistance> load = response.getData().mutationsAnalysis().drugResistance();
                                progDialog.dismiss();
                                if (load != null){
                                    Paper.book().write("MUTRESULT",load);
                                    Paper.book().write("DATA",response.getData());
                                }
                                myMutations.clear();
                                Intent intent = new Intent(getActivity(), MutationResultActivity.class);
                                startActivity(intent);
                                  getActivity().finish();
                            }
                        }
                );
            }
            @Override
            public void onFailure(@NotNull ApolloException e) {
                getActivity().runOnUiThread(
                        new Runnable() {
                            @Override
                            public void run() {
                                myMutations.clear();
                                progDialog.dismiss();
                                errorDialog.show();
                            }
                        }
                );
            }
        });
```

Figure 14. ARVPredictor: Sequence Analysis Results.

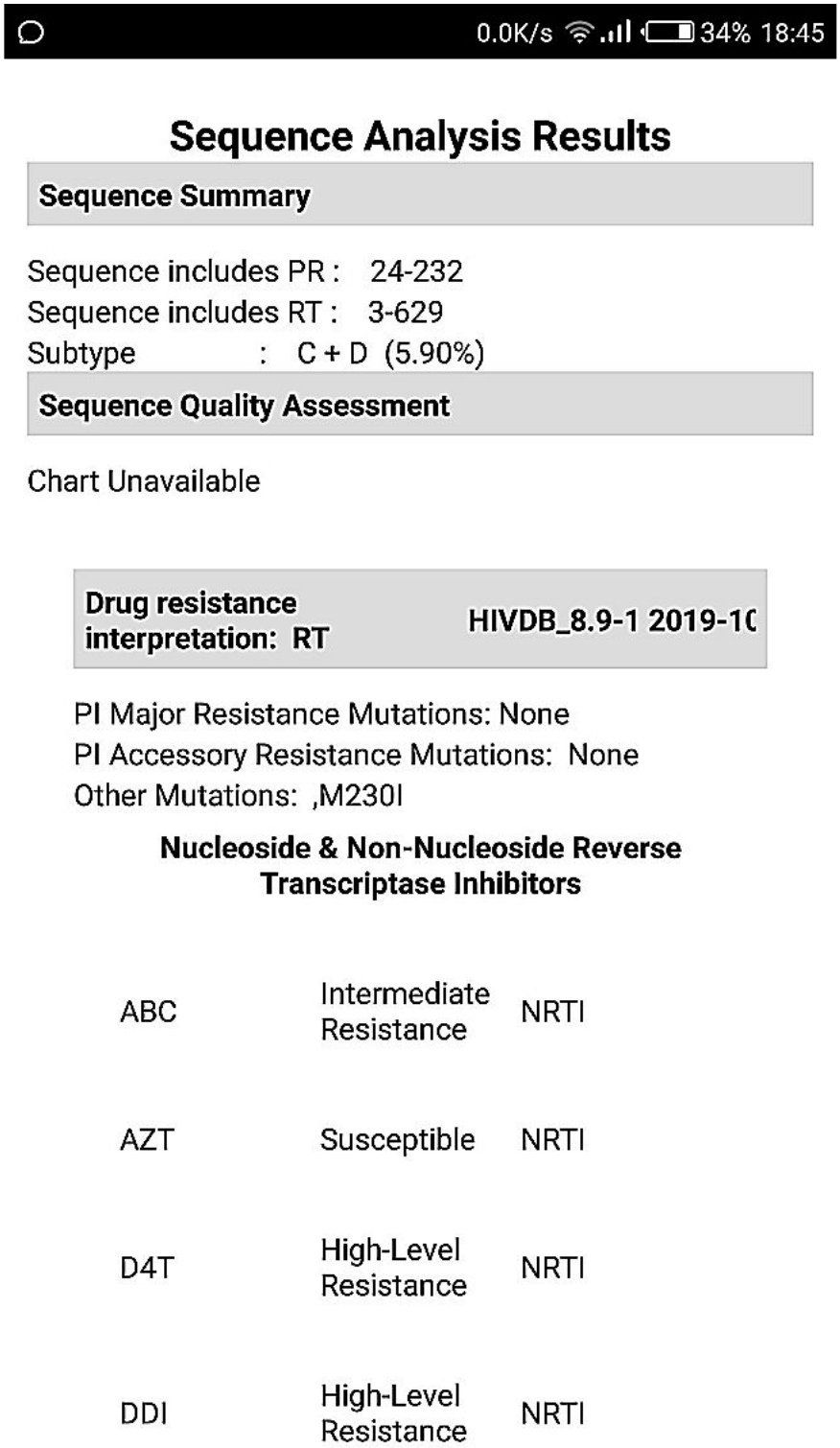


Figure 15. ARVPredictor sequence analysis code.

```
myApolloClient.query(SequenceAnalysisRequestQuery.builder().sequences(actualInputs).build())
        .enqueue(new ApolloCall.Callback<SequenceAnalysisRequestQuery.Data>() {
            @Override
            public void onResponse(@NotNull Response<SequenceAnalysisRequestQuery.Data> response) {
                getActivity().runOnUiThread(new Runnable() {
                    @Override
                    public void run() {
                          sequence.setText(response.getData().toString());
                        Paper.book().write("SEQDATA",response.getData().sequenceAnalysis());
                        Paper.book().write("RAWDATA",response.getData());
                        progDialog.dismiss();
                        header.setText("");
                        sequence.setText("");
                        actualInputs.clear();
                        Intent intent = new Intent(getActivity(), SequenceAnalysisResults.class);
                        startActivity(intent);
                    }
                });
            }
            @Override
            public void onFailure(@NotNull ApolloException e) {
                  sequence.setText(e.getMessage());
                getActivity().runOnUiThread(new Runnable() {
                    @Override
                    public void run() {
                        header.setText("");
                        sequence.setText("");
                        header.setText("");
                        sequence.setText("");
                        progDialog.dismiss();
                        errorDialog.show();
                    }
                });
            }
        });
```

Discussion

Principal Findings

Timely and accurate assistance to health care providers in the management of patients with HIV has been a challenge, especially in resource-limited environments. We presume that with the introduction of ARVPredictor, the much needed solution to health care providers at the point of care will be achieved. The planning and choice of tools used in the development of this app were done with the expected work environment and data output in mind. The choice supported heavy and quick analysis of the mutations or sequences to be uploaded by the health care provider. The overall aim is to provide a solution that offers accurate and easily accessible management strategy to HIV health care providers at their fingertips with a short turnaround time. A great effort was made to minimize the expected input variables for accuracy and time management. To achieve this, the mutation screen presents a dropdown and choice-enabled entries, which include a complete list of all currently documented mutations. The sequence analysis window allows keying in, pasting, and uploading from a remote file.

The app, which is now freely downloadable from Google Play Store or App Store and enlisted as ARVPredictor, is expected to be free for use and technically supported by developers. Its security is enhanced through controlled and authenticated log-in process, while critical and back end data are only accessible to authorized individuals.

Conclusions

In evaluating the functionality of the ARVPredictor in comparison to the replicated Stanford HIV Database in this study, there is strong evidence that several benefits, including but not limited to concurrence, convenience, and simplicity, are realized. ARVPredictor can therefore be used to determine HIV-1 drug-resistance mutations in the HIV *pol* gene with ease and convenience in mobile devices. The app equally has the advantages of high-speed data networks and smartphone accessibility. It therefore adds to other available and upcoming mHealth interventions in the area of HIV and antiretroviral use among different health care providers. Mobile technology has enabled faster and efficient communication among health care service providers and their patients. However, additional benefits, such as improved diagnosis, accuracy, and better coordination, will still contribute highly to the biomedical field.

Conflicts of Interest

None declared.

Multimedia Appendix 1
ARVPredictor User Guide.
[PDF File (Adobe PDF File), 1286 KB - formative_v6i2e26891_app1.pdf]

Multimedia Appendix 2
ARVPredictor Test Performance Sequences (Accession Numbers: KX505314-KX505372 and MK588680-MK588752) using Stanford HIV database as the gold standard.
[DOCX File , 27 KB - formative_v6i2e26891_app2.docx]

Multimedia Appendix 3
Test Performance Output: Performance results of ARVPredictor using the Stanford HIV database as the gold standard in identification of HIV subtypes and mutations.
[DOCX File , 29 KB - formative_v6i2e26891_app3.docx]

References

1. Griffin H. Has Mobile Phone Technology Had an Impact on the Quality of Life in the Developing World? IJARBSS 2013 Jun 18;3(7):7-14. [doi: 10.6007/ijarbss/v3-i7/5]
2. Quade M, Leimstoll U. Mobile business with smartphones and tablets: effects of mobile devices in SMEs. In: BLED 2015 Proceedings. 2015 Presented at: 28 th Bled eConference; June 7-10; Bled, Slovenia p. 290-305 URL: https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1025&context=bled2015
3. 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]
4. Lippi G, Plebani M. Laboratory applications for smartphones: risk or opportunity? Clin Biochem 2011 Mar;44(4):273-274. [doi: 10.1016/j.clinbiochem.2010.12.016] [Medline: 21223958]
5. mobiThinking. Global mobile statistics 2014 Part B: Mobile Web; mobile broadband penetration; 3G/4G subscribers and networks; mobile search. mobiForge. 2014. URL: https://mobiforge.com/research-analysis/global-mobile-statistics-2014-part-b-mobile-web-mobile-broadband-penetration-3g4g-subscribers-and-ne [accessed 2020-12-20]
6. Divall P, Camosso-Stefinovic J, Baker R. The use of personal digital assistants in clinical decision making by health care professionals: a systematic review. Health Informatics J 2013 Mar 13;19(1):16-28 [FREE Full text] [doi: 10.1177/1460458212446761] [Medline: 23486823]
7. Miralles C, Granell C, Díaz-Sanahuja L, Van Woensel W, Bretón-López J, Mira A, et al. Smartphone Apps for the Treatment of Mental Disorders: Systematic Review. JMIR Mhealth Uhealth 2020 Apr 02;8(4):e14897-e14810 [FREE Full text] [doi: 10.2196/14897] [Medline: 32238332]
8. Steinkamp JM, Goldblatt N, Borodovsky JT, LaVertu A, Kronish IM, Marsch LA, et al. Technological Interventions for Medication Adherence in Adult Mental Health and Substance Use Disorders: A Systematic Review. JMIR Ment Health 2019 Mar 12;6(3):e12493-e12424 [FREE Full text] [doi: 10.2196/12493] [Medline: 30860493]
9. Reins JA, Boß L, Lehr D, Berking M, Ebert DD. The more I got, the less I need? Efficacy of Internet-based guided self-help compared to online psychoeducation for major depressive disorder. J Affect Disord 2019 Mar 01;246:695-705 [FREE Full text] [doi: 10.1016/j.jad.2018.12.065] [Medline: 30611913]
10. Killikelly Z, He Z, Reeder C, Wykes T. Improving Adherence to Web-Based and Mobile Technologies for People With Psychosis: Systematic Review of New Potential Predictors of Adherence. JMIR Mhealth Uhealth 2017 Jul 20;5(7):e94-e22 [FREE Full text] [doi: 10.2196/mhealth.7088] [Medline: 28729235]
11. Fiol-DeRoque MA, Serrano-Ripoll MJ, Jiménez R, Zamanillo-Campos R, Yáñez-Juan AM, Bennasar-Veny M, et al. A Mobile Phone-Based Intervention to Reduce Mental Health Problems in Health Care Workers During the COVID-19 Pandemic (PsyCovidApp): Randomized Controlled Trial. JMIR Mhealth Uhealth 2021 May 18;9(5):e27039-e27020 [FREE Full text] [doi: 10.2196/27039] [Medline: 33909587]
12. Lee H. Future of the Smartphone for Patients and Healthcare Providers. Healthc Inform Res 2016 Jan;22(1):1-2 [FREE Full text] [doi: 10.4258/hir.2016.22.1.1] [Medline: 26893944]
13. About HIV and AIDS. 2020. URL: https://www.avert.org/about-hiv-aids [accessed 2020-12-20]
14. HIV Drug Resistance. Avert. 2019. URL: https://www.avert.org/professionals/hiv-programming/treatment/drug-resistance [accessed 2020-12-26]
15. AIDSinfo. URL: https://aidsinfo.unaids.org/ [accessed 2020-12-26]
16. Tchiakpe R, Keke RK, Vidal N, Ahoussinou C, Sekpe O, Dagba HG, et al. Moderate rate of transmitted resistance mutations to antiretrovirals and genetic diversity in newly HIV-1 patients diagnosed in Benin. BMC Res Notes 2020 Jul 02;13(1):314-315 [FREE Full text] [doi: 10.1186/s13104-020-05151-w] [Medline: 32616057]
17. 90-90-90: An Ambitious Treatment Target to Help End the Aids Epidemic. UNAIDS. 2014. URL: https://www.unaids.org/en/resources/909090 [accessed 2020-10-26]

18. Hirsch MS, Brun-Vézinet F, D'Aquila RT, Hammer SM, Johnson VA, Kuritzkes DR, et al. Antiretroviral drug resistance testing in adult HIV-1 infection: recommendations of an International AIDS Society-USA Panel. JAMA 2000 May 10;283(18):2417-2426. [doi: 10.1001/jama.283.18.2417] [Medline: 10815085]
19. Ehrenkranz PD, Baptiste SL, Bygrave H, Ellman T, Doi N, Grimsrud A, et al. The missed potential of CD4 and viral load testing to improve clinical outcomes for people living with HIV in lower-resource settings. PLoS Med 2019 May;16(5):e1002820 [FREE Full text] [doi: 10.1371/journal.pmed.1002820] [Medline: 31141516]
20. ARVPredictor - Apps on Google Play. 2021 Aug 05. URL: https://play.google.com/store/apps/details?id=co.ke.ikocare&hl=en&gl=KE [accessed 2021-01-02]
21. Rhee SY, Gonzales MJ, Kantor R, Betts BJ, Ravela J, Shafer RW. Human immunodeficiency virus reverse transcriptase and protease sequence database. Nucleic Acids Res 2003 Jan 01;31(1):298-303 [FREE Full text] [doi: 10.1093/nar/gkg100] [Medline: 12520007]
22. Stanford Database (HIVdb). URL: https://hivdb.stanford.edu/hivdb [accessed 2020-05-10]
23. Bennett DE, Camacho RJ, Otelea D, Kuritzkes DR, Fleury H, Kiuchi M, et al. Drug resistance mutations for surveillance of transmitted HIV-1 drug-resistance: 2009 update. PLoS One 2009 Mar 6;4(3):e4724 [FREE Full text] [doi: 10.1371/journal.pone.0004724] [Medline: 19266092]
24. Choudhury A. Waterfall Model/SDLC. 2011. URL: http://www.sdlc.ws/waterfall-model/ [accessed 2020-12-26]
25. Beynon-Davies P, Carne C, Mackay H, Tudhope D. Rapid application development (RAD): an empirical review. European Journal of Information Systems 2017 Dec 19;8(3):211-223 [FREE Full text] [doi: 10.1057/palgrave.ejis.3000325]
26. Pawar RP. A comparative study of Agile software development methodology and traditional waterfall model. IOSR Journal of Computer Engineering 2015:1-8 [FREE Full text]
27. Chung I, Ko I. Data-Sharing Method for Multi-Smart Devices at Close Range. Mobile Information Systems 2015;2015:1-11 [FREE Full text] [doi: 10.1155/2015/931765]
28. Paetsch F, Eberlein A, Maurer F. Requirements engineering and agile software development. New York, NY: IEEE; 2003 Presented at: 12th IEEE International Workshops on Enabling Technologies; June 11, 2003; Linz, Austria p. 308-313. [doi: 10.1109/enabl.2003.1231428]
29. Gomaa H, Mason G. Software Modeling and Design: UML, Use Cases, Patterns, and Software Architectures. Cambridge, MA: Cambridge University Press; 2011.
30. Smyth N. Android Studio 4.1 Development Essentials: Java Edition. Cary, NC: Payload Media Inc; 2020.
31. Potts D. Java Programming Language Handbook (First Edition). Scottsdale, AZ: Coriolis Group; Apr 17, 1996.
32. IntelliJ IDEA Handbook. Cambridge, MA: Exelixis Media; 2017.
33. Java Programming Language - An Overview. 2003. URL: https://www.sciencedirect.com/topics/computer-science/java-programming-language [accessed 2020-12-18]
34. Dusch V. Nginx - the webserver you might actually like. 2012 Presented at: PHP Unconference; August 1, 2012; Hamburg, Germany URL: https://www.researchgate.net/publication/231400849_Nginx_-_The_Webserver_you_might_actually_like/link/5a71c5300f7e9ba2e1cc43a4/download
35. DigitalOcean - The developer cloud. 2011. URL: https://www.digitalocean.com/ [accessed 2020-12-18]
36. MySQL. 1995. URL: https://www.mysql.com/ [accessed 2020-12-18]
37. GraphQL (Microservices) Architecture by Apollo. ITNEXT. 2021. URL: https://itnext.io/graphql-microservices-architecture-by-apollo-8b6eb557c5e2 [accessed 2021-06-02]
38. Drohan D. Android Retrofit. 2017. URL: https://square.github.io/retrofit/ [accessed 2020-12-26]
39. McHugh M. Interrater reliability: the kappa statistic. Biochem Med 2012:276-282 [FREE Full text] [doi: 10.11613/bm.2012.031]
40. Macharia VM, Ngugi C, Lihana R, Ngayo MO. Transmitted HIV-1 drug resistance and the role of herpes simplex virus-2 coinfection among fishermen along the shores of Lake Victoria, Kisumu, Kenya. J HIV Retrovirus 2016;02(03):26 [FREE Full text] [doi: 10.21767/2471-9676.100026]
41. ARV_Predictor_App_OngadiBA. GitHub. 2021 Aug 02. URL: https://github.com/bongadi/ARV_Predictor_App_OngadiBA [accessed 2021-01-02]

Abbreviations

HIVdb: HIV Drug Resistance Database
HIVDR: HIV drug resistance
ICT: information and communications technology
IDE: Integrated Development Environment
NAHCC: National Advanced HIV Clinical Course
NCBI: National Center for Biotechnology Information
NNRTI: non-nucleoside reverse transcriptase inhibitors
NRTI: nucleoside reverse transcriptase inhibitor
RAD: rapid application development

UNAIDS: United Nations Programme on HIV/AIDS

Edited by G Eysenbach; submitted 02.01.21; peer-reviewed by R Poluru, E Andrikopoulou; comments to author 10.05.21; revised version received 18.07.21; accepted 27.11.21; published 02.02.22.

Please cite as:
Ongadi B, Lihana R, Kiiru J, Ngayo M, Obiero G
An Android-Based Mobile App (ARVPredictor) for the Detection of HIV Drug-Resistance Mutations and Treatment at the Point of Care: Development Study
JMIR Form Res 2022;6(2):e26891
URL: https://formative.jmir.org/2022/2/e26891
doi:10.2196/26891
PMID:35107425

Original Paper

Quantifying the Quality of Web-Based Health Information on Student Health Center Websites Using a Software Tool: Design and Development Study

Anagha Kulkarni[1], PhD; Mike Wong[1], MSc; Tejasvi Belsare[1], MSc; Risha Shah[1], MSc; Diana Yu Yu[1], BSc; Bera Coskun[1], BSc; Carrie Holschuh[2*], PhD; Venoo Kakar[3*], PhD; Sepideh Modrek[3*], PhD; Anastasia Smirnova[4*], PhD

[1]Department of Computer Science, San Francisco State University, San Francisco, CA, United States

[2]School of Nursing, San Francisco State University, San Francisco, CA, United States

[3]Department of Economics, San Francisco State University, San Francisco, CA, United States

[4]Department of English Language and Literature, San Francisco State University, San Francisco, CA, United States

[*]these authors contributed equally

Corresponding Author:
Anagha Kulkarni, PhD
Department of Computer Science
San Francisco State University
1600 Holloway Ave
San Francisco, CA, 94132
United States
Phone: 1 4153381008
Email: ak@sfsu.edu

Abstract

Background: The internet has become a major source of health information, especially for adolescents and young adults. Unfortunately, inaccurate, incomplete, or outdated health information is widespread on the web. Often adolescents and young adults turn to authoritative websites such as the student health center (SHC) website of the university they attend to obtain reliable health information. Although most on-campus SHC clinics comply with the American College Health Association standards, their websites are not subject to any standards or code of conduct. In the absence of quality standards or guidelines, monitoring and compliance processes do not exist for SHC websites. Thus, there is no oversight of the health information published on SHC websites by any central governing body.

Objective: The aim of this study is to develop, describe, and validate an open-source software that can effectively and efficiently assess the quality of health information on SHC websites in the United States.

Methods: Our cross-functional team designed and developed an open-source software, QMOHI (Quantitative Measures of Online Health Information), that assesses information quality for a specified health topic from all SHC websites belonging to a predetermined list of universities. The tool was designed to compute 8 different quality metrics that quantify various aspects of information quality based on the retrieved text. We conducted and reported results from 3 experiments that assessed the QMOHI tool in terms of its scalability, generalizability in health topics, and robustness to changes in universities' website structure.

Results: Empirical evaluation has shown the QMOHI tool to be highly scalable and substantially more efficient than manually assessing web-based information quality. The tool's runtime was dominated by network-related tasks (98%), whereas the metric computations take <2 seconds. QMOHI demonstrated topical versatility, evaluating SHC website information quality for four disparate and broad health topics (COVID, cancer, long-acting reversible contraceptives, and condoms) and two narrowly focused topics (hormonal intrauterine device and copper intrauterine device). The tool exhibited robustness, correctly measuring information quality despite changes in SHC website structure. QMOHI can support longitudinal studies by being robust to such website changes.

Conclusions: QMOHI allows public health researchers and practitioners to conduct large-scale studies of SHC websites that were previously too time- and cost-intensive. The capability to generalize broadly or focus narrowly allows a wide range of applications of QMOHI, allowing researchers to study both mainstream and underexplored health topics. QMOHI's ability to robustly analyze SHC websites periodically promotes longitudinal investigations and allows QMOHI to be used as a monitoring

tool. QMOHI serves as a launching pad for our future work that aims to develop a broadly applicable public health tool for web-based health information studies with potential applications far beyond SHC websites.

(JMIR Form Res 2022;6(2):e32360) doi:10.2196/32360

KEYWORDS

online health information quality; information quality metrics; automated quantification tool; student health center websites; digital health; health information; health information websites; adolescents; online health; infodemiology; public health; health websites

Introduction

Background

Since the early 1990s, internet has been a major source of health information, and its adoption among health care providers and patients has been growing ever since [1-3]. Health information provided on various internet sites often varies greatly in terms of the quality and reliability of the content [1,4-7]. Common assessments include that the information is too technical or difficult to read, the website is difficult to use (ie, search or navigate), or is unreliable. Assessment instruments have been proposed to help users navigate the high variability in the quality of web-based health information [8-14]. For instance, DISCERN uses questionnaires to help users assess the quality of health information [9]. Similarly, guides for publishing quality health care information have been proposed [10,15,16]. However, the assessment instruments have to be applied manually, typically by field experts, and the implementation of guidelines is not enforced [17,18]. Therefore, the adoption and implementation of the proposed best practices for web-based health information have been limited and nonsystematic.

Adolescents and young adults are particularly vulnerable to the risks arising from inaccurate, incomplete, or outdated web-based health information because they tend to rely heavily on the internet for their information needs [19-24]. Studies have found that adolescents and young adults are savvy internet users who are aware of the problems with the quality of web-based information and thus prefer to use authoritative websites for health information [22]. In a qualitative study with focus groups, usability tests, and in-depth interviews, participants preferred institutional sources of health information over private websites [25]. One such prominent institutional source of health information is the student health center (SHC) websites at higher education institutes (HEIs).

In 2016, approximately 41% of the students aged 18-24 years were enrolled in an HEI with a higher proportion of female attendees than male attendees (43% female attendees vs 38% male attendees) and growing racial and ethnic diversity of the student population, as reported by the National Center for Education Statistics [26]. On the basis of a national study of universities and their SHCs, 85% of the 214 participating higher education institutions in the United States had an SHC website and on-campus clinic in 2015 [27]. SHC websites are commonly perceived as an extension of the SHC clinics and thus are regarded as an authoritative and credible source of health information by adolescents and young adults [27-29]. Rather than physically visiting an SHC clinic on a university campus, most students now make their first contact with an SHC through their website. As such, SHC websites are a leading accessible source of high-quality health information for adolescents and young adults in the United States.

Most on-campus SHC clinics that students visit in person comply with the American College Health Association (ACHA) standards [30]. More than 800 HEIs in the United States have ACHA membership, which provides a healthy campus framework, health and wellness consulting, patient satisfaction assessment service, and national college health assessment to improve overall health status on campus. However, ACHA is limited to on-campus SHC clinics and does not extend its services to SHC websites. As a result, the quality of health information on SHC websites is not monitored by any central governing body.

Objectives

On the basis of these observations, this study aims to develop, describe, and validate an open-source software that can effectively and efficiently assess the quality of health information on SHC websites in the United States. The tool QMOHI (Quantitative Measures of Online Health Information) provides a suite of quantitative measures of information quality, which can be used by health care administrators, researchers, and practitioners to assess, monitor, and improve the information posted on SHC websites.

Methods

QMOHI System Design and Implementation

Overview

A cross-functional team consisting of computer scientists, a software developer, a public health researcher, a nurse practitioner, an economist, and a linguist outlined the framework and capabilities necessary to assess information quality on SHC websites. The team identified exemplars of high-quality SHC websites and then worked with subject matter experts to identify distinct attributes of web-based information that modeled quality, such as topic coverage, accessibility, navigation, readability, sentimentality, and polarity. The team iteratively refined the initial framework and incorporated key measures of quality into the QMOHI software tool. For the development of the QMOHI software, the Agile methodology was adopted to facilitate iterative design, development, testing, and feedback workflow [31]. We evaluated the QMOHI for the following key properties:

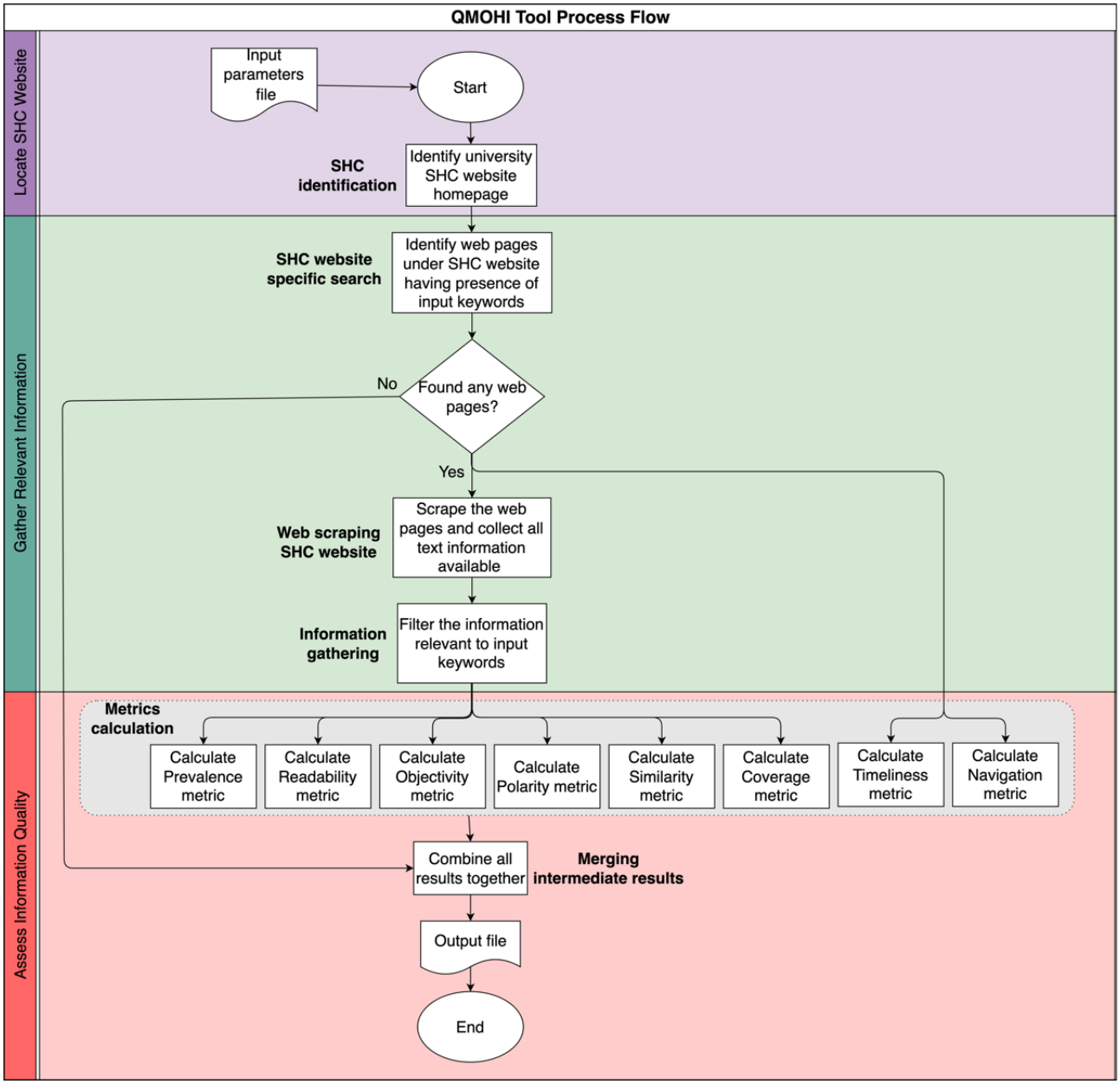
- Scalable—ability to provide quality assessment for a large number of SHC websites efficiently

- Generalizable—ability to conduct a quality assessment of any topic of information on SHC websites
- Robust—ability to be redeployed periodically on SHC websites while adapting to changes in website content and structure
- Fully automated—ability to perform the quality analysis without any human intervention

The QMOHI tool was designed with the assumption that the user would specify two key pieces of input information: (1) the list of universities of interest and (2) the topics of interest, in the form of keywords. These user inputs guided the information gathering and analysis conducted by QMOHI. At a high level, the QMOHI tool was organized into three key phases: phase 1—locate SHC website, phase 2—gather the related information and specific text on the topic of interest from the SHC website, and phase 3—assess the quality of information.

Figure 1 provides a flowchart for the QMOHI tool, in which the 3 phases are delineated using different background colors.

Figure 1. Process flow of the QMOHI (Quantitative Measures of Online Health Information) tool. SHC: student health center.



Phase 1: Locate SHC Website

QMOHI first found the SHC website, more specifically, the web address (URL) of the SHC website for each of the universities specified by the user. We designed and developed an algorithmic approach for this task that consists of four simple steps:

1. Constructing a search query by joining the given university name with the phrase *student health center* (eg, *Texas A & M University Central Texas student health center*).
2. Running the search query using a commercial search engine (eg, Google Custom Search application programming interface).
3. Retrieving URL of the first result; if the URL was not from the *.edu* domain, the next URL was retrieved. This was

repeated until the third URL was processed. If none of the top 3 URLs were from *.edu* domain, it was then concluded that the SHC website could not be found for this university and ended.

4. Sanitizing the retrieved URLs to obtain the definitive URL for the SHC home page by checking whether the URL redirected another URL with the help of the Selenium WebDriver; if yes, then the new URL was used, and the sub-URLs such as*/contacts*, */appointments*, and */location* from the URL were removed.

This multistep approach was necessary because of the large variability found in the web addresses of SHC websites. There are no standards or even commonly accepted conventions for SHC website naming or hosting structures. The following 6 California State universities illustrate the scope of variability among sister HEIs. All six California State universities mentioned here have a different approach for formulating their SHC web address:

1. California Polytechnic State University, San Luis Obispo: hcs.calpoly.edu
2. California State University, Bakersfield: www.csub.edu /healthcenter
3. California State University, Stanislaus: www.csustan.edu /health-center
4. California State University, San Bernardino: www.csusb.edu/student-health-center
5. California State Polytechnic University, Pomona: www.cpp.edu/~health
6. San Francisco State University (SFSU): health.sfsu.edu/

Phase 2: Gather Topical Information

The core task of this phase was to download all the textual information related to the topics of interest from the SHC website identified in the previous phase. To operationalize this process, we used the following approach:

1. Constructing a disjunctive search query from all the topic keywords specified by the user. (Example query: *Corona, coronavirus, COVID; site: health.sfsu.edu*)
2. Using a commercial search engine (eg, Google Custom Search application programming interface) to conduct a site-specific search with the above query against the SHC website. (Site-specific search returns only those webpages that are hosted under the specified site, in our case, the SHC website.)
3. Downloading all webpages in the search result. In addition, the URLs of these webpages were saved. The URLs would be required to compute one of the quality metrics in phase 3.
4. From each webpage, every sentence containing any of the input keywords (anchor sentences) and 5 sentences before and after it (context sentences) were extracted. This step filtered out nonrelevant content by anchoring and localizing the extraction process around the topic keywords.
5. Consolidating all the information extracted in step 4 from all the result webpages of SHC. The duplicate sentences from the consolidated information were filtered.

The data gathered by this approach formed the basis for the analysis conducted in the next phase.

Phase 3: Quantify Information Quality

Overview

The QMOHI tool computed an array of quantitative measures of quality for the gathered information in this phase. Eight quality metrics—readability, coverage, prevalence, objectivity, polarity, navigation, timeliness, and similarity—were implemented in QMOHI. Each quality metric captured a unique aspect of web-based information that was important in the context of health care information dissemination and reflected the multidimensional nature of information quality [32,33]. Every metric was designed and developed such that its computation was completely automated to facilitate large-scale studies. These metrics and the motivations behind them are described as follows.

Metric 1: Readability (Reading Level)

If the information provided on an SHC website used a simple, easy-to-understand language, then it was more likely to be understood and correctly applied. In contrast, if the information on the SHC website used specialized medical terminology, then an average college student would be unlikely to find it accessible. There is an extensive body of research in the context of physician–patient communication that transfers over to web-based health information communication [34-37]. We referred to this concept as *information understandability* and quantified it using the Flesch–Kincaid readability tests [38]. The Flesch–Kincaid readability tests consist of two metrics: Flesch Reading Ease (FRE) and Flesch–Kincaid Grade Level (FKGL), which use linguistic properties of the textual content to score its readability, as follows:

1. Counting the number of syllables, words, and sentences for the consolidated content gathered in the previous phase
2. Computing the FRE metric:

3. Computing the FKGL metric:

A higher score for the FRE metric indicated that the text is easy to read, and a lower score indicated that the material is difficult to read. The scores computed by the FKGL metric corresponded to US grade levels. We applied these metrics to assess the understandability of the information provided on SHC websites.

Metric 2: Prevalence

The volume of relevant information was a crucial aspect of information quality [39]. When relevant information was mentioned in passing and never repeated, it was likely to be overlooked or misunderstood [40]. Therefore, the quantity of relevant information provided on SHC websites was also important. One SHC website may provide just a sentence about the topic of interest, whereas the other may include a detailed post, along with additional reading pointers. The prevalence metric captured this intuition by computing the cumulative

frequency of all input keywords present in the information gathered from the SHC website:

Metric 3: Coverage

Some health care topics required several keywords to be completely expressed. If an SHC website contained more of these keywords, then it provided more in-depth and complete information about the given health topic. Here, we defined our next metric, coverage, to model this intuition as the ratio of the number of keywords found on the SHC website to the total number of input keywords:

The coverage metric ranged from 0 to 100 based on the number of input keywords found, where 0% indicated that none of the input keywords were found on the SHC website and 100% indicated the presence of all input keywords on the SHC website. Although the coverage metric can provide the percentage overlap between input keywords and information on SHC, this metric alone should not be considered as completeness of the information on the health topic. This is because input keywords might be only a subset of all keywords related to a particular health topic. As such, the utility of both prevalence and coverage metrics depended on the comprehensiveness of the input keywords for a specific health topic.

Metric 4: Sentiment—Objectivity

High-quality health information is high in factual information content and low in unsupported opinions. A measure of these 2 directly opposing qualities can be expressed as *objectivity* and *subjectivity*, respectively. Objectivity is an information quality metric that quantifies the extent to which information is impartial [40]. TextBlob [41] provided sentiment analysis, including subjectivity scoring algorithms based on a curated weighted lexicon approach. The subjectivity scores were bounded in 0 and 1, where 1 is the most subjective and 0 is the most objective. In this work, we computed the subjectivity score of the information gathered in the previous phase of QMOHI using TextBlob, and then defined the objectivity metric as 1 (*subjectivity*).

Metric 5: Sentiment—Polarity

Along with the objectivity measure, polarity is important for assessing the quality of the information on the SHC website. The same information about the evidence on health effects can be framed either positively or negatively [42], for example, “This disease can be difficult to cure entirely if detected in later stages” and “This disease can be easy to cure entirely if detected in early stages.” Both sentences express similar meanings, but their polarities are contrary. Critically, different positive and negative framing can shift people’s preferences, even if the options are objectively equivalent [43]. Polarity of the health information on the SHC website may affect people’s decisions about health services.

The polarity metric quantified the positivity, negativity, or neutrality of the health information on the SHC website. For this tool, the polarity score was computed using TextBlob’s sentiment analysis feature [41] on the health information collected from the SHC website. This score ranged between –1 and 1, where 1 indicated a strongly positive statement, –1 indicated a strongly negative statement, and 0 indicated a neutral statement, for example,

1. “They have the best available doctors, equipment and treatment facilities.” This sentence shows affirmation. It has a polarity score of 0.7.
2. “If the cancer is located only in the breast, the 5-year relative survival rate of people with breast cancer is 99%.” This sentence is neutral; it has a polarity score of 0.
3. “The service of health center AAA is atrocious for XYZ.” This sentence shows negative expressions, it has a polarity score of –0.39.

Metric 6: Navigation (Number of Clicks)

Well-designed websites make it easy for users to find the information they need, minimizing the demand for users’ time and effort. This intuition was modeled by the navigation metric that computed the minimum number of clicks needed to reach the desired content when starting from the SHC home page. At a high level, the algorithm for computing this metric was designed to simulate the website traversal path a human would follow when looking for specific information on SHC websites. To find the content closest to the SHC home page (minimum number of clicks), this exploration was conducted in a breadth-first search. To operationalize this logic, a customized tree data structure with a queue was used to prioritize the webpages (URLs) that had to be checked iteratively.

As shown in Textbox 1, the expected input by the navigation algorithm is the SHC home page URL and the URLs for webpages retrieved by site-specific search in phase 2. The first node to be created in the tree data structure was for the SHC home page (line 1) and was added to the queue (line 2). At each iteration, a node at the head of the queue was obtained (line 4). The program was terminated if the level of the current node was >10 (line 5).

It was assumed that keywords’ content was not present on this SHC website, and a special value of –1 and empty trace was returned for the navigation metric to indicate the same (line 6). If the current node’s URL matched with any of the target webpage URLs (line 8), then the program ended by returning the level (number of clicks) and trace of the current node (line 9). When a current node’s URL did not match any of the target URLs, all hyperlinks on the current page were extracted (line 11). The hyperlinks that were external to the SHC web domain were filtered out. For the remaining hyperlinks, a new tree node was created that was attached to the current node as a child node (lines 13-19). This process was repeated until the queue was empty (line 3) or until either of the other 2 stopping criteria were met (lines 6 or 9).

Textbox 1. Algorithm for the navigation metric.

Input: (1) SHC home page URL and (2) Target pages: URLs for webpages retrieved by site-specific search in phase 2

Output: (1) Minimum number of clicks (Navigation metric) and (2) Trace (An ordered list of URLs—path from SHC home page to closest target page)

1. Initialize: Tree data structure with one node (root) containing:
   URL: SHC home page URL
   level: 0
   trace: SHC home page URL
2. Add root node to the queue
3. while queue is not empty do
4. Pop the node at the head of queue
5. if level of current node >10
6. return –1 and empty trace
7. end if
8. if current node's URL matches with any of the target page URLs
9. return level and trace of the current node
10. else
11. Extract all the hyperlinks from the contents of the node's URL
12. Filter out the hyperlinks that are outside of SHC web domain
13. For each hyperlink *h*
14. Create a new child node where:
15. URL: *h*,
16. level: parentNode.level+1,
17. trace: append (*h* to parentNode.trace)
18. Add the new node to the queue
19. End for
20. end if
21. end while

Metric 7: Timeliness

Health care information is dynamic in which new or improved treatments are brought to the market, sometimes replacing existing treatments, or relegating them to be used only under specific conditions. SHC websites should be regularly checked and revised to stay current with the latest health information, removing deprecated information, and reorganizing existing information to reflect critical health care priorities, such as vaccine availability during a pandemic. Outdated information, without the advice of a trained health care provider, can lead to suboptimal decisions. Therefore, the timeliness of information is an important aspect of information quality. Webpages on a particular university's SHC website may be modified at different times, and certain webpages may be updated more often than others. It is important to know when the information was last updated on the SHC webpage from which the information of a certain health topic is referred.

The timeliness metric quantified this insight through the *last revised timestamps* on SHC webpages that contain the input keywords. These timestamps were fetched from the webpage headers with the *Last-Modified* tag, and if absent, they were marked as –1. For webpages with a –1 timeliness metric score, the recency of content could not be determined. With more recent timestamps, the probability of the latest information increased.

Metric 8: Relevancy or Similarity

Relevancy describes the extent to which information is applicable and helpful to users' information needs [40]. Relevancy of health information on the SHC website is contextual and subjective; as such, it is difficult to assess directly. We can approximate relevancy by calculating the lexical similarity between the information on the SHC website and an *ideal* reference document, which is a document, manually created by experts, containing all the information relevant to the health topic of interest (perhaps using Centers for Disease Control and Prevention references, for example). To operationalize this intuition, we used a cosine similarity function, which is defined as follows:

where  is a numeric vector representation of the collated SHC website content gathered in phase 2 and  is a numeric vector representation of the ideal reference document. Similarity values closer to 1 indicated that the relevance of the topical information on the SHC website is high, whereas values closer to 0 implied low relevance.

Experimental Setup

Overview

The following set of experiments provide an empirical evaluation of the QMOHI tool on three key performance metrics:

1. Scalability—measured by timed trials versus human annotators to navigate an SHC website with a specific information goal and performance benchmarking trails
2. Generalizability—evaluated by comparing results with varying information specificity and looking for poor performance
3. Robustness—evaluated by computing quality metrics over time as SHC websites change in both content and structure

Experiment 1: Scalability and Efficacy

Overview

The first experiment investigated the scalability of the QMOHI tool using two methods: (1) by comparing the time needed by human annotators to find topically relevant information on the SHC website to that by QMOHI and (2) through performance benchmarking of the QMOHI tool by measuring its end-to-end runtime and studying the breakdown of the runtime.

Method 1

For the first method, 200 universities were chosen at random from a larger set of all 4-year, public, bachelor's granting universities in the United States (N=549). The list of 200 universities was shuffled and partitioned into 20 equal groups to allow for timing the task at the group level rather than at the university level to smooth out any individual university-level idiosyncrasies. Two annotators conducted the task on all 20 groups for the health topic of long-acting reversible contraceptives (LARC), which was represented by the following keywords: *IUD*, *intrauterine device*, *IUI*, *intrauterine implant*, *contraceptive implant*, *contraceptive shot*, *contraceptive injection*, *Depo Provera*, *and Depo-provera*. The annotators were instructed to perform the following steps:

1. Find the SHC website of the university with Google search.
2. Calculate the minimum number of clicks needed to reach the first mention of any of the given keywords from the SHC home page. The starting point is the SHC website's home page, with the number of clicks as 0.
3. Indicate *No mention* if none of the keywords were found on the SHC website.
4. Record the time required to perform the whole task on every group of 10 universities.

The task of finding the first topically relevant webpage on the SHC website could be considered equivalent to computing the navigation metric using the QMOHI tool. Thus, the time required by QMOHI to compute the navigation metric was compared with the annotation time.

Method 2

The authors also conducted performance benchmarking for the QMOHI tool by measuring its end-to-end runtime. Specifically, 20 (10%) universities, selected at random from the subset of 200 universities known to have SHC websites, were searched using the QMOHI tool for 2 health topics (topic 1: *pap smear* and topic 2: *all contraception*) on cloud servers. The *pap smear* topic mapped to 6 keywords query, each a variation of *pap smear* and *pap test*. The *all contraception* topic was represented using 37 keywords, including the following: *birth control*, *contraceptive implant*, *hormonal IUD*, and others. The runtime of the tool for every university and every topic was recorded. In addition, the time spent by the tool gathering the information (network time) versus processing the information (compute time) was recorded to facilitate a thorough performance analysis.

Cloud servers provide accessible and reliable networks and dedicated infrastructure, as opposed to local student laptops or university infrastructure, which may be multipurpose or have unreliable networks. Aside from the operating system itself, the cloud server was set up to exclusively run QMOHI during benchmarking. The cloud server used was an Amazon EC2 *t2.large* instance, featuring dual central processing units with 8 GB memory, and the network throughput was profiled using *iperf* from an EC2 instance in Virginia to a local *iperf* server in California and measured an average 51.1 MiB/s over 3 connections. As the universities were all within the United States, transcontinental communications approached the upper bounds of network traversal.

Experiment 2: Generalizability

The second experiment examined the QMOHI tool's ability to compute information quality metrics for a wide range of health topics. Specifically, the quality of information for the 4 health topics—COVID, cancer, LARC, and condoms—on the SHC website of SFSU was evaluated for this experiment.

The other part of this experiment tested QMOHI's ability to work with narrowly focused health topics. The relevant information for such topics can be sparse and spread on SHC websites. Whether QMOHI can tackle these data challenges was examined by this experiment with the following two fine-grained health topics: hormonal intrauterine device (IUD) and copper IUD (Paragard), which are searched for on SHC website of the SFSU.

The set of input keywords used with QMOHI for each of the above health topics is given in the Multimedia Appendix 1.

Experiment 3: Robustness

Robustness is the ability of a software system to recover from errors, such as unexpected changes in input. Public health studies are often longitudinal, and data collection and analysis must be conducted periodically over a longer period of time. During this time, SHC websites might change both webpage content and structure (ie, file names, directories, and even complete URL changes). QMOHI can analyze website content regardless of

changes in the website structure. To evaluate the ability of the QMOHI tool to extract content from moving targets, a longitudinal study was conducted on multiple health topics over a period of 3 months for a large set of universities. Specifically, the QMOHI tool was run on July 14, August 14, and September 14, 2020, on the SHC websites of 549 public universities in the United States for the following five health topics: pap smear, condoms, LARC, all forms of contraception, and superset of all above keywords.

Universities, topics, and keywords were kept consistent in all 3 executions of the QMOHI tool over 3 months.

Results

Experiment 1: Scalability

Method 1

The results for method 1 (navigation task) are listed in Table 1. The task was to find the first topically relevant information webpage on the SHC website. In the fastest scenario, QMOHI completed the task for 10 universities in 1 minute, which is an order of magnitude faster than the manual approach. However, the small difference between the maximum task times for the 2 approaches (approximately 41 minutes vs 34 minutes) was puzzling. To understand the underlying reason, Figure 2 provides zoomed-in data: group-level task times for each of the 20 groups. These data reveal two outliers: groups 6 and 12. QMOHI's task times for these 2 groups were exceptionally high compared with the other groups.

The unresponsiveness of SHC websites for one of the universities in each of the 2 groups was detected to be the root cause behind this disparity. For most universities (179/200, 89.5%), QMOHI's task time was less than a minute. For fewer universities (16/200,8.5%), the task time was under 2 minutes, and a handful of universities (3/200,1.5%) required 6 minutes or less. However, for 2 universities, the task times were 25 minutes and 32 minutes because of the unresponsiveness of the SHC websites.

We isolated the 2 outliers and compared the manual approach with QMOHI for the remaining 99% (198/200) universities. The average task time per university for the manual approach was 2.52 (SD 0.67) minutes and for QMOHI, it was 0.32 (SD 0.18) minutes. The QMOHI tool was more than 7 times faster than human annotators at the task of finding the first webpage with relevant information on the SHC website.

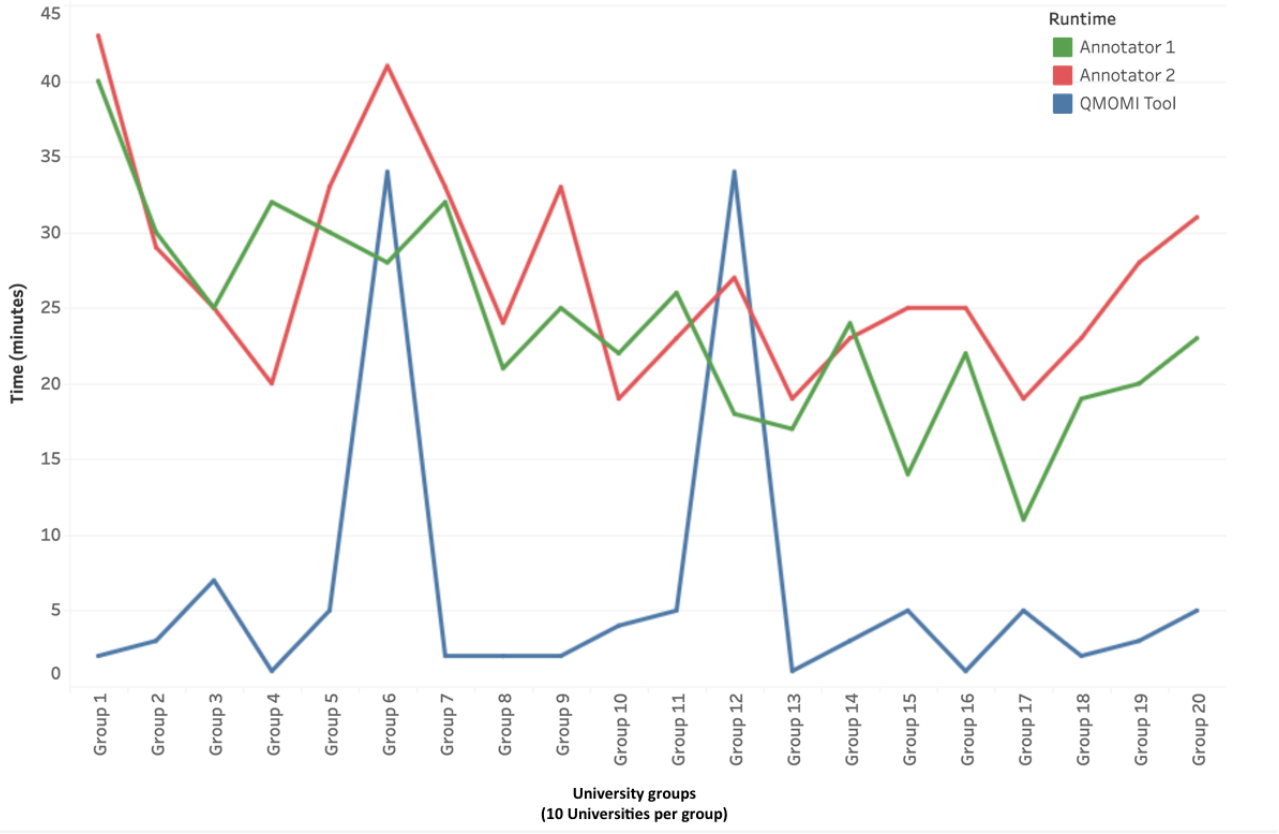
It is worth noting that this experiment studied the simplest of the metrics for both the tool and human annotators. Other quality metrics, such as readability and prevalence, which are more difficult for humans to assess, would likely increase the time for manual approach substantially. Empirical benchmarking of these task times will be a part of future work.

Table 1. Experiment 1—scalability. Aggregate group-level task times.

Method	Time (minutes)			
	Minimum time per group	Maximum time per group	Average time per group (SD)	Total time for all 20 groups
Annotator 1	11	40	23.95 (6.85)	479
Annotator 2	19	43	27.15 (6.85)	543
QMOHI[a] tool	1	34	6.30 (9.62)	126

[a]QMOHI: Quantitative Measures of Online Health Information.

Figure 2. Scalability experiment: Runtime comparison chart for navigation metric. Group-level task times, with 10 universities per group. QMOHI: Quantitative Measures of Online Health Information.



Method 2

Benchmarking revealed that QMOHI's mean end-to-end runtime per university was 77.06 (SD 97.65) seconds for topic 1 (*pap smear;* number of hits=11) and 114.06 (SD 138.85) seconds for topic 2 (*all contraception;* number of hits=13). No relevant content was found in 9 sites for topic 1 and in 7 sites for topic 2.

The runtime of the tool was dominated by network-related tasks (ie, retrieving webpages). For topic 1, the network time accounted for 98.33% (75.78/77.06 seconds) of the total runtime. For topic 2, the network time accounted for 98.23% (112.03/114.06 seconds) of the total runtime. The tool's processing time accounted only for 1.67% (1.29/77.06 seconds) and 1.77% (2.02/114.06 seconds) for topic 1 and topic 2, respectively.

The network times were less interesting to compare, as a human annotator would also experience similar latency retrieving the pages using their browser. However, the quality metric computation was consistently performed in a few seconds by the QMOHI tool, with only approximately 1 second slower performance for queries with 6-fold more keywords. This was in contrast to human annotation, which required a few minutes to read the content, and many more to perform the quality assessments manually. Overall, these results showed that the QMOHI tool is highly scalable and substantially more efficient than the manual approach.

Experiment 2: Generalizability

Table 2 provides 6 quality metrics computed by QMOHI for information posted on the SHC website at SFSU for 4 distinct health topics (COVID, Cancer, LARC, and Condoms) and 2 closely related health topics (Hormonal IUD and Copper IUD). These results illustrated QMOHI tool's versatility in terms of being applicable to any given topic as long as it was represented as a set of keywords. As such, the QMOHI tool could be used to study the information quality of a wide variety of topics.

Some of the observations from the metric values are as follows: navigation metric value of 0 for COVID aligns with the current trend on public health websites, which is to post a message related to COVID on the home page. The higher coverage of the *Condoms* topic compared with the other topics is expected as information dissemination on condoms is one of the focus areas for most SHC websites.

If the aforementioned results showcase the *breadth* ability of QMOHI, then the results in Table 2, group B demonstrate the *depth* ability of the tool. Table 2, group B provides information quality metrics for two closely related contraception methods: hormonal IUD and copper IUD (Paragard).

Overall, these results suggest that the QMOHI tool is capable of generating information quality metrics for any given topic. Users can customize the input keywords to the QMOHI tool for a particular topic of any granularity, making it a generic tool with broad applicability.

Table 2. Experiment 2—generalizability. Results showing the QMOHI (Quantitative Measures of Online Health Information) tool's ability to compute information quality metrics for 4 diverse health topics (group A) and closely related health topics (group B).

Health topic	Readability (Flesch–Kincaid)		Navigation (number of clicks from home page)	Coverage (0–100)	Objectivity (0.0–1.0)	Polarity (-1.0 to 1.0)
	Reading ease score (0–100)	Grade level (K–12)				
Group A: 4 diverse health topics						
COVID	77.61	5.18	0	37.50	0.639	0.091
Cancer	76.06	5.81	1	11.11	0.542	0.182
LARC[a]	73.27	7.30	1	33.33	0.496	0.183
Condoms	75.42	6.12	1	100.00	0.373	0.111
Group B: 2 closely related health topics						
Hormonal IUD[b]	73.19	7.98	1	42.86	0.486	0.161
Copper IUD	70.51	8.14	1	25.00	0.486	0.161

[a]LARC: long-acting reversible contraceptives.

[b]IUD: intrauterine device.

Experiment 3: Robustness

Figure 3 illustrates the robustness of results in terms of the correlation between the metric values across the 3 reruns of QMOHI for the 5 health topics. For every metric, the pairwise correlation for the 3 time points (July, August, and September 2020) was computed.

As shown in Figure 3, most of the correlation values were close to 1, indicating high fidelity in the reproduction of the results across multiple time points. The absence of perfect correlation scores suggested that the metric values were time-varying because of the dynamic nature of information on the internet. Our analysis revealed two types of changes that had happened to the SHC websites between the tool's reruns: (1) the content of the website had been updated and (2) the directory structure of the SHC website itself had changed. Table 3 provides a few examples of the second type of change.

Figure 3. Experiment 3—robustness of QMOHI (Quantitative Measures of Online Health Information) tool. Pairwise correlation scores between metric values across the 3 monthly reruns of QMOHI (from July 2020 until September 2020) for the 5 health topics. LARC: long-acting reversible contraceptive.

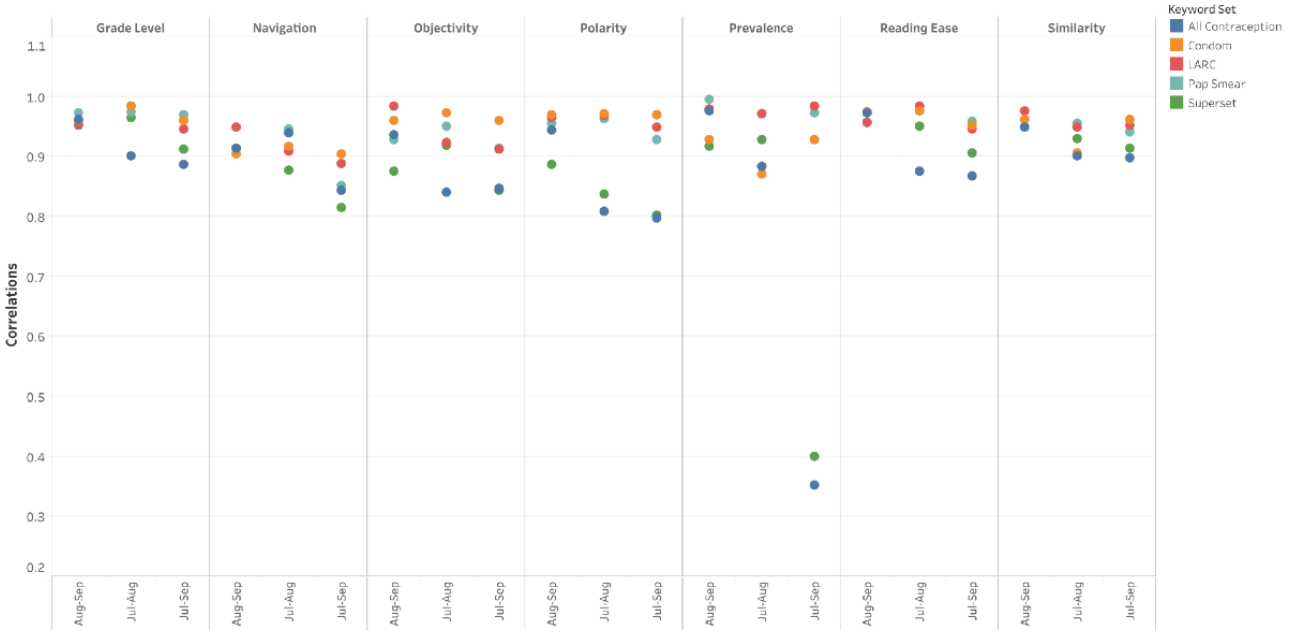


Table 3. Experiment 3—robustness. Examples of QMOHI's (Quantitative Measures of Online Health Information) ability to adapt to changing university student health center (SHC) website structure over 3 reruns.

University name	Old SHC website structure	New SHC website structure
University of Maryland Baltimore County	/studenthealth/services—hours/student-health-center/	/studenthealth/student-health-center
West Chester University of Pennsylvania	/_services/stu.inf/	/_services/studentHealthServices
Francis Marion University	/studenthealthservices/	/studentservices
Concord University	/student-life/node/35	/studenthealth
Coastal Carolina University	/services/studenthealth/	/health

Discussion

Principal Findings

In this study, we described a new open-source software tool, QMOHI, which has been designed and developed to quantify the quality of health information available on SHC websites. We then conducted an empirical evaluation of the QMOHI tool along three key performance metrics: scalability, generalizability, and robustness.

In our first evaluation, we showed that the navigation capabilities of QMOHI are at least seven times more efficient than the manual approach in determining web-based information quality. The runtime of the tool was dominated by network-related tasks. Once the relevant webpages are found, the processing times for computing the quality metrics are trivial. In contrast, human annotators would likely spend most of their time ascertaining information quality. In the second evaluation, we used a tool to retrieve quality metrics on broad and narrow health topics. We showed that once the user selects appropriate keywords, the tool can be adapted to any health topic, thereby establishing the generalizability and versatility of the tool. In the final evaluation, we redeployed QMOHI across 3 periods and showed that the tool is not vulnerable to typical structural changes to SHC websites, thereby allowing users to conduct longitudinal studies.

Limitations

Currently one of the main limitations of QMOHI is its reliance on the keywords provided by users for the health topic of interest. The data gathered by the tool are entirely dependent on these keywords. The ability of the keywords to represent the health topic accurately and completely directly affects the accuracy of the information quality metrics provided by QMOHI. The keywords can also become a source of bias and thus influence the outcomes and conclusions drawn from studies in unexpected ways. One of the future directions of this work will explore automated keyword-generation approaches that require only the name of the health topic from the user and thus remove the dependence on user-provided keywords.

The data-gathering phase of QMOHI currently only collates textual information containing the keywords. This limits the information *visible* to the tool as relevant information is sometimes embedded in images and pdfs. To overcome this limitation, we plan to leverage recent advancements in computer vision to extract text from images and scanned documents.

The QMOHI project's codebase can be downloaded and installed by following step-by-step instructions on the project webpage. In the future, we seek to take this a step further by providing a *plug-and-play* setup where minimal installation is needed. For this, we leverage the virtualization frameworks (eg, Docker) that are being increasingly adopted to lower the barriers for users with any background.

The applicability of QMOHI is currently restricted to the SHC websites of universities. This narrow focus was beneficial in terms of providing guardrails during the first cycle of project development. However, our goal is to lift this restriction and allow other web-based health information dissemination platforms to also use the quality assessment provided by QMOHI.

Comparison With Previous Work

Health information quality assessment is an active field of research [9,11-13]. Nearly all existing approaches, including DARTS [11], study by Dobbins et al [12], DISCERN [9], and Lida [13], use surveys crafted by experts as the central tool for information quality assessment. These approaches can produce high-quality assessments, but are costly, time-consuming, and prone to human errors. QMOHI automates quality assessments by using natural language processing techniques in lieu of survey takers.

Table 4 shows how QMOHI fits in a sampling of the ecosystem of health information quality assessment tools. For a fair comparison, we combine QMOHI's prevalence and coverage metrics as part of relevancy and QMOHI's sentiment and polarity as part of reliability. DARTS, study by Dobbins et al [12], and DISCERN start with the assumption that the user has found a webpage of health information relevant to their interests. QMOHI and Lida start with the assumption that the user has (1) a specific health information need and (2) access to the internet. Lida does a superb job for assessing web usability, far more extensively than QMOHI (which assesses navigability only) and has a battery of automated tests to achieve those goals. Lida then relies on manual surveying to conduct information quality assessments. Many of these tools lack readability assessment and are used in conjunction with an external Flesch reading level analyzer [4-6]. QMOHI offers integrated Flesch-readability metrics. DARTS (Finland), study by Dobbins et al (Canada) [12], DISCERN (United Kingdom), and Lida (United Kingdom) were all developed outside of the United States; these tools rely on human survey takers, and are compatible with content in any language, provided that the survey is accessible to the survey takers. For example, DARTS

specifically accommodates health care information in Finnish and English [44]. QMOHI focuses on university SHC websites in the United States. We believe that QMOHI offers a well-balanced and larger feature set than the existing tools. An empirical comparative analysis with tools such as AutoDISCERN [14,45] is part of future work.

Table 4. Comparison of information quality assessment tools.

Tool	Is fully automated	Is freely available	Assessments and metrics				
			Web usability	Readability	Reliability	Timeliness	Relevancy
DARTS[a]		✓			✓	✓	
Dobbins et al [12]		✓			✓	✓	✓
DISCERN		✓			✓	✓	✓
Lida			✓		✓	✓	✓
QMOHI[b]	✓	✓	✓	✓	✓	✓	✓

[a]Mnemonic for Date, Author, References, Type, Sponsors.

[b]QMOHI: Quantitative Measures of Online Health Information.

Conclusions

This work introduced a new tool for public health research, QMOHI, that facilitates the scale monitoring of the quality of web-based information available on university SHC websites. QMOHI provides a suite of 8 metrics that quantify different aspects of information quality. Longitudinal studies that require periodic reexamination of the same information can be effectively facilitated by QMOHI. Such a tool can assist college health administrators in monitoring the recency and relevancy of the information provided on the SHC website. QMOHI can also be instrumental for centrally operated bodies, such as the ACHA, to help with the evaluation and standardization of health information on SHC websites of universities across the country. Overall, QMOHI is a powerful tool that can accelerate public health research based on web-based health information. QMOHI is an open-source project that is publicly available for nonprofit use [46].

Conflicts of Interest

None declared.

Multimedia Appendix 1
Input keywords used with QMOHI (Quantitative Measures of Online Health Information) tool for the various health topics used in this study.
[DOCX File , 14 KB - formative_v6i2e32360_app1.docx]

References

1. Millman A, Lee N, Kealy K. ABC of medical computing. The internet. BMJ 1995 Aug 12;311(7002):440-443 [FREE Full text] [doi: 10.1136/bmj.311.7002.440] [Medline: 7503911]
2. Cline R, Haynes K. Consumer health information seeking on the internet: the state of the art. Health Educ Res 2001 Dec;16(6):671-692. [doi: 10.1093/her/16.6.671] [Medline: 11780707]
3. Lee K, Hoti K, Hughes JD, Emmerton L. Dr Google and the consumer: a qualitative study exploring the navigational needs and online health information-seeking behaviors of consumers with chronic health conditions. J Med Internet Res 2014 Dec 02;16(12):e262 [FREE Full text] [doi: 10.2196/jmir.3706] [Medline: 25470306]
4. Grewal P, Williams B, Alagaratnam S, Neffendorf J, Soobrah R. Quality of vascular surgery web sites on the internet. J Vasc Surg 2012 Nov;56(5):1461-1467 [FREE Full text] [doi: 10.1016/j.jvs.2012.04.058] [Medline: 22801107]
5. Sabharwal S, Badarudeen S, Unes Kunju S. Readability of online patient education materials from the AAOS web site. Clin Orthop Relat Res 2008 May;466(5):1245-1250 [FREE Full text] [doi: 10.1007/s11999-008-0193-8] [Medline: 18324452]
6. Eltorai AE, Ghanian S, Adams CA, Born CT, Daniels AH. Readability of patient education materials on the American association for surgery of trauma website. Arch Trauma Res 2014 Jun;3(2):e18161 [FREE Full text] [doi: 10.5812/atr.18161] [Medline: 25147778]
7. Al-Bahrani A, Plusa S. The quality of patient-orientated internet information on colorectal cancer. Colorectal Dis 2004 Sep;6(5):323-326. [doi: 10.1111/j.1463-1318.2004.00604.x] [Medline: 15335363]
8. 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;34(4):231-243. [doi: 10.3109/17538150903359030] [Medline: 19919300]

9. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. J Epidemiol Community Health 1999 Feb;53(2):105-111 [FREE Full text] [doi: 10.1136/jech.53.2.105] [Medline: 10396471]
10. Risk A, Dzenowagis J. Review of internet health information quality initiatives. J Med Internet Res 2001 Dec 26;3(4):E28 [FREE Full text] [doi: 10.2196/jmir.3.4.e28] [Medline: 11772543]
11. Närhi U, Pohjanoksa-Mäntylä M, Karjalainen A, Saari JK, Wahlroos H, Airaksinen MS, et al. The DARTS tool for assessing online medicines information. Pharm World Sci 2008 Dec;30(6):898-906. [doi: 10.1007/s11096-008-9249-9] [Medline: 18791806]
12. Dobbins M, Watson S, Read K, Graham K, Yousefi Nooraie R, Levinson AJ. A tool that assesses the evidence, transparency, and usability of online health information: development and reliability assessment. JMIR Aging 2018 May 07;1(1):e3 [FREE Full text] [doi: 10.2196/aging.9216] [Medline: 31518240]
13. Category Archives: Lida. Minervation. URL: https://www.minervation.com/category/lida/ [accessed 2022-01-23]
14. Al-Jefri M, Evans R, Lee J, Ghezzi P. Automatic identification of information quality metrics in health news stories. Front Public Health 2020;8:515347 [FREE Full text] [doi: 10.3389/fpubh.2020.515347] [Medline: 33392124]
15. Boyer C, Selby M, Scherrer J, Appel R. The health on the net code of conduct for medical and health websites. Comput Biol Med 1998 Sep;28(5):603-610. [doi: 10.1016/s0010-4825(98)00037-7] [Medline: 9861515]
16. Eysenbach G, Diepgen TL. Towards quality management of medical information on the internet: evaluation, labelling, and filtering of information. BMJ 1998 Nov 28;317(7171):1496-1500 [FREE Full text] [doi: 10.1136/bmj.317.7171.1496] [Medline: 9831581]
17. Cisu TI, Mingin GC, Baskin LS. An evaluation of the readability, quality, and accuracy of online health information regarding the treatment of hypospadias. J Pediatr Urol 2019 Feb;15(1):40-46. [doi: 10.1016/j.jpurol.2018.08.020] [Medline: 30449679]
18. Bompastore NJ, Cisu T, Holoch P. Separating the wheat from the chaff: an evaluation of readability, quality, and accuracy of online health information for treatment of peyronie disease. Urology 2018 Aug;118:59-64. [doi: 10.1016/j.urology.2018.02.054] [Medline: 29723588]
19. Escoffery C, Miner KR, Adame DD, Butler S, McCormick L, Mendell E. Internet use for health information among college students. J Am Coll Health 2005;53(4):183-188. [doi: 10.3200/JACH.53.4.183-188] [Medline: 15663067]
20. Castrén J, Huttunen T, Kunttu K. Users and non-users of web-based health advice service among Finnish university students - chronic conditions and self-reported health status (a cross-sectional study). BMC Med Inform Decis Mak 2008 Jan 31;8:8 [FREE Full text] [doi: 10.1186/1472-6947-8-8] [Medline: 18237414]
21. Hanauer D, Dibble E, Fortin J, Col NF. Internet use among community college students: implications in designing healthcare interventions. J Am Coll Health 2004;52(5):197-202. [doi: 10.3200/JACH.52.5.197-202] [Medline: 15029941]
22. Buhi ER, Daley EM, Oberne A, Smith SA, Schneider T, Fuhrmann HJ. Quality and accuracy of sexual health information web sites visited by young people. J Adolesc Health 2010 Aug;47(2):206-208. [doi: 10.1016/j.jadohealth.2010.01.002] [Medline: 20638014]
23. Harris K, Byrd K, Engel M, Weeks K, Ahlers-Schmidt CR. Internet-based information on long-acting reversible contraception for adolescents. J Prim Care Community Health 2016 Apr;7(2):76-80 [FREE Full text] [doi: 10.1177/2150131915621058] [Medline: 26643120]
24. Walsh-Buhi ER, Helmy HL. Trends in long-acting reversible contraceptive (LARC) use, LARC use predictors, and dual-method use among a national sample of college women. J Am Coll Health 2018;66(4):225-236. [doi: 10.1080/07448481.2017.1399397] [Medline: 29111911]
25. Eysenbach G, Köhler C. How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. BMJ 2002 Mar 09;324(7337):573-577 [FREE Full text] [doi: 10.1136/bmj.324.7337.573] [Medline: 11884321]
26. Indicator 19: College Participation Rates. National Center for Education Statistics. URL: https://nces.ed.gov/programs/raceindicators/indicator_rea.asp [accessed 2022-01-23]
27. Hicks RW, Dinkel S. The hidden gems in student health centers. J Am Assoc Nurse Pract 2016 Apr;28(4):178-180. [doi: 10.1002/2327-6924.12364] [Medline: 27098401]
28. Judson K, Goldsack J, Sonnad SS. Availability of e-information for women's health services: assessment of California State University Student Health Centers. J Womens Health (Larchmt) 2010 Dec;19(12):2219-2225. [doi: 10.1089/jwh.2010.1936] [Medline: 20964492]
29. Stanley SJ, Yan K, Jiao J, Lutovsky BR, Aubrey JS, Pitts MJ. Communicating about sex when it matters: a content analytic investigation of sexual health information on college student health center websites. J Applied Commun Res 2019 Oct 10;47(5):591-610. [doi: 10.1080/00909882.2019.1675895]
30. About ACHA. American College Health Association. URL: https://www.acha.org/ACHA/About/About_ACHA/ACHA/About/About_ACHA.aspx [accessed 2022-01-23]
31. Highsmith J, Cockburn A. Agile software development: the business of innovation. Computer 2001 Sep;34(9):120-127. [doi: 10.1109/2.947100]

32. Fadahunsi KP, Akinlua JT, O'Connor S, Wark PA, Gallagher J, Carroll C, et al. Protocol for a systematic review and qualitative synthesis of information quality frameworks in eHealth. BMJ Open 2019 Mar 05;9(3):e024722 [FREE Full text] [doi: 10.1136/bmjopen-2018-024722] [Medline: 30842114]
33. Knight S, Burn J. Developing a framework for assessing information quality on the world wide web. InSITE Informing Sci IT Educ 2005;5 [FREE Full text] [doi: 10.28945/2854]
34. Hadlow J, Pitts M. The understanding of common health terms by doctors, nurses and patients. Soc Sci Med 1991;32(2):193-196. [doi: 10.1016/0277-9536(91)90059-I] [Medline: 2014414]
35. Koch-Weser S, Dejong W, Rudd R. Medical word use in clinical encounters. Health Expect 2009 Dec;12(4):371-382 [FREE Full text] [doi: 10.1111/j.1369-7625.2009.00555.x] [Medline: 19709316]
36. Thompson CL, Pledger LM. Doctor-patient communication: is patient knowledge of medical terminology improving? Health Commun 1993 Apr;5(2):89-97. [doi: 10.1207/s15327027hc0502_2]
37. Fields AM, Freiberg CS, Fickenscher A, Shelley KH. Patients and jargon: are we speaking the same language? J Clin Anesth 2008 Aug;20(5):343-346. [doi: 10.1016/j.jclinane.2008.02.006] [Medline: 18761241]
38. Flesch R. A new readability yardstick. J Appl Psychol 1948 Jun;32(3):221-233. [doi: 10.1037/h0057532] [Medline: 18867058]
39. Foley O, Helfert M. Information quality and accessibility. In: Sobh T, editor. Innovations and Advances in Computer Sciences and Engineering. Dordrecht, Netherlands: Springer; 2010.
40. Wang RY, Strong DM. Beyond accuracy: what data quality means to data consumers. J Manag Inf Syst 2015 Dec 11;12(4):5-33. [doi: 10.1080/07421222.1996.11518099]
41. TextBlob: simplified text processing. TextBlob. URL: https://textblob.readthedocs.io/en/dev/ [accessed 2022-01-23]
42. Akl E, Oxman A, Herrin J, Vist G, Terrenato I, Sperati F, et al. Framing of health information messages. Cochrane Database Syst Rev 2011 Dec 07(12):CD006777. [doi: 10.1002/14651858.CD006777.pub2] [Medline: 22161408]
43. Tversky A, Kahneman D. The framing of decisions and the psychology of choice. Science 1981 Jan 30;211(4481):453-458. [doi: 10.1126/science.7455683] [Medline: 7455683]
44. Prusti M, Lehtineva S, Pohjanoksa-Mäntylä M, Bell JS. The quality of online antidepressant drug information: an evaluation of English and Finnish language Web sites. Res Social Adm Pharm 2012;8(3):263-268. [doi: 10.1016/j.sapharm.2011.03.003] [Medline: 21712148]
45. Kinkead L, Allam A, Krauthammer M. AutoDiscern: rating the quality of online health information with hierarchical encoder attention-based neural networks. BMC Med Inform Decis Mak 2020 Jun 09;20(1):104 [FREE Full text] [doi: 10.1186/s12911-020-01131-z] [Medline: 32517759]
46. QMOHI: tool to provide quantitative measures of online health information. GitHub. URL: https://github.com/aksfsu/QMOMI_tool [accessed 2022-01-23]

Abbreviations

ACHA: American College Health Association
FKGL: Flesch–Kincaid Grade Level
FRE: Flesch Reading Ease
HEI: higher education institute
IUD: intrauterine device
LARC: long-acting reversible contraceptives
QMOHI: Quantitative Measures of Online Health Information
SFSU: San Francisco State University
SHC: student health center

Edited by A Mavragani; submitted 24.07.21; peer-reviewed by K Fadahunsi, I Mircheva; comments to author 19.09.21; revised version received 29.10.21; accepted 28.11.21; published 02.02.22.

Please cite as:
Kulkarni A, Wong M, Belsare T, Shah R, Yu Yu D, Coskun B, Holschuh C, Kakar V, Modrek S, Smirnova A
Quantifying the Quality of Web-Based Health Information on Student Health Center Websites Using a Software Tool: Design and Development Study
JMIR Form Res 2022;6(2):e32360
URL: https://formative.jmir.org/2022/2/e32360
doi:10.2196/32360
PMID:35107423

Original Paper

A Digital Therapeutic Intervention Delivering Biofeedback for Panic Attacks (PanicMechanic): Feasibility and Usability Study

Ellen McGinnis[1], PhD; Aisling O'Leary[2], BSc; Reed Gurchiek[2], PhD; William E Copeland[1], PhD; Ryan McGinnis[2], PhD

[1]Department of Psychiatry, University of Vermont Medical Center, Burlington, VT, United States

[2]M-Sense Research Group, University of Vermont, Burlington, VT, United States

Corresponding Author:
Ryan McGinnis, PhD
M-Sense Research Group
University of Vermont
33 Colchester Avenue
Burlington, VT, 05401
United States
Phone: 1 2025095783
Email: ryan.mcginnis@uvm.edu

Abstract

Background: Panic attacks (PAs) are an impairing mental health problem that affects >11% of adults every year. PAs are episodic, and it is difficult to predict when or where they may occur; thus, they are challenging to study and treat.

Objective: The aim of this study is to present PanicMechanic, a novel mobile health app that captures heart rate–based data and delivers biofeedback during PAs.

Methods: In our first analysis, we leveraged this tool to capture profiles of real-world PAs in the largest sample to date (148 attacks from 50 users). In our second analysis, we present the results from a pilot study to assess the usefulness of PanicMechanic as a PA intervention (N=18).

Results: The results demonstrate that heart rate fluctuates by about 15 beats per minute during a PA and takes approximately 30 seconds to return to baseline from peak, cycling approximately 4 times during each attack despite the consistently decreasing anxiety ratings. Thoughts about health were the most common trigger and potential lifestyle contributors include slightly worse stress, sleep, and eating habits and slightly less exercise and drug or alcohol consumption than typical.

Conclusions: The pilot study revealed that PanicMechanic is largely feasible to use but would be made more so with modifications to the app and the integration of consumer wearables. Similarly, participants found PanicMechanic useful, with 94% (15/16) indicating that they would recommend PanicMechanic to others who have PAs. These results highlight the need for future development and a controlled trial to establish the effectiveness of this digital therapeutic for preventing PAs.

(JMIR Form Res 2022;6(2):e32982) doi:10.2196/32982

KEYWORDS

mental health; mHealth; biofeedback; panic attack; digital medicine; app; mobile health; application; biofeedback; mobile phone

Introduction

Background

Almost one-third (>28%) of adults have experienced a panic attack (PA) and >11% have experienced one in the past 12 months [1]. People experiencing PAs exhibit impairment in physical and emotional health as well as in occupational and financial functioning. They also have increased use of health care facilities, emergency departments, and psychoactive drugs [2]. A PA is defined as “an abrupt surge of intense fear or discomfort that reaches a peak within minutes,” characterized by rapid activation of stress-related physiology and fear-related cognitions [3]. Although the body’s ability for rapid physiological activation is evolutionarily advantageous (eg, to provide a bodily response to a direct predatory attack [4]), the generation of this response during PAs, which are inherently *unexpected*, impairs daily functioning [3].

Quantifying PAs

Given their critical evolutionary role in the body's acute response to stress [5] and measurement feasibility, heart rate (HR) and respiration variables have been studied as key metrics in PA research. For instance, HR was observed to increase significantly 1 minute before patient-identified PA onset and remain at a level significantly higher than the resting state for the approximately a 10-minute duration of the PA [6]. However, other studies measuring HR during real-world PAs demonstrated mixed results such that only a moderate percentage of examined PAs (37%-68%) were found to exhibit rapid HR increases compared with non-PA periods. Similarly, respiration changes have been observed before PA and during the PA as characterized by changes in tidal volume and end-tidal PCO_2, but these PA-related values did not differ significantly from typical resting states [7-9].

As PAs are unexpected and episodic, they are difficult to capture in real-world situations. To collect and present sufficient data, studies on PA physiology have (1) presented spontaneous PA data as case studies [10,11], (2) induced PAs medically in the laboratory [12], or (3) focused on a subsample of 4% of people having PAs who have panic disorder (PD) [6,13], which is characterized by recurrent PAs with persistent worry between PAs. There is evidence to suggest that PA physiology differs in severity depending on how the PA is induced [12] and participant symptomatology [6]; thus, generalizations to real-world PAs for individuals without PD are limited. The sample size across studies is also relatively low. In a study of individuals with PD, 43 participants were monitored for 6 continuous days and 13 unexpected PAs were observed, demonstrating that high effort and participation is necessary to collect data on even a small number of episodic PAs. In the 3 studies that aimed to capture the average physiological changes during real-world PAs in individuals without PD, the combined number of PAs was only 50 [7-9] including several PAs within individuals. Thus, there is a clear need for additional studies of PA physiology, particularly in individuals without PD.

A key factor that has likely limited previous studies of PA physiology is the availability of technologies for providing noninvasive measurements in real-world environments. For instance, one of the HR studies described above conducted continuous ambulatory monitoring with 9 wearable sensors connected by wires to a data monitor worn in a fanny pack [6]. This approach imposes a significant burden on researchers and participants alike. Less cumbersome equipment does exist [14-16], yet these specialized devices can be expensive and burdensome to carry around, limiting accessibility. To better capture a representative picture of real-world PAs, more feasible and accessible tools are needed to measure PAs wherever and whenever they occur.

Treating PAs

Although there is limited evidence demonstrating the real-time physiological response to PAs, there is more research on their treatment. It is critical that individuals experiencing PAs have access to feasible and evidence-based interventions as they tend to seek help at significantly higher rates (46%) than individuals with any other axis 1 disorder each year [2,17]. Despite increased treatment-seeking, only 18% of services delivered to these individuals are based on evidence specific to PA reduction [2], as most treatments are intended for general mental health impairment (ie, anxiety). Existing PA treatments can be categorized by their mechanism of action as follows: (1) actively attempting to prevent physiological symptoms (avoidance techniques) or (2) reframing the patient's perspective of their physiological symptoms (approach techniques).

Avoidance techniques, such as meditation or progressive muscle relaxation, focus on preventing physiological symptoms at onset. In individuals with general anxiety, avoidance techniques have been found to induce (not reduce) anxiety in a significant number of patients (17%-54%) [18]. For PAs specifically, prescribed relaxation demonstrates poorer effectiveness and significantly higher attrition than other treatment options [19]. Psychopharmacological intervention (eg, benzodiazepines) is another potential avoidance technique for treating PAs. Although this approach effectively stops the physiological symptoms of PAs, it does not decrease the likelihood of future PAs and has serious clinical side effects (eg, dependence, rebound anxiety, and memory impairment) that reduce quality of life [20]. Thus, there is little evidence demonstrating that avoidance techniques are effective in addressing the underlying mechanisms that drive PAs and some of these approaches have significant side effects.

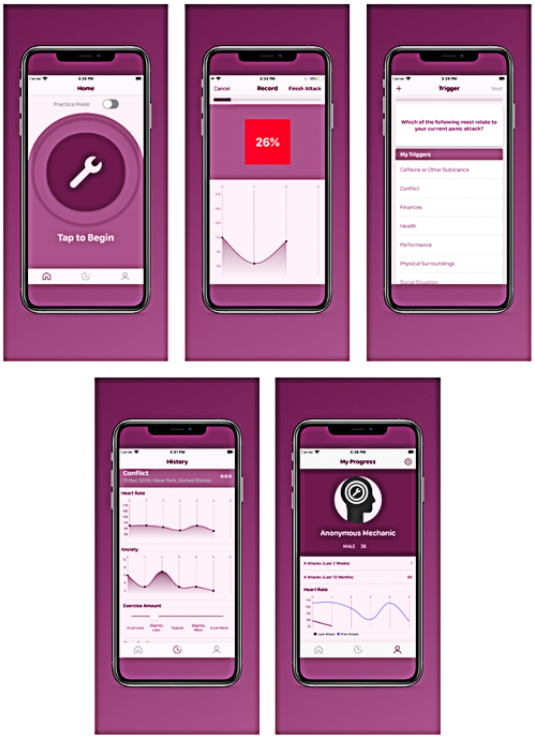
Approach techniques aim to help patients think about their physiological symptoms differently and confront them. Psychotherapies such as cognitive behavioral therapy with interoceptive exposure involve reframing panic-related thoughts and learning to experience symptoms with less panic via in vivo exposure (eg, by inducing PA symptoms, such as dizziness, in a session by spinning the patient in a chair). These therapies are very effective in reducing PA frequency and severity [19] but require a weekly 1-hour visit with a trained and licensed clinician for 12 consecutive weeks. Unfortunately, access to these evidence-based therapies is limited [1], which motivates the development of alternative approaches for treating PAs that do not require the presence of a licensed mental health professional.

Biofeedback is an approach technique in which some form of involuntary physiology (eg, respiratory rate, electroencephalogram, or HR) is continuously measured over time and simultaneously displayed back to the user. During biofeedback, the user is trained to improve their health by learning to regulate internal bodily processes that typically occur involuntarily [21]. Biofeedback with breathing training (BT) is one of the most commonly studied biofeedback techniques for treating PAs. This approach, which requires a specialized device (ie, capnometer) to measure breathing data and display it back to the patient in real time, helps patients learn how to raise or lower their end-tidal PCO_2. In 2 studies of patients with PD, a 4-week biofeedback with BT intervention (two 17-minute sessions per day) was shown to be effective at reducing PA frequency and severity [15,22] at 1 year follow-up. However, studies have shown that individuals with PD exhibit breathing irregularities even when not experiencing a PA [23] and thus the effectiveness of this treatment may be limited to those with PD [24]. Other forms of biofeedback without BT have been

shown to be effective for treating more general anxiety [25,26] and thus could also be effective for treating PAs in those without PD. For example, a meta-analysis of 24 studies investigating the impact of HR variability biofeedback on stress and anxiety revealed a rather large effect (Hedges *g*=0.83) [26]. Although these data suggest that biofeedback can improve anxiety-related symptoms across physiology, treatment effectiveness specific to PAs in persons without PD remains unknown. Biofeedback is thought to provide effective intervention for PAs because it allows users to "feel more in control of their bodily reactions and react less fearfully to them," thus ending the cycle of panic by acquiring a sense of mastery [17]. In support of this theory, *perceived control* was found to be associated with the effectiveness of both biofeedback with BT and cognitive behavioral therapy [27]. HR, but not respiration variables, has been significantly related to patient-reported rating of *fear of losing control* [6], indicating that HR may be an ideal choice for such interventional strategies. HR is significantly higher during PAs [6] than in a resting state for those who have PA, but there is no evidence to suggest that individuals with PAs (but not PD) experience differences in HR during typical resting states. Approach techniques that demonstrate long-term reduction in panic and anxiety focus on confronting symptoms when they are active, even attempting to reproduce similar symptoms while in a therapist's office (eg, interoceptive therapy). Now that we have the technology to provide the tools and guidance during an episodic PA, biofeedback could be especially effective during a PA, when an individual can immediately observe their HR fluctuations; however, this has not yet been examined in PAs.

To address the unmet needs of (1) quantifying PAs in individuals in more representative and larger samples and with less burden and (2) offering an accessible, biofeedback-based treatment option for PAs, our research team has developed a digital therapeutic called PanicMechanic (Figure 1). This mobile health (mHealth) app can accurately and feasibly collect HR data during a PA using only a smartphone and use these data to provide HR-based biofeedback to users during their PAs [28]. This tool can profile remote PAs, which was limited by technology so far, but the usefulness of this novel intervention still needs to be assessed to inform future efficacy trials. Thus, in this study, we aim to present our novel digital therapeutic PanicMechanic and leverage it as a feasible and scalable data collection tool to capture profiles of real-world PAs in a sample of help-seeking PanicMechanic users (analysis 1) and conduct a pilot study to assess the feasibility and usefulness of PanicMechanic as a PA intervention in a convenience sample of university students (analysis 2).

Figure 1. PanicMechanic mobile health app screens. The PanicMechanic mHealth app is available wherever and whenever a user experiences a PA (screen 1). It provides biofeedback through objective measurement of HR during the PA (screen 2) and allows users to capture their perceived anxiety throughout the PA and identify their behavioral and thought triggers (screens 3 and 4). These data are aggregated over time to allow users to track their progress and identify trends that may be helpful for preventing future PAs (screens 4 and 5).



Methods

To address these aims, we first introduce PanicMechanic and then describe the data collection and analysis procedures used to capture the profiles of real-world PAs. Finally, we describe the data collection and analysis procedures used to pilot PanicMechanic as a PA intervention.

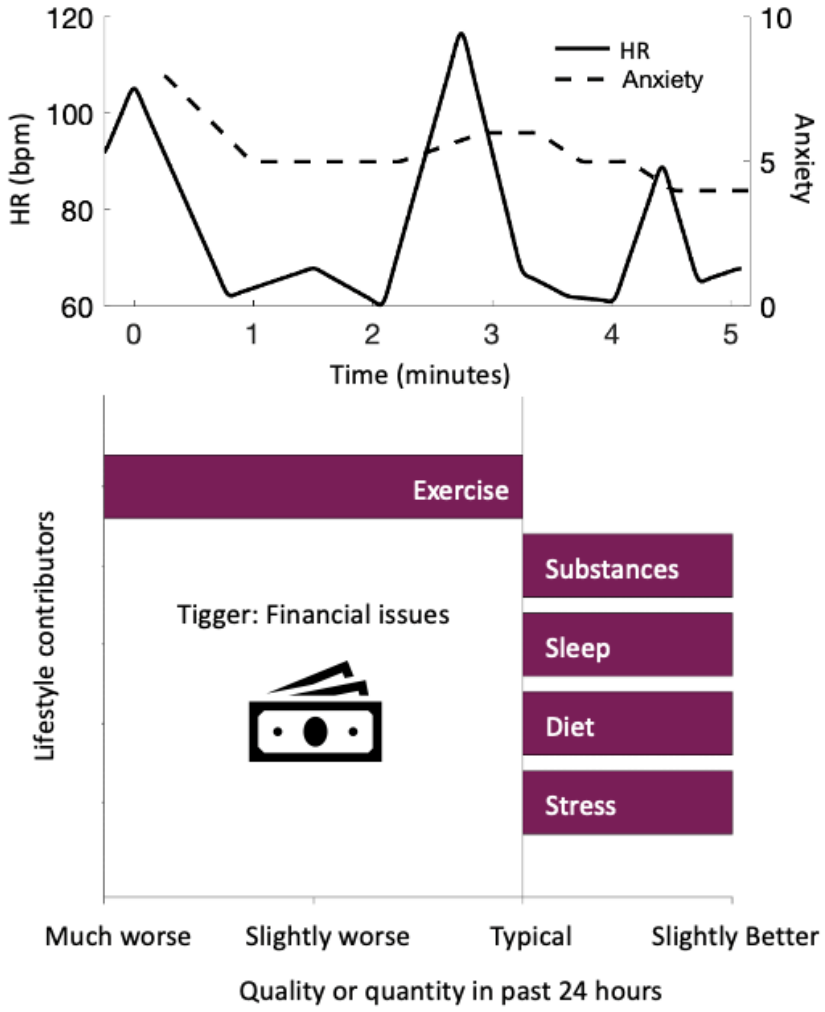
PanicMechanic

PanicMechanic is a commercially available mHealth app (available on Android and iPhone operating system) developed by our team and released in April 2020 that guides users through

their PAs (Figure 1). The app provides biofeedback through objective measurements of HR during the PA, which are enabled via an analog to reflective photoplethysmography provided by the smartphone's camera and validated algorithms [29]. For photoplethysmography, the camera is used to assess changes in the color of the fingertip associated with blood being pumped by the heart through the capillary bed of the fingertip. The app also allows users to rate their anxiety levels throughout the PA and choose from a selection of and record information about lifestyle contributors and triggers (see Figure 2 for data from an example user's PA). App features were developed by a clinical psychologist with experience in treating PAs, and leveraged best practices in clinical psychology while considering input from a variety of individuals who undergo or who have undergone PAs. Our previous work has demonstrated the feasibility of this measurement modality during PAs [29,30].

Upon download, users complete a brief tutorial and are encouraged to open the app whenever they start to experience a PA. During a PA, users are first instructed to place an index finger on the lens of their phone camera with the flash activated to record their HR for the first time. The app cycles through screens that display a graph of HR over time (biofeedback) and prompts the users to make additional HR measurements, rate their anxiety level, and answer questions about lifestyle contributors (exercise, sleep, nutrition, stress, and substance use) and triggers for their PA. The screens provide encouraging messages, such as "You got this!" throughout the attack and using the data from previously logged attacks, the app also provides an estimate of the time that remains in the user's attack. PanicMechanic aggregates user data over time to allow identification of trends (eg, HR, anxiety ratings, and most common triggers) that may help users prevent future attacks.

Figure 2. Example of panic attack case recorded with PanicMechanic. bpm: beats per minute; HR: heart rate. Data from a PA tracked by a PanicMechanic user are reported here. The app tracked HR (bpm) and anxiety ratings (on a scale of 0 to 10; low to high) during attacks (top). This attack lasted for approximately 5 minutes, during which time HR showed 3 distinct peaks and anxiety ratings generally decreased. The app also allowed users to identify lifestyle factors and triggers that may have contributed to the attack (bottom). Data were tracked over time and reported back to the user (Figure 1) to help them identify trends that may be helpful for preventing future PAs.



Capturing Profiles of Real-world PAs (Analysis 1)

To capture profiles of real-world PAs, we considered data from PanicMechanic users in the 1-year period from April 2020 to April 2021. All the users of the app agreed to its terms and conditions, which included the statement, "Your User Content might be anonymized and used for research purposes." As of April 2021, PA data were available from 148 anonymized PanicMechanic users.

Data from the first PA tracked by these users were used to examine the average profile of HR and anxiety ratings throughout an attack. The inclusion criteria for considering the first PA from a given user were that the HR time series had to have at least 4 HR samples (80/148, 54.1% rejected) and a clear peak (18/148, 12.2% rejected). These criteria led us to reject, for example, instances when app use was terminated before the end of the PA (before a peak HR occurred) or commenced after the peak HR had occurred. These data were leveraged to compute the ensemble average time series of HR and subjective anxiety ratings to capture the physiological profile of real-world PAs and explore their relationships during the PA.

We cannot control when users begin using the app during their PA; hence, to compute the ensemble time series, it was necessary to express the time series relative to a standardized instant in time. Thus, all time stamps were expressed relative to the first peak in the HR signal as this was an easily identifiable physiological feature, was expected to characterize a typical PA [6], and was common across many of the tracked PAs. As the time stamps did not exactly match across the time series, it was necessary to linearly interpolate the HR data over a uniform grid (0.01-second intervals) for the time interval (−15 and 60 seconds, ie, only data 15 seconds before and 60 seconds after the first HR peak were considered) before averaging. Time series were then smoothed using a low-pass filter (zero-phase, fourth order, Butterworth filter, and 0.1 Hz cutoff frequency) and averaged. If a time series contained data only for time stamps within (but not the entire) a time interval, it was included in the ensemble average only for those time stamps for which data were recorded or could be interpolated. For example, if a time series had data at time stamps −17, −5, 0, 20, and 33 seconds, it would contribute to the ensemble average only for the subinterval (−15 and 33 seconds). Ensemble average time series of subjective anxiety rating data (n=48) were processed in the same way as the HR data (n=50) for PAs that met the inclusion criteria.

On capturing the profiles of real-world first recorded PAs, we also extracted information from each PA about the typical duration of recordings, severity (in terms of peak HR and anxiety rating), and information about lifestyle contributors and triggers. To provide a more comprehensive picture of these factors, we considered all the data from the first 4 PA recordings that met the inclusion criteria from 148 PanicMechanic users.

PanicMechanic as a PA Intervention (Analysis 2)

To assess the feasibility and usefulness of PanicMechanic as a PA intervention, we considered data from a pilot study of PanicMechanic use among university students. This demographic was chosen as a convenience sample because the majority of university students have smartphones and are open to using mental health apps [31] and because PAs have been shown to rise significantly during the transition to adulthood [32]. Email, social media, and flyer advertisements were used to recruit 18 participants at a northeastern public university in March 2020, April 2020, and May 2020, notably (and coincidently) just after the COVID-19 pandemic–related stay-at-home order was issued in the northeastern United States. Interested persons were screened via phone for eligibility with the following inclusion criteria: must own an Apple iPhone (Android version was not bug-free at the time of the study), be at least 18 years old, report experiencing a PA in the past month, and have university student status.

The participants provided written informed consent before completing a baseline assessment survey. They were given access to PanicMechanic on their personal smartphones and viewed the app's tutorial. They were instructed to use the app whenever they experienced a PA during the following 12 weeks. Throughout the study period, the participants completed 2-minute weekly web-based surveys about their PA and were administered a 15-minute web-based follow-up survey about their experience with using the app 1 week after the study period. Upon study completion, participants were compensated with Amazon gift cards worth up to US $50. All data collection activities were approved by the institutional review board of the University of Vermont (CHRBSS 00000747).

The baseline assessment survey included 42 items that captured participant demographics; PA symptom assessment using the Structured Clinical Interview for Diagnostic and Statistical Manual for Mental Disorders, which is a standardized semistructured interview; and previous treatment and mental health history. During the 2-minute weekly survey, if the participant self-reported capturing a PA with PanicMechanic, they also rated how difficult it was to use the app and indicated the ways (if any) in which the app was helpful. The follow-up survey included open-ended questions about helpfulness and challenges of app use, an indication of whether the participant would use the app in the future, and if they would recommend the app to others who experience PAs.

Descriptive data for participants in the pilot study were available, including demographic characteristics (N=18) and their PanicMechanic use (n=16; of the 18 participants, 2 [11%] participants did not participate beyond the first week of the study period and were lost to follow-up). Content analysis was conducted on qualitative data responses to 2 open-ended items in the follow-up survey at the conclusion of the 12-week intervention: "Overall, in what ways did you find app use helpful" and "Overall, what challenges did you face in using the app during panic attacks?" Content analysis followed the steps detailed in a previous work [33], including first inductive and then deductive analyses to understand the data. First, 2 raters independently conducted inductive content analysis that involved reading all the participant responses, making notes on their content, grouping them, and organizing the responses into categories [33]. Next, the raters jointly conducted deductive content analysis that involved developing a structured matrix of categories, coding the data into those categories, and discussing how those data compare to expectations [33]. Any

discrepancies between the 2 raters were blindly analyzed by a third rater and majority ruled. Consensus categories, descriptions, percentages of total responses, and example responses are presented.

Results

Profiles of Real-world PAs (Analysis 1)

A total of 148 PanicMechanic users tracked at least some HR data from at least one PA. In accordance with the iOS app privacy policy, no demographic data were required to download and use the app; thus, we were not able to report any demographic data besides the knowledge that all 148 users had unique and verified email addresses. PanicMechanic had been featured on news media outlets in April 2020 when it became available but had no paid advertising to promote it. Figure 2 provides an example of the data recorded by PanicMechanic during a PA tracked by a user, including the time series of HR and anxiety ratings (top), lifestyle contributors, and the PA trigger (bottom). This example PA lasted just over 5 minutes, during which time you could see a decreasing trend in anxiety rating and 3 clear instances when the user's HR peaked (at between 90 and 115 beats per minute [bpm]) and then returned to baseline (at between 60 and 70 bpm). Before this attack, triggered by financial issues, the user reported much worse exercise but slightly better diet, sleep, substance use, and stress than typical. It should be noted that significant heterogeneity was observed in HR patterns across PanicMechanic users, including in the timing, amplitude, and duration of HR fluctuations and particularly following the relatively consistent pattern of initial HR peak and return to baseline.

Figure 3 (top) shows the mean HR for the first PA recorded by 50 users with at least 4 HR measurements demonstrating activation (an increase to peak) and recovery (a decrease from peak) slopes. The mean peak HR was approximately 98 (SD 21.56) bpm. It appeared to take approximately 30 seconds from peak HR for a significant recovery down to 85 bpm, which was maintained for an average of 30 seconds. Subjective anxiety rating per minute (range 0-10) maintained a weak yet significant recovery slope (E=−0.43, SE 0.001; $P<.001$) from peak HR to PA end, which appears to be insensitive to specific changes in HR. Figure 3 (bottom) also demonstrates individual heterogeneity across the demeaned 50 recorded first PAs. Several individual PAs appeared to show a secondary smaller peak HR, which could indicate a cyclical HR pattern for a subset of users that was also exhibited in the example PA presented in Figure 2. The average PA recording lasted 4.64 (SD 6.27) minutes. Exercise amount, stress level, sleep, and eating habits were all *slightly worse* to *typical* and substance use was *typical* to *slightly better* during the 24 hours before a PA (Figure 4). The most commonly identified triggers out of the given choices were *health*, *conflict*, *performance*, and *workload* (Figure 4).

Figure 3. Mean heart rate (HR) and anxiety level during panic attacks. bpm: beats per minute. This figure shows ensemble average HR (solid, left axis, N=50) and subjective anxiety rating (dashed, right axis, N=48) responses for the first PA measured by PanicMechanic users (top). Both time series are expressed relative to the first peak in the HR signal. Demeaned HR recordings (gray, bottom) used for computing the ensemble average (black, bottom) demonstrate the variety in HR trajectories observed.

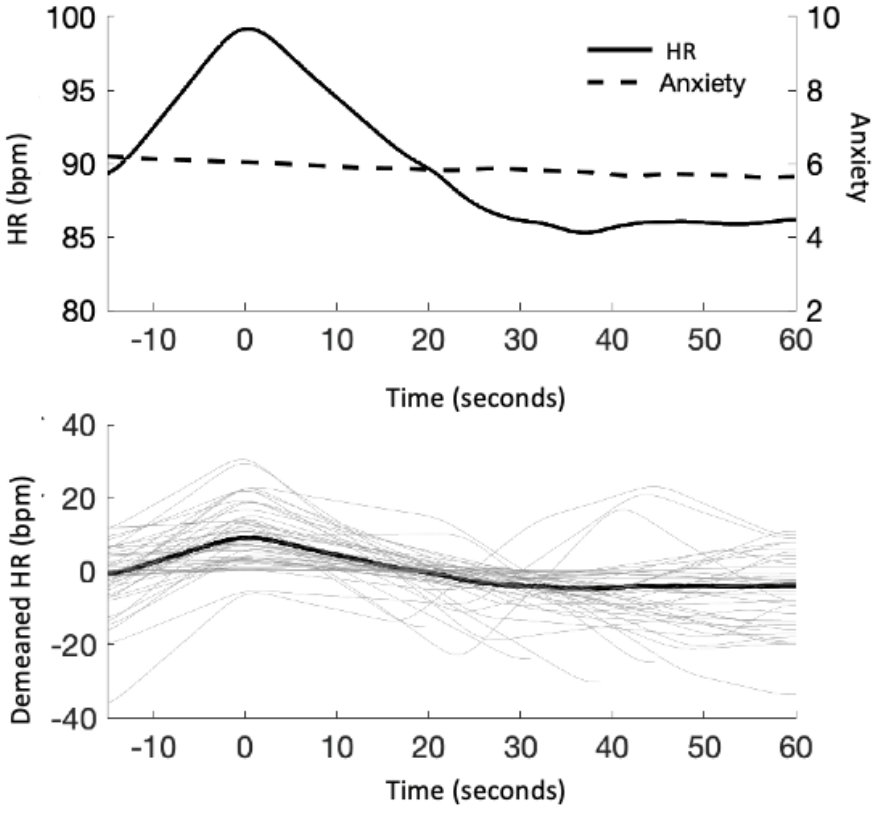
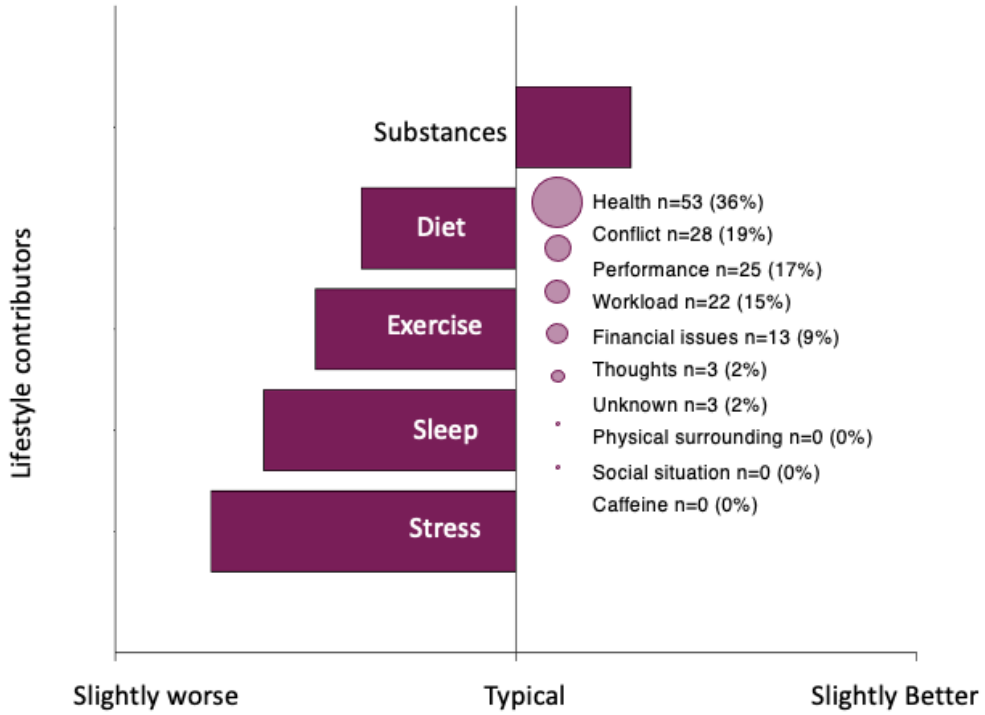


Figure 4. Frequencies of panic attack data. This figure shows participant-reported PA data including mean scores for potential lifestyle contributors that may impact the likelihood of experiencing a PA and the frequency of triggers identified as being responsible for inducing the PA. Lifestyle contributor responses were on a 5-point scale from *a lot worse or less* to *a lot better or more* compared with *typical*. Stress, diet, sleep quality, and exercise were all reported as slightly worse, whereas substance use was reported as slightly better immediately before the PA. Thoughts about health were the most common trigger followed by conflict, performance, and workload.



PanicMechanic as a PA Intervention (Analysis 2)

The demographics of the 18 study participants aged 19-35 years (Table 1) indicate that most of them identified as women, were White, had diagnosed mental health disorder, and had previously sought services for their panic. None of the participants had received biofeedback previously. Participants cited an increase in PA frequency, severity, and duration (mean 5.94, SD 2.38 on a scale of 1-10; 1=not at all increased, 5=somewhat increased, 10=extremely increased) owing to the COVID-19 pandemic, which began shortly before the study recruitment commenced. The most commonly reported panic symptom was a racing, pounding, or skipping heart (16/16, 100%) followed by shortness of breath (13/16, 83%), trembling or shaking (12/16, 78%), and feeling crazy or a loss of control (12/16, 78%; Figure 5). Over the 3-month intervention, the following information was collected in weekly surveys: 94% (15/16) of the participants self-reported to have experienced 123 attacks, of which 39 (31.7%) were self-reported as recorded in the app. The reasons for missing PA recordings are listed in Table 2. Of the 10 participants, 9 (90%) participants with a valid recorded attack (recorded at least four HR data points, which would take about 60 seconds to record) indicated that it was helpful in at least one of the following ways: (1) reducing the severity or duration of the PA, (2) learning about their personalized fear response, or (3) feeling more in control of their body. The participants who used the app rated the user difficulty as 4.73 (SD 2.57) out of 10.

Table 2 presents a qualitative analysis of participant comments from the final survey (n=16). In response to an open-ended question, “Overall, in what ways did you find app use helpful?” the most frequent responses were about observing symptom patterns, which aided understanding of body response (symptom pattern recognition) and the structure of the app helping to redirect user attention (guided attention). In response to the open-ended question, “Overall, what challenges did you face in using the app during panic attacks?” the most frequent responses were about remembering or being motivated enough to open the app (forgot or unmotivated to use) while experiencing panic and app problems owing to technology glitch or user error (such as difficulty placing finger properly for HR to be detected) being a barrier to use (technology difficulties). Representative testimonials and additional response categories are presented in Table 2. When asked if they would use the app again in spite of difficulties, 56% (9/16) of the participants indicated they would. When asked if they would recommend the app to someone else experiencing panic, 94% (15/16) of the participants indicated they would recommend it. In examining the discrepancy between the prevalence of future use and recommendation to others, it appeared that some participants believed the app could be helpful to others who may better remember to open the app at the start of panic symptoms.

Figure 5. Symptoms of panic attack among the participants. This figure shows that the most common PA symptom is racing HR, which was experienced by 100% (16/16) of participants and supports our choice to consider HR biofeedback for PAs. Shortness of breath was the next most common symptom, experienced by 83% (13/16) of participants. The number of responses range from 17 to 18 owing to some missing responses.

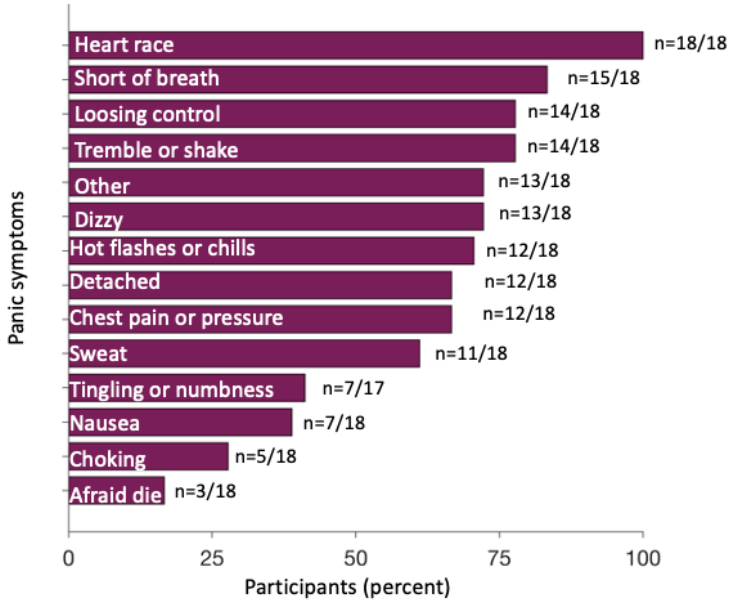


Table 1. Pilot study participant demographics (N=18).

Pilot study demographics	Values
Gender, n (%)	
Female	13 (72)
Male	3 (17)
Other	2 (11)
Age (years), mean (SD)[a]	24 (5)
Race, n (%)	
White	15 (83)
Asian American	1 (6)
Other	2 (11)
Ethnicity, n (%)	
Non-Hispanic	16 (89)
Hispanic	2 (11)
Mental health diagnosis, n (%)	
Any diagnosis	13 (72)
Type of diagnosis	
Anxiety	10 (56)
Depression	7 (39)
OCD[b]	3 (17)
PTSD[c]	3 (17)
ADHD[d]	2 (11)
Panic	2 (11)
Eating disorder	2 (11)
Bipolar disorder	1 (6)
Adjustment disorder	1 (6)
Personality disorder	1 (6)
Panic treatment sought, n (%)	
Currently in therapy	12 (67)
Ever in therapy	17 (94)
Physician	13 (72)
Emergency room or urgent care	5 (28)
Prescribed medication	10 (56)
Self-medicated (alcohol or drugs)	5 (28)
Biofeedback	0 (0)

[a]Range: 19-35 years.

[b]OCD: obsessive compulsive disorder.

[c]PTSD: posttraumatic stress disorder.

[d]ADHD: attention-deficit/hyperactivity disorder.

Table 2. Pilot study content analysis (n=16).

Qualitative category	Values, n (%)	Example testimonial 1	Example testimonial 2
Overall, in what ways did you find app use helpful?			
Symptom pattern recognition (observing symptom patterns aided the understanding of body response)	6 (38)	"Rather than freaking out and feeling like I was dying, I saw that my heart rate was just slightly elevated and fluctuating, and I knew that the attack was temporary and I could work through it."	"I liked how it kept track of my heart rate because seeing it decrease was calming."
Guided attention (the structure redirected attention)	6 (38)	"Watching my heart rate during the panic attack helped me focus more on what was going on."	"Seeing my pulse change was really helpful in giving me something to focus on to calm down during a panic attack."
Accessibility (knowing it was accessible)	3 (19)	"The app provides instant assistance with attacks, instead of waiting to get help."	"The app was accessible from my pocket, easy to read, and easy to follow."
Physiological validation (the personalized data objectively acknowledged the experience as a panic attack)	3 (19)	"The app was most helpful in getting me to acknowledge experiences that I've had for a long time as real, manageable symptoms of a known disorder - rather than just terrifying feelings."	"I found the information provided by the app regarding panic attacks to be calming and affirming."
Affirmations (the words of encouragement)	2 (13)	"The positive affirmations it gives you is helpful."	"It helped encourage me through it and stay in tune with myself."
Triggers (being asked to identify triggers)	2 (13)	"Identifying the trigger and watching your body calm down as you calm down was helpful."	"The app has many different triggers that we could choose from."
Overall, what challenges did you face in using the app during panic attacks?			
Forgot or unmotivated to use (remembering or being motivated to open the app owing to panic)	8 (50)	"I often lacked the presence of mind or motivation to get my phone and start tracking it."	"It is not my first instinct to use an app when I am having an attack."
Technology difficulties (owing to glitch or user error)	5 (31)	"Sometimes I felt like it wasn't recording my pulse right which I fixated on."	"The app never gave me an average length of my panic attacks so it always said I had 0 min left."
Symptom barrier (panic symptoms impacted app use once it was opened)	4 (25)	"Physically shaking made it hard for me to keep my finger on long enough to read my heart rate."	"I had trouble answering because I was freaking out."
Not accessible (did not have access to phone)	3 (19)	"I didn't end up having my phone with me during most of my panic attacks."	"It's too inconvenient for me to use considering my panic attacks often happen while driving."
Repetitive guidance (structure of app was repetitive)	1 (6)	"Questions too repetitive, especially when tracking more than one attack a day."	__[a]
Would you use the app again?			
Yes	9 (57)	"To monitor myself and keep myself in touch with my body and reality."	"It was helpful so I would continue using it."
Would you recommend the app?			
Yes	15 (94)	"I would say yes because it made me feel more educated on my physical well-being."	"Although it doesn't work for me [Forgot to Use], I definitely recognize the benefit of real time biofeedback, and I feel like this is a great option for people who struggle with anxiety and panic attacks."

[a]Second testimonial is not available.

Discussion

Summary

Overview

In this study, we examined the profile of real-world PAs recorded by users of the mHealth app, PanicMechanic, in the largest observation of PA physiology yet considered in this research area (N=50). We also assessed the feasibility and usefulness of the app as a PA intervention in a pilot study of university students (N=18). We discuss the results of these studies in the following sections, including an analysis of the characteristic changes in HR and anxiety ratings during PAs, common lifestyle contributors and triggers for PAs, and the

observations of the pilot study participants after using PanicMechanic for the 12-week intervention period.

Profiles of Real-world PAs

During recorded PAs, HR appears to fluctuate by approximately 15 bpm (Figure 3), which is consistent with several studies that reported significant HR change at the onset of a PA [6]. However, other studies with null results identified a HR increase only when HR was *disproportionate to activity levels* [8]. Our results show that the average HR varied between 85 and 100 bpm, which are within normal limits of resting HR for adults [34] and therefore would not qualify as HR increase under that study criterion. It may be noted that although there is a significant difference between PA HR peak and baseline, it is the *relative* HR increase that indicates a PA. For reference, an increase of approximately 15 bpm is seen in 30 seconds of stair climbing in a healthy individual aged 44 years [35].

On the basis of data from the PanicMechanic users, there appears to be an average duration from peak to baseline HR of approximately 30 seconds during PAs (Figure 3). Given that previous works indicate that PAs last approximately 10 minutes, these results imply that there are several cycles of HR peaks and recovery to baseline within a PA. Although the frequency of HR changes during PAs has not been reported previously, a study appears to show that 4 HR cycles occur within a typical 10-minute PA (Figure 2) [6]. Similarly, the data we report in Figure 2 show 3 clear HR cycles during a 5-minute PA recording. This cyclical HR pattern could be beneficial to biofeedback intervention, allowing multiple chances for patients to observe and expect recovery patterns. From the data presented in Figure 2, we note that significant heterogeneity was observed in the pattern of HR during the recorded PAs across users. Further exploration of this heterogeneity remains an interesting potential area of future study as these HR patterns could further inform our understanding of how and when PAs induce cycles of physiological arousal.

Subjective anxiety did not follow a cyclical pattern, but slowly decreased across the HR monitoring period (Figure 3). This is consistent with research demonstrating that physiological data does not necessarily correlate with subjective anxiety ratings during fear tasks (eg, tasked with giving a spontaneous speech in front of judges [36]). In previous literature, a significant HR increase was seen to precede self-reported PA onset by approximately 1 minute [6]; thus, there may be a sequential nonlinear relationship between HR and subjective anxiety outside the scope of our recorded PA profile. Observing moment-by-moment HR metrics alongside self-reported anxiety may allow users, or, importantly, help clinicians to observe and expect panic patterns that they may not be able to intervene on otherwise.

We also derived profiles of PA triggers (of the available choices), which were not previously examined. As expected, because this mHealth app became available during the COVID-19 pandemic in the United States (released in April 2020), health was the primary trigger in more than one-third of PAs (Figure 4). This is supported by the pilot study data in which the participants indicated that the pandemic increased the frequency, severity, and duration of their PAs. On days where PAs were recorded, the users reported slightly worse stress, sleep, and eating habits; slightly less exercise than typical; and slightly less drug or alcohol consumption in the past 24 hours (Figure 4). These results could inform future studies focused on predictive models of PA occurrence by providing objective monitoring of these parameters or behavioral interventions to reduce PA risk.

PanicMechanic as a PA Intervention

The pilot study results provide insight into the populations that are most interested in participating in an mHealth intervention for PAs. Results show that most participants had multiple mental health disorders (Table 1), consistent with previous literature on individuals who experience PAs [37]. Of the 18 participants, 2 (11%) participants had PD (observed in 3.8% of persons with PAs in the general public [1]). In addition, consistent with previous literature, the participants (all experienced a PA in the month before study participation) had also sought several forms of help services before the study [2] but, importantly, had never tried biofeedback. Interestingly, all the participants reported experiencing subjective HR increases during PAs (Figure 5), suggesting the validity of HR-based biofeedback.

We assessed the subjective feasibility of using PanicMechanic as a biofeedback intervention for PAs. Quantitatively, the participants rated the difficulty of use as a 4.73 out of 10. Qualitatively, half of the participants indicated that the main difficulty of app use was not remembering or being motivated to open the app during the onset of a PA. Despite difficulties, 56% (9/16) of the participants indicated that they would use the app again in the future. Future development work to incorporate hearables and wrist-worn or other wearable devices is indicated to capture HR continuously and more seamlessly. We envision a system in which a push notification asks the user to open the PanicMechanic app if a HR increase has occurred outside of exercise, which may precede a PA [6]. This type of system modification could, for example, reduce the challenge of not remembering to open the app at the beginning of a PA.

However, despite these challenges, 94% (15/16) of the participants would recommend the PanicMechanic app to others who have PAs. Interestingly, several participants indicated that they would not use the app themselves but would recommend it. These participants cited that they believed in the theory of the intervention, but that their use challenges such as forgetting to use the app or their own circumstances such as only experiencing PAs while driving prevented them from benefitting from the intervention. Quantitatively and qualitatively, the participants found this biofeedback app useful. In weekly reports, 90% (14/16) of the participants who had valid use of the app (at least 4 HR data points, which would take at least 60 seconds to record) indicated that PanicMechanic was helpful in one of the following three ways: (1) reducing the severity or duration of the PA, (2) learning about their personalized fear response, or (3) feeling more in control of their body. Qualitatively, they reported that observing their personalized symptom patterns aided their own understanding of their body and also that the app helped them focus by guiding them through their panic in a structured way. These mechanisms are consistent with previous theory on how biofeedback works [17] and

indicate that PanicMechanic shows promise as an accessible PA intervention.

There are several limitations to our study. First, we cannot be certain of the validity and duration of the recorded PAs. It is possible that users recorded HR data outside of a PA, started to record data after PA onset, and completed data recording before the end of a PA. We attempted to mitigate these problems by providing a step-by-step tutorial and *practice* mode on the app upon download and only analyzing PAs that had at least 4 HR measurements with a visible peak. Importantly, these steps yielded exclusion of 66% of our HR data. Thus, although collection of unsupervised data allows for much larger sample sizes than previously studied, the available data diminishes as it is cleaned [6-9]. As the app provided biofeedback during the first PA recording, it is possible that PAs may be attenuated in intensity owing to biofeedback, as the effective number of sessions studied ranged from 1 to 50 [26]. In addition, we report the prevalence of triggers of remote PAs. However, it is important to note that in this iteration of PanicMechanic, there were specific choices of triggers and no ability to input other options. Users were also able to track a PA without indicating a trigger. These prelisted triggers likely do not encompass all the triggers of PAs and thus the results can only represent the prevalence of the given triggers. Given the likely need for multiple sessions for PA improvement and the fact that only 5 (31%) of the 16 participants in our pilot study used the app more than twice, we were not able to evaluate the objective effectiveness of PanicMechanic. This should be analyzed in future studies of app users and specifically in a controlled trial. Overall, our pilot study sample size was small and they represented a subsample of all the people who have PAs. Future studies should include a larger sample with a more diverse range of people who have PAs to inform more generalizable conclusions.

Conclusions

Overall, we have presented a small window into the profile of real-world PAs in the largest sample of people who have PAs to date. We also demonstrated promising preliminary results from a pilot study with participants indicating that PanicMechanic is useful, albeit with some feasibility challenges, which can be addressed with simple app improvements that leverage existing technologies.

Acknowledgments

This study was supported by the University of Vermont Ventures funding. This study was also supported by the National Institute of Mental Health Career Development Award to EM (MH123031-01A1).

Conflicts of Interest

RM and EM are cofounders of PanicMechanic Inc, which has commercialized the PanicMechanic mHealth app currently offered on the Apple App Store and Google Play Store. This conflict is being managed by the University of Vermont. RM: stock ownership for Epicore Biosystems, Inc, and Impellia, Inc; consulting for HX Innovations Inc, University of Washington, and Happy Health, Inc; research funding from MC10, Inc, Epicore Biosystems, Inc, Medidata, Inc. These conflicts are being managed by the University of Vermont.

References

1. Kessler RC, Chiu WT, Jin R, Ruscio AM, Shear K, Walters EE. The epidemiology of panic attacks, panic disorder, and agoraphobia in the National Comorbidity Survey Replication. Arch Gen Psychiatry 2006 Apr;63(4):415-424 [FREE Full text] [doi: 10.1001/archpsyc.63.4.415] [Medline: 16585471]
2. Klerman GL, Weissman MM, Ouellette R, Johnson J, Greenwald S. Panic attacks in the community. Social morbidity and health care utilization. JAMA 1991 Feb 13;265(6):742-746. [Medline: 1990190]
3. Diagnostic and statistical manual of mental disorders. Oxford Bibliographies. URL: https://www.oxfordbibliographies.com/view/document/obo-9780199828340/obo-9780199828340-0022.xml [accessed 2021-12-18]
4. Mobbs D, Marchant JL, Hassabis D, Seymour B, Tan G, Gray M, et al. From threat to fear: the neural organization of defensive fear systems in humans. J Neurosci 2009 Sep 30;29(39):12236-12243 [FREE Full text] [doi: 10.1523/JNEUROSCI.2378-09.2009] [Medline: 19793982]
5. Hall J, Hall M. Guyton and Hall Textbook of Medical Physiology E-Book. Amsterdam: Elsevier Health Sciences; 2020.
6. Meuret AE, Rosenfield D, Wilhelm FH, Zhou E, Conrad A, Ritz T, et al. Do unexpected panic attacks occur spontaneously? Biol Psychiatry 2011 Nov 15;70(10):985-991 [FREE Full text] [doi: 10.1016/j.biopsych.2011.05.027] [Medline: 21783179]
7. Barr Taylor C, Telch MJ, Havvik D. Ambulatory heart rate changes during panic attacks. J Psychiatr Res 1982;17(3):261-266. [doi: 10.1016/0022-3956(82)90004-8] [Medline: 7187688]
8. Taylor CB, Sheikh J, Agras WS, Roth WT, Margraf J, Ehlers A, et al. Ambulatory heart rate changes in patients with panic attacks. Am J Psychiatry 1986 Apr;143(4):478-482. [doi: 10.1176/ajp.143.4.478] [Medline: 3953890]
9. Cameron OG, Lee MA, Curtis GC, McCann DS. Endocrine and physiological changes during "spontaneous" panic attacks. Psychoneuroendocrinology 1987;12(5):321-331. [doi: 10.1016/0306-4530(87)90061-8] [Medline: 3432497]
10. Cohen AS, Barlow DH, Blanchard EB. Psychophysiology of relaxation-associated panic attacks. J Abnorm Psychol 1985 Feb;94(1):96-101. [doi: 10.1037//0021-843x.94.1.96] [Medline: 3980861]

11. Lader M, Mathews A. Physiological changes during spontaneous panic attacks. J Psychosom Res 1970 Dec;14(4):377-382. [doi: 10.1016/0022-3999(70)90004-8] [Medline: 5495258]
12. Goetz R, Klein D, Gorman J. Consistencies between recalled panic and lactate-induced panic. Anxiety 1994;1(1):31-36. [doi: 10.1002/anxi.3070010108] [Medline: 9160544]
13. Freedman RR, Ianni P, Ettedgui E, Puthezhath N. Ambulatory monitoring of panic disorder. Arch Gen Psychiatry 1985 Mar;42(3):244-248. [doi: 10.1001/archpsyc.1985.01790260038004] [Medline: 3977546]
14. Spire launches new app with breath monitoring, contextual feedback. MobiHealthNews. 2016. URL: https://www.mobihealthnews.com/content/spire-launches-new-app-breath-monitoring-contextual-feedback [accessed 2021-12-18]
15. Kaplan A, Mannarino AP, Nickell PV. Evaluating the impact of Freespira on panic disorder patients' health outcomes and healthcare costs within the allegheny health network. Appl Psychophysiol Biofeedback 2020 Sep;45(3):175-181. [doi: 10.1007/s10484-020-09465-0] [Medline: 32342249]
16. Rubin J, Eldardiry H, Abreu R, Ahern S, Du H, Pattekar A, et al. Towards a mobile and wearable system for predicting panic attacks. In: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2015 Presented at: UbiComp '15: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing; Sep 7 - 11, 2015; Osaka Japan. [doi: 10.1145/2750858.2805834]
17. Bouton ME, Mineka S, Barlow DH. A modern learning theory perspective on the etiology of panic disorder. Psychol Rev 2001 Jan;108(1):4-32. [doi: 10.1037/0033-295x.108.1.4] [Medline: 11212632]
18. Wells A. Panic disorder in association with relaxation induced anxiety: an attentional training approach to treatment. Behav Ther 1990;21(3):273-280. [doi: 10.1016/s0005-7894(05)80330-2]
19. Craske MG, Brown TA, Barlow DH. Behavioral treatment of panic disorder: a two-year follow-up. Behav Ther 1991;22(3):289-304. [doi: 10.1016/s0005-7894(05)80367-3]
20. Pull C, Damsa C. Pharmacotherapy of panic disorder. Neuropsychiatr Dis Treat 2008 Aug;4(4):779-795 [FREE Full text] [doi: 10.2147/ndt.s1224] [Medline: 19043522]
21. Bhat P. Efficacy of Alfa EEG wave biofeedback in the management of anxiety. Ind Psychiatry J 2010 Jul;19(2):111-114 [FREE Full text] [doi: 10.4103/0972-6748.90341] [Medline: 22174533]
22. Meuret AE, Wilhelm FH, Ritz T, Roth WT. Feedback of end-tidal pCO2 as a therapeutic approach for panic disorder. J Psychiatr Res 2008 Jun;42(7):560-568 [FREE Full text] [doi: 10.1016/j.jpsychires.2007.06.005] [Medline: 17681544]
23. Abelson JL, Weg JG, Nesse RM, Curtis GC. Persistent respiratory irregularity in patients with panic disorder. Biol Psychiatry 2001 Apr 01;49(7):588-595. [doi: 10.1016/s0006-3223(00)01078-7] [Medline: 11297716]
24. Wollburg E, Roth WT, Kim S. Effects of breathing training on voluntary hypo- and hyperventilation in patients with panic disorder and episodic anxiety. Appl Psychophysiol Biofeedback 2011 Jun;36(2):81-91. [doi: 10.1007/s10484-011-9150-5] [Medline: 21373936]
25. Moore NC. A review of EEG biofeedback treatment of anxiety disorders. Clin Electroencephalogr 2000 Jan;31(1):1-6. [doi: 10.1177/155005940003100105] [Medline: 10638346]
26. Goessl VC, Curtiss JE, Hofmann SG. The effect of heart rate variability biofeedback training on stress and anxiety: a meta-analysis. Psychol Med 2017 Nov;47(15):2578-2586. [doi: 10.1017/S0033291717001003] [Medline: 28478782]
27. Meuret AE, Rosenfield D, Seidel A, Bhaskara L, Hofmann SG. Respiratory and cognitive mediators of treatment change in panic disorder: evidence for intervention specificity. J Consult Clin Psychol 2010 Oct;78(5):691-704 [FREE Full text] [doi: 10.1037/a0019552] [Medline: 20873904]
28. Biofeedback and anxiety. Psychiatric Times. 1999. URL: https://www.psychiatrictimes.com/view/biofeedback-and-anxiety [accessed 2021-12-18]
29. McGinnis RS, McGinnis EW, Petrillo C, Ferri J, Scism J, Price M. Validation of smartphone based heart rate tracking for remote treatment of panic attacks. IEEE J Biomed Health Inform 2021 Mar;25(3):656-662. [doi: 10.1109/JBHI.2020.3001573] [Medline: 32750933]
30. McGinnis R, McGinnis E, Petrillo C, Price M. Mobile biofeedback therapy for the treatment of panic attacks: a pilot feasibility study. In: Proceedings of the IEEE 16th International Conference on Wearable and Implantable Body Sensor Networks (BSN). 2019 Presented at: IEEE 16th International Conference on Wearable and Implantable Body Sensor Networks (BSN); May 19-22, 2019; Chicago, IL, USA. [doi: 10.1109/BSN.2019.8771040]
31. Kern A, Hong V, Song J, Lipson SK, Eisenberg D. Mental health apps in a college setting: openness, usage, and attitudes. Mhealth 2018 Jun 30;4:20 [FREE Full text] [doi: 10.21037/mhealth.2018.06.01] [Medline: 30050916]
32. Copeland WE, Angold A, Shanahan L, Costello EJ. Longitudinal patterns of anxiety from childhood to adulthood: the Great Smoky Mountains Study. J Am Acad Child Adolesc Psychiatry 2014 Jan;53(1):21-33 [FREE Full text] [doi: 10.1016/j.jaac.2013.09.017] [Medline: 24342383]
33. 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]
34. Umetani K, Singer DH, McCraty R, Atkinson M. Twenty-four hour time domain heart rate variability and heart rate: relations to age and gender over nine decades. J Am Coll Cardiol 1998 Mar 01;31(3):593-601 [FREE Full text] [doi: 10.1016/s0735-1097(97)00554-8] [Medline: 9502641]

35. Teh KC, Aziz AR. Heart rate, oxygen uptake, and energy cost of ascending and descending the stairs. Med Sci Sports Exerc 2002 Apr;34(4):695-699. [doi: 10.1097/00005768-200204000-00021] [Medline: 11932581]
36. Takahashi T, Ikeda K, Ishikawa M, Kitamura N, Tsukasaki T, Nakama D, et al. Anxiety, reactivity, and social stress-induced cortisol elevation in humans. Neuro Endocrinol Lett 2005 Aug;26(4):351-354. [Medline: 16136010]
37. Kessler RC, Stang PE, Wittchen H, Ustun TB, Roy-Burne PP, Walters EE. Lifetime panic-depression comorbidity in the National Comorbidity Survey. Arch Gen Psychiatry 1998 Sep;55(9):801-808. [doi: 10.1001/archpsyc.55.9.801] [Medline: 9736006]

Abbreviations

bpm: beats per minute
BT: breathing training
HR: heart rate
mHealth: mobile health
PA: panic attack
PD: panic disorder

Edited by G Eysenbach; submitted 17.08.21; peer-reviewed by J Callan, A Teles; comments to author 05.11.21; revised version received 17.11.21; accepted 23.11.21; published 03.02.22.

Please cite as:
McGinnis E, O'Leary A, Gurchiek R, Copeland WE, McGinnis R
A Digital Therapeutic Intervention Delivering Biofeedback for Panic Attacks (PanicMechanic): Feasibility and Usability Study
JMIR Form Res 2022;6(2):e32982
URL: https://formative.jmir.org/2022/2/e32982
doi:10.2196/32982
PMID:35113031

Original Paper

Consistency and Sensitivity Evaluation of the Saudi Arabia Mental Health Surveillance System (MHSS): Hypothesis Generation and Testing

Nora A Althumiri[1,2], MSc; Mada H Basyouni[1,2,3], MPhil; Nasser F BinDhim[1,2,4], PhD

[1]Sharik Association for Research and Studies, Riyadh, Saudi Arabia
[2]Decision Support for Research and Studies (IDM), Riyadh, Saudi Arabia
[3]Kingdom of Saudi Arabia Ministry of Health, Riyadh, Saudi Arabia
[4]College of Medicine, Alfaisal University, Riyadh, Saudi Arabia

Corresponding Author:
Nasser F BinDhim, PhD
Sharik Association for Research and Studies
Alyasmin, Abi Baker Rd
Riyadh, 13313
Saudi Arabia
Phone: 966 580033299
Email: nd@idm.sa

Abstract

Background: Public health surveillance systems should be evaluated periodically, and the evaluation should include recommendations for improving the system's quality and efficiency. Each surveillance system may have a unique situation in which evaluating its quality depends on its methodology, aims, and other factors, such as the frequency of repeating the survey in the case of survey-based surveillance.

Objective: As the consistency of the surveillance system to capture demographic data and its sensitivity to monitor the intended health-related event are important indicators of the quality of the surveillance system, the aim of this article is to evaluate the Saudi Arabia Mental Health Surveillance System (MHSS) in terms of consistency and sensitivity via the scientific hypothesis testing process.

Methods: The quality of the MHSS was assessed by examining (1) the consistency of the main demographic variables and (2) the sensitivity to changes in score between the 2 mental health screening tools used in the MHSS and between the 3 waves collected in 3 consecutive months. The assessment uses all data collected via the MHSS between May 2020 and July 2020. The first null hypothesis predicted there were differences between the distributions of the demographic variables between the 3 waves. The second predicted there were no differences between the scores of the Patient Health Questionnaire 9 (PHQ-9) and the Generalized Anxiety Disorder 7-item scale (GAD-7) between the 3 waves.

Results: In terms of sampling variables (age, gender, and region), there were no significant differences between the 3 waves in age, using one-way ANOVA, nor in gender and region, using the chi-square test. In addition, there were no significant differences between the 3 waves in all other demographic variables, except in the income variable. However, in terms of the PHQ-9 score, the one-way ANOVA ($F_{2,12334}$=8.05; P<.001) showed significant differences between waves. Similarly, significant differences between waves were found in the GAD-7 score ($F_{2,12334}$=7.09; P=.001).

Conclusions: The MHSS showed a consistent distribution of the sample demographic variables, while being sensitive to the changes in mental health scores across waves. The MHSS can generate an acceptable level of consistency and sensitivity to monitor mental health trends.

International Registered Report Identifier (IRRID): RR2-10.2196/23748

(JMIR Form Res 2022;6(2):e23965) doi:10.2196/23965

KEYWORDS

mental health; evaluation; screening; surveillance; data quality; surveillance system quality; surveillance system evaluation

Introduction

Public health surveillance is the ongoing, systematic collection, analysis, interpretation, and dissemination of health data essential to the planning, implementation, and evaluation of public health programs for use in public health action to reduce morbidity and mortality and to improve health [1,2]. A surveillance system, however, is a collection of processes and components that enable a successful surveillance process, including data collection, data quality monitoring, data management, data analysis, interpretation of analytical results, information dissemination, and application of the information to public health programs [2]. Methodologically, the most popular surveillance systems in public health repeat cross-sectional surveys on a regular basis in the form of waves, which allow the data to be clustered by periods. Eventually, public health surveillance systems should generate information to inform decision-makers in many areas, including prevention program planning and management, health promotion, quality improvement, and resource allocation [2].

Mental disorders account for more collective disability burden than any other group of illnesses, including cancer and heart disease [3]. Disability can be caused by the effect of mental illness on emotions, thoughts, and daily function and the link between mental illness and general health, especially chronic diseases [4]. Historically, surveillance focused on infectious diseases, then broadened to other areas, such as chronic diseases [5]. Currently, mental health is increasingly recognized as a field in public health surveillance [5-7]. Mental health screenings are now included in established health surveillance surveys, such as the US Centers for Disease Control and Prevention's (CDC) National Health Interview Survey (NHIS), National Health and Nutrition Examination Survey (NHANES), Behavioral Risk Factor Surveillance System (BRFSS), and the Substance Abuse and Mental Health Services Administration's (SAMHSA) National Survey on Drug Use and Health (NSDUH) [5].

In April 2020, the Sharik Association for Research and Studies, formerly known as the Sharik Association for Health Research, established the Saudi Arabia Mental Health Surveillance System (MHSS) in collaboration with the Saudi Health Council (SHC), which is the highest national authority in the health domain in Saudi Arabia. The MHSS is a monthly phone-based cross-sectional survey conducted across the 13 administrative regions of Saudi Arabia, with a monthly sample of approximately 4000 participants. At the time of writing this manuscript, 3 waves have been successfully completed with a sample size of more than 12,000 participants. The full details of the MHSS's scientific approach were published as a protocol paper [8]. The MHSS disseminates its results in an electronic dashboard developed for this project to inform decision-makers of the results as soon as possible and to compare the results with previous waves. A subcommittee under the SHC governs the MHSS, with members representing the main stakeholders across the SHC, including the Ministry of Health.

Although the need for public health surveillance systems has long been recognized, there is increasing pressure to improve the effectiveness of these surveillance systems via appropriate evaluation [9]. According to the CDC Guidelines for Evaluating Public Health Surveillance Systems, evaluation ensures that problems of public health importance are being monitored efficiently and effectively [1]. A recent systematic review focusing on existing approaches to surveillance system evaluation found only 10 originated from the public health surveillance field [9]. However, most of the approaches (13/15) could be defined as either frameworks or guidelines, as they provided a general or structured roadmap for the evaluation process, while fewer provided systematic information about how the evaluation should be carried out and, therefore, could be defined as methods [9]. However, all the assessment methods shared some common aspects and none suggested that all of their attributes would be relevant to each evaluation; instead, they could be selected according to the context and objectives of the evaluation at hand [1,9]. In terms of data quality evaluation, most guidelines suggest "data quality reflects the completeness and validity of the data recorded in the surveillance system" [1]. The first issue is that, with the advances of electronic data collection systems, the completeness of data is no longer an issue, as the data are always complete when an electronic system enforces it, which is the case with the MHSS. The second issue is that all of the guidelines focus on clinical data collection, although none focus on cross-sectional survey–based surveillance. Nevertheless, each surveillance system could have a unique approach and parameters for evaluating data quality depending on its methodology, aims, and other factors, such as the frequency of repeating the survey in the case of survey-based surveillance. Nevertheless, to ensure the quality of the MHSS, we considered all relevant attributes by evaluating the public health surveillance systems issued by the CDC [1].

Looking outside the public health domain for a more practical guide to assess surveillance system quality in general, the US Environmental Protection Agency (EPA) has published several versions of the Guidance for Data Quality Assessment: Practical Methods for Data Analysis, which is described as a "toolbox" of useful techniques for assessing the quality of data [10]. This guidance encourages the use of assumption and hypothesis testing to assess data quality and includes a well-detailed process for doing so [10].

In spite of that, public health surveillance systems should be evaluated periodically, and the evaluation should include recommendations for improving the system's quality and efficiency [1,11]. As the consistency of the surveillance system in capturing demographic data and its sensitivity to monitor the intended health-related event are important indicators of the quality of the surveillance system, the aim of this article is to evaluate the MHSS in terms of consistency and sensitivity via the scientific hypothesis testing process by following the EPA Guidance for Data Quality Assessment.

Methods

This study assesses the quality of the MHSS by examining (1) the consistency of the main demographic variables and (2) the changes between the scores of the 2 mental health screening

tools used in the MHSS and between the 3 waves collected in 3 consecutive months.

Null Hypotheses

With the assumption that the 3 waves were conducted in 3 consecutive months using the same data collection process and sampling methodology, which is further controlled by an automated sampling system [12], and the assumption that major demographic variables (eg, age, gender, education level, and marital status) will not change significantly over a short period of time, there will be no significant difference between the distribution of the demographic variables between the 3 waves. Thus, the first null hypothesis for this study is as follows: there are differences between the distributions of the demographic variables between the 3 waves.

With the assumption that the screening tools, the Patient Health Questionnaire 9 (PHQ-9) [13,14] and Generalized Anxiety Disorder 7-item scale (GAD-7) [15], are very sensitive to detecting changes in depressive and anxiety symptoms and that depressive and anxiety symptoms could vary significantly between individuals and within the same individual over a short period of time, as both tools measure symptoms within the prior 2 weeks, and because we are assessing different groups of individuals in each wave who will produce different scores, we assume that there will be significant changes in both scores across the 3 waves. Thus, the second null hypothesis for this study will be as follows: there is no difference between the scores of PHQ-9 and GAD-7 between the 3 waves.

In addition, internal consistency and test-retest reliability for both the PHQ-9 and GAD-7 were performed with the assumption that the system will generate an acceptable level of internal consistency and test-retest reliability.

Variables

For the first experiment, the demographic variables to be tested are age, gender, region, education level, income level, marital status, and work status. Age, gender, and region are part of the sampling variables and are used to determine the completion of the sampling quota. Therefore, they will generate evidence about the quality of the sampling system because they are expected to be proportional according to the MHSS methodology [8].

For the second and third experiments, the total score of the PHQ-9 and GAD-7 will be used.

Data

This study will use all the data generated by the MHSS between May 2020 to July 2020, which includes 12,337 participants (4004 participants in wave 1, 4180 participants in wave 2, and 4153 participants in wave 3). Each of the participants participated in one wave only. For the test-retest reliability, only 22 participants, 11 of whom (50%) were female, were interviewed twice in a pilot study before initiating the surveillance system, with one week between the two interviews.

Data Analysis

For continuous variables (age and PHQ-9 and GAD-7 scores), a one-way ANOVA test will be used; box plots were also included to show the overall trend of the data. For other categorical variables, the chi-square test will be used. Internal consistency was assessed using the Cronbach α coefficient, and the test-retest reliability was assessed with the intraclass correlation coefficient. We used the SPSS statistical software, version 20 (IBM Corp).

Results

Assumption 1: Consistency of the Main Demographic Variables Between the 3 Waves Collected in 3 Consecutive Months

In terms of sampling variables (age, gender, and region), there were no significant differences between the 3 waves in age ($F_{2,12334}$=0.71; P=.49), in gender (χ^2_2=0.1; P=.97), and in regions (χ^2_{24}= 4.8; P>.99).

In terms of other demographic variables, as shown in Table 1, there were no significant differences, except for income level.

Table 1. Participant demographics.

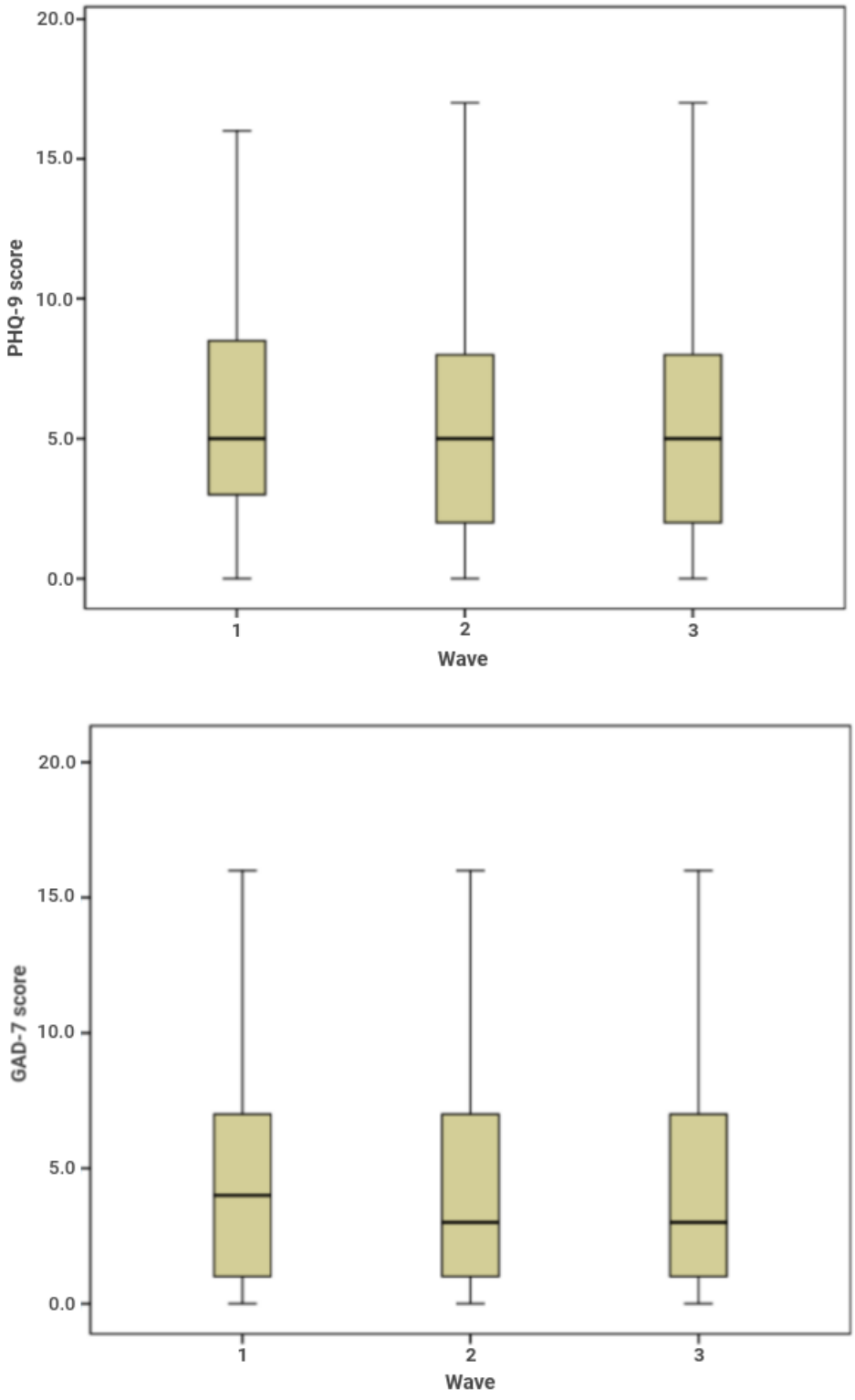
Characteristic	Wave 1, n=4004	Wave 2, n=4180	Wave 3, n=4153	All waves, N=12,337	*P* value
Gender, n (%)					
Male	1989 (49.7)	2083 (49.8)	2058 (49.6)	6130 (49.7)	.97
Female	2015 (50.3)	2097 (50.2)	2095 (50.4)	6207 (50.3)	
Education level, n (%)					
High school or less	1444 (36.1)	1457 (34.9)	1465 (35.3)	4366 (35.4)	.27
Undergraduate diploma	484 (12.1)	456 (10.9)	466 (11.2)	1406 (11.4)	
Bachelor's degree	1846 (46.1)	2022 (48.4)	1955 (47.1)	5823 (47.2)	
Postgraduate degree (eg, master's/PhD)	230 (5.7)	245 (5.9)	267 (6.4)	742 (6)	
Income level, n (%)					
Less than 5000 SAR[a] (<US $1331.47)	629 (15.7)	604 (14.4)	689 (16.6)	1922 (15.6)	<.001
Between 5001 to 8000 SAR (US $1331.74 to $2130.36)	687 (17.2)	595 (14.2)	668 (16.1)	1950 (15.8)	
Between 8001 to 11,000 SAR (US $2130.62 to $2929.24)	619 (15.5)	610 (14.6)	664 (16)	1893 (15.3)	
Between 11,001 to 13,000 SAR (US $2929.51 to $3461.83)	486 (12.1)	551 (13.2)	502 (12.1)	1539 (12.5)	
Between 13,001 to 16,000 SAR (US $3462.10 to $4260.71)	542 (13.5)	628 (15)	559 (13.5)	1729 (14)	
More than 16,000 SAR (>US $4260.71)	1041 (26)	1192 (28.5)	1071 (25.8)	3304 (26.8)	
Region, n (%)					
Asir	321 (8)	322 (7.7)	321 (7.7)	964 (7.8)	>.99
Baha	316 (7.9)	311 (7.4)	314 (7.6)	941 (7.6)	
Eastern Region	314 (7.8)	322 (7.7)	323 (7.8)	959 (7.8)	
Hail	293 (7.3)	326 (7.8)	320 (7.7)	939 (7.6)	
Jazan	312 (7.8)	321 (7.7)	324 (7.8)	957 (7.8)	
Al Jouf	288 (7.2)	318 (7.6)	320 (7.7)	926 (7.5)	
Madinah	321 (8)	325 (7.8)	316 (7.6)	962 (7.8)	
Makkah	325 (8.1)	325 (7.8)	323 (7.8)	973 (7.9)	
Najran	303 (7.6)	322 (7.7)	321 (7.7)	946 (7.7)	
Northern Border	318 (7.9)	318 (7.6)	321 (7.7)	957 (7.8)	
Qassim	309 (7.7)	328 (7.8)	320 (7.7)	957 (7.8)	
Riyadh	301 (7.5)	323 (7.7)	320 (7.7)	944 (7.7)	
Tabuk	382 (7.1)	319 (7.6)	310 (7.5)	912 (7.4)	
Marital status, n (%)					
Never married	1548 (38.7)	1641 (39.3)	1611 (38.8)	4800 (38.9)	.77
Married	2196 (54.8)	2269 (54.3)	2279 (54.9)	6744 (54.7)	
Divorced/separated	169 (4.2)	165 (3.9)	152 (2.7)	486 (3.9)	
Widowed	91 (2.3)	105 (2.5)	111 (2.7)	307 (2.5)	
Employment status, n (%)					
Employed	1579 (39.4)	1723 (41.2)	1638 (39.4)	4940 (40.0)	.29
Self-employed	179 (4.5)	189 (4.5)	170 (4.1)	538 (4.4)	
Unemployed	1121 (28)	1081 (25.9)	1117 (26.9)	3319 (26.9)	
Student	816 (20.4)	853 (20.4)	907 (21.8)	2576 (20.9)	
Retired	309 (7.7)	334 (8)	321 (7.7)	964 (7.8)	

[a]1 SAR=US $0.27.

Assumption 2: The Changes Between the 2 Mental Health Screening Tools Used in the MHSS and Between the 3 Waves Collected in 3 Consecutive Months

In terms of the PHQ-9 score, the one-way ANOVA ($F_{2,12334}$=8.05; P<.001) showed significant differences between waves. Similarly, significant differences between waves were found for the GAD-7 score ($F_{2,12334}$=7.09; P=.001). Figure 1 shows the box plots for both PHQ-9 and GAD-7.

Figure 1. Box plots of the Patient Health Questionnaire 9 (PHQ-9) and Generalized Anxiety Disorder 7-item scale (GAD-7) across the 3 waves of data collection.



Assumption 3: Internal Consistency and Test-Retest Reliability

Internal consistency measures for both scales (α=.838 for the PHQ-9, α=.881 for the GAD-7) were good. In the analysis of test-retest reliability, the intraclass correlation coefficient was 0.941.

Discussion

Principal Findings

This study has used assumption and hypothesis testing to evaluate the quality of the MHSS, which uses repeated cross-sectional surveys conducted in a systematic process as a surveillance system. The results showed that, for the first assumption, the null hypothesis was rejected for 6 out of 7 variables. This finding confirms the assumption that there will be no significant differences between the 3 consecutive waves in the main demographic variables if the surveillance system can generate high-quality data. Similarly, for the second assumption, the null hypothesis was rejected. This finding confirms the assumption that there will be significant differences in PHQ-9 and GAD-7 scores between the 3 consecutive waves if the surveillance system can generate high-quality data. In addition, the internal consistency and test-retest reliability showed an acceptable level of data consistency and reliability.

Although income level showed a significant difference on the chi-square test, a closer look reveals that the within-wave variability is still small, with the largest variability at 3% between the highest and lowest proportions. The age, gender, and region variables were used by the electronic data collection system to control the sample, to provide proportional distribution in gender and region and a similar average age as the general adult population in Saudi Arabia. The lack of differences in these 3 variables confirms the quality of the electronic sampling system and the assumption that it will generate the sample as planned.

The results showed that the MHSS generated a consistent composition of demographic variables in each wave, while showing sensitivity to the changes in the mental health screening measures (PHQ-9 and GAD-7). This finding provides an acceptable level of confidence that the MHSS is measuring what it is intended to measure. However, it is important to clarify that the PHQ-9 and GAD-7 have shown a high level of sensitivity and specificity across the literature, and that is part of the reason we selected them for mental health screening in this mental health surveillance system, to build on the strength of these well-established screening tools [13-15]. We acknowledge that it is hard to separate the effect of the screening tools from the surveillance system sensitivity; however, knowing that these two screening tools are sensitive to change and knowing that the surveillance system can generate consistent demographic segments each wave, we assumed that the change in PHQ-9 and GAD-7 scores will be significant if the surveillance system is functioning well overall. Confirming this assumption via hypothesis-based testing is not a solid confirmation of the sensitivity and consistency of the system, but it can generate some confidence until further evidence can be found.

This study used assumption and hypothesis testing to investigate the ability of the MHSS to generate an acceptable level of data consistency and sensitivity to monitor mental health trends, as defined by the EPA Guidance for Data Quality Assessment, although hypothesis-based evaluation is not an EPA or environmental research specific method. However, the accuracy of the results depends on the validity of the assumptions, which is a limitation for hypothesis testing in general. Another limitation is that the results came from internal comparison data from the same surveillance system, not from external data generated by another surveillance system or a cross-sectional survey using the same or similar methodology to provide external validity. Furthermore, although the ANOVA results used to test the second assumption show statistical significance, they might not be practically significant, especially with the short time period between waves. Finally, this study looked at the surveillance system quality as a broader concept than just data completion or data accuracy when comparing the same record to another record using the same participants. It showed that surveillance system quality could be investigated independently with various scientific testing methods.

Conclusion

The MHSS showed a consistent distribution of the sample demographics, while being sensitive to the changes in mental health scores across waves. The MHSS can generate an acceptable level of consistency and sensitivity to monitor mental health trends.

Data Availability Statement

Data will be available upon request via the Saudi National Health Research Center through NFB.

Funding

This project is funded by the KACST (Grant 5-20-01-000-0001). The funder had no role in data collection, data analysis, data interpretation, or writing of the report. NFB had full access to all study data and had final responsibility for the decision to submit this research for publication.

Acknowledgments

The authors wish to express their gratitude and appreciation to King Abdulaziz City for Science and Technology (KACST) for funding this study.

Authors' Contributions

All authors provided major contributions to the design, development, and revision of this study. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. German R, Lee LM, Horan JM, Milstein B, Pertowski C, Waller MN, Guidelines Working Group Centers for Disease Control and Prevention (CDC). Updated guidelines for evaluating public health surveillance systems: recommendations from the Guidelines Working Group. MMWR Recomm Rep 2001 Jul 27;50(RR-13):1-35; quiz CE1. [Medline: 18634202]
2. Groseclose SL, Buckeridge DL. Public health surveillance systems: recent advances in their use and evaluation. Annu Rev Public Health 2017 Mar 20;38:57-79. [doi: 10.1146/annurev-publhealth-031816-044348] [Medline: 27992726]
3. World Health Organization. Promoting mental health: concepts, emerging evidence, practice: a report of the World Health Organization, Department of Mental Health and Substance Abuse in collaboration with the Victorian Health Promotion Foundation and the University of Melbourne. World Health Organization: Institutional Repository for Information Sharing. 2005. URL: https://apps.who.int/iris/bitstream/handle/10665/43286/9241562943_eng.pdf?sequence=1&isAllowed=y [accessed 2022-01-21]
4. Institute of Medicine (US) Committee on Crossing the Quality Chasm: Adaptation to Mental Health and Addictive Disorders. Improving the Quality of Health Care for Mental and Substance-Use Conditions: Quality Chasm Series. Washington, DC: National Academies Press; 2006.
5. Freeman EJ, Colpe LJ, Strine T, Dhingra S, McGuire LC, Elam-Evans LD, et al. Public health surveillance for mental health. Prev Chronic Dis 2010 Jan;7(1):A17 [FREE Full text] [Medline: 20040232]
6. Gater R, Chisholm D, Dowrick C. Mental health surveillance and information systems. East Mediterr Health J 2015 Sep 28;21(7):512-516 [FREE Full text] [doi: 10.26719/2015.21.7.512] [Medline: 26442892]
7. Reeves W, Strine TW, Pratt LA, Thompson W, Ahluwalia I, Dhingra SS, Centers for Disease Control and Prevention (CDC). Mental illness surveillance among adults in the United States. MMWR Suppl 2011 Sep 02;60(3):1-29. [Medline: 21881550]
8. BinDhim NF, Althumiri NA, Basyouni MH, Alageel AA, Alghnam S, Al-Qunaibet AM, et al. A mental health surveillance system for the general population during the COVID-19 pandemic: protocol for a multiwave cross-sectional survey study. JMIR Res Protoc 2020 Nov 26;9(11):e23748 [FREE Full text] [doi: 10.2196/23748] [Medline: 33156802]
9. Calba C, Goutard FL, Hoinville L, Hendrikx P, Lindberg A, Saegerman C, et al. Surveillance systems evaluation: a systematic review of the existing approaches. BMC Public Health 2015 May 01;15:448 [FREE Full text] [doi: 10.1186/s12889-015-1791-5] [Medline: 25928645]
10. Guidance for data quality assessment: practical methods for data analysis. United States Environmental Protection Agency. 2000. URL: https://www.epa.gov/sites/default/files/2015-06/documents/g9-final.pdf [accessed 2022-01-21]
11. Azofeifa A, Stroup DF, Lyerla R, Largo T, Gabella BA, Smith CK, Behavioral Health Surveillance Working Group. Evaluating behavioral health surveillance systems. Prev Chronic Dis 2018 May 10;15:E53 [FREE Full text] [doi: 10.5888/pcd15.170459] [Medline: 29752804]
12. Do not wait — Automate your health research project now!. Smart Health. 2012. URL: https://shproject.net/ [accessed 2022-01-11]
13. Kroenke K, Spitzer RL. The PHQ-9: a new depression diagnostic and severity measure. Psychiatr Ann 2002 Sep;32(9):509-515. [doi: 10.3928/0048-5713-20020901-06]
14. 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]
15. 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]

Abbreviations

BRFSS: Behavioral Risk Factor Surveillance System
CDC: US Centers for Disease Control and Prevention
EPA: US Environmental Protection Agency
GAD-7: Generalized Anxiety Disorder 7-item scale
KACST: King Abdulaziz City for Science and Technology
MHSS: Saudi Arabia Mental Health Surveillance System
NHANES: National Health and Nutrition Examination Survey
NHIS: National Health Interview Survey
NSDUH: National Survey on Drug Use and Health

PHQ-9: Patient Health Questionnaire 9
SAMHSA: Substance Abuse and Mental Health Services Administration
SHC: Saudi Health Council

Edited by A Mavragani; submitted 30.08.20; peer-reviewed by L Guo, J Lau; comments to author 07.12.20; revised version received 09.12.20; accepted 09.01.22; published 03.02.22.

Please cite as:
Althumiri NA, Basyouni MH, BinDhim NF
Consistency and Sensitivity Evaluation of the Saudi Arabia Mental Health Surveillance System (MHSS): Hypothesis Generation and Testing
JMIR Form Res 2022;6(2):e23965
URL: https://formative.jmir.org/2022/2/e23965
doi:10.2196/23965
PMID:35113027

Original Paper

Interactive Mobile Phone HIV Adherence Support for Men Who Have Sex With Men in the Philippines Connect for Life Study: Mixed Methods Approach to Intervention Development and Pilot Testing

Cara O'Connor[1,2], MPH; Katerina Leyritana[2], MD; Aoife M Doyle[1,3], PhD; James J Lewis[4], PhD; Randeep Gill[5], MD; Edsel Maurice Salvaña[6,7], MD

[1]Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, United Kingdom

[2]Sustained Health Initiatives of the Philippines (SHIP), Mandaluyong, Philippines

[3]Medical Research Council International Statistics and Epidemiology Group, London School of Hygiene and Tropical Medicine, London, United Kingdom

[4]Y Lab, The Public Services Innovation Lab for Wales, School of Social Sciences, Cardiff University, Cardiff, United Kingdom

[5]Johnson & Johnson Global Public Health, London, United Kingdom

[6]Institute of Molecular Biology and Biotechnology, National Institutes of Health, University of the Philippines Manila, Ermita, Philippines

[7]Division of Infectious Diseases (Global Health), Department of Medicine, University of Pittsburgh, Pittsburgh, PA, United States

Corresponding Author:
Cara O'Connor, MPH
Faculty of Epidemiology and Population Health
London School of Hygiene and Tropical Medicine
Keppel Street
London, WC1E 7HT
United Kingdom
Phone: 44 791336248
Email: caraoc@gmail.com

Abstract

Background: The HIV epidemic in the Philippines is one of the fastest growing epidemics globally, and infections among men who have sex with men are rising at an alarming rate. The World Health Organization recommends the use of mobile health (mHealth) technologies to engage patients in care and ensure high levels of adherence to antiretroviral therapy (ART). Existing mHealth interventions can be adapted and tailored to the context and population served.

Objective: This study aims to create a locally tailored intervention using a mobile phone platform to support treatment adherence for HIV patients on ART in the Philippines.

Methods: A mixed methods approach guided by the Behavior Change Wheel framework was used to adapt an existing mHealth adherence support platform for the local setting and target population. A literature review, retrospective clinical record review, and focus group discussions with patients were conducted to understand the drivers of ART adherence and tailor the intervention accordingly. The resulting intervention was pilot-tested for 8 weeks, followed by focus group discussions with patients who received the intervention to assess the acceptability of the design.

Results: Key issues contributing to nonadherence included side effects, lack of behavioral skills for pill taking, social support, mental health, and substance use. Patients identified mHealth as an acceptable mode of intervention delivery and wanted mHealth services to be highly personalizable. The study team, clinicians, and software developers integrated these findings into the intervention, which included a menu of services as follows: pill reminders, health tips, adherence feedback, appointment reminders, and symptom reporting. During the pilot phase, technical issues in the interactive voice response system (IVRS) were identified and addressed. Patients who participated in the pilot phase expressed a preference for SMS text messaging over the IVRS. Patients responded positively to the appointment reminders and health tips, whereas patient feedback on daily and weekly pill reminders and adherence feedback was mixed.

Conclusions: The mobile phone–based SMS text messaging and IVRS intervention was acceptable to men who have sex with men in Manila, the Philippines, and qualitative analysis suggested that the intervention helped promote ART adherence and appointment attendance.

(JMIR Form Res 2022;6(2):e30811) doi:10.2196/30811

KEYWORDS

mHealth; adherence; HIV; antiretroviral therapy; intervention development; mobile phone

Introduction

HIV on the Rise in the Philippines

The Philippines has the fastest growing HIV epidemic in the Asia-Pacific region [1-3]. National surveillance data show that the number of new HIV cases in the Philippines has increased at an alarming rate during the past decade, with an increase from 311 cases identified in 2007 to 12,778 cases identified in 2019—a 41-fold increase in new HIV diagnoses [4]. According to the surveillance reports by the Joint United Nations Program on HIV/AIDS, the Philippines' progress toward reaching HIV/AIDS 90-90-90 goals has been slow, with 73% of people living with HIV being aware of their status, 44% on treatment, and low coverage of viral load testing (<50%) [5,6].

The group most impacted by HIV in the Philippines is men who have sex with men (MSM), representing 84% of new diagnoses since 2015. The median age of new cases is 28 years, and >80% of people living with HIV/AIDS in the Philippines are aged <35 years [4,7]. In 2015, a national surveillance survey found that HIV prevalence among MSM who practiced anal sex was 6%—an increase from 3.3% in 2013 [4-6,8-10].

As the burden of HIV increases, it is imperative that as many HIV-infected people as possible are diagnosed, started on treatment, and successfully retained in care. Achieving adequate viral suppression through the use of antiretroviral therapy (ART) will be one of the key tools in ending the HIV epidemic in the Philippines. Unfortunately, widespread stigma, lack of knowledge, and barriers to accessing care pose a challenge to engaging patients in testing and ensuring high levels of adherence to ART and retention in care [8,10,11]. As in many low- and middle-income countries, high rates of first-line treatment failure, loss to follow-up, and suboptimal treatment adherence led to poor outcomes for many HIV patients in the Philippines [12,13].

Evidence-based public health interventions are required. However, a 2015 report by the World Health Organization highlights that the body of HIV research conducted in the Philippines has been limited [14], and a systematic review of HIV risk studies in the Philippines through April 2018 found only 3 publications that included data about the group most affected by HIV—MSM [15].

Mobile Health for Adherence

As mobile phone technologies have become widespread in low- and middle-income countries, mobile phone interventions have become increasingly popular in the global health and development sectors as a potentially inexpensive and efficient way to communicate with and deliver services to people.

Mobile phones are multifunctional tools that can be used for a variety of functions that range from simple alarm functions, SMS text messaging, calls, interactive voice response systems (IVRSs), and complex apps and games. People usually have their mobile devices with them; therefore, using mobile technologies allows the timing of the intervention delivery to be synchronized with the most relevant time to claim the attention of the recipient [16]. Moreover, mobile phones can be used almost anywhere. As wireless coverage is improving globally, mobile communications can be provided even in remote areas.

In the Philippines, 99% of the population is reached by mobile cellular network coverage and mobile phone use is among the highest in the world, with 155 mobile connections per 100 people [17,18]. Although the coverage of mobile networks is high, smartphone coverage and mobile internet (Long-Term Evolution) speeds are lower in the Philippines than in other countries in the region [19], which limits the potential reach of mobile internet and app-based solutions.

The 2016 World Health Organization Consolidated Guidelines on the Use of ART for the Treatment and Prevention of HIV Infection promoted the use of SMS text messaging to improve adherence to therapy [20]. Research has shown that mobile health (mHealth) interventions have potential benefits for a wide variety of health issues, including antiretroviral adherence, smoking cessation, diabetes control, maternal health, and vaccination programs [16,21].

Mobile phone interventions have proven successful in improving ART adherence in Africa, South Asia, and Latin America [22-28]. Several systematic reviews have been published regarding mHealth for ART adherence specifically [29,30]. A variety of mHealth approaches to improve adherence to antiretroviral medications have been studied globally, including daily and weekly short text messages [22,31-33], weekly long text messages [31], weekly voice messages [34], and fortnightly phone calls [35,36]. The outcome measures of ART adherence interventions vary; outcome measures include self-reported adherence, objective measures of adherence (ie, pill count, pharmacy refill, and medication monitors), biological end points (ie, viral load suppression), and quality-of-life measures.

In Kenya, 2 important examples of successful ART adherence interventions were implemented. At the WelTel Kenya1 multisite randomized clinical trial of HIV-infected adults initiating ART, adherence to ART was reported in 61.5% (168/273) of patients receiving the SMS text messaging intervention compared with 49.8% (132/265) of patients in the control group (relative risk for nonadherence 0.81; *P*=.006). Suppressed viral loads were reported in 57.1% (156/273) of

patients in the SMS text messaging group and 48.3% (128/265) of patients in the control group (relative risk for virologic failure 0.84; *P*=.04) [22].

Another randomized trial of 131 adult patients who had initiated ART less than 3 months before enrollment found that 53% of participants receiving weekly SMS text messaging reminders achieved adherence of ≥90% during the 48 weeks of the study, compared with 40% of participants in the control group (*P*=.03). Participants in groups receiving weekly reminders were also significantly less likely to experience treatment interruptions exceeding 48 hours during the 48-week follow-up period compared with participants in the control group [31].

Multiple reviews of the literature on adherence programs suggest that mobile phones are a feasible, acceptable, and effective mode of delivery for HIV interventions targeting young MSM [37-40]. There is also evidence that daily reminders can support habit forming over 2-3 months and that weekly reminders effectively support adherence [23,25,29,41,42]. It is not clear whether improvements in adherence are sustained if reminders are stopped once a habit is formed. Some evidence suggests that weekly messages with interactive elements that elicit a response from the user may be the most effective SMS text messaging adherence interventions, but many questions remain unanswered [23,24].

Aim and Objective

The investigators aim to create a locally tailored intervention using a mobile phone platform to support treatment adherence for HIV patients on ART at the study clinic in Metro Manila, Philippines.

The objective of the formative research phase of the study is to adapt an existing technology platform (Connect for Life) for the local context. We seek to answer the following questions:

- What is the level of adherence in the study clinic population and similar populations in the country and region?
- What are the barriers to and determinants of ART adherence among the study clinic population?
- What components should an mHealth intervention include to address these barriers and determinants?

Methods

Setting

The Sustained Health Initiatives of the Philippines (SHIP) Clinic is a public-private partnership that opened in 2012. It is a low-cost, private facility in Metro Manila, a city of approximately 13 million people in the predominantly Catholic country of the Philippines and the most densely populated city in the world. As of April 2021, the SHIP clinic provided HIV primary care and wraparound services to approximately 900 patients. Between 2012 and 2018, SHIP was a satellite partner clinic of the Sexually Transmitted Infection/AIDS Guidance Intervention and Prevention Unit at the Philippine General Hospital, the largest public hospital in the country.

Approximately 98% of SHIP's clients are MSM, with an average age of 30 years at the initial consultation. Most are employed full time or part time. The patients come from all regions of Metro Manila, and some live outside of Metro Manila in other provinces. SHIP currently enrolls approximately 4 new patients each month.

Ethical clearance for the study was obtained from the University of the Philippines Manila Research Ethics Board (protocol number 2016-265-01) and from the London School of Hygiene and Tropical Medicine (reference number 11631). All patients provided written consent before inclusion in the study.

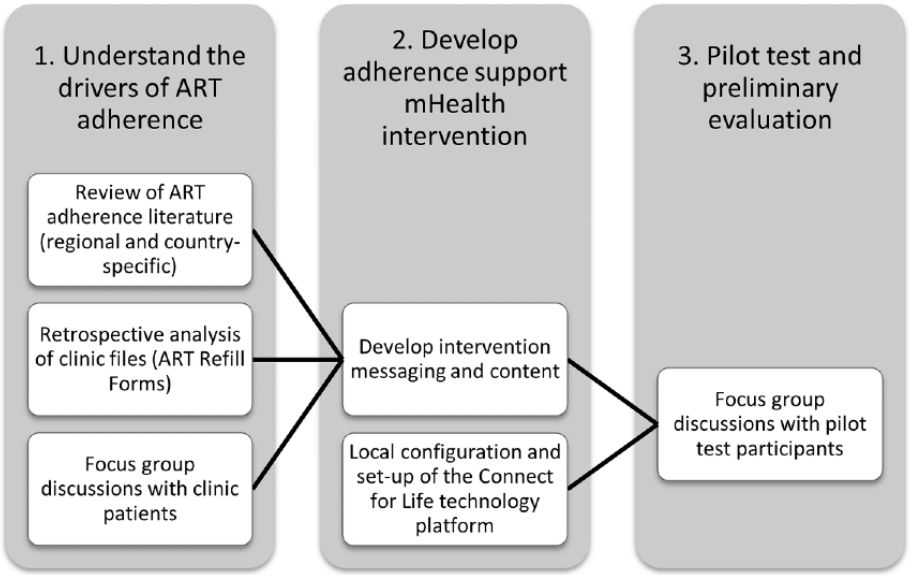
Intervention Development Approach

Overview

To determine the best configuration of mobile phone support services for patients in the study site, we used the following methodology (Figure 1):

1. Formative research to understand the drivers of ART adherence: literature review of factors associated with ART adherence (global, regional, and country-specific data), retrospective analysis of clinic files (ART refill forms), and focus group discussions (FGDs) with patients in the SHIP clinic.
2. Development of mHealth intervention: we adapted an existing mHealth adherence support platform, tailored it to the setting and target population, and pilot-tested the platform, guided by the Behavior Change Wheel (BCW) approach.
3. Pilot test and preliminary evaluation: we piloted the intervention with a subset of clinic patients for 8 weeks and then conducted FGDs with patients who received the intervention in the pilot phase.

Figure 1. Intervention development process. ART: antiretroviral therapy, mHealth: mobile health.



Formative Research to Understand the Drivers of ART Adherence

Review of ART Adherence Literature

A literature review of regional, country-specific, and site-level routine clinical data on ART adherence was conducted by the investigators. The literature provided point estimates for adherence that could serve as a comparison with our study population and outline some of the main facilitators of and barriers to adherence in the Philippines context.

Retrospective Analysis of Clinic Files

A record review of all pharmacy refill forms from the study clinic was conducted. Data were captured from 3381 pharmacy refill forms collected during routine clinical care for 682 patients between May 2012 and August 2016. The pharmacy refill forms included basic demographic information, dispensing data, pill count, and self-report of the number of doses missed in the past 30 days. Data quality for these forms was poor, with missing forms, fields left blank, and inconsistent or conflicting data in a large proportion of the records. Owing to these limitations in data quality, only the most recent refill form for each patient was included in the analysis, as data were much more complete in the recent forms. The estimate of adherence was calculated as follows:

> Adherence percentage = 1 − (number of pills reported missed since last visit/number of pills dispensed at the last visit) **(1)**

FGDs With Clinic Patients

During the formative research stage, the study team conducted FGDs to explore adherence challenges and possible approaches to support adherence. The specific topics covered during the discussions were adherence challenges, use of mobile phones, attitudes toward receiving adherence reminders, priority health education topics for mHealth tips, and acceptability of receiving an adherence score as a feedback mechanism.

Focus group participants were recruited through convenience sampling of clinic patients as identified by the SHIP clinic physician. Patients were eligible to participate if they were aged ≥18 years, HIV-positive and on ART at the SHIP clinic, and willing to participate in a group discussion setting. Privacy around HIV-positive status was the biggest barrier to recruitment, and only patients who were publicly open about their HIV-positive status participated in FGDs.

Each FGD was facilitated by a qualified HIV test counselor who was experienced in qualitative methods. A second staff member took detailed notes throughout the session, and immediately after the discussion, the notetaker and facilitator debriefed and recorded their initial observations. The FGDs were conducted in a mix of English and Filipino, which is common in Metro Manila. The discussions were held in a hired conference room located in the building next to the SHIP clinic, selected for the convenience of the participants. Before the discussion, the participants completed the informed consent process and provided demographic data and ART adherence data using a short questionnaire. The FGDs were audio recorded on 2 devices, and the discussions ranged from 60 to 105 minutes.

The FGDs were transcribed, and a framework-guided rapid analysis was conducted. Transcripts were manually coded using a deductive coding methodology in which initial coding grouped responses into overarching themes as per the topic areas included in the FGD guide. Following initial coding, line-by-line coding was used to assign the subthemes. Qualitative data were consolidated in a structured template based on the a priori research questions. The template enabled the consolidation of data into matrices by each category to identify salient themes.

Develop Adherence Support mHealth Intervention

Develop Intervention SMS Text Messaging and Content

The intervention development process was broadly guided by the BCW developed by Michie et al [43-45]. Behavior change techniques (BCTs) related specifically to ART adherence were informed by the information-motivation-behavioral skills (IMB) model of ART adherence [46].

The BCW is a method for characterizing and developing behavior change interventions based on a comprehensive causal analysis of behavior (Figure 2) [45]. In the BCW approach, the intervention design process consists of 3 stages.

The first stage of intervention development is to understand the behavior. In this case, the specific target behavior is optimal adherence to ART, defined as taking at least 95% of the prescribed ART doses on time. To understand behavior, the BCW approach starts with the question, *What conditions internal to individuals and in their social and physical environment need to be in place for a specific behavioral target to be achieved?* On the basis of the formative research findings, the components of capability, opportunity, and motivation that interact to account for behavior in the BCW approach were summarized [43-45]. Using the capability, opportunity, motivation, behavior (COM-B) model, we aimed to understand the challenges faced by patients and identify opportunities to address specific behaviors through the provision of BCTs.

The second stage of intervention development in the BCW model is to identify the intervention options. In this case, we planned to use an mHealth platform that would be tailored to the setting and population.

The third stage is to identify the content and implementation options, including BCTs and mode of delivery. To better understand the most appropriate BCTs, we referenced the BCW taxonomy of BCTs [47] and IMB skills model of ART adherence (Figure 3) [46]. IMB is a useful behavioral theory for exploring factors that lead to adherence and is supported by robust evidence [46,48,49]. It posits that adherence-related information, motivation, and behavioral skills are the fundamental determinants of adherence to ART. The model's mediational assumption asserts that ART adherence information and motivation generally work through ART adherence behavioral skills to affect adherence behavior. We used the IMB skills model to identify the aspects of motivation, information, and behavioral skills that our intervention might target.

The intervention services were tailored based on input from the IMB skills model, input by SHIP patients during FGDs, and information from clinical service providers at the study site. The study team and clinicians worked together to write 210 health tips, script the reminder messages, and map the call flows. The lead clinic physician created a symptom-reporting algorithm. A local voice talent agency was engaged to record the content. Figure 4 provides examples of tips in each of the health tip categories.

Figure 2. The Behavior Change Wheel [45].

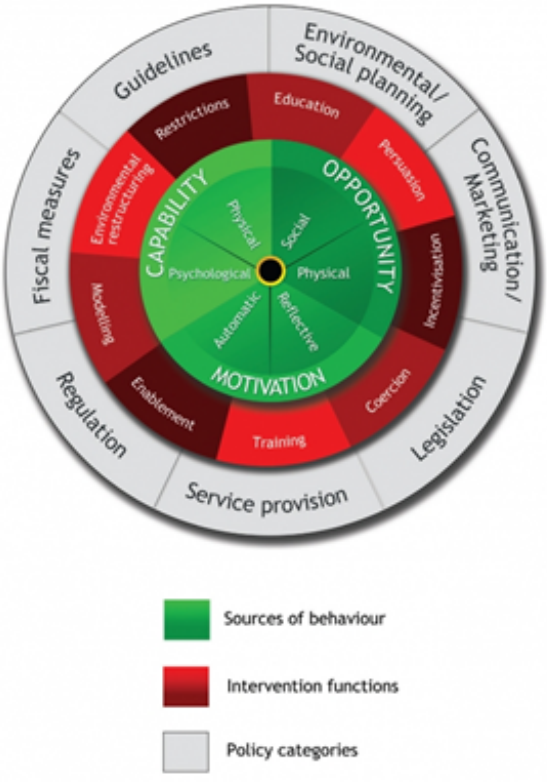


Figure 3. The information-motivation-behavioral skills model of highly active antiretroviral therapy adherence [46]. HAART: highly active antiretroviral therapy.

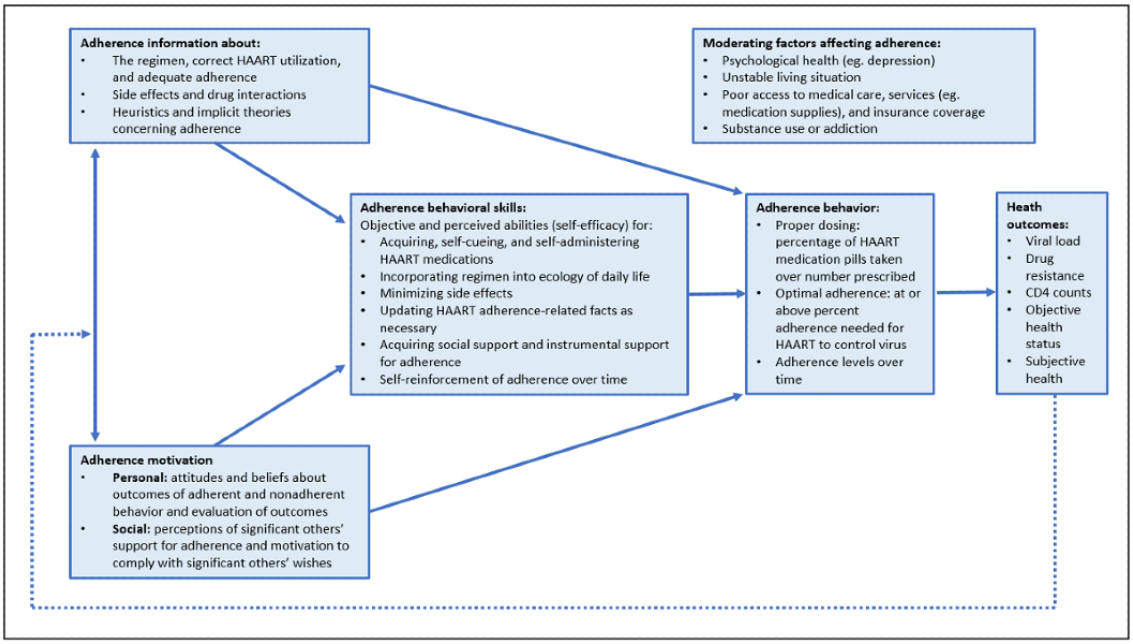
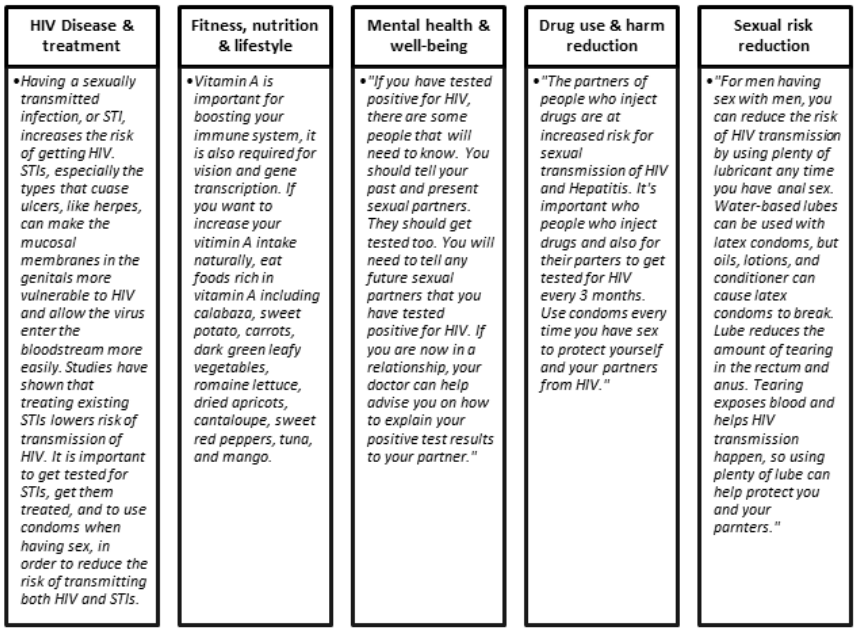


Figure 4. Health tip topic areas and sample tips. STI: sexually transmitted infection.



Local Configuration and Setup of Connect for Life Technology Platform

From 2015 to 2016, SHIP staff worked with internet technology specialists and public health professionals from the study sponsor Janssen Global Public Health to adapt the Connect for Life platform for use at the SHIP clinic. Connect for Life is a technology built on the Mobile Technology for Community Health (MOTECH) open-source software platform [50]. It enables health facilities to connect to patients via their cell phones or feature phones through IVRS or SMS text messaging. It was piloted in India and Uganda before roll out in the Philippines [51,52].

The platform has the following functionalities: pill reminders, visit reminders, symptom reporting, health tips, and adherence feedback messages. The study team collaborated with clinic physicians and software developers to adapt the various functions of the Connect for Life platform to align with the needs of the patients, as documented in the formative research phase.

Pilot Test and Preliminary Evaluation

Overview

During the first 8 weeks of piloting the intervention, 62 patients were enrolled in the study. These patients received adherence reminder calls and health tips and reported their adherence via IVRS. During this pilot test phase, the feasibility and

acceptability of the intervention were analyzed before moving to a larger scale implementation phase.

Feasibility

To assess the feasibility of the intervention, use data from the Connect for Life platform were analyzed. This included the number of calls generated from the platform, the number of calls answered by the participants, and the outcomes of those calls.

Acceptability

To assess acceptability, 2 FGDs were held to assess user experience. All eligible study participants were invited to participate in the focus groups, of which only 5 agreed to participate in a FGD (the major barrier to participation in these FGDs was the difficulty of transport owing to the traffic congestions in Metro Manila). Participants discussed their experience with Connect for Life; their reactions to the reminders, health tips, and adherence feedback; their feedback on the call length and call frequency; and their suggestions for improving the system.

Results

Understanding Drivers of ART Adherence

Review of ART Adherence Literature (Regional and Country-Specific)

Globally, approximately 40% of patients report suboptimal adherence to ART [53,54]. In the regional Therapeutics, Research, Education, and AIDS Training (TREAT) in Asia cohort (which includes 12 clinical sites from Thailand, Hong Kong, Malaysia, the Philippines, and Indonesia) of the 1316 patients, 421 (31.99%) self-reported suboptimal adherence of <100% [13]. Similar to our Connect for Life study cohort, majority of the TREAT Asia group comprised a male (67%) population and was aged <40 years (66%); however, most participants of the TREAT Asia cohort were exposed to HIV via heterosexual contact (69%), whereas our study group was primarily homosexual. The study found that the adherence rate was the lowest during the first 6 months on ART and the rate improved the longer the patient was on treatment [13].

Several key factors influencing ART adherence are well documented in the literature, including medication side effects, substance abuse, presence of social support, and time spent on treatment [13,28,54-57]. In the Philippines context, issues of stigma and discrimination also emerged as a major barrier to medication adherence [1,58,59].

Retrospective Analysis of Clinic Files

On the basis of the pharmacy refill forms for SHIP clinic patients, 67.7% (317/468) of patients reported perfect adherence in the 30 days before their most recent refill, 31.8% (149/468) reported suboptimal adherence <100%, and 20% (94/468) reported adherence <95%. A retrospective review of pharmacy refill data is summarized in Table 1.

Table 1. Sustained Health Initiatives of the Philippines clinic adherence data from pharmacy refill forms (N=682).

	Value
Demographic information (n=542), median (IQR; range)	
Age (years)	32 (28.6-35.9; 21-72)
HIV history (years)	
Time since diagnosis	3 (1.8-5; 0-25)
Time from diagnosis to antiretroviral therapy initiation	0.2 (0.1-0.9; 0-21)
Time on antiretroviral therapy	2.4 (1.5-3.9; 0-10)
Adherence estimates for patients with 30-day adherence reported at the last pharmacy refill (n=468), n (%)	
100% adherence	317 (67.7)
Missed 1 dose—adherence (95%-100%)	55 (11.8)
Missed ≥2 doses—suboptimal adherence (<95%)	94 (20.1)

FGDs With Clinic Patients

We also conducted FGDs with 1.8% (12/682) of the participants regarding their adherence challenges. All participants were male, 75% (9/12) were homosexual, 25% (3/12) were bisexual, and 67% (8/12) had full-time employment. The time patients had been on ART ranged from 5 months to 6 years, with a median time of 4 years. Overall, 83% (10/12) of the participants reported that they sometimes forgot to take their medications and 42% (5/12) had missed a dose within the past 2 weeks.

FGD findings on the causal factors for ART adherence are summarized in Table 2.

Table 2. Causal analysis of antiretroviral therapy adherence behavior.

Reason for nonadherence	Illustrative quotes from FGD[a] participants
Inconsistent daily routines, change in habits, behavioral skills, or difficulties with timing of dose	
Common reasons that patients report missing doses include simply forgetting, being busy, being away from home, and changes in routine [57].	• "You usually...take it at home, not in the office; there are some instances when you calculated the time...so you have to be in the office to take it properly. Then when you are there, you forget to take it, it's because you're busy already working." • "The challenge that I faced with ARV[b]...I think it's very essential for those working in BPO [business process outsourcing], is adjusting the time when your schedule shifts, because it has to be taken during your sleeping time...And, you know, you can't disclose, 'I was late because I overslept because I was really high with my ARVs'."
Low social support	
Patients who have a treatment support person are more likely to be adherent [60]. Having a good relationship with the HIV primary care physician and clinic staff was an important factor.	• "My partner is really helping me a lot to adhere to the schedule in taking the medications...When my partner gets too busy, the tendency is that we both forget that I need to take the medications." • "The reason why most of the patients are lost to follow up is because they feel like they are treated like patients in other [HIV treatment] hubs. The reason why we continue going to SHIP is because we feel welcome, we feel like it's like an extension of our family. Unlike in other hubs – they feel they have to wait; they don't know if they are going to die on that day or that hour. They feel that they are not that important."
Medication side effects or type of regimen	
Experiencing an adverse drug reaction is associated with poor adherence [61]. Furthermore, a large cohort study in Southeast Asia found that patients taking an NRTI[c] + NNRTI[d] regimen had poorer adherence than those who initiated on an NRTI + protease inhibitor regimen [13]. This is most likely because of difficulty tolerating the central nervous system side effects of efavirenz, a theme that was noticed throughout our focus groups.	• "If we open a fresh bottle of ARV sometimes it feels kind of strong...It's like the first time. You feel all the side effects of the ARV." • "For me it really is the headache, especially this first few weeks." • "Especially when I was having a pneumonia, especially with interactions with antibacterials – It's really hard to actually take the ARV together with the other medicines because you will be getting a really, really painful stomach, even if you ate something. So sometimes in order for me to finish the whole course of the meds that's been described I have to skip if I really can't tolerate anymore."
Shorter time on ART	
Some studies show that longer duration on ART[e] is associated with better adherence [13]. Treatment-experienced FGD participants insisted that *newbies* would benefit most from the intervention.	• "For the newbies this would be a big help because for a while it's a way for them to adjust. Not all of them are still open in discussing their status with people, and this is a first step for them to accept the fact that they have this situation that they need to cope with. And to do that, it's like the IVR is helping [them]. So, it's a big help."
Substance use or abuse	
Patients who use illicit drugs or abuse alcohol may be less likely to adhere to their medication regimen [54,62]. Among our focus group participants, use of methamphetamine in the context of *Partee n Play* emerged as a theme.	• "[When you are high on drugs] You tend to delay it more and more. When you are high you are more carefree, it's like 'I'll take it later, then later, then later'…" • "I make it a point of, I have been with my friends taking drugs, and then I know that some of them have that schedule of taking the ARV. So I make it a point that I remind them to take ARV. It's like a sisterly bond, like 'Friend, it's your time…' You have to insist. It's like a responsibility within friends."
Stress, coping abilities, or poor mental health	

Reason for nonadherence	Illustrative quotes from FGD[a] participants
People living with HIV are more likely to be affected by depression and anxiety [54,55]. Focus group participants stated that coping with a new diagnosis can be overwhelming. Interruptions in treatment for patients who have been in care for several years may be caused by episodes of depression.	• "The only reason why we really skip for days is like when you are really depressed. And drugs, with your serotonin and dopamine levels really low and you're really emotional. You tend to be like 'my life sucks and I don't want to take my meds.'" • "You mentioned harm reduction – okay, yes. Could be. Another thing we are not really addressing is mental wellness...It's one reason why we consciously skip our medication, is our mental wellness."
Stigma	
Many people living with HIV are fearful of the repercussions of disclosing their status (or having it disclosed inadvertently) to family, friends, or employers [33,55,63,64]. Focus group participants shared their fears and their experiences that disclosing their diagnosis could result in personal rejection, losing their housing (multi-generation family homes are the norm in Philippines), or being dismissed from their jobs. They do not want to be seen taking medicines around other people. The psychological challenges of coping with and accepting an HIV diagnosis during the early stages is a major factor for nonadherence.	• "For me I've been battling this on my own for 6 years. None of my relatives know that I'm positive. The only people that know that I'm positive are my friends. So, I think this reminder thing...the IVR thing, the health tips, is really good." • "I think there is one point that I when I consciously, not really skipped, but delayed it 4 to 6 hours, just because when that alarm went really crazy everyone was looking at me...There's this thing now that gay people are being judged when we take our meds in public...That's why it's hard to have that really loud alarm now."

[a]FGD: focus group discussion.

[b]ARV: antiretroviral.

[c]NRTI: nucleoside reverse transcriptase inhibitor.

[d]NNRTI: nonnucleoside reverse transcriptase inhibitor.

[e]ART: antiretroviral therapy.

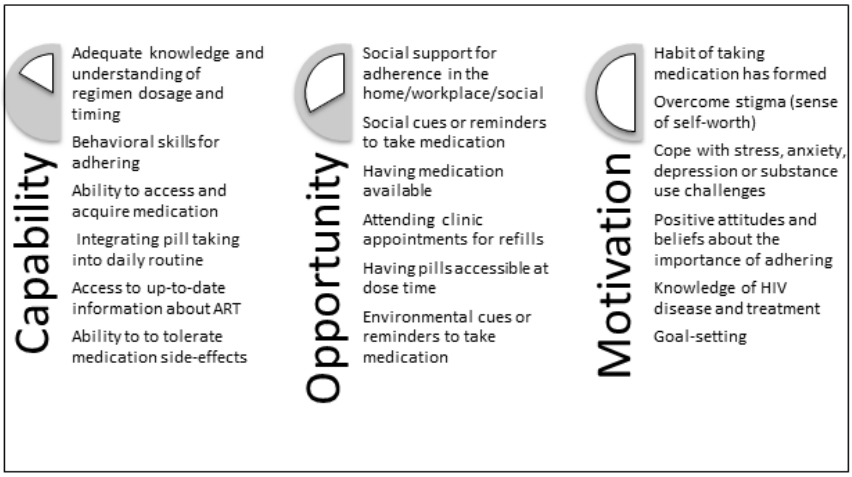
Development of the mHealth Intervention

Overview

The intervention services were tailored based on the input from SHIP patients and clinicians during formative research.

On the basis of the formative research findings, the various components of the COM-B framework were summarized, incorporating the aspects of the IMB skills model of adherence (Figure 5).

Figure 5. Summary of the components contributing to optimal antiretroviral therapy adherence based on the capability, opportunity, motivation, and behavior and information, motivation, and behavioral models. ART: antiretroviral therapy.



The focus group findings suggested that mobile phones would be an acceptable mode of delivery for HIV interventions targeting young MSM in the Philippines. During the FGDs, participants provided detailed input about the acceptability of various intervention aspects, including pill reminders, health tips, visit reminders, adherence feedback, and symptom reporting.

mHealth Intervention Preferences

Patients reported that they would like mHealth services to be personalizable. For example, patients requested to be able to select whether they receive pills and visit reminders via SMS text messages or via calls, as well as to determine the frequency and the time of the day that they receive reminders. Participants believed that newer patients who recently started ART would benefit most from daily pill reminders and experienced patients

would prefer less frequent reminders. They stated that it would be important to be able to opt in or out of any call or SMS text messaging services at any time.

Patients expressed a strong interest in health tips that covered a variety of topics, not strictly HIV disease, and they wanted to personally select which categories of health tips they would hear.

In summary, patients suggested that an ideal mHealth service should be personalized based on the following factors: (1) call or SMS text messages, (2) timing of calls or messages, (3) frequency of calls or messages, and (4) content or topic areas for health education messages.

mHealth Intervention Configuration

The investigators created a standard service scheme (Table 3) that could be adjusted at the patient's request. The recommended service scheme included *pill reminders* for all patients. As focus groups and literature review suggested that more intensive adherence support is required in the early stages of HIV treatment while forming of habits, ART-naïve patients and patients in their first 24 weeks of ART received reminders daily for the first 24 weeks and weekly for the next 24 weeks. Patients who were experienced with ART received weekly reminders for 24 weeks and no reminders after 48 weeks.

During IVRS pill reminder calls, patients were prompted to *report symptoms or side effects* of medications using an IVRS touch-tone menu. The patients received SMS text message recommendations for over-the-counter medications and advice depending on the algorithm outcome. The system automatically generated an alert for the clinician of any symptom reports that required urgent attention.

All patients received SMS *appointment reminders* at 2 set times in advance of their scheduled clinic visit date.

All patients who received IVRS calls for their pill reminders could receive a weekly *adherence feedback* message informing them of their *score*—from 0 to 7—based on the number of days they reported taking their doses in the prior week via the IVRS platform. The adherence feedback score was followed by a short motivational message to encourage improvement among patients with low adherence or support continued good adherence among patients with high adherence scores.

Patients would automatically receive audio *health tips* when they received pill reminder calls, or they could opt to receive health tips via SMS text messages. For patients new on ART, there was a tailored set of health tips that explained the basics of HIV and ART. In addition, we created tips on a variety of other health topics based on the suggestions of patients from the FGDs. The following five broad categories were selected for health tips: (1) HIV disease and treatment, which include tips about HIV testing and diagnosis, transmission of HIV, coinfections, and laboratory tests for people living with HIV; (2) fitness, nutrition, or lifestyle, which included tips for exercise and eating healthy; (3) mental health or well-being, which included tips on acceptance and disclosure of HIV status, and approaches for understanding and dealing with depression, anxiety, and stress; (4) drug use and harm reduction, which included medical information about common recreational drugs, safer injection, and hepatitis C; and (5) sexual risk reduction, which included tips on condoms and lubricants and tips on leading a healthy sex life with HIV.

The investigators worked with 2 clinic providers and a local voice talent agency to write and record 210 health tips that related to the common questions and issues raised by patients and tips that incorporated the themes that emerged from the focus groups. The messages were crafted ensuring that they not only provided didactic information related to the health topic but also ended with a specific action or behavior that the patient could adopt to improve or to minimize the impact of a specific behavior.

The system was configured to protect patient privacy and prevent unintended disclosure of health information. Upon answering any call from the system, the patient would immediately hear a *jingle*, a song that was associated with the Connect for Life system. Upon hearing the *jingle*, they would enter a personal identification number to advance to the next step of the call. No health-related information would be transmitted unless the personal identification number was keyed in, to protect patient privacy and confidentiality.

Services in the intervention package address the 3 main components of the COM-B model. Capability is addressed through health tips, which aim to improve knowledge regarding ART and HIV disease and improve behavioral skills. Opportunity is addressed through the pill reminder service, which provides an external prompt or cue for pill taking and supports habit forming through the appointment reminder service, which prompts attendance at the clinic for refill, thereby increasing accessibility and availability of medications; and through the symptom-reporting algorithm, which addresses the medical barriers to pill taking by expediting a response to side effects or medication reactions. Motivation is addressed through health tips (eg, messages designed to help with stress, overcome stigma, and inform positive attitude toward pill taking) and adherence feedback messages, which reward and reinforce high adherence and encourage improvement for low adherence.

Table 3 presents the proposed service scheme. However, the services were flexible and a patient could opt out of any call or SMS text messaging service that they did not wish to receive or opt into services depending on their preference and the clinician's judgment. The clinician could reactivate or extend the pill reminders for patients who needed additional support.

Table 3. Connect for Life services scheme.

Patient characteristics	Pill reminder and adherence feedback messages (voice or SMS text messages)	Health tips (voice or SMS text messages)	Appointment reminders (voice or SMS text messages)	Symptom reporting (voice calls only)
Treatment naïve and recently initiated (<6 months on antiretroviral therapy), or treatment experienced more than 6 months with adherence <80% in the 30 days before enrollment	• Daily reminders from 0 to 24 weeks • Weekly reminders from 25 to 48 weeks	Health tips play during all pill reminder calls; health tips topics tailored to new patients	Yes	Yes (during all pill reminder calls)
Treatment experienced >6 months with adherence ≥80% in the 30 days before enrollment	• Weekly reminders from 0 to 24 weeks • No reminders from 24 to 48 weeks	Health tips frequency and topics selected based on the preference of clinician and patient	Yes	Yes (during all pill reminder calls)

Pilot Test and Preliminary Evaluation Findings

A pilot test phase was conducted from October 2017 to January 2018, in which 62 patients were enrolled in the service. During the pilot test period, we received reports of several technical issues that affected the functionality of the system. In all, 2 FGDs were held in January and February 2018, after approximately 3 months of the pilot project implementation. There were 5 participants—3 in one discussion and 2 in the next. FGD findings on the themes that emerged from the pilot test are summarized in Table 4.

On the basis of the findings from the pilot phase FGDs, enrollment in the study was suspended because of pending solutions to technical issues. The study team worked with software developers to trace the source of the technical issues. It was determined that the platform was functioning well and that the technical failures were because of issues within the local telecommunications infrastructure (ie, poor call quality). After the team addressed all the technical issues on the software development side, enrollment continued with SMS text messaging services only.

Since mid-2019, we have found that as telecom services improved in the Philippine setting, voice calls in the Connect for Life system can now be delivered with fewer technical issues. Following the initial pilot study, the intervention was scaled up at the SHIP clinic and currently serving 1491 patients at 2 HIV clinics. The platform is being further developed to move from the MOTECH base to open medical record system. We plan to pilot test the new version in several HIV treatment sites across the Philippines.

Table 4. Themes emerging from the pilot test evaluation.

Themes	Illustrative quotes from FGD[a] participants
CfL[b] technical issues or functionality	
• PIN[c] issues: DTMF[d] is the signal to the phone company that is generated when a user presses a telephone's touch keys. All FGD participants reported instances in which they attempted to enter their PIN code and the code was not recognized. The frequency of DTMF problems varied widely among the participants. • Call origin: Calls are generated from an interactive voice response platform using a United States–based telecom provider. The interactive voice response service provider sets the incoming call number to be displayed as the patient's own phone number. However, patients reported that this was inconsistent and that some of the incoming calls from the CfL system that they answered displayed phone numbers originating in different countries.	• "Actually, I just experienced that issue last night [DTMF malfunction]. Sometimes I have been able to enter [my PIN code] and sometimes I haven't. The jingle kept going on, so I kept entering the PIN again and nothing happened, so I just hung up." • "In my case, I think I received thrice already from various locations an unknown number that's why I didn't bother answering. One from South Korea, one from US and one from China. The problem is if the number is unknown basically I don't answer it. I'm just guessing that the number came from CfL."
SMS vs voice call preferences	
• There was mixed feedback about whether SMS text messages or voice calls were more effective or acceptable. Some participants said that the frequency and length of voice calls were too much. Several FGD participants requested to be changed from voice calls to SMS text messages, as texting is more convenient and less intrusive. Others preferred to stay on voice calls as they are more difficult to ignore.	• "I think SMS would be nice to have as an option. If at the time the program calls you but you didn't answer an SMS reminder would be good just to keep in touch." • "I hated the call because I've been receiving the calls especially when I'm on my way home in an Uber. If I mistakenly answer it without the headset, the voice will be loud and basically everyone in the Uber would know." • "If you are going to put the schedule of the consultation, I'd rather those to be in text because there's too much information that I need to remember." • "I think it's also cultural when people don't like answering calls. Mostly Asians I know don't like answering calls. I'm not good at answering calls and most of the people I know don't also like answering calls especially if the number is unknown and overseas and then you hear this very gloomy guy voice."
Adherence feedback *gamification*	
• Participants did not like receiving adherence feedback scores because it was inaccurate and made them feel stressed.	• "[The adherence feedback score] has no effect. It has no significance to me." • "For me I don't even care about it because it just stresses me out."
Pill reminders	
• Daily pill reminder calls were not as used as expected by the study team based on the findings from the first 2 focus groups. After the pilot phase, patients reported that, although they like the idea of regular reminder calls, in practice, they are often too busy to answer the calls and report their adherence. • The issue of poor uptake of pill reminder calls was further compounded by the technical issues with the entering their PIN code (DTMF issue). • Some participants said that the pill reminders did not make a big difference for them as they already had other systems in place to remind them to take their medications.	• "It helped. Sometimes I would forget but it would help to remind me because I usually take my pill after work, and after work I'm just so tired, I don't check the time and sometimes I almost forget because I'm so sleepy." • "I hate to be reminded that I have this condition every single day. I know I need to take it but I don't need to be reminded every single day that I have to." • "If you call seven times a week that's a bit irritating for the patient. What the patient can do is have the option to get reminded through text."
Health tips	
• The content of the health tips was useful and informative. All participants wanted to continue to receive health tips. • Some participants would prefer SMS text messages rather than voice recordings for the health tips. Some thought the voice recording spoke too slowly; therefore, they would prefer to read it by SMS text messages. • One technical issue reported was that sometimes the same health tips were received for multiple days instead of receiving a new tip each day, as intended.	• "The health tips are super helpful. Those are the tips about alcohol, and that say you can have sex, you are not prevented but protected. There are even those great tips on eating and what you should eat." • "Just the voice. The girl answering the questions in the health tips is okay. The guy is very depressing."

Themes	Illustrative quotes from FGD[a] participants
Other findings	
• Participants were enthusiastic about receiving the automated reminders for their clinic appointments. • Participants stated they would have liked a more in-depth orientation or onboarding process at the outset of the intervention. They emphasized the importance of onboarding, setting expectations, and a thorough explanation of the intervention. • Not all participants understood they could change or adapt the service model. • Peer support: Participants mentioned that they found participating in a FGD with other people living with HIV very helpful and asked if there could be an opportunity for the clinic to organize in person support groups.	• "But what I noticed was that it helped with the appointment. That was a big help as I was reminded that I had to go to the clinic. That's a big deal to me. But about missing the meds, it's still human." • "I think the program's good. I could recommend that for the newbies. I think the program should be laid on properly. For example, scheduling, the time, reminders, and the tips. Maybe after a month if the patient has already established a routine so maybe it could lessen the reminders." • "Besides, the importance of the support group is for patients who have not disclosed to family members. There you can get support or have conversations like this. If there were a support group now, I'd want to be a part of it because I would like to share what I have experienced before with others."

[a]FGD: focus group discussion.

[b]CfL: Connect for Life.

[c]PIN: personal identification number.

[d]DTMF: dial tone multifrequency.

Discussion

Principal Findings

The intervention development approach resulted in an mHealth intervention tailored to the information needs and communication preferences of MSM in the Philippines. The intervention was designed to address various aspects of capability, opportunity, and motivation to achieve optimal adherence to ART.

The formative research found that mobile phone use is widespread in the Philippines and that mobile phones are an acceptable mode of communication for health information and adherence support. The literature review and FGDs revealed that in our patient population, key behavioral barriers to adherence included challenges around forming consistent routines and habits, low social support, stress and mental health issues, substance use, and social stigma of living with HIV. Focus group participants strongly emphasized the need for social and family support to enable and encourage good adherence to ART. Key clinical issues affecting adherence included medication side effects (especially among efavirenz-based regimens) and shorter duration on ART.

Following the pilot test, recipients of the intervention reported that the tone, frequency, and content of the voice messages were acceptable and appropriate. In the prepilot focus groups, participants preferred the male voice actor whose voice sounded more "attractive" according to several participants, whereas in the postpilot groups, several participants mentioned that they preferred the female voice actor because her tone was warmer and she came across as a trusted friend. This finding indicates that more iterations of recording should be tested in future implementations before a full-scale roll out and that budgets and project work plans should allow for several rounds of recording. The accounts of the focus group participants indicated that the intervention increased their knowledge and adherence behaviors. However, a large-scale cohort is needed to assess the intervention's effectiveness.

In the prepilot focus groups, participants were enthusiastic about receiving voice reminders via phone calls, following their experience of participating in the pilot most participants expressed a preference for SMS text messaging over voice calls. This preference may have been related to the inconvenience of answering phone calls, and it may also have been related to the technical problems experienced with the IVRS. These technical challenges posed a significant challenge to the feasibility of the intervention, and delivery would need to be adapted to allow for SMS text messaging options to achieve full-scale implementation.

Strengths and Limitations

The strength of the intervention development process was the participatory approach, which included the beneficiaries or users of the potential intervention, clinical service providers, and developers of the technology platform. The views of the target audience were collected during focus groups, which informed the tone, style, frequency, duration, and content of the intervention.

The BCW is a robust intervention development approach that provides a comprehensive understanding of the sources of a behavior, spectrum of intervention functions, and environment in which the behavior occurs. A strength of this approach is the COM-B model at the hub of the BCW. By identifying the capabilities, opportunities, and motivations behind a behavior, we can clearly identify the most relevant intervention approaches and BCTs. This approach allowed us to develop a solid intervention plan that described the technique, mode, and content to address each identified barrier to or enabler of ART adherence.

A weakness of our approach was the sampling and recruitment strategy for the participants in the focus groups. It was a challenge to identify patients who were willing to participate

in a group where everyone had a HIV-positive status because many patients were not publicly *out* as people living with HIV, indicating that individual interviews may be an option in future studies. The patients who agreed to participate may not be representative of the wider patient population, introducing a degree of selection bias to the process. The study only included patients >18 years and the patients were almost exclusively male; thus, the findings do not address the distinct needs and challenges of adolescents and women living with HIV. In addition, there was low attendance among those who confirmed their intention to participate in the focus groups. This is reflective of the larger need to provide differentiated models of care in the Philippines, as transportation to the clinic site is not easy in Metro Manila because of traffic congestion.

Another weakness of our approach was that we had an intervention mode in mind—mobile phone—at the outset of the intervention development process. Although there are several key determinants of adherence that the Connect for Life platform can address (ie, knowledge, habit forming, and environmental cues to take medication), there are other factors that the mHealth approach does not address (ie, physical availability of medication and social support).

Notably, the technical challenges experienced in delivering the intervention during the pilot phase made it difficult to assess the true acceptability and feasibility of the planned intervention. Feedback received after the pilot phase focused largely on the mobile phone functionality issues, which then limited the discussion regarding the content and design of the intervention as it was intended to be delivered. Conducting a small pilot phase with a few participants allowed us to identify the problems with functionality and adapt the intervention before scaling up the intervention to the larger cohort; however, a more iterative process with several pilot stages would have been advantageous if budget and timeline had allowed us to do so.

Comparison With Prior Work

Research on ART adherence has shown that less time on ART is associated with an increased risk of poor adherence [13,65-67]. With this in mind, the intervention was designed with more frequent (daily) pill reminders for patients during their first 6 months on ART and less frequent (weekly) reminders for patients with longer than 6 months on ART. However, after the intervention design was completed and pilots, an analysis of the Philippine cohort found a different trend within the study population, observing that, even before receiving the intervention, newer patients in the Connect for Life cohort tended to be more adherent compared with patients who had taken ART for longer and showed signs of *treatment fatigue* [68]. This highlights the importance of the ability of clinicians to tailor the reminder frequency and other intervention functions based on individual patient needs.

Before this study, 2 other projects using the same technology that the Connect for Life program was built on were implemented and evaluated in India and Uganda. First, a program called Treatment Advice by Mobile Alerts (TAMA), provided people living with HIV in India with daily or weekly pill reminders, adherence feedback, automated algorithms for managing clinical events for patients being initiated on ART, health tips, appointment reminders, and real-time reporting to the clinics of patient interaction with TAMA. Evaluation of the TAMA pilot found that patients gave the platform a high system usability score and gave generally positive feedback about their experience with using the technology. In TAMA, patients could call a toll-free number to access health tips and a clinical event algorithm. Health tips were used by 76% (42/55) of the patients, and automated clinical advice was accessed by 64% (35/55) of the participants in the pilot study. In the Philippines, these functions were available only through outgoing system-generated calls and SMS text messaging because of the prohibitive cost of toll-free inbound telephone lines in the Philippine setting [51,69].

The second project, the Call for Life Uganda program, also found good uptake, acceptability, and positive response to the system. In Uganda, there was a strong preference for interactive voice response over SMS text messages, which was different from the Philippines where participants preferred SMS text messages [70,71].

Conclusions

Our research found that a mobile phone–based SMS text messaging intervention and IVRS intervention were acceptable to MSM in Manila, the Philippines, and the FGDs suggested that it helped promote ART adherence and appointment attendance. A randomized controlled trial is required to establish the effects of the intervention on the clinical outcomes of HIV care and treatment.

Acknowledgments

This study received sponsorship from Johnson & Johnson Global Public Health. The authors would like to thank the following collaborators in the Johnson & Johnson team for developing the Connect for Life platform and for their extensive work to tailor the technology to the Philippine setting: Paula McKenna, Piet Knaepen, Avinash Agrawal, and Jurgen de Beckker. In addition, they would like to thank the implementing partner, that is, the Sexually Transmitted Infection/AIDS Guidance Intervention and Prevention Unit at the Philippine General Hospital, without whose efforts the study would not have been possible.

Conflicts of Interest

RG is an employee of Janssen Pharmaceutica NV. Connect for Life platform was developed and funded by Janssen Pharmaceutica NV as part of its commitment to Global Public Health to actively build health communities worldwide through innovative and impactful health solutions and partnership. This project was funded by Janssen Pharamceutica NV under that project.

References

1. Gangcuangco L. HIV crisis in the Philippines: urgent actions needed. Lancet Public Health 2019 Feb;4(2):e84. [doi: 10.1016/s2468-2667(18)30265-2]
2. Geronimo J. WHO: PH has fastest growing HIV epidemic in the world. Rappler. 2016. URL: http://www.rappler.com/nation/93839-who-ph-response-hiv [accessed 2022-01-12]
3. The Lancet. Losing the fight against HIV in the Philippines. Lancet 2017 Aug;390(10095):626. [doi: 10.1016/s0140-6736(17)32211-0]
4. HIV/AIDS and ART Registry of the Philippines (HARP) Report April-June 2020. National HIV/AIDS & STI Surveillance and Strategic Information Unit, Manila, Philippines. 2020. URL: http://www.doh.gov.ph/sites/default/files/statistics/EB_HIV_Mar-AIDSreg2016.pdf [accessed 2022-01-12]
5. Joint United Nations Program on HIV/AIDS (UNAIDS). UNAIDS Data 2017. 2017 Jul 20. URL: http://www.unaids.org/sites/default/files/media_asset/20170720_Data_book_2017_en.pdf [accessed 2022-01-16]
6. Philippines Country Data 2020. Joint United Nations Programme on HIV/AIDS (UNAIDS). 2020. URL: https://www.aidsdatahub.org/sites/default/files/resource/2020-aids-data-book-ph.pdf [accessed 2022-01-12]
7. HIV/AIDS and ART Registry of the Philippines (HARP) Report February 2021. National HIV/AIDS & STI Surveillance and Strategic Information Unit, Manila, Philippines. 2021. URL: http://www.doh.gov.ph/sites/default/files/statistics/EB_HIV_Mar-AIDSreg2016.pdf [accessed 2022-01-12]
8. 2014 Global AIDS Response Progress Reporting Country Progress Report Philippines. UNAIDS. 2014. URL: https://www.unaids.org/sites/default/files/country/documents/PHL_narrative_report_2014.pdf [accessed 2022-01-12]
9. 2015 Integrated HIV Behavioral and Serologic Survey (IHBSS) Fact Sheets. National HIV/AIDS & STI Surveillance and Strategic Information Unit. 2016. URL: http://www.doh.gov.ph/sites/default/files/publications/2015IntegratedHIVBehavioralandSerologicSurveillanceFactsheets.pdf [accessed 2022-01-12]
10. 2013 Integrated Behavioral and Serologic Surveillance (IHBSS) Report. National Epidemiology Center Department of Health Philippines. 2014. URL: http://www.doh.gov.ph/sites/default/files/publications/2013IHBSSFinalReport(XII.2014).pdf [accessed 2022-01-12]
11. de Lind van Wijngaarden JW, Ching AD, Settle E, van Griensven F, Cruz RC, Newman PA. "I am not promiscuous enough!": exploring the low uptake of HIV testing by gay men and other men who have sex with men in Metro Manila, Philippines. PLoS One 2018 Jul 6;13(7):e0200256 [FREE Full text] [doi: 10.1371/journal.pone.0200256] [Medline: 29979766]
12. Leyritana K, Alejandria M, Tenorio A, Salvaña E. The evolving epidemiology of people living with HIV in a developing country in the midst of an epidemic: a descriptive cohort. Poster Abstract Session: Global HIV. 2013. URL: https://idsa.confex.com/idsa/2013/webprogram/Paper40015.html [accessed 2022-01-12]
13. Jiamsakul A, Kumarasamy N, Ditangco R, Li PC, Phanuphak P, Sirisanthana T, TREAT Asia Studies to Evaluate Resistance Monitoring Study (TASER-M). Factors associated with suboptimal adherence to antiretroviral therapy in Asia. J Int AIDS Soc 2014;17:18911 [FREE Full text] [doi: 10.7448/IAS.17.1.18911] [Medline: 24836775]
14. Republic of the Philippines, Manila, Philippines : external review of the national health sector response to HIV and sexually transmitted infections 2013. World Health Organization. 2015. URL: http://iris.wpro.who.int/bitstream/handle/10665.1/11217/9789290617068_eng.pdf [accessed 2022-01-12]
15. Restar A, Nguyen M, Nguyen K, Adia A, Nazareno J, Yoshioka E, et al. Trends and emerging directions in HIV risk and prevention research in the Philippines: a systematic review of the literature. PLoS One 2018 Dec 5;13(12):e0207663 [FREE Full text] [doi: 10.1371/journal.pone.0207663] [Medline: 30517178]
16. Free C, Phillips G, Galli L, Watson L, Felix L, Edwards P, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. PLoS Med 2013;10(1):e1001362 [FREE Full text] [doi: 10.1371/journal.pmed.1001362] [Medline: 23349621]
17. Baller S, Dutta S, Lanvin B. The Global Information Technology Report 2016. World Economic Forum. 2016. URL: https://www3.weforum.org/docs/GITR2016/WEF_GITR_Full_Report.pdf [accessed 2022-01-12]
18. World Telecommunication/ICT Indicators Database. International Telecommunication Union (ITU). 2019. URL: https://doi.org/10.7910/DVN/IP5UBV [accessed 2022-01-12]
19. Fenwick S. Philippines : Mobile Network Experience Report. Open Signal. 2020. URL: https://www.opensignal.com/reports/2020/11/philippines/mobile-network-experience [accessed 2022-01-12]
20. World Health Organization. Consolidated Guidelines on the Use of Antiretroviral Drugs for Treating and Preventing HIV Infection: Recommendations for a Public Health Approach, 2nd Ed. Geneva: World Health Organization; 2016.
21. Kanters S, Park JJ, Chan K, Socias ME, Ford N, Forrest JI, et al. Interventions to improve adherence to antiretroviral therapy: a systematic review and network meta-analysis. Lancet HIV 2017 Jan;4(1):31-40. [doi: 10.1016/S2352-3018(16)30206-5]
22. Lester RT, Ritvo P, Mills EJ, Kariri A, Karanja S, Chung MH, et al. Effects of a mobile phone short message service on antiretroviral treatment adherence in Kenya (WelTel Kenya1): a randomised trial. Lancet 2010 Nov 27;376(9755):1838-1845. [doi: 10.1016/S0140-6736(10)61997-6] [Medline: 21071074]

23. Finitsis DJ, Pellowski JA, Johnson BT. Text message intervention designs to promote adherence to antiretroviral therapy (ART): a meta-analysis of randomized controlled trials. PLoS One 2014;9(2):e88166 [FREE Full text] [doi: 10.1371/journal.pone.0088166] [Medline: 24505411]
24. Mbuagbaw L, van der Kop ML, Lester RT, Thirumurthy H, Pop-Eleches C, Ye C, et al. Mobile phone text messages for improving adherence to antiretroviral therapy (ART): an individual patient data meta-analysis of randomised trials. BMJ Open 2013;3(12):e003950 [FREE Full text] [doi: 10.1136/bmjopen-2013-003950] [Medline: 24345901]
25. Mbuagbaw L, Sivaramalingam B, Navarro T, Hobson N, Keepanasseril A, Wilczynski NJ, Patient Adherence Review (PAR) Team. Interventions for Enhancing Adherence to Antiretroviral Therapy (ART): a systematic review of high quality studies. AIDS Patient Care STDS 2015 May;29(5):248-266 [FREE Full text] [doi: 10.1089/apc.2014.0308] [Medline: 25825938]
26. Bigna JJ, Noubiap JJ, Kouanfack C, Plottel CS, Koulla-Shiro S. Effect of mobile phone reminders on follow-up medical care of children exposed to or infected with HIV in Cameroon (MORE CARE): a multicentre, single-blind, factorial, randomised controlled trial. Lancet Infect Dis 2014 Jul;14(7):600-608. [doi: 10.1016/S1473-3099(14)70741-8] [Medline: 24932893]
27. da Costa TM, Barbosa BJ, Gomes EC, Sigulem D, de Fátima MH, Filho AC, et al. Results of a randomized controlled trial to assess the effects of a mobile SMS-based intervention on treatment adherence in HIV/AIDS-infected Brazilian women and impressions and satisfaction with respect to incoming messages. Int J Med Inform 2012 Apr;81(4):257-269 [FREE Full text] [doi: 10.1016/j.ijmedinf.2011.10.002] [Medline: 22296762]
28. Dowshen N, Kuhns LM, Johnson A, Holoyda BJ, Garofalo R. Improving adherence to antiretroviral therapy for youth living with HIV/AIDS: a pilot study using personalized, interactive, daily text message reminders. J Med Internet Res 2012 Apr 05;14(2):e51 [FREE Full text] [doi: 10.2196/jmir.2015] [Medline: 22481246]
29. Horvath T, Azman H, Kennedy GE, Rutherford GW. Mobile phone text messaging for promoting adherence to antiretroviral therapy in patients with HIV infection. Cochrane Database Syst Rev 2012;3:CD009756. [doi: 10.1002/14651858.CD009756] [Medline: 22419345]
30. Chib A, van Velthoven MH, Car J. mHealth adoption in low-resource environments: a review of the use of mobile healthcare in developing countries. J Health Commun 2015 Mar;20(1):4-34. [doi: 10.1080/10810730.2013.864735] [Medline: 24673171]
31. Pop-Eleches C, Thirumurthy H, Habyarimana JP, Zivin JG, Goldstein MP, de Walque D, et al. Mobile phone technologies improve adherence to antiretroviral treatment in a resource-limited setting: a randomized controlled trial of text message reminders. AIDS 2011 Mar 27;25(6):825-834 [FREE Full text] [doi: 10.1097/QAD.0b013e32834380c1] [Medline: 21252632]
32. Hardy H, Kumar V, Doros G, Farmer E, Drainoni M, Rybin D, et al. Randomized controlled trial of a personalized cellular phone reminder system to enhance adherence to antiretroviral therapy. AIDS Patient Care STDS 2011 Mar;25(3):153-161 [FREE Full text] [doi: 10.1089/apc.2010.0006] [Medline: 21323532]
33. Mbuagbaw L, Bonono-Momnougui RC, Thabane L. Considerations in using text messages to improve adherence to highly active antiretroviral therapy: a qualitative study among clients in Yaoundé, Cameroon. HIV AIDS (Auckl) 2012;4:45-50 [FREE Full text] [doi: 10.2147/HIV.S29954] [Medline: 22570574]
34. Shet A, De Costa A, Kumarasamy N, Rodrigues R, Rewari BB, Ashorn P, HIVIND Study Team. Effect of mobile telephone reminders on treatment outcome in HIV: evidence from a randomised controlled trial in India. Br Med J 2014 Oct 24;349:g5978 [FREE Full text] [doi: 10.1136/bmj.g5978] [Medline: 25742320]
35. Maduka O, Tobin-West CI. Adherence counseling and reminder text messages improve uptake of antiretroviral therapy in a tertiary hospital in Nigeria. Niger J Clin Pract 2013;16(3):302-308 [FREE Full text] [doi: 10.4103/1119-3077.113451] [Medline: 23771450]
36. 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]
37. Krishnan A, Ferro EG, Weikum D, Vagenas P, Lama JR, Sanchez J, et al. Communication technology use and mHealth acceptance among HIV-infected men who have sex with men in Peru: implications for HIV prevention and treatment. AIDS Care 2015;27(3):273-282 [FREE Full text] [doi: 10.1080/09540121.2014.963014] [Medline: 25285464]
38. Uhrig JD, Lewis MA, Bann CM, Harris JL, Furberg RD, Coomes CM, et al. Addressing HIV knowledge, risk reduction, social support, and patient involvement using SMS: results of a proof-of-concept study. J Health Commun 2012 May 02;17 Suppl 1(sup1):128-145. [doi: 10.1080/10810730.2011.649156] [Medline: 22548606]
39. UNAIDS Meeting Report: Information and Communications Technologies : engaging the private sector and communities in HIV programmes with gay men and other men who have sex with men. UNAIDS. 2016. URL: https://www.unaids.org/en/resources/documents/2016/20160504_JC2824_ICTs_en.pdf [accessed 2022-01-12]
40. Schnall R, Travers J, Rojas M, Carballo-Diéguez A. eHealth interventions for HIV prevention in high-risk men who have sex with men: a systematic review. J Med Internet Res 2014;16(5):e134 [FREE Full text] [doi: 10.2196/jmir.3393] [Medline: 24862459]

41. Zogg JB, Woods SP, Sauceda JA, Wiebe JS, Simoni JM. The role of prospective memory in medication adherence: a review of an emerging literature. J Behav Med 2012 Feb 13;35(1):47-62 [FREE Full text] [doi: 10.1007/s10865-011-9341-9] [Medline: 21487722]
42. Gwadz M, Applegate E, Cleland C, Leonard NR, Wolfe H, Salomon N, Heart to Heart Collaborative Research Team. HIV-infected individuals who delay, decline, or discontinue antiretroviral therapy: comparing clinic- and peer-recruited cohorts. Front Public Health 2014;2:81 [FREE Full text] [doi: 10.3389/fpubh.2014.00081] [Medline: 25077137]
43. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
44. Michie S. Which behaviour change approach should I choose? An introduction to the Behaviour Change Wheel. University College London. 2012. URL: https://docplayer.net/amp/27630436-The-behaviour-change-wheel-a-system-for-designing-effective-interventions.html [accessed 2022-01-12]
45. Michie S, Atkins L, West R. The Behaviour Change Wheel: A Guide to Designing Interventions. Sutton, United Kingdom: Silverback Publishing; 2014.
46. Fisher JD, Amico KR, Fisher WA, Harman JJ. The information-motivation-behavioral skills model of antiretroviral adherence and its applications. Curr HIV/AIDS Rep 2008 Nov;5(4):193-203. [doi: 10.1007/s11904-008-0028-y] [Medline: 18838059]
47. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol 2008 May;27(3):379-387. [doi: 10.1037/0278-6133.27.3.379] [Medline: 18624603]
48. Amico KR, Toro-Alfonso J, Fisher JD. An empirical test of the information, motivation and behavioral skills model of antiretroviral therapy adherence. AIDS Care 2005 Aug 20;17(6):661-673. [doi: 10.1080/09540120500038058] [Medline: 16036253]
49. Starace F, Massa A, Amico KR, Fisher JD. Adherence to antiretroviral therapy: an empirical test of the information-motivation-behavioral skills model. Health Psychol 2006 Mar;25(2):153-162. [doi: 10.1037/0278-6133.25.2.153] [Medline: 16569106]
50. About MOTECH. Grameen Foundation. 2015. URL: http://docs.motechproject.org/en/latest/intro.html [accessed 2022-01-12]
51. Joshi A, Saple D, Sen K, Veldeman E, Rutten R, Rane M, et al. Supporting treatment of people living with HIV/AIDS in resource limited settings with IVRs. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2014 Presented at: CHI '14: CHI Conference on Human Factors in Computing Systems; April 26 - May 1, 2014; Toronto Ontario Canada p. 1595-1604. [doi: 10.1145/2556288.2557236]
52. Joshi R. Uniting mobile, education and pharma to combat HIV. EnterpriseInnovation. 2015. URL: https://www.janssen.com/apac/news-center/janssen-in-the-news/uniting-mobile-education-and-pharma-to-combat-HIV [accessed 2016-04-29]
53. Bezabhe WM, Chalmers L, Bereznicki LR, Peterson GM. Adherence to antiretroviral therapy and virologic failure: a meta-analysis. Medicine (Baltimore) 2016 Apr;95(15):e3361 [FREE Full text] [doi: 10.1097/MD.0000000000003361] [Medline: 27082595]
54. Ortego C, Huedo-Medina TB, Llorca J, Sevilla L, Santos P, Rodríguez E, et al. Adherence to highly active antiretroviral therapy (HAART): a meta-analysis. AIDS Behav 2011 Oct;15(7):1381-1396. [doi: 10.1007/s10461-011-9942-x] [Medline: 21468660]
55. Safren SA, Biello KB, Smeaton L, Mimiaga MJ, Walawander A, Lama JR, PEARLS (ACTG A5175) Study Team. Psychosocial predictors of non-adherence and treatment failure in a large scale multi-national trial of antiretroviral therapy for HIV: data from the ACTG A5175/PEARLS trial. PLoS One 2014 Aug 25;9(8):e104178 [FREE Full text] [doi: 10.1371/journal.pone.0104178] [Medline: 25153084]
56. Grierson J, Koelmeyer RL, Smith A, Pitts M. Adherence to antiretroviral therapy: factors independently associated with reported difficulty taking antiretroviral therapy in a national sample of HIV-positive Australians. HIV Med 2011 Oct;12(9):562-569 [FREE Full text] [doi: 10.1111/j.1468-1293.2011.00928.x] [Medline: 21554524]
57. Chesney MA, Ickovics JR, Chambers DB, Gifford AL, Neidig J, Zwickl B, et al. Self-reported adherence to antiretroviral medications among participants in HIV clinical trials: the AACTG adherence instruments. Patient Care Committee and Adherence Working Group of the Outcomes Committee of the Adult AIDS Clinical Trials Group (AACTG). AIDS Care 2000 Jun;12(3):255-266. [doi: 10.1080/09540120050042891] [Medline: 10928201]
58. Santos A. Dying of shame and AIDS in the Philippines. Deutsche Welle. 2016 May 30. URL: http://www.dw.com/en/dying-of-shame-and-aids-in-the-philippines/a-19292896 [accessed 2022-01-12]
59. Santos AP. Sex and sensibilities podcast: HIV epidemic in the PH. Rappler. 2015. URL: https://www.rappler.com/moveph/111434-hiv-epidemic-philippines-unaids/ [accessed 2022-01-12]
60. Fisher JD, Fisher WA, Amico KR, Harman JJ. An information-motivation-behavioral skills model of adherence to antiretroviral therapy. Health Psychol 2006 Jul;25(4):462-473. [doi: 10.1037/0278-6133.25.4.462] [Medline: 16846321]
61. Rajesh R, Sudha V, Varma D, Sonika S. Association between medication adherence outcomes and adverse drug reactions to highly active antiretroviral therapy in Indian human immunodeficiency virus-positive patients. J Young Pharm 2012 Oct;4(4):250-260 [FREE Full text] [doi: 10.4103/0975-1483.104369] [Medline: 23492531]
62. Jiamsakul A, Sirivichayakul S, Ditangco R, Wong K, Li PC, Praparattanapan J, TREAT Asia Studies to Evaluate Resistance – Surveillance Study (TASER-S). Transmitted drug resistance in recently infected HIV-positive Individuals from four

urban locations across Asia (2007-2010) - TASER-S. AIDS Res Ther 2015;12:3 [FREE Full text] [doi: 10.1186/s12981-015-0043-1] [Medline: 25685169]
63. Rodrigues R, Poongulali S, Balaji K, Atkins S, Ashorn P, De Costa A. 'The phone reminder is important, but will others get to know about my illness?' Patient perceptions of an mHealth antiretroviral treatment support intervention in the HIVIND trial in South India. BMJ Open 2015 Nov 02;5(11):e007574 [FREE Full text] [doi: 10.1136/bmjopen-2015-007574] [Medline: 26525717]
64. Lester R, Gelmon L, Plummer F. Cell phones: tightening the communication gap in resource-limited antiretroviral programmes? AIDS 2006 Nov 14;20(17):2242-2244. [doi: 10.1097/QAD.0b013e3280108508] [Medline: 17086071]
65. Jiamsakul A, Kerr SJ, Kiertiburanakul S, Azwa I, Zhang F, Chaiwarith R, TREAT Asia HIV Observational Database (TAHOD). Early suboptimal ART adherence was associated with missed clinical visits in HIV-infected patients in Asia. AIDS Care 2018 Dec 18;30(12):1560-1566 [FREE Full text] [doi: 10.1080/09540121.2018.1499859] [Medline: 30021450]
66. Protopopescu C, Carrieri MP, Raffi F, Picard O, Hardel L, Piroth L, ANRS CO8 APROCO-COPILOTE Study Group. Brief report: prolonged viral suppression over a 12-year follow-up of HIV-infected patients: the persistent impact of adherence at 4 months after initiation of combined antiretroviral therapy in the ANRS CO8 APROCO-COPILOTE Cohort. J Acquir Immune Defic Syndr 2017 Mar 01;74(3):293-297. [doi: 10.1097/QAI.0000000000001249] [Medline: 27861235]
67. Moosa A, Gengiah TN, Lewis L, Naidoo K. Long-term adherence to antiretroviral therapy in a South African adult patient cohort: a retrospective study. BMC Infect Dis 2019 Sep 05;19(1):775 [FREE Full text] [doi: 10.1186/s12879-019-4410-8] [Medline: 31488063]
68. O Connor C, Leyritana K, Calica K, Gill R, Doyle AM, Lewis JJ, et al. Risk factors affecting adherence to antiretroviral therapy among HIV patients in Manila, Philippines: a baseline cross-sectional analysis of the Philippines Connect for Life Study. Sex Health 2021 Mar;18(1):95-103. [doi: 10.1071/SH20028] [Medline: 33648625]
69. Health at your fingertips: a collection of mHealth initiatives from the research-based pharmaceutical industry. International Federation of Pharmaceutical Manufacturers & Associations. 2013. URL: https://www.ifpma.org/wp-content/uploads/2016/01/IFPMA_Health_at_your_fingertips_July2013.pdf [accessed 2022-01-12]
70. Byonanebye DM, Nabaggala MS, Naggirinya AB, Lamorde M, Oseku E, King R, et al. An interactive voice response software to improve the quality of life of people living with HIV in Uganda: randomized controlled trial. JMIR Mhealth Uhealth 2021 Feb 11;9(2):e22229 [FREE Full text] [doi: 10.2196/22229] [Medline: 33570497]
71. Parkes-Ratanshi R, Nabaggala M, Bwanika A, Lamorde M, King R, Owarwo N, et al. Call for life Uganda TM: an RCT using interactive voice response for PLHIV on ART. In: Proceedings of the Conference on Retroviruses and Opportunistic Infections. 2019 Presented at: Conference on Retroviruses and Opportunistic Infections; March 4-7, 2019; Seattle, Washington URL: https://www.croiconference.org/abstract/call-life-uganda-tm-rct-using-interactive-voice-response-plhiv-art/

Abbreviations

ART: antiretroviral therapy
BCT: behavior change technique
BCW: Behavior Change Wheel
COM-B: capability, opportunity, motivation, behavior
FGD: focus group discussion
IMB: information-motivation-behavioral skills
IVRS: interactive voice response system
mHealth: mobile health
MOTECH: Mobile Technology for Community Health
MSM: men who have sex with men
SHIP: Sustained Health Initiatives of the Philippines
TAMA: Treatment Advice by Mobile Alerts
TREAT: Therapeutics, Research, Education, and AIDS Training

Edited by A Mavragani; submitted 30.05.21; peer-reviewed by D Pförringer, C Reis; accepted 08.12.21; published 03.02.22.

Please cite as:
O'Connor C, Leyritana K, Doyle AM, Lewis JJ, Gill R, Salvaña EM
Interactive Mobile Phone HIV Adherence Support for Men Who Have Sex With Men in the Philippines Connect for Life Study: Mixed Methods Approach to Intervention Development and Pilot Testing
JMIR Form Res 2022;6(2):e30811
URL: https://formative.jmir.org/2022/2/e30811
doi:10.2196/30811
PMID:35113030

Original Paper

Use of a Smartphone App Versus Motivational Interviewing to Increase Walking Distance and Weight Loss in Overweight/Obese Adults With Peripheral Artery Disease: Pilot Randomized Trial

Tracie Collins[1], MS, MPH, MD; Mugur Geana[2*], MD, PhD; Kathryn Overton[3*], MS; Mary Benton[4*], MA, PhD; Liuqiang Lu[4*], MS; Faarina Khan[4*], MPH; Mason Rohleder[4*], MPH; Jasjit Ahluwalia[5*], MS, MPH, MD; Ken Resnicow[6*], PhD; Yiliang Zhu[7], PhD

[1]College of Population Health, University of New Mexico, Albuquerque, NM, United States

[2]School of Journalism and Mass Communications, University of Kansas, Lawrence, KS, United States

[3]University of New Mexico, Albuquerque, NM, United States

[4]University of Kansas School of Medicine, Wichita, KS, United States

[5]Brown University, Providence, RI, United States

[6]University of Michigan, Ann Arbor, MI, United States

[7]University of New Mexico, School of Medicine, Albuquerque, NM, United States

[*]these authors contributed equally

Corresponding Author:
Tracie Collins, MS, MPH, MD
College of Population Health
University of New Mexico
MSC09 5070
1 University of New Mexico
Albuquerque, NM, 87131-0001
United States
Phone: 1 505 272 4979
Email: tccollins@salud.unm.edu

Abstract

Background: Walking therapy improves functional outcomes in patients with peripheral artery disease (PAD). Less is known about the additive benefit of a dietary intervention.

Objective: Our objectives were to develop a smartphone app and, as a pilot, explore its potential efficacy as compared to motivational interviewing (MI) to increase walking distance and promote weight loss in overweight/obese adults with PAD.

Methods: We conducted a 3-month, 2-arm randomized pilot study at the University of Kansas. Inclusion criteria were BMI >27 kg/m^2 and symptomatic PAD, defined by an ankle-brachial index <0.9. Patients were randomized into 2 groups: MI, delivered through in-person and telephone counseling, and app, a mobile smartphone app. Both interventions encouraged walking for exercise and healthy dietary habits (increasing fruits and vegetables and whole grains while reducing fat and sugary drinks). We assessed medical history at baseline. At baseline and 3 months, participants completed an assessment of 6-minute walking distance, weight, quality of life, exercise behaviors, and dietary habits. The primary outcome was 3-month change in walking distance. Secondary outcomes were changes in weight, quality of life, exercise behaviors, and dietary habits. We used a Wilcoxon rank-sum test to analyze the primary and secondary outcomes at 3 months within the MI and app groups and to compare the changes between the groups with adjustment for baseline.

Results: We randomized 29 participants with a mean age of 66.03 (SD 8.12) years; 25 participants completed the trial. At baseline, mean walking distance among completers was 260.40 (SD 94.32) meters and 326.15 (SD 69.28) meters for MI and app participants, respectively. At 3 months, the mean walking distance was 298.67 (SD 101.20) meters and 331.19 (SD 58.63) meters for MI and app participants, respectively (group difference *P*=.03, adjusting for baseline). Increase in walking distance at 3 months was 40.5 meters (95% CI 6.77 to 61.34; *P*=.02) in MI group. At baseline, mean body weight was 253.10 (SD 59.45) lbs and 225.13 (SD 58.93) lbs for MI and app participants, respectively. At 3 months, mean body weight was 242.14 (SD 58.54) lbs and 223.44 (SD 59.54) lbs for MI and app, respectively (group difference *P*=.006, adjusting for baseline). Pre-post study decrease in

weight was 10.1 lbs (95% CI –17.9 to –3.0) and 2.3 lbs (95% CI –3.4 to –0.7) in MI and app group, respectively. Comparing baseline to 3 months, there were no statistically significant differences in quality of life, exercise behaviors, or dietary habits.

Conclusions: Our study demonstrates that MI can promote walking and weight loss in overweight/obese adults with PAD. The smartphone app showed a small weight loss but no statistically significant increase in walking distance. As this was a pilot study, future large-scale studies are needed to replicate the efficacy of MI to promote weight loss in overweight or obese adults with PAD.

Trial Registration: ClinicalTrials.gov NCT03694652; https://clinicaltrials.gov/ct2/show/NCT03694652

(JMIR Form Res 2022;6(2):e30295) doi:10.2196/30295

KEYWORDS

mobile health; smartphone app; peripheral artery disease; motivational interviewing

Introduction

Peripheral artery disease (PAD) is atherosclerosis of the abdominal aorta and arteries of the lower extremities that causes stenosis or occlusion [1]. Patients with PAD are at increased risk for cardiovascular events and functional decline [2-5]. As such, treatment for PAD has focused not only on risk factor modification but also on improving walking distance [6-10].

Community-based walking therapy is an effective noninvasive treatment option for PAD [11-13]. Motivating adherence to community-based walking therapy can be achieved through traditional, in-person approaches. Alternatively, motivating adherence can be achieved through mobile health or counseling [8,9]. We explored the potential efficacy of a smartphone app versus a well-known counseling approach, motivational interviewing (MI), to improve walking distance and promote weight loss among overweight/obese adults with PAD.

Although walking therapy is an important component of improving outcomes in patients with PAD, less is known about the efficacy of dietary interventions for disease management—in particular, to promote weight loss in overweight/obese patients with PAD.

Our pilot study is novel insofar as it focuses on walking therapy combined with dietary intervention. We delivered this combined intervention using 2 approaches: a smartphone app versus MI, delivered both in person and by phone. The app offers an innovative approach to behavior change for patients with PAD, and the use of MI builds upon our prior work involving patients with PAD [14,15]. The app was developed by MG and his team in the School of Journalism and Mass Communications, University of Kansas, Lawrence, Kansas. The theoretical framework for the app includes the transtheoretical model, social cognitive theory, and self-determination theory. Similarly, for MI, the theoretical framework includes the self-determination theory and transtheoretical model. In addition, MI is based on client-centered therapy [16-19].

In this pilot, we explored the potential efficacy of our app versus MI to increase walking distance and promote weight loss among overweight/obese adults with PAD.

Methods

Participants

We conducted a 2-arm, pilot randomized trial. Inclusion criteria were age 50 years or older, overweight/obese (BMI >27), and symptomatic PAD—leg symptoms were captured by a survey, and PAD was confirmed with the use of the ankle-brachial index, the ratio of the systolic blood pressure in the ankle to that in the arm. Our cut point for the ankle-brachial index was <0.9.

Individuals were excluded if they demonstrated at least one of the following conditions: intolerance to fruits, vegetables, fiber, or a low-fat diet; restricted water intake; pregnancy; prior major ischemia or critical leg ischemia; use of 24-hour supplemental oxygen; heart attack within the last 3 months; inability to walk for exercise; or currently walking at least 3 days per week for at least 30 minutes each day. The University of Kansas Medical Center institutional review board approved this study, and participants gave informed consent. This clinical trial was registered at ClinicalTrials.gov [NCT03694652].

Recruitment

We recruited study participants using a variety of modalities including the University of Kansas hospital system and unaffiliated health centers and community-based clinics; area clinicians and administrators voluntarily distributed flyers and mailings to eligible patients. We obtained university institutional review board approval and permission from participating hospitals and clinics prior to each phase of the recruitment process.

Interventions

Motivational Interviewing

Participants randomized to MI participated in an initial 1-hour, face-to-face session. Following the initial visit, we conducted four 20-minute telephone calls: 1 call every 2 weeks for 1 month followed by 2 monthly phone calls.

Author KR provided MI counselor training through a series of workshops. The counselor was a medical school graduate and a student in the master's in public health program at University of Kansas School of Medicine–Wichita who was also applying for residency in primary care. Although the counselor had no experience in MI prior to this trial, author KR has trained multiple research staff in the delivery of MI. There were

standardized patients and ongoing review of sessions through direct observation. The counselor's goals were to elicit and reinforce change talk about increasing walking behavior and healthy dietary habits (Table 1).

During each counseling session, the counselor assessed participant willingness to walk for exercise using the Patient-Centered Assessment and Counseling for Exercise (PACE) survey, which assesses readiness to exercise [20,21]. Following a discussion of their current PACE score, the counselor queried patients about the importance of exercising and their current motivation to increase walking behavior. The counselor concluded each session by offering patients the option to set a walking goal for the next 2 weeks.

The counselor followed a similar process with respect to the nutritional portion of the interview. She guided patients to choose 1 of 6 nutritional options to work on in the following 2 weeks: (1) reducing sodium, (2) increasing fruits, (3) increasing vegetables, (4) decreasing sugary drinks, (5) increasing whole grains, or (6) mindful eating. Similar to the exercise therapy, the counselor discussed the importance of change, patient confidence levels, and individual values. After patients established exercise and nutritional goals, counselors would arrange for subsequent follow-up sessions at 2-week intervals.

Table 1. Behavior change techniques: nutrition and physical activity.

Technique	MI[a]	App
1.1 Goal setting (behavior): agreed on weekly walking goals	✓	✓
1.2 Problem solving: identifying triggers to eating unhealthy foods or avoiding walking for exercise	✓	✓
1.4 Action planning: setting aside time to exercise and planning meals in advance	✓	✓
1.6 Discrepancy between current behavior and goal: recorded walking goals or self-reported dietary goals were not met	—[b]	✓
3.1 Social support (unspecified): participants received recommendations on the value of having a buddy to walk with	✓	✓
4.1 Instruction on walking therapy to improve walking distance	✓	✓
4.3 Re-attribution: if a participant attributed their desire for food to boredom, we provided guidance on mindful eating	—	✓
5.1 Information about health consequences: participants were provided information about the potential for disease progression in the absence of a walking intervention	✓	✓
5.4 Monitoring of emotional consequences: participants were queried about satisfaction with their weekly dietary and walking goals	—	✓
8.2 Behavior substitution: participants were provided with guidance on substituting unhealthy dietary choices with healthy dietary choices	—	✓
9.1 Credible source: participants viewed videos in which the principal investigator, who is board-certified in internal and vascular medicine, shared the importance of walking for exercise for persons with peripheral artery disease	—	✓
9.2 Pros and cons: participants were queried regarding the pros and cons of eating a healthy diet and walking for exercise	✓	✓
10.4 Social reward: participants were congratulated for achieving their weekly goals	—	✓
13.2 Framing/reframing: participants were provided with cognitive structuring to think of tasks to reduce sedentary behavior	✓	✓
15.1 Verbal persuasion about capability: participants were told they can walk for exercise despite leg discomfort	✓	✓

[a]MI: motivational interviewing.

[b]Not applicable.

Smartphone App

Development

We created the PAD mobile app with support from the Center for Excellence in Health Communication to Underserved Populations at the University of Kansas School of Journalism and Mass Communication. Author MG supervised the design, operational functionality, and development of the app. We employed iterative feedback from experts and research team members to design the user interface, content, and functionality.

The app was designed for use with Android OS smartphones (patients without this type of phone received it as part of the study). We tailored nutritional content for clarity and lower health literacy levels. We designed the mobile app to allow users to input walking plans, track walking intervals, and record episodic pain. We designed in-app notifications to encourage patient use and for diet and exercise management.

The app was completely automated and adaptive to patient use. The study team was available to troubleshoot and answer questions as needed. Safeguards permitted recovery of previously saved data in the event of accidental malfunction. Participant data were encrypted and saved locally to the phone. Once initialized, the app sent secure, depersonalized data to an online backup server via the internet.

Implementation

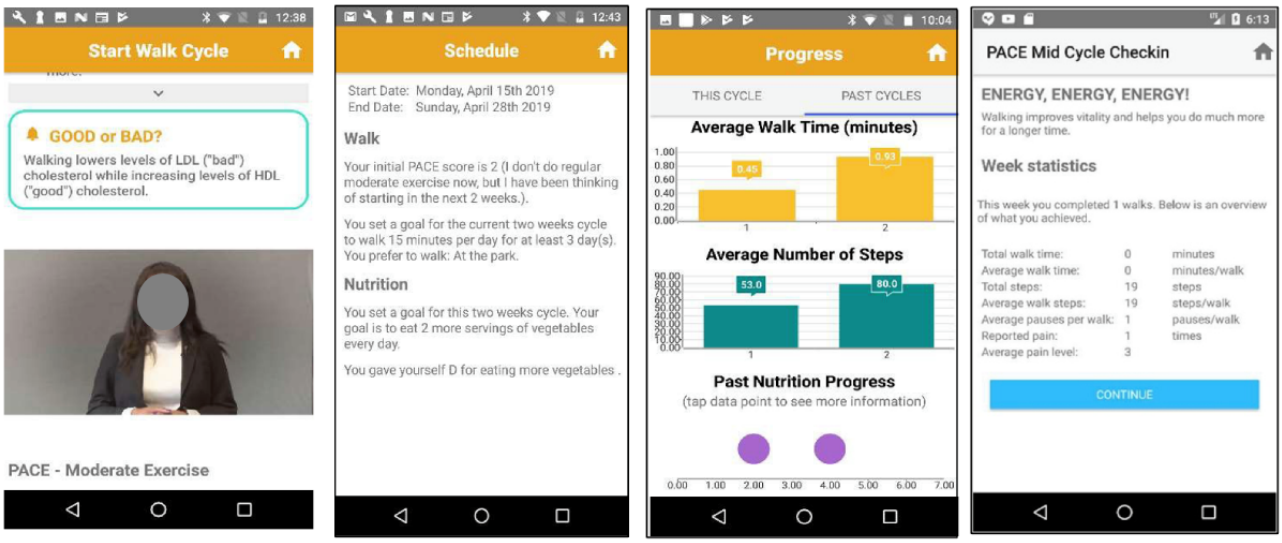
We developed the PAD mobile app specifically for this pilot study based on the PACE and MI principles described previously. The PAD mobile app had 2 distinct components: a nutritional component and a walking component. Similar to the MI group content, the nutritional component of the mobile app contained 6 modules, with each of the topics designed to be

completed over a 2-week period. We designed each module to reflect nutritional issues specific to PAD patients. Patients were permitted to choose the order the modules for completion, and completion was self-initiated. The app employed identical measures for assessing patient perceptions of the importance of change, confidence levels, and individual values. The app used the PACE format to set exercise goals, and self-assessment to set nutritional goals over a 12-week period. Similar to the MI group content, nutritional modules were further divided into six 2-week cycles.

On the first day of the study, patients installed the app, and it prompted them to enter their demographic, health, and identity information. The app then presented patients with 6 possible nutritional changes that they could choose to address during the first 2-week cycle. Once participants chose a dietary modification, they were prompted to perform a self-assessment; the app would then adjust goals according to patients' responses.

We made explicit efforts to ensure that patients would not have trouble installing, using, or interacting with the app. Several safeguards were installed to provide patients with continuous support and resources. The app provided extensive exercise tracking (Figure 1) as well as a comprehensive reporting menu with information about patient progress (Figure 1). Automatic reminders prompted participants to engage with the app and to select their next PAD intervention cycle at 2-week intervals.

Figure 1. PAD mobile app: components of walking and nutrition modules.



Measures

Ankle-Brachial Index

The ankle-brachial index was used to define the presence or absence of PAD. During this assessment, a participant rested for 5 minutes after which a 5 MHz hand-held Doppler with an attached stethoscope was used to measure systolic blood pressures in both brachial arteries and in both ankles (ie, the dorsalis pedis and posterior tibial arteries) [22].

Medical History

The principal investigator and colleagues developed the Lifestyle and Clinical Survey (Multimedia Appendix 1) to obtain patients' medical history, including smoking status and sociodemographic and comorbidity data. It has a summary k-statistic for reliability of 0.81 (95% CI 0.78 to 0.84) and a summary k-statistic for validity of 0.58 (95% CI 0.52 to 0.64) [23].

Stage of Readiness to Engage in Exercise

The PACE score was used to identify a participant's readiness for exercise. To obtain a PACE score, a participant chose 1 of 8 graded statements that best described their current level of and interest in physical exercise [24]. We assessed a participant's PACE score at baseline and 6 weeks; we employed these assessments to tailor the intervention to each patient's respective stage of readiness.

Outcomes

The primary outcome was the 3-month change in walking distance as measured by the 6-minute walk test, which provides information on a patient's ability to walk in the community. The test is conducted by placing 2 cones 100 feet apart in a marked hallway and instructing patients to walk as many laps around the cones as possible. Patients were permitted to stop walking during the test; however, time was recorded during the rest period. We recorded time and distance to onset of leg discomfort and total distance walked. In a prior study involving 64 patients with PAD, the reliability coefficient for distance during 6-minute walk tests performed 1 week apart was 0.94 with a coefficient of variation of 11.7% [25].

Secondary outcomes included weight loss, quality of life, exercise behaviors, and dietary habits. For weight loss, we used a standardized scale to measure body weight. For health-related quality of life, we used the Vascular Quality of Life Questionnaire (VascuQoL) [26]. The intraclass (reliability) coefficient for the VascuQoL was 0.94 (CI >90%) and Cronbach α ranged between .7 and .9, indicating good internal consistency within the 5 domains. The VascuQoL total score has a correlation with the Fontaine classification of disease severity [27] of r=–.79 (P<.001) and with treadmill walking distances of r=.36 (P<.05). For self-reported physical activity, we used the Stanford Patient Education Research Center Exercise

Behavior Survey [28,29]. The Exercise Behavior Survey retest reliability scores between 0.56-0.72. We employed the Fat-Related Diet Habits Questionnaire [30] to assess basic consumption as well as healthier choices, such as substitution of low-calorie or low-fat food alternatives. We modified the survey to a 3.9 reading level. The Fat-Related Diet Habits Questionnaire has high test-retest and internal consistency reliabilities and correlations with percentage of calories from fat ranging from 0.34 to 0.57 (P<.01). The correlation of the sum of the 5 scales with percentage of calories from fat was 0.68 (P<.001) and, in multiple regression models, the multiple R^2 using all factors to predict percentage of calories from fat was 0.47. All outcomes were measured at baseline and 3 months.

Study Flow

Screening

We prescreened potential participants via phone. The prescreen assessment captured (1) physical activity readiness (Physical Activity Readiness Questionnaire) to identify conditions that would preclude participation in the study [31] and (2) leg symptoms (San Diego Claudication Questionnaire) [32]. Based on eligibility criteria, we scheduled candidates for an in-person visit.

Baseline In-Person Visit and Randomization

After obtaining informed consent, a blinded assessor measured individuals' ankle-brachial index, weight, and height. We referred eligible candidates to graded submaximal treadmill testing, as per American College of Cardiology guidelines.

Upon completion of treadmill testing and without evidence of coronary ischemia, participants completed questionnaires to assess quality of life, exercise behaviors, and dietary habits. We then randomly assigned participants to the app or MI. We used a random number generator to blindly assign patients to one of the 2 groups.

Analysis Plan

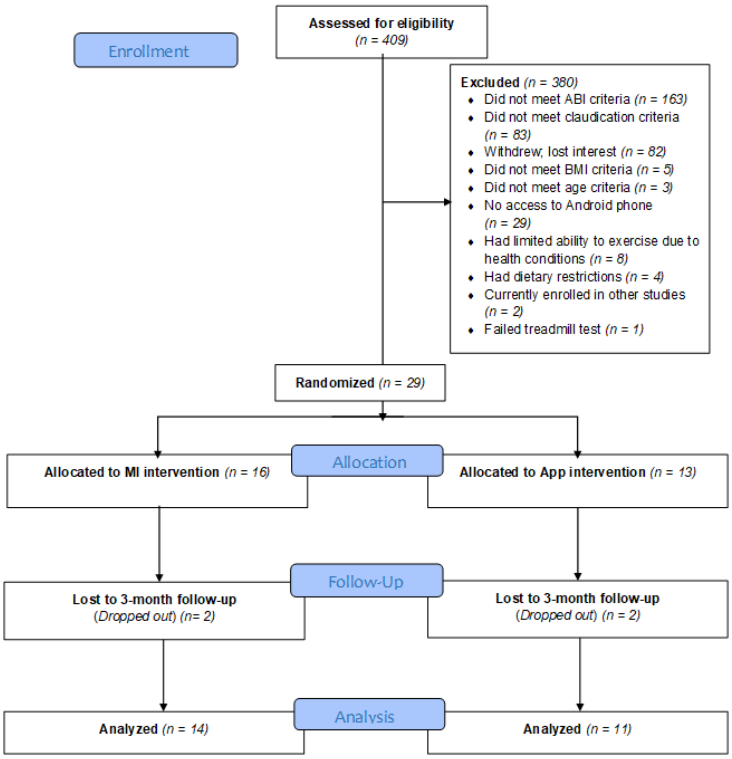
We used a Wilcoxon rank-sum test to analyze the primary outcome of change in 6-minute walking distance at 3 months within the MI and app groups and compare the changes between the groups with adjustment for baseline (difference in difference). We used this same method for other outcomes. For comparing participants' baseline characteristics between the MI and app groups, we used a Wilcoxon test for continuous measurements and a Fisher test for dichotomous measurements. We use R (version 4.03, R Foundation for Statistical Computing) for all data analyses.

Thirteen participants per group were required to provide 80% power at a 0.05 error level and to detect a difference between the 2 groups of 50 meters walking distance change.

Results

We randomized a total of 29 participants and assigned 16 to MI and 13 to the app (Figure 2); 25 participants completed the 3-month assessment (MI=14 and app=11), which retained >75% power.

Figure 2. CONSORT flow diagram. ABI: ankle-brachial index; MI: motivational interviewing



Baseline characteristics of all 29 participants are shown for the entire sample and by treatment group in Table 2. For the overall cohort, the mean age was 66.0 (SD 8.12) years. There was no group difference in prevalence of cardiovascular risk factors or alcohol use.

At baseline, mean walking distance by group was 260.40 (SD 94.32) meters for MI and 326.15 (SD 69.28) meters for the app. At 3 months, the mean walking distance was 298.67 (SD 101.20) meters for MI and 331.19 (SD 58.63) meters for the app. The median increase in walking distance was 40.51 meters (95% CI 6.77 to 61.34, *P*=.02) in MI and 8.04 meters (95% CI –15.86 to 22.92; *P*=.41) in app. Group difference in change of walking distance was 33.60 meters (95% CI 8.69 to 58.39, *P*=.03).

At baseline, mean body weight by group was 253.10 (SD 59.45) lbs for MI and 225.13 (SD 59.83) lbs for the app. At 3 months, mean body weight by group was 242.14 (SD 58.54) lbs for MI and 223.44 (SD 59.54) lbs for the app. The median weight decrease was 10.1 lbs (95% CI –17.9 to –3.0; *P*=.01) in MI and 2.3 lbs (95% CI –3.4 to –0.7; *P*=.02) in app. Group difference in weight change was –7.48 lbs (95% CI –14.60 to –3.20; *P*=.006).

Table 2. Baseline participants characteristics.

Characteristics	Overall (n=29)	MI[a] (n=16)	App (n=13)	*P* value
Age (years), mean (SD)	66.03 (8.12)	68.39 (7.36)	63.12 (8.34)	.01
BMI, mean (SD)	38.25 (9.25)	39.69 (10.18)	36.47 (8.00)	.37
Weight (lbs), mean (SD)	236.67 (59.05)	244.30 (60.38)	227.28 (58.36)	.53
Waist, mean (SD)	48.12 (5.78)	48.41 (6.18)	47.77 (5.49)	.63
Income ($), mean (SD)	50,805.43 (39,571.16)	50,534.5 (47,581.8)	51,166.67 (30,022.77)	.59
Female, n (%)	20 (70)	11 (69)	9 (69)	>.99
Educations (≥ high school), n (%)	28 (97)	16 (100)	12 (92)	.45
Worried about housing loss, n (%)	3 (10)	1 (6)	2 (15)	.57
Employment, n (%)				
Full-time work	12 (41)	5 (31)	7 (54)	.22
Unemployed but not seeking work	15 (52)	10 (63)	5 (38)	.27
Insurance, n (%)				
None/uninsured	4 (14)	2 (13)	2 (15)	>.99
Medicare	16 (55)	11 (69)	5 (38)	.14
Private insurance	8 (28)	3 (19)	5 (38)	.41
Health history, n (%)				
Myocardial infarction	8 (28)	4 (25)	4 (31)	>.99
Cardiac catheterization	12 (41)	9 (56)	3 (23)	.13
Claudication	9 (31)	5 (31)	4 (31)	>.99
Hypertension	23 (79)	12 (75)	11 (85)	.66
Hypercholesterolemia	19 (68)	10 (63)	9 (69)	>.99
Diabetes	13 (45)	8 (50)	5 (38)	.71
Diabetes-related complications	8 (28)	7 (44)	1 (8)	.04
Arthritis other than rheumatoid	12 (41)	9 (56)	3 (23)	.13
At least 100 cigarettes during lifetime	18 (62)	10 (63)	8 (62)	>.99
Ethanol use	18 (62)	8 (50)	10 (77)	.25

[a]MI: motivational interviewing.

In addition to walking distance and weight loss, we also assessed changes in quality of life, exercise behaviors, and dietary habits (Multimedia Appendix 2). Overall, there were no differences between the app and MI groups in quality of life, exercise behaviors, and dietary habits, comparing baseline to 3 months.

To encourage patients to complete the nutritional components of the app, researchers sent out reminders both midcycle and at the end of each cycle. Despite these nudges, patient participation with the nutritional component of the app was highly variable. Out of 14 participants, only 3 individuals started all 6 nutrition modules. Among patients randomized to the PAD mobile app, 29% completed all 6 modules, 24% completed 5 modules, and 5% completed only 1 module. A total of 72% (10/14) of participants started 4 nutrition modules or more. Overall, there

was a high degree of variability in terms of engagement with the nutritional component of the app.

Included with the nutritional component of the app were guided assessments designed to measure patients' confidence and personal importance in achieving nutritional goals. Across all of the cycles, patients' initial average confidence in achieving nutritional goals was 6.4 on a scale from 1 (lowest confidence) to 10 (highest confidence). In addition, the average perceived personal importance of the nutrition topics was 6.5 on a scale from 1 to 10.

The app was designed to prompt self-evaluation of progress at 2 time points during each nutritional module. We designed interactions with the app to mimic the dialogue that occurred at 2-week intervals in the MI group. The use of the app varied across the group of participants, with some using it as little as 8 times during the 12-week period and others using it over 200 times during the same interval. There were a total of 1137 interactions with the app, of which 111 were generated by directly responding to a prompt from an app notification.

Patients entered midcycle data in only 13 out of the 52 started modules (25%). Patients entered end-of-cycle data in only 18 out of the 52 modules initiated. Thus, there was a huge attrition rate associated with participation in the nutritional component of the app. This attrition occurred despite the presence of multiple, automatic system notifications—9, on average—sent to support nutritional goals and to remind participants to log in and enter feedback.

Participants demonstrated substantially greater interaction with the walking component of the PAD mobile app than the nutritional modules. The app tracked a total of 355 walking episodes over the duration of the study; this represented an average of 25 walking episodes per participant.

Patients reported variable levels of pain throughout the 12-week study period. A total of 55% (6/11) of participants recorded less pain on their last cycle compared with the first one. A total of 36% (4/11) recorded more intense pain at their last recorded cycle in comparison to their first.

The app recorded average walking time per recorded walking episode as 00:17:57. Based on limited data from 3 devices, we found that patients walked at an average pace of 2.15 seconds per step. Thus, participants averaged approximately 500 steps per session.

Discussion

Principal Findings

We found that MI but not the app promoted increased walking distance and both interventions promoted somewhat significant weight loss among overweight or obese adults with PAD, although the loss in the app group was rather small in size. To our knowledge, this is the first study to demonstrate the potential efficacy of MI to promote weight loss in overweight or obese patients with PAD. Applications of MI to improve outcomes have been demonstrated in prior work. Cunningham and colleagues [33] found that MI was successful in motivating PAD patients to increase walking distances. This pilot confirmed that MI was efficacious to improve walking distance. However, in prior work by Collins et al [15] involving African Americans with PAD, MI was not efficacious. Findings from that study highlight the need among some populations for more physician-directed counseling to motivate adherence to walking therapy. However, understanding the potential efficacy of MI on a larger scale is still warranted.

MI was more efficacious than our app for improving weight loss. The reason for the efficacy of MI versus our app is likely multifactorial. One consideration is the human contact that is afforded by MI as compared to the app (or mHealth). Both human contact and mHealth have their advantages and disadvantages, as noted by Santarossa et al [34]. mHealth should create accountability and social support to successfully motivate behavior change. Although we designed our app to provide support, we can certainly address this more in future work. The theoretical framework underlying MI includes the self-determination theory and transtheoretical model. Thus, the potential mechanisms by which MI was efficacious include intrinsic and extrinsic motivation as well as stages of change. However, there were no statistically significant improvements in dietary habits. A larger study with assessment of potential mediators and moderators for intervention efficacy could shed more light.

Also, some technological design challenges may explain our limited findings with respect to walking distance and weight loss among patients randomized to the app. For example, the current version of the app requires that patients deliberately initiate it during every use. In addition to technological challenges, we observed that patients' overall use of the mobile app was highly variable. This finding could be improved through additional in-app notifications.

Despite midcycle and end-of-cycle notifications, patient participation with the nutritional component of the app was also highly variable. Among the participants randomized to the app, very few started all 6 nutrition modules. We speculate that complexity of the nutritional components of the app may have contributed to decreased participation. This decreased participation may have in turn contributed to the modest findings among individuals randomized to the app.

Limitations

The small sample of participants is an obvious limitation of the study as it precludes generalizability and has limited power. We readily acknowledge these constraints as we prepare to move forward with a larger clinical trial in the future.

Additionally, because of licensing constraints imposed by the Apple Corporation, we designed the mobile app for use with Android phones (an Apple version is in development). Relaxing these constraints in the future will allow for a more streamlined, feature-rich, and automated app. Our ultimate goal is to expand accessibility of the app as an mHealth alternative among the broader PAD population. We did not complete in-depth interviews with app participants regarding reasons for or against use of the app. Finally, improvements are needed in the app to increase use. Again, another complication associated with the app is that patients were prompted to provide progress data, but

these data needed to be entered voluntarily into the app in order to chart progress. The app was designed to prompt midcycle for an initial evaluation of the progress as well as at the end of the module. The purpose of these data points was to assess and record completion, estimate goal achievement, and identify barriers. Researchers designed interactions with the mobile app to mimic the dialogue that occurred over the phone in the first arm of the study. Thus, the notifications were intended to be similar to conversations that a trained evaluator would have with the patient by telephone every 2 weeks.

Conclusion

MI significantly improved walking distance and promoted weight loss for participants randomized to MI versus the app. As the use of mobile devices is increasing, the need for development of mobile apps to promote walking therapy and health dietary habits cannot be ignored. The potential of MI to promote weight loss in overweight/obese adults with PAD warrants further study. Additionally, the development of a mobile app that is compatible with IOS as well as Android phones with content available in multiple languages will be the focus of a future study. This pilot project represents one of the first pilot studies to include behavioral interventions targeting both exercise and diet in overweight/obese adults with PAD.

Acknowledgments

Research reported in this publication was supported by the National Heart, Lung, and Blood Institute of the National Institutes of Health under award number 1R56HL138244-01. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Authors' Contributions

TC, MB, FK, and MR collected data. TC, MG, KO, MB, and LL wrote the first draft of the manuscript. MG, LL, KR, JA, and YZ provided critical review and language for the final draft of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Lifestyle and clinical survey.
[DOCX File , 275 KB - formative_v6i2e30295_app1.docx]

Multimedia Appendix 2
Study outcomes by intervention groups for completers only.
[DOCX File , 19 KB - formative_v6i2e30295_app2.docx]

Multimedia Appendix 3
CONSORT-eHEALTH checklist (V 1.6.2).
[PDF File (Adobe PDF File), 99 KB - formative_v6i2e30295_app3.pdf]

References

1. Shalan A, Abdulrahman A, Habli I, Tew G, Thompson A. YORwalK: desiging a smartphone exercise application for people with intermittent claudication. Stud Health Technol Inform 2018;247:311-315. [Medline: 29677973]
2. Heald CL, Fowkes FGR, Murray GD, Price JF, Ankle Brachial Index Collaboration. Risk of mortality and cardiovascular disease associated with the ankle-brachial index: systematic review. Atherosclerosis 2006 Nov;189(1):61-69. [doi: 10.1016/j.atherosclerosis.2006.03.011] [Medline: 16620828]
3. Ankle Brachial Index Collaboration, Fowkes FGR, Murray GD, Butcher I, Heald CL, Lee RJ, et al. Ankle brachial index combined with Framingham Risk Score to predict cardiovascular events and mortality: a meta-analysis. JAMA 2008 Jul 09;300(2):197-208 [FREE Full text] [doi: 10.1001/jama.300.2.197] [Medline: 18612117]
4. McDermott MM, Guralnik JM, Tian L, Ferrucci L, Liu K, Liao Y, et al. Baseline functional performance predicts the rate of mobility loss in persons with peripheral arterial disease. J Am Coll Cardiol 2007 Sep 04;50(10):974-982 [FREE Full text] [doi: 10.1016/j.jacc.2007.05.030] [Medline: 17765125]
5. Collins TC, Petersen NJ, Suarez-Almazor M. Peripheral arterial disease symptom subtype and walking impairment. Vasc Med 2005 Aug;10(3):177-183. [doi: 10.1191/1358863x05vm615oa] [Medline: 16235770]
6. Collins TC, Beyth RJ. Process of care and outcomes in peripheral arterial disease. Am J Med Sci 2003 Mar;325(3):125-134. [doi: 10.1097/00000441-200303000-00005] [Medline: 12640288]

7. Eason SL, Petersen NJ, Suarez-Almazor M, Davis B, Collins TC. Diabetes mellitus, smoking, and the risk for asymptomatic peripheral arterial disease: whom should we screen? J Am Board Fam Pract 2005;18(5):355-361 [FREE Full text] [doi: 10.3122/jabfm.18.5.355] [Medline: 16148245]
8. Kim M, Kim C, Kim E, Choi M. Effectiveness of mobile health-based exercise interventions for patients with peripheral artery disease: systematic review and meta-analysis. JMIR Mhealth Uhealth 2021 Feb 15;9(2):e24080 [FREE Full text] [doi: 10.2196/24080] [Medline: 33587042]
9. Paldán K, Simanovski J, Ullrich G, Steinmetz M, Rammos C, Jánosi RA, et al. Feasibility and clinical relevance of a mobile intervention using TrackPAD to support supervised exercise therapy in patients with peripheral arterial disease: study protocol for a randomized controlled pilot trial. JMIR Res Protoc 2019 Jun 26;8(6):e13651 [FREE Full text] [doi: 10.2196/13651] [Medline: 31244477]
10. Campia U, Gerhard-Herman M, Piazza G, Goldhaber SZ. Peripheral artery disease: past, present, and future. Am J Med 2019 Oct;132(10):1133-1141. [doi: 10.1016/j.amjmed.2019.04.043] [Medline: 31150643]
11. Collins T, Ghidei W, Ahluwalia J. Benefits of walking therapy in patients with diabetes mellitus and peripheral arterial disease who are limited by leg pain or fatigue. J Gen Intern Med 2011;26:S176.
12. Gardner AW. Supervised exercise therapy provided by local physiotherapists improves walking distance in patients with claudication. Evid Based Med 2011 Apr 21;16(2):43-44. [doi: 10.1136/ebm1151] [Medline: 21427053]
13. McDermott MM, Liu K, Guralnik JM, Criqui MH, Spring B, Tian L, et al. Home-based walking exercise intervention in peripheral artery disease: a randomized clinical trial. JAMA 2013 Jul 03;310(1):57-65 [FREE Full text] [doi: 10.1001/jama.2013.7231] [Medline: 23821089]
14. Quirk F, Dickinson C, Baune B, Leicht A, Golledge J. Pilot trial of motivational interviewing in patients with peripheral artery disease. Int Angiol 2012 Oct;31(5):468-473. [Medline: 22990510]
15. Collins TC, Lu L, Ahluwalia JS, Nollen NL, Sirard J, Marcotte R, et al. Efficacy of community-based exercise therapy among African American patients with peripheral artery disease: a randomized clinical trial. JAMA Netw Open 2019 Feb 01;2(2):e187959 [FREE Full text] [doi: 10.1001/jamanetworkopen.2018.7959] [Medline: 30768192]
16. Miller WR. Motivational interviewing with problem drinkers. Behav Psychother 2009 Jun 16;11(2):147-172. [doi: 10.1017/s0141347300006583]
17. Miller WR, Rollnick S. Meeting in the middle: motivational interviewing and self-determination theory. Int J Behav Nutr Phys Act 2012;9:25 [FREE Full text] [doi: 10.1186/1479-5868-9-25] [Medline: 22385872]
18. Rollnick S, Allison J. Motivational interviewing. In: Heather N, Stockwell T, editors. The Essential Handbook of Treatment and Prevention of Alcohol Problems. West Sussex: John Wiley & Sons; 2004:105-116.
19. Rollnick S, Miller WR. What is motivational interviewing? Behav Cogn Psychother 2009 Jun 16;23(4):325-334. [doi: 10.1017/S135246580001643X]
20. Patrick K, Sallis JF, Long B, Calfas KJ, Wooten W, Heath G, et al. A new tool for encouraging activity. Phys Sportsmed 1994 Nov;22(11):45-55. [doi: 10.1080/00913847.1994.11947706] [Medline: 29275663]
21. Calfas KJ, Sallis JF, Oldenburg B, Ffrench M. Mediators of change in physical activity following an intervention in primary care: PACE. Prev Med 1997;26(3):297-304. [doi: 10.1006/pmed.1997.0141] [Medline: 9144753]
22. Collins TC, Petersen NJ, Suarez-Almazor M, Ashton CM. The prevalence of peripheral arterial disease in a racially diverse population. Arch Intern Med 2003 Jun 23;163(12):1469-1474. [doi: 10.1001/archinte.163.12.1469] [Medline: 12824097]
23. Collins T, O'Malley K, Petersen N, Suarez-Almazor M. The Lifestyle and Clinical Survey: a pilot study to validate a medical history questionnaire. Fed Pract 2005;22(5):25-46 [FREE Full text]
24. Prochaska JO, DiClemente CC. Stages and processes of self-change of smoking: toward an integrative model of change. J Consult Clin Psychol 1983;51(3):390-395. [doi: 10.1037//0022-006x.51.3.390] [Medline: 6863699]
25. Montgomery PS, Gardner AW. The clinical utility of a six-minute walk test in peripheral arterial occlusive disease patients. J Am Geriatr Soc 1998 Jun;46(6):706-711. [doi: 10.1111/j.1532-5415.1998.tb03804.x] [Medline: 9625185]
26. Morgan MB, Crayford T, Murrin B, Fraser SC. Developing the Vascular Quality of Life Questionnaire: a new disease-specific quality of life measure for use in lower limb ischemia. J Vasc Surg 2001 Apr;33(4):679-687 [FREE Full text] [doi: 10.1067/mva.2001.112326] [Medline: 11296317]
27. Norgren L, Hiatt WR, Dormandy JA, Nehler MR, Harris KA, Fowkes FGR, TASC II Working Group. Inter-society consensus for the management of peripheral arterial disease (TASC II). J Vasc Surg 2007 Jan;45 Suppl S:S5-S67 [FREE Full text] [doi: 10.1016/j.jvs.2006.12.037] [Medline: 17223489]
28. 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]
29. Lorig K, Stewart A, Ritter P, Gonzalez V, Laurent D, Lynch J. Outcome Measures for Health Education and Other Health Care Interventions. Thousand Oaks: Sage Publications; 1996.
30. Shannon J, Kristal AR, Curry SJ, Beresford SA. Application of a behavioral approach to measuring dietary change: the fat- and fiber-related diet behavior questionnaire. Cancer Epidemiol Biomarkers Prev 1997 May;6(5):355-361 [FREE Full text] [Medline: 9149896]
31. Thomas S, Reading J, Shephard RJ. Revision of the Physical Activity Readiness Questionnaire (PAR-Q). Can J Sport Sci 1992 Dec;17(4):338-345. [Medline: 1330274]

32. Criqui MH, Denenberg JO, Bird CE, Fronek A, Klauber MR, Langer RD. The correlation between symptoms and non-invasive test results in patients referred for peripheral arterial disease testing. Vasc Med 1996;1(1):65-71. [doi: 10.1177/1358863X9600100112] [Medline: 9546918]
33. Cunningham MA, Swanson V, O'Carroll RE, O'Caroll RE, Holdsworth RJ. Randomized clinical trial of a brief psychological intervention to increase walking in patients with intermittent claudication. Br J Surg 2012 Jan;99(1):49-56. [doi: 10.1002/bjs.7714] [Medline: 22038532]
34. Santarossa S, Kane D, Senn CY, Woodruff SJ. Exploring the role of in-person components for online health behavior change interventions: can a digital person-to-person component suffice? J Med Internet Res 2018 Apr 11;20(4):e144 [FREE Full text] [doi: 10.2196/jmir.8480] [Medline: 29643048]

Abbreviations

MI: motivational interviewing
PACE: Patient-Centered Assessment and Counseling for Exercise
PAD: peripheral artery disease
VascuQoL: Vascular Quality of Life Questionnaire

Edited by G Eysenbach; submitted 17.05.21; peer-reviewed by K Knittle, C Cheng; comments to author 30.07.21; revised version received 01.10.21; accepted 27.11.21; published 03.02.22.

Please cite as:
Collins T, Geana M, Overton K, Benton M, Lu L, Khan F, Rohleder M, Ahluwalia J, Resnicow K, Zhu Y
Use of a Smartphone App Versus Motivational Interviewing to Increase Walking Distance and Weight Loss in Overweight/Obese Adults With Peripheral Artery Disease: Pilot Randomized Trial
JMIR Form Res 2022;6(2):e30295
URL: https://formative.jmir.org/2022/2/e30295
doi:10.2196/30295
PMID:35113020

Original Paper

Preferences for Technology-Mediated Behavioral Lifestyle Interventions With Different Levels of Coach and Peer Support Among Latino Men: Comparative Study Within One Arm of a Randomized Controlled Trial

Lisa G Rosas[1], MPH, PhD; Nan Lv[2], PhD; Lan Xiao[1], PhD; Kristen MJ Azar[3]; Steven P Hooker[4], PhD; Elizabeth M Venditti[5,6], PhD; Megan A Lewis[7], PhD; Patricia Zavella[8], PhD; Jun Ma[2], PhD

[1]Department of Epidemiology and Population Health, Stanford University, Palo Alto, CA, United States
[2]Department of Medicine, University of Illinois, Chicago, IL, United States
[3]Sutter Health, Palo Alto, CA, United States
[4]College of Health and Human Service, San Diego State University, San Diego, CA, United States
[5]Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA, United States
[6]Department of Epidemiology, University of Pittsburgh, Pittsburgh, PA, United States
[7]RTI, Seattle, WA, United States
[8]Latin American and Latino Studies, University of California Santa Cruz, Santa Cruz, CA, United States

Corresponding Author:
Lisa G Rosas, MPH, PhD
Department of Epidemiology and Population Health
Stanford University
1701 Page Mill Rd.
2nd Floor
Palo Alto, CA, 94304
United States
Phone: 1 6505759519
Email: lgrosas@stanford.edu

Abstract

Background: Although Latino men have the highest prevalence (45%) of obesity among all men in the United States, traditional weight loss interventions have not effectively engaged this hard-to-reach and diverse group. Offering choices among technology-mediated weight loss interventions may offer advantages.

Objective: The aim of this study is to examine Latino men's preferences among 3 weight loss intervention options. We also examined whether attendance in group sessions (videoconference and in person) and weight loss differed according to intervention choice.

Methods: Latino men (n=200; mean age 47.3, SD 11.8 years) participated in a comparative effectiveness trial based on primary care and were randomized to receive the 1-year HOMBRE (Hombres con Opciones para Mejorar su Bienestar para Reducir Enfermedades Crónicas; English translation: Men With Options to Improve Their Well-being and Reduce Chronic Disease) intervention. HOMBRE is a weight loss intervention that offers 3 delivery options. During an orientation session, a trained bilingual coach helped men select 1 of the 3 intervention options that differed in coach, peer support, and available language. We used canonical discriminant analysis to assess multivariate associations of demographic, clinical, employment, cultural, and technology use and access factors with men's intervention choices. We used generalized linear models to estimate weight loss at 6, 12, and 18 months for men in each intervention option.

Results: Among Latino men, 28% (56/200) chose videoconference groups, 31% (62/200) chose web-based videos, and 41% (82/200) chose in-person groups. The canonical discriminant analysis identified 1 orthogonal dimension that distinguished between men who chose an in-person group and men who chose web-based videos. Men who were older, spoke Spanish, and did not use a computer frequently had a higher probability of choosing in-person groups versus web-based videos. For men who selected a group delivery option, 86.9% (107/123) attended ≥25% of the sessions, 83.7% (103/123) attended ≥50% of the sessions, and 73.2% (90/123) attended ≥75% of the sessions, with no differences by type of group (videoconference or in person). Men who

chose videoconference and in-person group sessions lost significantly more weight at 6 months (both $P<.001$) and 18 months ($P=.02$ and $P=.04$, respectively) than those who chose web-based videos. Men who chose in-person group sessions also lost significantly more weight at 12 months ($P=.008$) than those who chose web-based videos.

Conclusions: There were significant differences according to demographic, employment, cultural, and technology use factors between men who chose 1 of the 3 intervention options. Men who chose one of the group-based options (videoconference or in person) lost significantly more weight than those who chose web-based videos. Providing options that accommodate the diversity of Latino men's preferences is important for increasing engagement in behavioral interventions.

Trial Registration: ClinicalTrials.gov NCT03092960; https://clinicaltrials.gov/ct2/show/NCT03092960

(JMIR Form Res 2022;6(2):e29537) doi:10.2196/29537

KEYWORDS

Latino men's health; technology-mediated behavioral interventions; weight management; mobile phone

Introduction

Background

Obesity is a major contributor to the leading causes of death among all men in the United States, such as heart disease, cancer, and type 2 diabetes [1,2]. As part of the largest minority group in the United States, Latino men are disproportionately represented among men with obesity compared with all other races and ethnicities [3]. The US Preventive Services Task Force recommends intensive behavioral lifestyle interventions for obesity treatment [4]. However, research to identify effective behavioral lifestyle interventions has derived primarily from research on non-Hispanic White women [5]. Minority men and Latino men specifically have been particularly underrepresented in this research [5]. Thus, it is imperative to develop and scale interventions tailored to Latino men.

Technology, including web-based and smartphone apps, offers opportunities for extending the reach and engagement of behavioral lifestyle interventions to priority populations such as Latino men. Technology can help overcome prevalent barriers to engaging in interventions because of competing priorities from family, inflexible work schedules, and unreliable transportation. In addition, technology-mediated approaches became essential during the COVID-19 pandemic, when in-person meetings were not allowed or encouraged. Latinos increasingly have access to technology in general and smartphones specifically, which makes this a promising approach to maximize reach and engagement in this population [6-9]. However, little is known about Latino men's preferences for intervention delivery formats, especially those using technology [10].

Objectives

The HOMBRE (Hombres con Opciones para Mejorar su Bienestar para Reducir Enfermedades Crónicas; English translation: Men With Options to Improve Their Well-being and Reduce Chronic Disease) trial was designed to compare a culturally adapted behavioral lifestyle intervention for Latino men with minimal-intensity control. The culturally adapted behavioral lifestyle intervention offered men 3 options for engaging in the intervention sessions: coach-facilitated group sessions using web-based videoconferencing, prerecorded videos of group sessions available on the web, and coach-facilitated group sessions in person. The 3 choices differed in the used technology, the level of coach and peer support, and language options. The goal of this study is to examine Latino men's preferences among the 3 intervention options according to demographic, clinical, employment, cultural, and technology use and access factors. We also examined whether attendance (for the videoconference and in-person groups) and weight loss differed among the intervention options. Understanding Latino men's preferences according to key baseline characteristics can inform future implementation of technology-based interventions for this high-priority population.

Methods

Study Design

The institutional review board for Sutter Health, Northern California, and Stanford University approved the study. All participants provided written informed consent. The trial protocol has been previously published [11]. Participants' deidentified study data and identifiers were protected following the Protection of Human Subjects protocol. Study recruitment and intervention were not affected by the COVID-19 pandemic as they were completed before the pandemic.

Recruitment and Participants

A total of 424 Latino men who had a BMI of ≥27 kg/m^2 and ≥1 cardiometabolic risk factor (high waist circumference, high triglycerides, high blood pressure, high fasting plasma glucose, or low high-density lipoprotein cholesterol) were enrolled in the HOMBRE trial following a multistep process, as described in the trial protocol [11]. Patients with significant psychiatric (eg, bipolar or psychotic disorder) or medical comorbidities (eg, active cancer or organ failure) were excluded. Participants were randomly assigned to receive the 12-month HOMBRE behavioral lifestyle intervention adapted for Latino men (212/424, 50%) or a minimal-intensity intervention (212/424, 50%). This study included men who participated in the HOMBRE intervention and attended an orientation session to make a choice on intervention delivery (200/424, 47.2%). Of the 212 men assigned to the HOMBRE intervention, 200 (94.3%) attended an orientation session and made a choice on intervention session delivery.

Description of the Intervention

The HOMBRE intervention was based on the Group Lifestyle Balance (GLB) intervention, a group-based adaptation of the

original Diabetes Prevention Program intervention [12-14] grounded in social cognitive theory [15]. Social cognitive theory emphasizes a triadic, reciprocally deterministic relationship between the individual, environment, and behavior. The year-long HOMBRE intervention included 12 weekly sessions during the intensive phase (months 1-3) and 8 monthly contacts during the maintenance phase (months 4-12) either by phone or email. The HOMBRE intervention offered men 3 options for engaging in the 12 weekly sessions: coach-facilitated group sessions using web-based videoconferencing, prerecorded videos of group sessions available on the web, and coach-facilitated group sessions in person (Table 1). A trained health coach facilitated the group sessions on videoconference and in person using a cultural adaptation of the GLB. The Latino Patient Advisory Board adapted the GLB and made the following major changes: (1) added an orientation session before session 1 to provide a brief overview of the intervention to participants and family members; (2) incorporated the *MyPlate* visual in the orientation and in the early intervention sessions, given its effectiveness for communicating the types of food choices recommended by the intervention; and (3) invited family members to sessions 6 and 12, given the cultural importance of family support [16]. The prerecorded videos did not include any of these adaptations, except that health coaches encouraged men to watch the videos with a family member. In all 3 options, health coaches encouraged men to self-monitor weight using a study-provided digital scale, physical activity using a study-provided wearable activity tracker, and dietary intake using MyFitnessPal web or a smartphone app (available in Spanish and English).

Table 1. Session delivery options in the HOMBRE (Hombres con Opciones para Mejorar su Bienestar y Reducir Enfermedades Crónicas) intervention.

Characteristics	Videoconference	Web-based videos	In person
Description	• A bilingual, bicultural coach facilitated weekly sessions on a videoconferencing platform (Zoom)	• Men were given access to prerecorded web-based videos of coach-facilitated group sessions	• A bilingual, bicultural coach facilitated weekly sessions at the clinic where men were recruited
Coach support	• Coach-facilitated sessions • Feedback from the coach on diet and physical activity monitoring	• Self-directed sessions • Option to contact the coach for feedback on diet and physical activity monitoring	• Coach-facilitated sessions • Feedback from the coach on diet and physical activity monitoring
Peer support	• Support from other members of the group	• No peer support	• Support from other members of the group
Frequency of sessions	• Weekly	• Self-paced; weekly recommended	• Weekly
Language	• English or Spanish per preference	• English with Spanish subtitles	• English or Spanish per preference

In addition to the differences in delivery, the 3 intervention options differed in level and type of coach feedback (Table 1). The videoconferencing and in-person options had the highest level of *real-time* coach involvement. Using the tracking data from participants, the health coach provided individualized feedback on diet and physical activity goals during the intensive phase. Individualized feedback provided ample opportunity for tailoring based on cultural and other individual differences. During the maintenance phase, men who chose the videoconferencing and in-person options received monthly phone calls from the coach, which focused on supporting continued goal progress and problem solving for encountered barriers. The web-based video format had the lowest level of coach involvement, according to the protocol, as participants watched prerecorded videos of a coach facilitating the 12 sessions with a multiethnic group of men and women. Men in this option received standardized weekly messages in months 1 to 3, with reminders to watch the videos, use written materials, self-monitor, and reach out to their assigned coach with questions or requests for more individualized feedback. This was in alignment with this intervention choice, which included less proactive coach interaction. In months 4 to 12, men in the web-based video option received monthly standardized messages that included handouts on maintenance topics, reminders to self-monitor, and contact information for the coach.

Intervention Choice

Following randomization, all participants attended an in-person group orientation session. Trained health coaches offered the orientation sessions in English and Spanish at different times of the day throughout the 2 weeks following randomization to accommodate all participants. Group sessions were designed for approximately 10 participants but could accommodate varying sizes as needed. Participants were encouraged to bring their partners and other family members to the orientation session with the purpose of increasing understanding and social support among family members. Group orientation sessions followed a standardized protocol that featured a didactic component to provide information on the background and goals of the intervention and a small group discussion component specific to their randomization arm. The goal of the small group discussion for the HOMBRE arm was to support men in making a choice among the 3 intervention options. The small group discussion included 3 components to support men in making a choice: (1) description of the 3 intervention choices provided by a health coach, (2) a worksheet that helped men reflect on each of the 3 options, and (3) a small group discussion on the

pros and cons of each option. The worksheet prompted men to think about whether they liked participating in groups, the degree to which they would like support from a coach, their comfort with new technologies, and their availability for attending regular weekly sessions. In the small group, the coach asked men to discuss the pros and cons of each option to assist them in considering the choices. At the end of the session, men were asked to make their choice, given their intervention materials, and assisted with using the activity tracker and other technologies based on their choice.

The intervention participants' initial choices of intervention delivery were used to group them into the 3 options (videoconference, web-based videos, and in person). The participants were allowed to change their choice within 4 weeks. However, none of the participants elected to change. Coaches could transfer patients from the videoconference or in-person groups to the web-based videos if they did not attend the first 4 sessions; 15 men (n=11, 73% from the videoconferencing group and n=4, 27% from the in-person group) were transferred to the web-based videos group.

Measures

Baseline characteristics included demographic (eg, age, income, education, marital status, and household size) and clinical characteristics (eg, weight, waist circumference, blood pressure, depression symptoms, quality of life, and sleep function), employment (eg, employment status and occupation), cultural characteristics (eg, language, acculturation, and health literacy), and technology use and access. Weight, waist circumference, and blood pressure were measured in duplicate according to standard protocols [17-19]. Depression symptoms were measured using the 9-item Patient Health Questionnaire, with scores between 0 (best) and 27 (worst) [20,21]. Quality of life measures included the health-related quality of life EuroQol 5-dimension questionnaire [22,23], scored on 5 levels (no, slight, moderate, severe, or extreme problems) for 5 domains, and obesity-specific quality of life, with higher scores (range 0-100) indicating more obesity-related psychosocial problems [24,25]. Sleep function was measured using the Patient-Reported Outcomes Measurement Information System sleep disturbance and sleep-related impairment questionnaires [26]. Sleep disturbance and sleep-related impairment T scores ranged from 25 (high disturbance or impairment) to 80 (none at all), with a mean score of 50 (SD 10) representing the average of the calibration sample, which was generally more enriched for chronic illness. The level of acculturation was assessed using the Short Acculturation Scale for Hispanics, with higher scores (range 1-5) indicating higher acculturation to US society [27]. Health literacy was assessed using the newest vital sign, which uses a food nutrition label, with higher scores (range 0-6) indicating higher health literacy [28]. Technology use and access were assessed using a survey adapted from the Pew Hispanic Trust on technology access and use [29]. Session attendance was recorded in the videoconference and in-person groups only.

Statistical Analysis

Overview

We used 2 steps to identify the different profiles of demographic, clinical, employment, cultural, and technology use and access characteristics based on their choice of delivery options. First, we performed a bivariate analysis to choose a set of candidate variables. Second, we conducted multivariate analysis based on the variables identified in the bivariate analysis to derive the baseline characteristic profiles that significantly differentiated the men who chose 1 of the 3 intervention delivery options [30].

Bivariate Analysis

Percentages and means and SDs were used to describe the baseline characteristics among HOMBRE intervention participants overall and by intervention delivery option chosen. We used the Fisher least significant difference method [31] to test for significant differences among all 3 options, which included 2 steps. First, we used analysis of variance (for continuous variables) and chi-square tests (for categorical variables) to compare overall differences in demographic, clinical, employment, cultural, and technology use and access factors across the 3 intervention delivery options. Second, variables with $P<.05$ (2-tailed) from the first step were then further assessed for pairwise comparisons using Student *t* tests for continuous variables and chi-square tests for categorical variables.

Multivariate Analysis

Canonical discriminant analysis was used to derive linear combinations of the baseline characteristic profiles that significantly differentiated the men who chose 1 of the 3 intervention delivery options. Canonical discriminant analysis is a multivariate dimension reduction technique that derives a linear combination of explanatory variables that has the highest possible multiple correlation with the groups of a classification variable. The dimension defined by the linear combination is the first canonical dimension. This maximum multiple correlation is called the first canonical correlation. The coefficients of the linear combination are the canonical coefficients. The second canonical dimension is obtained by finding the linear combination with the next highest possible multiple correlation with the groups that is uncorrelated with the first canonical dimension. The process of extracting canonical dimensions can be repeated until the number of canonical dimensions equals the number of original variables or the number of groups minus 1, whichever is smaller. We included only the baseline characteristics with $P<.15$ from the bivariate analyses [32]. The categorical variables were coded as dummy variables in the canonical discriminant analysis. Standardized canonical coefficients measured the strength and direction of the correlation of each dimension with the characteristics. Participant scores on each dimension were calculated as the sum of the products of the canonical coefficients and the participant's individual values for the characteristics. These scores were then compared among the 3 intervention delivery options using analysis of variance.

We used generalized linear models to compare weight loss among men in different intervention options at 6, 12, and 18

months after randomization. We examined weight according to the men's initial and final intervention choices. Weight was measured by trained study staff at baseline. Weight was measured using a standard calibrated scale at baseline and 18 months at local clinic sites according to standard protocol [17]. Participants also self-reported weight using the study-provided digital scales at 6, 12, and 18 months (if no study-measured weight). According to the trial protocol, in the case of missing study-measured weight, the closest weight measurement from the electronic health record within 3 months of the due date of a missed 18-month visit or self-reported weight (if no electronic health record weight) was used.

Session attendance was calculated for weekly sessions and monthly phone calls combined and separately in the videoconference and in-person groups, excluding the 15 participants who were transferred to the web-based video option. Session attendance was then compared between the videoconference and in-person groups using Student *t* tests.

All analyses were conducted using SAS version 9.4 (SAS Institute Inc). Statistical significance was defined as $P<.05$ (2-sided).

Results

Baseline Characteristics

As shown in Tables 2 and 3, the participants who attended were middle-aged (mean age 47.3, SD 11.8 years), educated (150/195, 76.9% attended at least some college), employed full- or part-time (167/194, 86.1%), and had access to a computer (165/184, 89.7%) and smartphone (179/184, 97.3%). Among these men, 28% (56/200) chose web-based videoconference groups at the orientation, 31% (62/200) chose web-based videos, and 41% (82/200) chose in-person groups. In addition, most preferred to engage in the intervention in English (142/200, 71%) versus Spanish (58/200, 29%).

Table 2. Baseline characteristics overall and by initial choice of intervention delivery (N=200).

Characteristic	All	Videoconference (n=56)	Web-based videos (n=62)	In-person group (n=82)	*P* value
Demographic					
Age (years), mean (SD)	47.3 (11.8)	45.6 (10.9)[a]	45.4 (11.4)[a]	50.0 (12.3)[b]	.03
Income (US $; n=167), n (%)					.02
<75,000	49 (29.3)	8 (16)[a]	14 (25.9)[a,b]	27 (42.9)[b]	
75,000-<150,000	53 (31.7)	17 (34)[a]	17 (31.5)[a,b]	19 (30.2)[b]	
≥150,000	65 (38.9)	25 (50)[a]	23 (42.6)[a,b]	17 (27)[b]	
Education (n=195), n (%)					.03
High school, GED[c], or less	45 (23.1)	7 (12.7)[a]	12 (19.7)[a,b]	26 (32.9)[b]	
Some college	49 (25.1)	12 (21.8)[a]	19 (31.1)[a,b]	18 (22.8)[b]	
College graduate	58 (29.7)	17 (30.9)[a]	21 (34.4)[a,b]	20 (25.3)[b]	
More than college	43 (22.1)	19 (34.5)[a]	9 (14.8)[a,b]	15 (19)[b]	
Marital status (n=196), n (%)					.07
Married or living with a partner	156 (79.6)	38 (69.1)	52 (83.9)	66 (83.5)	
Single, separated, divorced, or widowed	40 (20.4)	17 (30.9)	10 (16.1)	13 (16.5)	
Household size (n=191), n (%)					.51
1-2	13 (6.8)	4 (7.3)	4 (6.7)	5 (6.6)	
3	35 (18.3)	10 (18.2)	10 (16.7)	15 (19.7)	
4	44 (23)	19 (34.6)	10 (16.7)	15 (19.7)	
5	49 (25.7)	10 (18.2)	18 (30)	21 (27.6)	
≥6	50 (26.2)	12 (21.8)	18 (30)	20 (26.3)	
Clinical					
BMI (kg/m^2), mean (SD)	33.1 (5.2)	33.7 (5.7)	32.6 (5.4)	33.0 (4.6)	.48
Weight (kg), mean (SD)	101.5 (33.7)	102.7 (22.3)	96.7 (17.3)	104.3 (46.8)	.39
Waist circumference (cm), mean (SD)	109.5 (12.3)	111.0 (14.0)	106.8 (11.8)	110.6 (11.2)	.11
Number of metabolic risks, mean (SD)	2.0 (0.9)	1.9 (0.9)	2.0 (0.8)	2.0 (0.9)	.62
SBP[d] (mm Hg), mean (SD)	122.4 (12.2)	121.2 (13.9)	121.5 (11.2)	123.9 (11.6)	.36
DBP[e] (mm Hg), mean (SD)	78.6 (9.4)	78.1 (10.0)	79.2 (9.0)	78.5 (9.4)	.81
PHQ-9[f] score (n=196), mean (SD)	3.5 (3.7)	3.6 (4.2)	3.6 (3.3)	3.2 (3.7)	.72
EQ-5D-5L[g] : mobility (n=196), n (%)					.63
No problems	170 (86.7)	50 (90.9)	52 (83.9)	68 (86.1)	
Slight problems	16 (8.2)	5 (9.1)	5 (8.1)	6 (7.6)	
Moderate problems	8 (4.1)	0 (0)	4 (6.5)	4 (5.1)	
Severe problems	2 (1)	0 (0)	1 (1.6)	1 (1.3)	
Extreme problems	0 (0)	0 (0)	0 (0)	0 (0)	
EQ-5D-5L: self-care (n=196), n (%)					.64
No problems	193 (98.5)	54 (98.2)	62 (100)	77 (97.5)	
Slight problems	2 (1)	1 (1.8)	0 (0)	1 (1.3)	
Moderate problems	1 (0.5)	0 (0)	0 (0)	1 (1.3)	
Severe problems	0 (0)	0 (0)	0 (0)	0 (0)	

Characteristic	All	Videoconference (n=56)	Web-based videos (n=62)	In-person group (n=82)	*P* value
Extreme problems	0 (0)	0 (0)	0 (0)	0 (0)	
EQ-5D-5L: usual activities (n=196), n (%)					.68
No problems	171 (87.2)	50 (90.9)	51 (82.3)	70 (88.6)	
Slight problems	19 (9.7)	4 (7.3)	8 (12.9)	7 (8.9)	
Moderate problems	6 (3.1)	1 (1.8)	3 (4.8)	2 (2.5)	
Severe problems	0 (0)	0 (0)	0 (0)	0 (0)	
Extreme problems	0 (0)	0 (0)	0 (0)	0 (0)	
EQ-5D-5L: pain and discomfort (n=196), n (%)					.15
No problems	103 (52.6)	25 (45.5)	29 (46.8)	49 (62)	
Slight problems	69 (35.2)	26 (47.3)	23 (37.1)	20 (25.3)	
Moderate problems	20 (10.2)	4 (7.3)	8 (12.9)	8 (10.1)	
Severe problems	4 (2)	0 (0)	2 (3.2)	2 (2.5)	
Extreme problems	0 (0)	0 (0)	0 (0)	0 (0)	
EQ-5D-5L: anxiety and depression (n=196), n (%)					.73
No problems	147 (75)	40 (72.7)	46 (74.2)	61 (77.2)	
Slight problems	35 (17.9)	9 (16.4)	12 (19.4)	14 (17.7)	
Moderate problems	13 (6.6)	5 (9.1)	4 (6.5)	4 (5.1)	
Severe problems	1 (0.5)	1 (1.8)	0 (0)	0 (0)	
Extreme problems	0 (0)	0 (0)	0 (0)	0 (0)	
Obesity-related problem raw score (n=196), mean (SD)	0.8 (0.7)	0.9 (0.8)	0.8 (0.7)	0.7 (0.7)	.29
PROMIS[h] sleep disturbance T score (n=196), mean (SD)	47.3 (8.9)	47.7 (9.1)	48.8 (7.5)	45.8 (9.7)	.13
PROMIS sleep impairment T score (n=195), mean (SD)	47.5 (9.2)	48.4 (9.4)	48.2 (8.8)	46.3 (9.3)	.34
Employment					
Employment status (n=194), n (%)					.24
Employed	167 (86.1)	51 (92.7)	55 (88.7)	61 (79.2)	
Unemployed	5 (2.6)	1 (1.8)	1 (1.6)	3 (3.9)	
Other (homemaker, student, retired, and disabled or not able to work)	22 (11.3)	3 (5.5)	6 (9.7)	13 (16.9)	
Work (hours per week; n=188), n (%)					.18
0 to <5	15 (8)	2 (3.8)	3 (4.9)	10 (13.5)	
5 to <30	9 (4.8)	2 (3.8)	2 (3.3)	5 (6.8)	
30 to <50	95 (50.5)	26 (49.1)	30 (49.2)	39 (52.7)	
≥50	69 (36.7)	23 (43.4)	26 (42.6)	20 (27)	
Occupation (n=186), n (%)					.04
Mostly sitting or standing	128 (68.8)	44 (81.5)[a]	39 (67.2)[a,b]	45 (60.8)[b]	
Mostly walking or heavy work	58 (31.2)	10 (18.5)[a]	19 (32.8)[a,b]	29 (39.2)[b]	
White- or blue-collar (n=152), n (%)					.02
White collar	108 (71.1)	39 (86.7)[a]	33 (68.8)[b]	36 (61)[b]	
Blue collar	44 (29)	6 (13.3)[a]	15 (31.3)[b]	23 (39)[b]	
Culture					

Characteristic	All	Videoconference (n=56)	Web-based videos (n=62)	In-person group (n=82)	*P* value
Language[i], n (%)					<.001
English	142 (71)	48 (85.7)[a]	61 (98.4)[j]	33 (40.2)[b]	
Spanish	58 (29)	8 (14.3)[a]	1 (1.6)[j]	49 (59.8)[b]	
Acculturation score, mean (SD)	3.4 (0.9)	3.7 (0.6)[a]	3.6 (0.7)[a]	3.1 (1.0)[b]	<.001
Health literacy score, mean (SD)	4.4 (1.9)	5.0 (1.5)[a]	4.5 (1.9)[a,b]	4.0 (2.0)[b]	.004
Health literacy category, n (%)					.02
Adequate literacy	147 (73.5)	49 (87.5)[a]	47 (75.8)[a,b]	51 (62.2)[b]	
Possibility of limited literacy	31 (15.5)	4 (7.1)[a]	8 (12.9)[a,b]	19 (23.2)[b]	
High likelihood (≥50%) of limited literacy	22 (11)	3 (5.4)[a]	7 (11.3)[a,b]	12 (14.6)[b]	

[a]Different superscripts denote statistically significant differences.

[b]Different superscripts denote statistically significant differences.

[c]GED: General Educational Development.

[d]SBP: systolic blood pressure.

[e]DBP: diastolic blood pressure.

[f]PHQ-9: Patient Health Questionnaire-9.

[g]EQ-5D-5L: EuroQol 5-dimension 5-level.

[h]PROMIS: Patient-Reported Outcome Measurement System.

[i]Language that the patient preferred for the intervention.

[j]Different superscripts denote statistically significant differences.

Table 3. Baseline technology use and access overall and by initial choice of intervention delivery (N=184).

Characteristic	All, n (%)	Videoconference (n=54), n (%)	Web-based videos (n=55), n (%)	In-person group (n=75), n (%)	*P* value
Desktop or laptop computer					.01
Yes	165 (89.7)	54 (100)[a]	50 (90.9)[b]	61 (81.3)[b]	
No	18 (9.8)	0 (0)[a]	5 (9.1)[b]	13 (17.3)[b]	
Declined to state	1 (0.5)	0 (0)[a]	0 (0)[b]	1 (1.3)[b]	
Smartphone (eg, iPhone or Android)					.51
Yes	179 (97.3)	54 (100)	53 (96.4)	72 (96)	
No	4 (2.2)	0 (0)	2 (3.6)	2 (2.7)	
Declined to state	1 (0.5)	0 (0)	0 (0)	1 (1.3)	
Cell phone but not smartphone					.43
Yes	18 (9.8)	3 (5.6)	9 (16.4)	6 (8)	
No	160 (87)	50 (92.6)	44 (80)	66 (88)	
I'm not sure	1 (0.5)	0 (0)	0 (0)	1 (1.3)	
Declined to state	5 (2.7)	1 (1.9)	2 (3.6)	2 (2.7)	
Tablet (eg, iPad or Kindle)					.07
Yes	139 (75.5)	48 (88.9)	39 (70.9)	52 (69.3)	
No	44 (23.9)	6 (11.1)	16 (29.1)	22 (29.3)	
Declined to state	1 (0.5)	0 (0)	0 (0)	1 (1.3)	
Dial-up internet service at home					.21
Yes	30 (16.3)	4 (7.4)	8 (14.6)	18 (24)	
No	149 (81)	49 (90.7)	46 (83.6)	54 (72)	
I'm not sure	1 (0.5)	0 (0)	0 (0)	1 (1.3)	
Declined to state	4 (2.2)	1 (1.9)	1 (1.8)	2 (2.7)	
Higher-speed broadband internet service such as DSL[c], cable, or fiber optic service at home					.09
Yes	173 (94)	53 (98.2)	52 (94.6)	68 (90.7)	
No	5 (2.7)	1 (1.9)	3 (5.5)	1 (1.3)	
I'm not sure	5 (2.7)	0 (0)	0 (0)	5 (6.7)	
Declined to state	1 (0.5)	0 (0)	0 (0)	1 (1.3)	
How often do you use a computer at work, school, home, or anywhere else?					.002
Never	12 (6.5)	0 (0)[a]	2 (3.6)[b]	10 (13.3)[b]	
Rarely	15 (8.2)	0 (0)[a]	6 (10.9)[b]	9 (12)[b]	
Sometimes	7 (3.8)	0 (0)[a]	2 (3.6)[b]	5 (6.7)[b]	
Often	26 (14.1)	7 (13)[a]	9 (16.4)[b]	10 (13.3)[b]	
Very often	124 (67.4)	47 (87)[a]	36 (65.5)[b]	41 (54.7)[b]	
How often do you communicate with others by email?					.003
Never	13 (7.1)	0 (0)[a]	4 (7.3)[a,b]	9 (12)[b]	
Rarely	15 (8.2)	1 (1.9)[a]	2 (3.6)[a,b]	12 (16)[b]	
Sometimes	31 (16.9)	7 (13)[a]	14 (25.5)[a,b]	10 (13.3)[b]	
Often	29 (15.8)	11 (20.4)[a]	7 (12.7)[a,b]	11 (14.7)[b]	
Very often	96 (52.2)	35 (64.8)[a]	28 (50.9)[a,b]	33 (44)[b]	

Characteristic	All, n (%)	Videoconference (n=54), n (%)	Web-based videos (n=55), n (%)	In-person group (n=75), n (%)	*P* value
How often do you talk to others using software or an app with video chat and voice call services?					.01
Never	38 (20.7)	7 (13)[a]	8 (14.6)[a,b]	23 (30.7)[b]	
Rarely	44 (23.9)	7 (13)[a]	15 (27.3)[a,b]	22 (29.3)[b]	
Sometimes	34 (18.5)	15 (27.8)[a]	9 (16.4)[a,b]	10 (13.3)[b]	
Often	24 (13)	7 (13)[a]	7 (12.7)[a,b]	10 (13.3)[b]	
Very often	43 (23.4)	18 (33.3)[a]	16 (29.1)[a,b]	9 (12)[b]	
Declined to state	1 (0.5)	0 (0)[a]	0 (0)[a,b]	1 (1.3)[b]	
How often do you access the internet on a cell phone, tablet, or other mobile handheld device?					.22
Never	6 (3.3)	0 (0)	2 (3.6)	4 (5.3)	
Rarely	4 (2.2)	0 (0)	1 (1.8)	3 (4)	
Sometimes	19 (10.3)	2 (3.7)	8 (14.6)	9 (12)	
Often	22 (12)	6 (11.1)	5 (9.1)	11 (14.7)	
Very often	132 (71.7)	46 (85.2)	39 (70.9)	47 (62.7)	
Declined to state	1 (0.5)	0 (0)	0 (0)	1 (1.3)	
How often do you use your cell phone to send or receive emails?					.10
Never	11 (6)	0 (0)	2 (3.6)	9 (12)	
Rarely	20 (10.9)	8 (14.8)	2 (3.6)	10 (13.3)	
Sometimes	30 (16.3)	8 (14.8)	11 (20)	11 (14.7)	
Often	41 (22.3)	11 (20.4)	16 (29.1)	14 (18.7)	
Very often	80 (43.5)	27 (50)	23 (41.8)	30 (40)	
Declined to state	2 (1.1)	0 (0)	1 (1.8)	1 (1.3)	
How often do you use your cell phone to send or receive SMS text messages?					.31
Never	4 (2.2)	0 (0)	2 (3.6)	2 (2.7)	
Rarely	5 (2.7)	0 (0)	1 (1.8)	4 (5.3)	
Sometimes	14 (7.6)	1 (1.9)	5 (9.1)	8 (10.7)	
Often	36 (19.6)	12 (22.2)	10 (18.2)	14 (18.7)	
Very often	122 (66.3)	41 (75.9)	36 (65.5)	45 (60)	
Declined to state	3 (1.6)	0 (0)	1 (1.8)	2 (2.7)	
How often do you use apps you downloaded to your mobile device (eg, smartphone or tablet)?					.10
Never	9 (4.9)	0 (0)	3 (5.5)	6 (8)	
Rarely	16 (8.7)	5 (9.3)	2 (3.6)	9 (12)	
Sometimes	32 (17.4)	6 (11.1)	14 (25.5)	12 (16)	
Often	43 (23.4)	13 (24.1)	11 (20)	19 (25.3)	
Very often	82 (44.6)	30 (55.6)	25 (45.5)	27 (36)	
Declined to state	2 (1.1)	0 (0)	0 (0)	2 (2.7)	
I feel completely comfortable using a web-based videoconferencing tool if someone shows me how to use it					.13
Strongly disagree	2 (1.1)	1 (1.9)	0 (0)	1 (1.3)	
Disagree	8 (4.4)	1 (1.9)	3 (5.5)	4 (5.3)	
Neither agree nor disagree	12 (6.5)	0 (0)	6 (10.9)	6 (8)	
Agree	62 (33.7)	15 (27.8)	16 (29.1)	31 (41.3)	
Strongly agree	99 (53.8)	37 (68.5)	30 (54.6)	32 (42.7)	

Characteristic	All, n (%)	Videoconference (n=54), n (%)	Web-based videos (n=55), n (%)	In-person group (n=75), n (%)	*P* value
Declined to state	1 (0.5)	0 (0)	0 (0)	1 (1.3)	
I feel completely comfortable using smartphone apps to track my diet or physical activity if someone shows me how to use them					.49
Strongly disagree	5 (2.7)	1 (1.9)	2 (3.6)	2 (2.7)	
Disagree	7 (3.8)	0 (0)	2 (3.6)	5 (6.7)	
Neither agree nor disagree	14 (7.6)	2 (3.7)	4 (7.3)	8 (10.7)	
Agree	58 (31.5)	16 (29.6)	18 (32.7)	24 (32)	
Strongly agree	99 (53.8)	35 (64.8)	29 (52.7)	35 (46.7)	
Declined to state	1 (0.5)	0 (0)	0 (0)	1 (1.3)	
I feel completely comfortable watching web-based videos using my electronic device, such as laptop or tablet					.12
Strongly disagree	5 (2.7)	1 (1.9)	1 (1.8)	3 (4)	
Disagree	5 (2.7)	0 (0)	1 (1.8)	4 (5.3)	
Neither agree nor disagree	10 (5.4)	0 (0)	4 (7.3)	6 (8)	
Agree	57 (31)	13 (24.1)	18 (32.7)	26 (34.7)	
Strongly agree	106 (57.6)	40 (74.1)	31 (56.4)	35 (46.7)	
Declined to state	1 (0.5)	0 (0)	0 (0)	1 (1.3)	

[a]Different superscripts denote statistically significant differences.

[b]Different superscripts denote statistically significant differences.

[c]DSL: digital subscriber line.

Bivariate Associations

Participants in the 3 intervention delivery options had similar clinical characteristics (eg, BMI, blood pressure, depression symptoms, and overall and obesity-specific quality of life) but differed significantly according to demographic, employment, cultural, and technology use factors (Tables 2 and 3). Men who chose videoconference were more likely to be younger (*P*=.04), have higher income (*P*=.005) and education (*P*=.03), have a white-collar job (*P*=.004), prefer English (*P*<.001), be more acculturated (*P*<.001), have higher health literacy (*P*<.001), have access to a computer (*P*=.004), and have higher technology skills (eg, computer, *P*<.001; email, *P*=.003; and video chat app use, *P*=.002) compared with men who chose the in-person group. Similarly, men who chose web-based videos were more likely to be younger (*P*=.02) and more acculturated (*P*=.001) than those who chose the in-person group. They were also more likely to have a blue-collar job compared with those who chose videoconference (*P*=.04) and more likely to speak English than men who chose the videoconference (*P*=.01) or in-person (*P*<.001) groups. Men who chose web-based videos had intermediate technology skills compared with those who chose either the videoconference or in-person group and were less likely to have (*P*=.02) and use a computer (*P*=.02) than those who chose videoconference but as likely to use emails (*P*=.08) and video chat apps (*P*=.33) as men who chose videoconference.

Multivariate Associations

Canonical discriminant analysis identified 1 orthogonal dimension representing statistically significant combinations of the baseline characteristics. The canonical variates of this single dimension explained 41% of the total variance of the choice of 3 intervention delivery options. Participants electing the in-person and web-based video options had the most extreme mean scores (0.98 vs −0.89; *P*<.001) on the canonical dimension 1. This signified that this dimension distinguished most significantly between these 2 intervention choices. According to characteristics with the highest positive or negative correlation coefficients and using 0.25 as a cutoff (Table 4), participants who were older, spoke Spanish, and did not use a computer frequently had a higher probability of choosing the in-person option versus the web-based videos option.

Table 4. Standardized coefficients from canonical discriminant analysis for individual baseline characteristics of participants in the HOMBRE (Hombres con Opciones para Mejorar su Bienestar para Reducir Enfermedades Crónicas) intervention (N=175)[a].

Characteristic	Dimension 1[b] coefficients
Demographic	
Age	0.33
Education (reference: high school, GED[c], or less)	
Some college	0.03
College graduate	0.12
More than college	0.10
Marital status (reference: single, separated, divorced, or widowed)	
Married or living with another person	−0.22
Clinical	
Waist circumference	0.11
Sleep disturbance T score	−0.20
Employment	
Occupation (reference: no job)	
Mostly sitting or standing	0.23
Mostly walking or heavy work	0.15
Culture	
Language (reference: English)	
Spanish	1.03
Short Acculturation Scale for Hispanics	0.09
Health literacy score	0.00
Technology use and access	
If they had a desktop or laptop computer (reference: no)	
Yes	0.15
If they had a tablet (eg, iPad or Kindle; reference: no)	
Yes	0.13
If they had higher-speed broadband internet service such as DSL[d], cable, or fiber optic service at home (reference: no)	
Yes	0.09
How often do you use a computer at work, school, home, or anywhere else? (reference: never, rarely, or sometimes)	
Often or very often	−0.27
How often do you communicate with others by email? (reference: never, rarely, or sometimes)	
Often or very often	−0.08
How often do you talk to others using software or an app with video chat and voice call services? (reference: never, rarely, or sometimes)	
Often or very often	−0.06
How often do you use your cell phone to send or receive emails? (reference: never, rarely, or sometimes)	
Often or very often	0.03
How often do you use apps you downloaded on your mobile device (eg, smartphone or tablet; reference: never, rarely, or sometimes)?	

Characteristic	Dimension 1[b] coefficients
Often or very often	−0.01
I feel completely comfortable using a web-based videoconferencing tool if someone shows me how to use it (reference: neither agree nor disagree, disagree, or strongly disagree).	
Agree or strongly agree	−0.08
I feel completely comfortable watching web-based videos using my electronic device, such as laptop or tablet (reference: neither agree nor disagree, disagree, or strongly disagree).	
Agree or strongly agree	0.13

[a]Results based on 175 HOMBRE intervention participants who had complete data for all baseline characteristics used in the canonical discriminant analysis.

[b]Dimension 1: canonical function $F_{44,302}$=2.82 (P<.001); R^2 of the canonical correlation=0.41.

[c]GED: General Educational Development.

[d]DSL: digital subscriber line.

Weight Loss

Compared with men who initially chose web-based videos, those who initially chose videoconference and in-person group sessions lost significantly more weight at 6 months (mean −3.9, SD 6.1 kg for videoconference and mean −4.3, SD 5.3 kg for in-person vs mean −0.3, SD 3.7 kg for web-based videos; P<.001 for both comparisons) and 18 months (mean −3.8, SD 8.4 kg for videoconference and mean −3.3, SD 6.0 kg for in-person vs mean −0.9, SD 4.6 kg for web-based videos; P=.02 and P=.04, respectively, for the comparisons), and those who chose in-person group sessions also lost significantly more weight at 12 months (mean −4.1, SD 6.0 kg vs mean −1.0, SD 4.8 kg; P=.008; Table 5). After accounting for men who were transferred to the web-based videos group at 4 weeks into the intervention after not attending group sessions in the videoconference (11/15, 73%) and in-person (4/15, 27%) group sessions, men in both of the group session formats (videoconference and in-person) lost significantly more weight than men in the web-based videos group at 6, 12, and 18 months.

Table 5. Weight change by initial and final choice of intervention delivery (N=200).

Weight change[a]	Initial intervention choice, mean (SD)			Final intervention choice, mean (SD)		
	Videoconference (n=56)	Web-based videos (n=62)	In-person group (n=82)	Videoconference (n=45)	Web-based videos (n=77)	In-person group (n=78)
Weight change at 6 months from baseline (kg)	−3.9 (6.1)[b]	−0.3 (3.7)	−4.3 (5.3)[b]	−4.8 (5.7)[b]	−0.4 (4.3)	−4.4 (5.3)[b]
Weight change at 12 months from baseline (kg)	−3.4 (7.9)	−1.0 (4.8)	−4.1 (6.0)[c]	−4.5 (8.0)[c]	−0.6 (4.8)	−4.2 (6.1)[c]
Weight change at 18 months from baseline (kg)	−3.8 (8.4)[d]	−0.9 (4.6)	−3.3 (6.0)[d]	−4.8 (8.6)[c]	−0.8 (4.6)	−3.2 (6.2)[d]

[a]Electronic health record–abstracted and self-reported weights were used for 6 and 12 months, and study-measured, electronic health record–abstracted, and self-reported weights were used for 18 months.

[b]Significant difference at P<.001 for comparing with the web-based video group.

[c]Significant difference at P<.01 for comparing with the web-based video group.

[d]Significant difference at P<.05 for comparing with the web-based video group. There was no difference between the in-person and videoconference groups at all time points.

Session Attendance

There was no significant difference (P>.05) in session attendance between the videoconference and in-person groups (Table 6). Overall, in both groups (excluding those who transferred to the web-based video option), the mean number of sessions attended was 16.5 (SD 6.5), with 86.9% (107/123) attending >25% sessions, 83.7% (103/123) attending >50% sessions, and 73.2% (90/123) attending >75% of the 21 sessions (1 orientation, 12 weekly sessions, and 8 monthly phone calls). Of the 12 weekly sessions, the participants attended an average of 10.1 (SD 3.8) sessions. Of the 8 monthly phone calls, they completed an average of 5.4 (SD 3.1) calls.

Table 6. Session attendance among participants in the web-based and in-person groups[a] (N=123).

Group	Number of total sessions attended (including make-up sessions, out of 21 sessions), mean (SD)	*P* value	Number of weekly core sessions attended (including makeup sessions, out of 12 sessions), mean (SD)	*P* value	Number of monthly contacts received (including makeup sessions, out of 8 calls), mean (SD)	*P* value
Overall (in person and videoconference)	16.5 (6.5)	N/A[b]	10.1 (3.8)	N/A	5.4 (3.1)	N/A
Videoconference	17.9 (4.7)	.07	10.9 (2.6)	.07	5.9 (2.6)	.13
In person	15.7 (7.3)	.07	9.7 (4.3)	.07	5.1 (3.3)	.13

[a]A total of 15 participants who were transferred to the web-based video option after not attending in-person or videoconference sessions were excluded.
[b]N/A: not applicable.

Discussion

Principal Findings

This study found that, when provided with a choice on how to engage in a weight loss intervention, 28% (56/200) of men chose videoconference groups, 31% (62/200) chose web-based videos, and 41% (82/200) chose in-person groups. There were significant differences in demographic, employment, cultural, and technology use and access factors among men who chose 1 of the 3 different options. For example, men who chose web-based videoconference groups were more likely to be younger, have higher income and education, have a white-collar type of job, prefer English, be more acculturated, have higher health literacy, have access to a computer, and have higher technology skills (eg, computer, email, and video chat app use) compared with men who chose in-person groups. Our multivariate analysis distinguished most significantly between those who chose in-person groups versus those who chose web-based videos. Men who were older, spoke Spanish, and did not use a computer frequently had a higher probability of choosing in-person groups versus web-based videos. In terms of weight loss, men who chose the videoconference and in-person group sessions lost more weight than those who chose web-based videos.

Comparison With Prior Work

Similar to this study, Piatt et al [33] tested 3 GLB-based lifestyle intervention delivery options (in-person groups, DVD, and internet education) among 555 adults (95% were White) and compared them with 1 arm where participants could choose the option (the preference arm). Their in-person groups and DVD options were similar to those offered in this study. In the preference arm, 60% chose the in-person option, none chose the DVD option, and 40% chose the internet education option. In our study, 41% (82/200) chose the in-person group, 31% (62/200) chose the web-based videos group, and 28% (56/200) chose the videoconference group. This difference may be attributed to the differences in sample demographics between the 2 studies—mean age of 52.3 (SD 12.7) years in the study by Piatt et al [33] versus 47.3 (SD 11.8) years in our study, 97% non-Latino White versus 100% (200/200) Latino sample, and 97% women versus 100% (200/200) men. Piatt et al [33] did not report on weight loss for each option but found that participants who were given a choice among the 3 options lost more weight at 6 months (−6.4 kg) than those who were randomly assigned to one of the options without a choice (−5.7 kg for in-person groups, −5.5 kg for DVD, and −6.2 kg for internet education; *P*<.001). In addition, Piatt et al [33] reported that the average session attendance was 8.4 for the preference arm participants who chose in-person groups and 9.2 for the preference arm participants who chose internet education. Similar to the study by Piatt et al [33], group session attendance in our study did not differ according to whether men initially chose the videoconference (mean 14.9, SD 7.4 sessions) or in-person (mean 15.0, SD 7.7 sessions) group. This could be because of the effectiveness of the protocol for supporting men in making a choice. The protocol included an explanation of the options from a health coach, included family members, and included peer discussion of the options. It is also possible that men could make an informed choice with less support, such as from a provider or staff member in the primary care setting. Future research could examine the most efficient and effective strategy for supporting men in making an informed choice about technology-mediated interventions.

Implications

These findings have important implications for translation to practice, especially given the COVID-19 pandemic, which has made technology-mediated options necessary in the short term and possibly more common in the long term. The fact that 59% (118/200) of Latino men in this study preferred a technology-mediated option (ie, videoconference or web-based videos) suggests that these can be acceptable delivery formats for chronic disease prevention programs for this important high-risk population. This is consistent with previous literature indicating that Latinos are comfortable with technology and open to technology-mediated interventions [34-36]. However, given that the web-based videos were less effective for weight loss than either of the group-based options, future research should focus on increasing the effectiveness of web-based video options. In addition, including the option of in-person interventions appears to be important for reaching other subgroups of Latino men, such as those with lower health literacy and limited access to technology. Recognizing the diversity of Latino men and understanding the specific sociodemographics is critical to designing effective and engaging interventions to reach this population. As in the HOMBRE intervention, offering a suite of choices may be important for reaching the diversity of Latino men.

These findings demonstrate that Latino men from lower socioeconomic backgrounds may need additional support for accessing and using technology-mediated interventions. Technology use and access differ along socioeconomic lines in the United States, including among Latinos, with lower socioeconomic-level subgroups having less reliable access to technology than their higher socioeconomic-level counterparts [37,38]. Additional support could include resources such as the provision of computers, tablets, or smartphones and reliable internet access for delivering technology-mediated interventions. Resources could also include capacity building in the specific technologies used in the intervention. Future research can test strategies to increase access to and familiarity with technology-mediated interventions for Latino men from lower socioeconomic backgrounds. Furthermore, the clinics that serve Latino men who may need additional support to access technology-mediated interventions tend to also be underresourced. Thus, policies that provide additional funding and support to enable these clinics to implement technology-mediated interventions will be needed. This is especially important during the COVID-19 pandemic or other circumstances when technology-mediated interventions are needed.

The findings from the HOMBRE trial may be specific to the type of technology-mediated interventions offered. The 2 technology-mediated options included smartphone apps for tracking diet and physical activity and either videoconferencing or web-based videos for the intervention sessions. Preferences may vary according to the type of technology used in the intervention. Other technologies such as SMS text messaging, social media, and digital voice assistants may be appealing to other groups of Latino men. For example, Latino migrant farm workers have high (81%-97%) access to mobile phones and prefer talking and SMS text messaging using their phones [36,39]. Therefore, behavioral interventions using mobile phones and SMS text messages may be more acceptable to this population than those using computers or laptops. Latinos are also highly engaged in social media (73% of adult internet users use Facebook, 34% use Instagram, and 25% use Twitter) [40], which could be leveraged for behavioral lifestyle interventions.

Limitations

This study has several limitations. First, all the technology use data were self-reported. Participants may have reported answers regarding their prior and current use of technology that they may have perceived as desirable [41]. Second, this study was conducted among a diverse sample of Latino men in the San Francisco Bay Area of California. The results of this investigation may not be generalizable to Latino men residing in other parts of the United States. Nevertheless, the findings appear to be consistent with previous findings on Latinos. Replication and assessment of generalizability in independent samples in other parts of the United States are needed to generalize the findings to a broader population. Third, we were not able to capture the frequency of videos viewed by the group that chose web-based videos. This would have enabled us to compare engagement in the intervention with the other choices. Fourth, we did not collect information on family member involvement, precluding an analysis of differences in family member involvement across delivery options. Fifth, caution should be used when interpreting the differences in weight loss among the intervention options, given that men were not randomized into these options. It is possible that the men who chose the videoconference and in-person groups were more motivated to lose weight than those who chose the web-based videos. Related, we were not able to determine the extent to which different levels of coach and peer support contributed to weight loss differences among the intervention options. The videoconference and in-person groups had the highest levels of coach and peer support. Future studies could examine the effectiveness of different delivery options and what aspects of these delivery options (eg, level of coach and peer support [42]) contribute to the varied effectiveness.

Conclusions

In conclusion, this study revealed that among a diverse group of Latino men recruited from primary care, 28% (56/200) chose videoconference groups, 31% (62/200) chose web-based videos, and 41% (82/200) chose in-person groups. Differences in demographic, employment, cultural, and technology use factors distinguished between men who chose each of the options, suggesting that when offering interventions in diverse groups of Latino men, choice of delivery may be recommended. We also found that men attending either of the group-based options (videoconference and in-person) lost more weight than men who chose the web-based videos.

Acknowledgments

The Patient Centered Outcomes Research Institute funded the study (AD 15-1503-29190), and the funder had no role in the review or approval of the manuscript for publication.

Authors' Contributions

All authors conceptualized and designed the study. NL and LX conducted the statistical analysis with input from LGR and JM. All authors interpreted the data. LGR drafted the manuscript. All authors critically revised the manuscript for important intellectual content. LGR, NL, LX, and JM obtained funding. LGR had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Conflicts of Interest

None declared.

References

1. Hales CM, Fryar CD, Carroll MD, Freedman DS, Ogden CL. Trends in obesity and severe obesity prevalence in US youth and adults by sex and age, 2007-2008 to 2015-2016. JAMA 2018 Apr 24;319(16):1723-1725 [FREE Full text] [doi: 10.1001/jama.2018.3060] [Medline: 29570750]
2. Chatterjee A, Gerdes MW, Martinez SG. Identification of risk factors associated with obesity and overweight-a machine learning overview. Sensors (Basel) 2020 May 11;20(9):2734 [FREE Full text] [doi: 10.3390/s20092734] [Medline: 32403349]
3. Ogden CL, Carroll MD, Fryar CD, Flegal KM. Prevalence of obesity among adults and youth: United States, 2011-2014. NCHS Data Brief 2015 Nov(219):1-8 [FREE Full text] [Medline: 26633046]
4. US Preventive Services Task Force, Curry SJ, Krist AH, Owens DK, Barry MJ, Caughey AB, et al. Behavioral weight loss interventions to prevent obesity-related morbidity and mortality in adults: US preventive services task force recommendation statement. JAMA 2018 Sep 18;320(11):1163-1171. [doi: 10.1001/jama.2018.13022] [Medline: 30326502]
5. Pagoto SL, Schneider KL, Oleski JL, Luciani JM, Bodenlos JS, Whited MC. Male inclusion in randomized controlled trials of lifestyle weight loss interventions. Obesity (Silver Spring) 2012 Jun;20(6):1234-1239 [FREE Full text] [doi: 10.1038/oby.2011.140] [Medline: 21633403]
6. Closing the digital divide: Latinos and technology adoption. Pew Research Center. 2013. URL: http://assets.pewresearch.org/wp-content/uploads/sites/7/2013/03/Latinos_Social_Media_and_Mobile_Tech_03-2013_final.pdf [accessed 2022-01-20]
7. Digital divide narrows for Latinos as more Spanish speakers and immigrants go online. Pew Research Center. 2016. URL: https://www.pewresearch.org/hispanic/2016/07/20/digital-divide-narrows-for-latinos-as-more-spanish-speakers-and-immigrants-go-online/ [accessed 2022-01-20]
8. Digital divide is shrinking for America's Hispanic population, NTIA data show. National Telecommunications and Information Administration. URL: https://www.ntia.doc.gov/blog/2019/digital-divide-shrinking-america-s-hispanic-population-ntia-data-show [accessed 2022-01-20]
9. Victorson D, Banas J, Smith J, Languido L, Shen E, Gutierrez S, et al. eSalud: designing and implementing culturally competent ehealth research with latino patient populations. Am J Public Health 2014 Dec;104(12):2259-2265. [doi: 10.2105/AJPH.2014.302187] [Medline: 25320901]
10. Valdez LA, Amezquita A, Hooker SP, Garcia DO. Mexican-origin male perspectives of diet-related behaviors associated with weight management. Int J Obes (Lond) 2017 Dec;41(12):1824-1830. [doi: 10.1038/ijo.2017.173] [Medline: 28757642]
11. Rosas LG, Lv N, Azar KM, Xiao L, Hooker SP, Lewis MA, et al. HOMBRE: a randomized controlled trial to compare two approaches to weight loss for overweight and obese Latino men (Hombres con Opciones para Mejorar el Bienestar y bajar el Riesgo de Enfermedades crónicas; men with choices to improve well-being and decrease chronic disease risk). Contemp Clin Trials 2018 May;68:23-34. [doi: 10.1016/j.cct.2018.02.019] [Medline: 29505867]
12. Kramer MK, Kriska AM, Venditti EM, Miller RG, Brooks MM, Burke LE, et al. Translating the Diabetes Prevention Program: a comprehensive model for prevention training and program delivery. Am J Prev Med 2009 Dec;37(6):505-511. [doi: 10.1016/j.amepre.2009.07.020] [Medline: 19944916]
13. Kramer MK, Kriska AM, Venditti EM, Semler LN, Miller RG, McDonald T, et al. A novel approach to diabetes prevention: evaluation of the Group Lifestyle Balance program delivered via DVD. Diabetes Res Clin Pract 2010 Dec;90(3):e60-e63. [doi: 10.1016/j.diabres.2010.08.013] [Medline: 20863586]
14. Diabetes prevention program group lifestyle balance™ materials. Diabetes Prevention Support Center. URL: https://www.diabetesprevention.pitt.edu/index.php/group-lifestyle-balance-materials/ [accessed 2022-01-20]
15. Prochaska JO, DiClemente CC. Stages and processes of self-change of smoking: Toward an integrative model of change. Journal of Consulting and Clinical Psychology 1983;51(3):390-395. [doi: 10.1037/0022-006x.51.3.390]
16. Rosas LG, Lv N, Lewis MA, Venditti EM, Zavella P, Luna V, et al. A Latino patient-centered, evidence-based approach to diabetes prevention. J Am Board Fam Med 2018;31(3):364-374 [FREE Full text] [doi: 10.3122/jabfm.2018.03.170280] [Medline: 29743220]
17. What is PhenX Toolkit? PhenX Toolkit. URL: https://www.phenxtoolkit.org/ [accessed 2022-01-20]
18. Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults: The Evidence Report. Rockville, MD: National Institutes of Health, National Heart, Lung, and Blood Institute; 1998.
19. Pickering TG, Hall JE, Appel LJ, Falkner BE, Graves J, Hill MN, Subcommittee of ProfessionalPublic Education of the American Heart Association Council on High Blood Pressure Research. Recommendations for blood pressure measurement in humans and experimental animals: Part 1: blood pressure measurement in humans: a statement for professionals from the Subcommittee of Professional and Public Education of the American Heart Association Council on High Blood Pressure Research. Hypertension 2005 Jan;45(1):142-161. [doi: 10.1161/01.HYP.0000150859.47929.8e] [Medline: 15611362]
20. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary care evaluation of mental disorders. Patient health questionnaire. JAMA 1999 Nov 10;282(18):1737-1744. [doi: 10.1001/jama.282.18.1737] [Medline: 10568646]

21. Diez-Quevedo C, Rangil T, Sanchez-Planell L, Kroenke K, Spitzer RL. Validation and utility of the patient health questionnaire in diagnosing mental disorders in 1003 general hospital Spanish inpatients. Psychosom Med 2001;63(4):679-686. [doi: 10.1097/00006842-200107000-00021] [Medline: 11485122]
22. EQ-5D User Guide. EQ-5D. URL: https://euroqol.org/publications/user-guides/ [accessed 2022-01-20]
23. Sobocki P, Ekman M, Agren H, Krakau I, Runeson B, Mårtensson B, et al. Health-related quality of life measured with EQ-5D in patients treated for depression in primary care. Value Health 2007;10(2):153-160 [FREE Full text] [doi: 10.1111/j.1524-4733.2006.00162.x] [Medline: 17391424]
24. Karlsson J, Taft C, Sjöström L, Torgerson JS, Sullivan M. Psychosocial functioning in the obese before and after weight reduction: construct validity and responsiveness of the Obesity-related Problems scale. Int J Obes Relat Metab Disord 2003 May;27(5):617-630. [doi: 10.1038/sj.ijo.0802272] [Medline: 12704406]
25. Bilbao A, Mar J, Mar B, Arrospide A, Martínez de Aragón G, Quintana JM. Validation of the Spanish translation of the questionnaire for the obesity-related problems scale. Obes Surg 2009 Oct;19(10):1393-1400. [doi: 10.1007/s11695-009-9800-3] [Medline: 19172366]
26. Yu L, Buysse DJ, Germain A, Moul DE, Stover A, Dodds NE, et al. Development of short forms from the PROMIS™ sleep disturbance and Sleep-Related Impairment item banks. Behav Sleep Med 2011 Dec 28;10(1):6-24. [doi: 10.1080/15402002.2012.636266] [Medline: 22250775]
27. Marin G, Gamba RJ. A new measurement of acculturation for Hispanics: the bidimensional acculturation scale for hispanics (BAS). Hispanic J Behav Sci 2016 Jul 25;18(3):297-316. [doi: 10.1177/07399863960183002]
28. The Newest Vital Sign homepage. The Newest Vital Sign. URL: https://pfe-pfizercom-prod.s3.amazonaws.com/health/nvs_flipbook_spanish_final.pdf [accessed 2022-01-20]
29. Technology adoption. Pew Research Center. URL: https://www.pewresearch.org/topics/technology-adoption/ [accessed 2022-01-20]
30. Chowdhury MZ, Turin TC. Variable selection strategies and its importance in clinical prediction modelling. Fam Med Community Health 2020;8(1):e000262 [FREE Full text] [doi: 10.1136/fmch-2019-000262] [Medline: 32148735]
31. Fisher R. The Design of Experiments. Edinburgh and London: Oliver & Boyd; 1935.
32. Hosmer D, Lemeshow S, Sturdivant R. Applied Logistic Regression. New York: John Wiley & Sons, Incorporated; 2013.
33. Piatt GA, Seidel MC, Powell RO, Zgibor JC. Comparative effectiveness of lifestyle intervention efforts in the community: results of the Rethinking Eating and ACTivity (REACT) study. Diabetes Care 2013 Feb;36(2):202-209 [FREE Full text] [doi: 10.2337/dc12-0824] [Medline: 22966092]
34. Villegas N, Cianelli R, de Tantillo L, Warheit M, Montano NP, Ferrer L, et al. Assessment of technology use and technology preferences for HIV prevention among Hispanic women. Hisp Health Care Int 2018 Dec;16(4):197-203 [FREE Full text] [doi: 10.1177/1540415318808828] [Medline: 31185747]
35. Chaet AV, Morshedi B, Wells KJ, Barnes LE, Valdez R. Spanish-language consumer health information technology interventions: a systematic review. J Med Internet Res 2016 Aug 10;18(8):e214 [FREE Full text] [doi: 10.2196/jmir.5794] [Medline: 27511437]
36. Price M, Williamson D, McCandless R, Mueller M, Gregoski M, Brunner-Jackson B, et al. Hispanic migrant farm workers' attitudes toward mobile phone-based telehealth for management of chronic health conditions. J Med Internet Res 2013 Mar 30;15(4):e76 [FREE Full text] [doi: 10.2196/jmir.2500] [Medline: 23624105]
37. Digital divide persists even as lower-income Americans make gains in tech adoption. Pew Research Center. 2021. URL: https://www.pewresearch.org/fact-tank/2019/05/07/digital-divide-persists-even-as-lower-income-americans-make-gains-in-tech-adoption/ [accessed 2022-01-20]
38. Weiss D, Rydland HT, Øversveen E, Jensen MR, Solhaug S, Krokstad S. Innovative technologies and social inequalities in health: A scoping review of the literature. PLoS One 2018;13(4):e0195447 [FREE Full text] [doi: 10.1371/journal.pone.0195447] [Medline: 29614114]
39. Sandberg JC, Spears Johnson CR, Nguyen HT, Talton JW, Quandt SA, Chen H, et al. Mobile and traditional modes of communication among male Latino farmworkers: implications for health communication and dissemination. J Immigr Minor Health 2016 Jun;18(3):522-531 [FREE Full text] [doi: 10.1007/s10903-015-0299-5] [Medline: 26463228]
40. Social media preferences vary by race and ethnicity. Pew Research Center. 2015. URL: https://www.pewresearch.org/fact-tank/2015/02/03/social-media-preferences-vary-by-race-and-ethnicity/ [accessed 2022-01-20]
41. Van de Mortel TF. Faking it: social desirability response bias in self-report research. Australian J Advanced Nurs 2008;25(4):40-48.
42. Chatterjee A, Gerdes M, Prinz A, Martinez S. Human coaching methodologies for automatic electronic coaching (eCoaching) as behavioral interventions with information and communication technology: systematic review. J Med Internet Res 2021 Mar 24;23(3):e23533 [FREE Full text] [doi: 10.2196/23533] [Medline: 33759793]

Abbreviations

GLB: Group Lifestyle Balance
HOMBRE: Hombres con Opciones para Mejorar su Bienestar para Reducir Enfermedades Crónicas

Edited by A Mavragani; submitted 11.04.21; peer-reviewed by D Griffith, R Krukowski, D West, A Chatterjee, H Belli; comments to author 04.06.21; revised version received 29.07.21; accepted 10.12.21; published 04.02.22.

Please cite as:
Rosas LG, Lv N, Xiao L, Azar KMJ, Hooker SP, Venditti EM, Lewis MA, Zavella P, Ma J
Preferences for Technology-Mediated Behavioral Lifestyle Interventions With Different Levels of Coach and Peer Support Among Latino Men: Comparative Study Within One Arm of a Randomized Controlled Trial
JMIR Form Res 2022;6(2):e29537
URL: https://formative.jmir.org/2022/2/e29537
doi:10.2196/29537
PMID:35119377

Original Paper

Improving Emergency Department Patient-Physician Conversation Through an Artificial Intelligence Symptom-Taking Tool: Mixed Methods Pilot Observational Study

Justus Scheder-Bieschin[1], MD; Bibiana Blümke[2], MD; Erwin de Buijzer[2], MD; Fabienne Cotte[2], MD; Fabian Echterdiek[3], MD; Júlia Nacsa[2], MFA; Marta Ondresik[2], MA; Matthias Ott[1], MD; Gregor Paul[4], MD; Tobias Schilling[1], MD; Anne Schmitt[2], BSc, MSc; Paul Wicks[2], PhD; Stephen Gilbert[2,5], PhD

[1]Department of Interdisciplinary Acute, Emergency and Intensive Care Medicine (DIANI), Klinikum Stuttgart, Stuttgart, Germany

[2]Ada Health, Berlin, Germany

[3]Department of Nephrology, Klinikum Stuttgart, Stuttgart, Germany

[4]Department of Infectious Diseases, Klinikum Stuttgart, Stuttgart, Germany

[5]The Else Kröner Fresenius Center for Digital Health, University Hospital Carl Gustav Carus Dresden, Technische Universität Dresden, Dresden, Germany

Corresponding Author:
Stephen Gilbert, PhD
The Else Kröner Fresenius Center for Digital Health
University Hospital Carl Gustav Carus Dresden
Technische Universität Dresden
Postfach 151 Fetscherstraße 74
Dresden, 01307
Germany
Phone: 49 17680396015
Email: stephen.gilbert@mailbox.tu-dresden.de

Abstract

Background: Establishing rapport and empathy between patients and their health care provider is important but challenging in the context of a busy and crowded emergency department (ED).

Objective: We explore the hypotheses that rapport building, documentation, and time efficiency might be improved in the ED by providing patients a digital tool that uses Bayesian reasoning–based techniques to gather relevant symptoms and history for handover to clinicians.

Methods: A 2-phase pilot evaluation was carried out in the ED of a German tertiary referral and major trauma hospital that treats an average of 120 patients daily. Phase 1 observations guided iterative improvement of the digital tool, which was then further evaluated in phase 2. All patients who were willing and able to provide consent were invited to participate, excluding those with severe injury or illness requiring immediate treatment, with traumatic injury, incapable of completing a health assessment, and aged <18 years. Over an 18-day period with 1699 patients presenting to the ED, 815 (47.96%) were eligible based on triage level. With available recruitment staff, 135 were approached, of whom 81 (60%) were included in the study. In a mixed methods evaluation, patients entered information into the tool, accessed by clinicians through a dashboard. All users completed evaluation Likert-scale questionnaires rating the tool's performance. The feasibility of a larger trial was evaluated through rates of recruitment and questionnaire completion.

Results: Respondents strongly endorsed the tool for facilitating conversation (61/81, 75% of patients, 57/78, 73% of physician ratings, and 10/10, 100% of nurse ratings). Most nurses judged the tool as potentially time saving, whereas most physicians only agreed for a subset of medical specialties (eg, surgery). Patients reported high usability and understood the tool's questions. The tool was recommended by most patients (63/81, 78%), in 53% (41/77) of physician ratings, and in 76% (61/80) of nurse ratings. Questionnaire completion rates were 100% (81/81) by patients and 96% (78/81 enrolled patients) by physicians.

Conclusions: This pilot confirmed that a larger study in the setting would be feasible. The tool has clear potential to improve patient–health care provider interaction and could also contribute to ED efficiency savings. Future research and development will extend the range of patients for whom the history-taking tool has clinical utility.

Trial Registration: German Clinical Trials Register DRKS00024115; https://drks.de/drks_web/navigate.do?navigationId=trial.HTML&TRIAL_ID=DRKS00024115

(JMIR Form Res 2022;6(2):e28199) doi:10.2196/28199

KEYWORDS

symptom assessment application; anamnesis; health care system; patient history taking; diagnosis; emergency department

Introduction

Background

The emergency department (ED) is, by definition, a high-stress environment. As it is so critical that the time of a health care provider (HCP) is used optimally, effective and empathetic communication with patients (and colleagues) can be challenging [1]. There is increasing recognition that hospital EDs face numerous challenges related to crowding, a problem likely to continue for the foreseeable future [2]. Obvious barriers to effective communication include time pressures caused by a full waiting room and urgent cases exceeding capacity [1]. More subtle and pervasive systemic factors include limitations of processes and interpersonal parameters such as societal and health disparities [3].

It has been proposed that appropriately designed artificial intelligence (AI)–based systems could reduce ED documentation errors, improve patient safety [4-8], and free up HCP time. Such time savings could potentially be used to improve efficiency and provide HCPs with more time to build rapport with patients [5].

One such AI-based digital symptom assessment tool is Ada (Ada Health GmbH), which uses a Bayesian probabilistic reasoning engine with an adaptive question flow to collect demographic information, medical history, and symptoms. Previous studies have demonstrated that the Ada tool was helpful and easy to use for patients in a primary care waiting room [9] and has the potential to reduce waiting times and increase efficiency at urgent care centers [10]. Its underlying technology has the highest condition-suggestion accuracy among tools of its class and has the same quality of safe urgency advice as general practitioners [11]. When integrated at a large US health care system, its urgency-advice accuracy recommendations were stated by independent researchers to be comparable to those of nurse-staffed telephone triage lines [12]. A number of studies [9,13,14] have described a range of tools for self-assessment of urgency and triage in the general practice and acute primary care setting, but studies evaluating the handover of history and symptom information to HCPs have not been carried out with real patients. If appropriately adapted to the ED setting, the benefits of a digital history-taking tool could assist nurse-led triage in the waiting room and assessment and treatment by ED physicians. This study used methodologies adopted in other evaluations of new digital technologies for patient self-reported history in the ED setting [15,16].

The Aim of This Study

In this study, we aim to evaluate a prototype digital history-taking and handover system, which includes a patient-facing tool for symptom and history taking. After patients inputted their data in the patient-facing tool, a clinical handover report was displayed to the physician or nurse using an HCP-facing tool. Our primary hypothesis was that such a tool might alleviate some of the challenges in building rapport and communication in a crowded ED. The secondary hypotheses were that the tool could alleviate documentation and time pressures. This pilot study, which was based on the approach reported in some studies [16-18], assessed in 2 study stages a prototype patient-facing system to assist communication. First, phase 1 involved initial implementation of the patient-facing and HCP-facing tools (version V1), followed by their evaluation by all users. Feedback on performance from phase 1 was then used to create a modified system (version V2), which was further evaluated in phase 2. Patients, physicians, and nurses quantitatively evaluated the 2 system versions in terms of their usability and usefulness in facilitating patient-HCP conversation and rapport formation in the ED setting. For HCP users, we also explored the helpfulness of the medical information provided at handover and its perceived potential for the system to save HCP time.

Methods

Design: Overall Study Approach and Study Type

This study used incremental mixed methods approaches and was designed such that phase 1 allowed initial learning about the tool [16-18] and co-design between researchers and HCPs. Modifications suggested by users that could be implemented within the project time frame were then used to develop a V2 prototype, and at a switchover point, the V1 tool was replaced by the V2 tool, which then remained unchanged during phase 2. There were no study design changes made between phase 1 and phase 2. The study was conducted between August 3, 2020, and August 21, 2020.

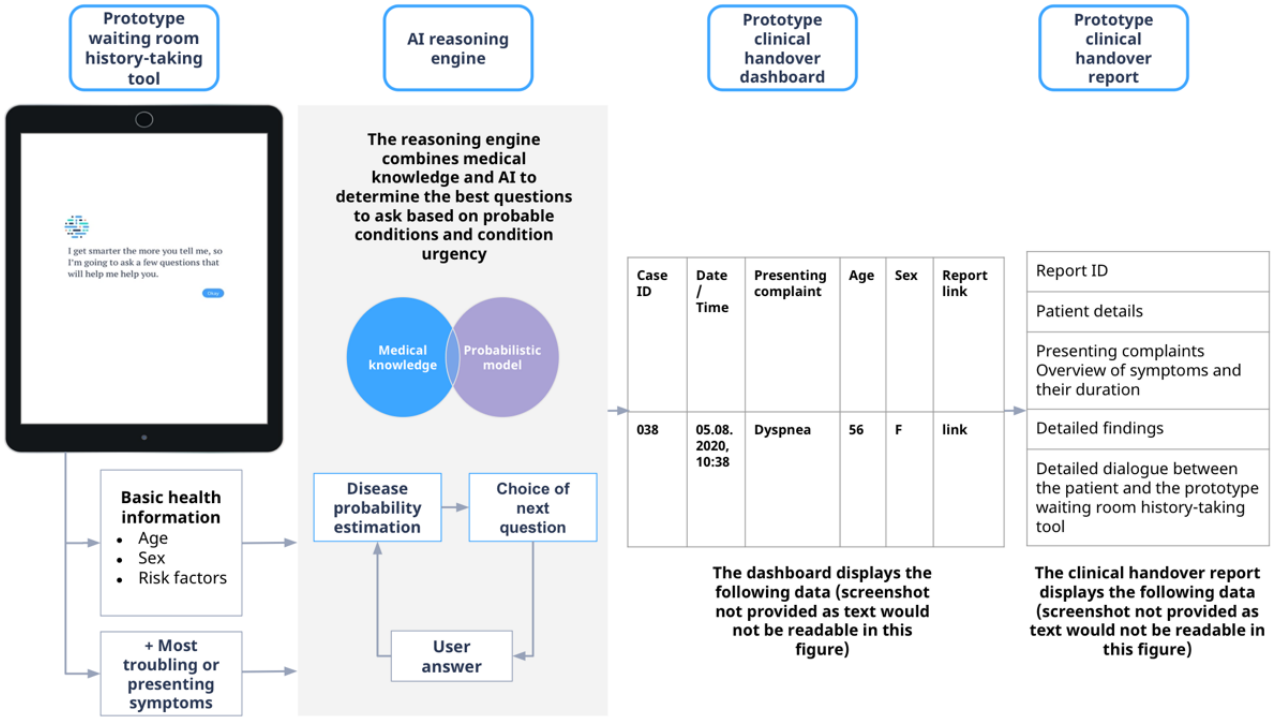
This study explored the potential for enhancing bidirectional patient-HCP communication in the ED through the introduction of an augmentative prototype history and symptom-taking tool. Questionnaire-based user perceptions on the potential for the system were collected alongside qualitative observations. The mixed methods approach, along with the full implementation and use of the tool for patients, provided real-world data in a manner often only achievable through interventional study designs. However, the overall study approach was observational because the system was not closely integrated into HCP clinical workflows during the study and did not replace standard practice in the ED. HCPs were carefully trained not to rely on the prototype system for formal or definitive symptom-taking support during the study. In this way, no interventional patient outcome measures were recorded. This study was conducted in accordance with requirements of the Standards for Quality Improvement Reporting Excellence [19], SPIRIT-AI (Standard

Protocol Items: Recommendations for Interventional Trials–Artificial Intelligence), and CONSORT-AI (Consolidated Standards of Reporting Trials–Artificial Intelligence) [20] guidelines.

Description of the Prototype Digital Symptom and History-Taking System

The system consisted of patient- and HCP-facing tools (Figure 1). The patient-facing tool was provided to recruited patients in the waiting room on a tablet computer (iPad, Apple Inc). The patient history–taking tool used a cloud-based Bayesian probabilistic reasoning engine combined with curated medical knowledge to ask the patient the optimal set of questions based on probability and urgency of conditions. The Ada reasoning engine versions used in the study were version 1.31.1 and version 1.31.2. The assessment was performed as an interaction-enabled question flow with options to confirm, deny, or skip each question. The tool asked successively tailored questions about the respondent's medical history and the main presenting complaints, as well as related attributes of their symptoms, such as severity and time course.

Figure 1. The prototype digital history and symptom-taking and handover system evaluated in this study showing the interactions between the patient-facing and health care provider–facing tools and describing how the artificial intelligence (AI) reasoning engine functions to sequentially ask the patient the most relevant question next. Although the screenshots presented are in English, the tool used in this study was in German.

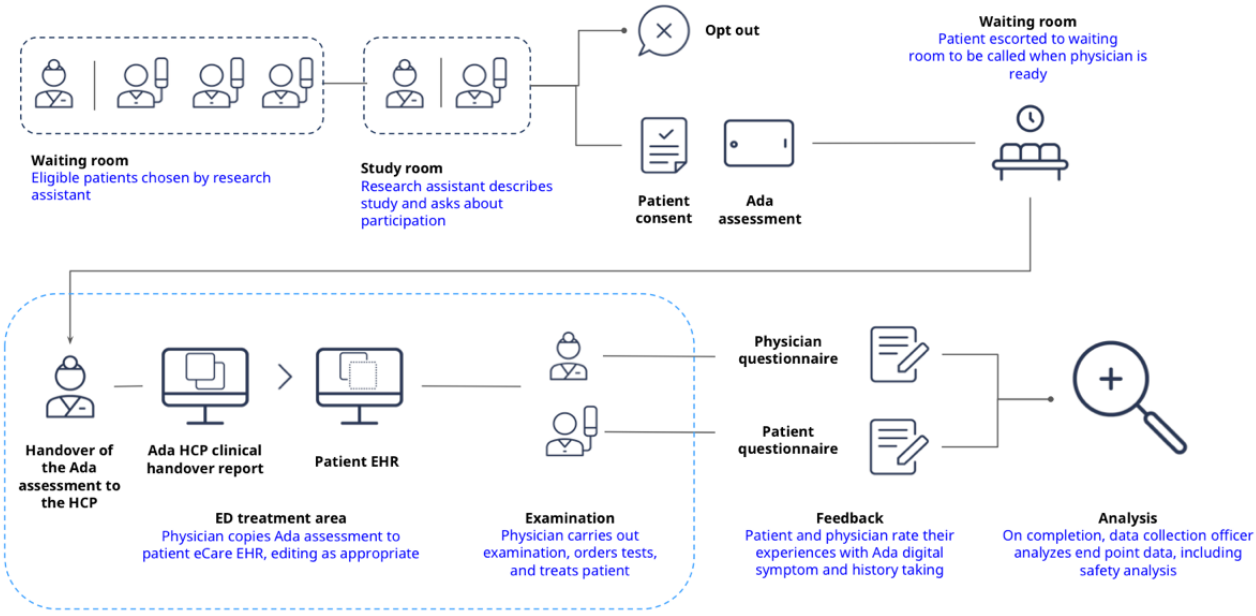


The HCP-facing tool provided a secure web-interface dashboard that listed all completed assessments to the ED clinical staff. The tool also provided a detailed handover report, designed to provide clinical information quickly and safely to HCPs. The handover report included the patient's basic information (sex and year of birth); basic medical history information (smoker, hypertension, diabetes, and pregnancy status); main presenting symptoms; details of these symptoms, including the specific questions asked by the tool; and answers provided by the patient.

Design: Study Procedure

The procedure of the study is described in Figure 2.

Figure 2. Overview of the study procedure. ED: emergency department; EHR: electronic health record; HCP: health care provider.



Recruitment, Inclusion, and Exclusion Criteria

The following inclusion criteria were applied: (1) patients aged >18 years, (2) attended the Stuttgart study site when recruitment was being undertaken, and (3) willing and able to provide consent. The exclusion criteria were as follows: (1) patients with severe injury or illness requiring immediate treatment, that is, Emergency Severity Index (ESI) levels 1-2 (patients classified as ESI level 3 were included if assessed as suitable) [21]; (2) patients with traumatic injury; (3) patients incapable of completing a health assessment, for example, because of illiteracy, mental impairment, or inebriation or other incapacity; and (4) patients aged <18 years. The clinician research assistant asked all potentially eligible patients in the waiting room if they would be interested in participating in a study evaluating a tool to record and pass on their symptoms to the ED physicians. Potentially eligible patients were informed that the study would not delay (or accelerate) their treatment at the ED. A single clinician research assistant carried out enrollment and consent.

Informed Consent and Study Data Management

If the patient agreed to be considered for inclusion, they were led to a separate private room adjacent to the ED where the nature, background, and scope of the study were explained and they were asked if they wanted to participate. If the patient consented, the pseudoanonymization procedure was followed. The patient's name was recorded alongside the next-in-sequence study patient enrollment number on the study enrollment (disambiguation) record, which is the only link between the study ID (study patient enrollment number) and patient name, which was kept securely by the principal investigator (TS). All quantitative and qualitative data were only accessible by the study team on secure systems. The study questionnaire data were recorded and stored at the study site in secure databases. Paper surveys were stored in secure hospital clinical trials file storage. The patient's enrollment number was entered into the patient-facing tool on the iPad, and the tool then asked the patient to agree to the terms and conditions and privacy policy. The study team was familiar with data privacy regulations and is committed to data protection principles. The study was approved by the local ethics committee at the University of Heidelberg (S-052-2020) and is registered in the German Clinical Trials Register (DRKS00024115).

Procedure for Patient Symptom Taking

The clinician research assistant answered any questions the patient had about the tool and helped them to use it if requested, recording the degree of help provided. The symptom-taking handover report was not provided to the patient and was automatically made available to the ED physicians in the ED treatment area through the HCP-facing tool.

Training and Study Procedure for HCPs in the ED

All the ED nurses and physicians were made aware of the study in advance through a presentation of the study and a written manual describing the system. Physicians were made aware of all patients who were enrolled in the study. The HCP logged onto the secure web interface using a secure ID to access the handover report.

Study Measurements

After examination by the ED physician, the patient and the ED physician completed separate paper-based questionnaires, with evaluation ratings of the tool on modified Likert scales. Nurses completed the same questionnaire as physicians when possible; however, it was recognized at the time of trial design that there would not always be a nurse–patient interaction after triage in which the handover report is relevant. All questionnaires were designed on the premise of optimizing completion (and therefore full participation) through highly simplified design. A validated usability tool design, for example, the System Usability Scale [22], was not used as the System Usability Scale addresses usability only, whereas it was important for the questionnaire in this study to address not only usability, but also usefulness of the tool. The Mobile App Rating Scale [23] is a validated tool for addressing usability and usefulness, but it would not

have been feasible for HCPs to complete this multidimensional instrument in real time without disruption to their clinical routine; nor would it have been feasible for all the HCPs to complete a training exercise in the Mobile App Rating Scale before its use, as recommended in the study by Stoyanov et al [23]. Considering these factors, there is currently no published validated instrument for addressing usability and usefulness for a tool specifically designed for increasing patient–HCP rapport around self-reported medical history and symptom taking. The primary hypothesis, that the current tool had the potential to facilitate bidirectional patient-clinician conversation and rapport building, was assessed through the patient questionnaire items (originally in German): (1*)* "Was the digital history-taking experience engaging?" (2) "Could you understand the questions asked by the tool?" (3) "Do you think that the tool could facilitate better treatment at the ED?" and (4) "Did you feel better understood when speaking to the physician because they were already aware of your medical problem?" Corresponding questions for nurses and physicians assessed the HCP perspectives of rapport, as described in the *Results* section. The secondary hypothesis, that the tool would have the potential to facilitate documentation and thereby save HCP time, was assessed through the following HCP questionnaire items: (1) "Would the tool provide medically helpful information?" and (2) "Would the tool (as currently implemented) save time?"

Qualitative insights were also collected in the study, with a focus on usability and user interface improvements. To collect data on physicians' interaction with the HCP handover, we applied methods of contextual inquiry [24]: observations, contextual interviews, and cognitive walkthroughs. This combination of methods was chosen for the primary purpose of enabling iterative co-design and development of products with users. The action-oriented study design allowed the rapid implementation of improvements based on phase 1 insights, resulting in a context-optimized patient tool and HCP interface in phase 2. The qualitative insights were collected by a multidisciplinary team consisting of a user experience researcher (AS), an interface designer in the product development team (JN), and the clinician research physician (JSB). Qualitative data were collected using multiple modalities (observational data, contextual interviews, and cognitive walkthroughs) to triangulate insights and reduce bias caused by any rapport built between participants and researchers throughout the study. Qualitative data were collected on a convenience sampling basis in 3 days of contextual interviews and cognitive walkthroughs at the start of phase 1 (conducted by AS, JN, and JSB) and at the start of phase 2 (conducted by JSB). Multimedia Appendix 1 contains the *User Research Guide*, which provides a detailed description of the qualitative methods applied.

Study Setting

The study was conducted in the ED of the Katharinenhospital, Klinikum Stuttgart, which is an adult tertiary referral and major trauma hospital in southwestern Germany. It provides interdisciplinary emergency treatment for between 100 and 120 patients per day. The center adopts the First View Concept [25], in which an emergency registrar or consultant sees each patient in an interdisciplinary approach. The center has 23 treatment rooms with central monitoring, a resuscitation room, a wound room, and a plaster room. It uses the internationally recognized ESI triage system to guide the treatment of ED patients according to medical urgency [21].

Sample Size Determination

This study was designed for real-world tool optimization (in phase 1), followed by a preliminary observation assessment of the tool's potential in the ED. It was also designed as a guide to a later larger trial and in line with literature on pilot study design [26,27]. Therefore, the sample size was estimated on the basis of having a sufficient number of patients to assess survey completion rate and to determine if there were any safety-related considerations that might be needed in a randomized controlled trial (RCT). The aspects that were piloted were as follows: (1) trialing of new procedures and enabling power calculations intended to be used in a later single- or multicenter RCT; (2) establishing how many patients and HCPs can be recruited and the feasible level of completed patient, physician, and nurse questionnaires; and (3) evaluating the general technical and logistical feasibility of a full-scale study, including questionnaire design and other data collection related issues.

Data Analysis

The quantitative data were analyzed using standard Python (version 3.7.4; Python Software Foundation) statistical modules (SciPy module version 1.3.0) using descriptive statistics and the Mann-Whitney U test, a nonparametric test of statistical significance suitable for categorical data [28]. For statistical significance testing, the value of a (here a=0.05) was adjusted using the Bonferroni correction, a/*m*, where *m* is the number of questions evaluated for each group. For patients, *m*=6 and $a_{corrected}$=0.0083; for HCPs, *m*=4 and $a_{corrected}$=0.0125; and for the degree of patient self-sufficiency, *m*=1 and $a_{corrected}$=0.05. The pairwise deletion method was used for handling missing data [29].

The qualitative data were analyzed using affinity diagrams, a clustering technique used in thematic analysis [30]. The categories of the affinity diagram were predefined as key areas of interest, outlined in the interview guide: usability, usefulness, comprehension, impact on patient-physician communication, clinical relevance of information, and fit with clinical workflows.

Results

Recruitment

The total number of recruited patients was 81 (41, 51%, women and 40, 49%, men), with 45 (56%) patients in phase 1 and 36 (44%) in phase 2 (Figure 3). Of the 81 recruited patients, there were 3 (4%) classified as ESI level 3, 77 (95%) classified as ESI level 4, and 1 (1%) classified as ESI level 5. A detailed description of the patient population is provided in Table 1, and the full results for each patient are presented in Table S1 of Multimedia Appendix 2.

Figure 3. Participant recruitment flowchart. ED: emergency department; ESI: Emergency Severity Index.

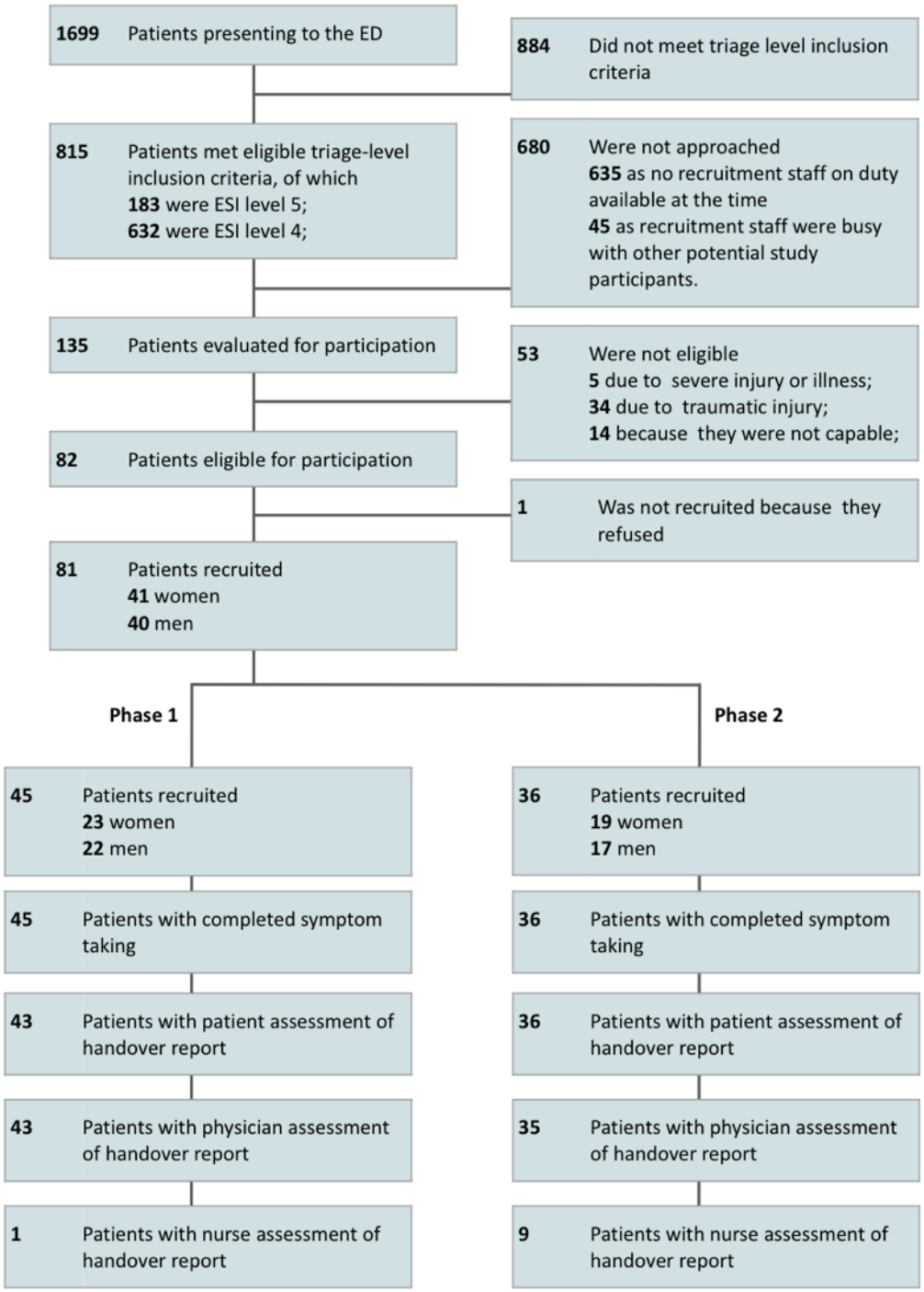


Table 1. Description of the data completeness and medical subdiscipline of the final main diagnosis stratified according to study phase (N=81).

Group by study phase	All patients (N=81), n (%)	Patients included in phase 1 (n=45), n (%)	Patients included in phase 2 (n=36), n (%)
Age (years), mean (SD)	38.7 (15)	36 (15.9)	42.1 (13.1)
Completed patient evaluation questionnaire	81 (100)	45 (100)	36 (100)
Completed physician evaluation questionnaire	78 (96)	43 (96)	35 (97)
Completed optional nurse evaluation questionnaire	10 (12)	1 (2)	9 (25)
Medical specialty classification (on the basis of emergency department discharge diagnosis)			
No diagnosis assigned	1 (1)	1 (2)	0 (0)
Orthopedics	12 (15)	9 (20)	3 (8)
Dermatology	5 (6)	5 (11)	0 (0)
Internal medicine, including specialties	28 (35)	13 (11)	15 (42)
Internal medicine: cardiovascular disease	4 (5)	1 (2)	3 (8)
Internal medicine: gastroenterology	4 (5)	1 (2)	3 (8)
Internal medicine: nephrology	1 (1)	1 (2)	0 (0)
Internal medicine: oncology	1 (1)	0 (0)	1 (3)
Internal medicine: rheumatology	1 (1)	1 (2)	0 (0)
Internal medicine, with no subspecialty	17 (21)	9 (20)	8 (22)
Neurology	20 (25)	8 (18)	12 (33)
Obstetrics and gynecology	1 (1)	1 (2)	0 (0)
Psychiatry	1 (1)	1 (2)	0 (0)
Surgery	10 (12)	6 (13)	4 (11)
Ear, nose, and throat	3 (4)	1 (2)	2 (6)

Baseline Data

Patient populations in the 2 study phases were similar (Figure 3 and Table 1): the mean age was 38.7 (SD 15.0) years for all patients, 36.0 (SD 15.9) years for patients included in phase 1, and 42.1 (SD 13.1) years for patients included in phase 2. The dominant medical classifications (by ED discharge diagnosis) were orthopedics, internal medicine, neurology, and surgery.

Data Exclusion and Missing Data

For the 81 patients enrolled in the study, questionnaires were completed by 100% (81/81) of the patients (of whom 73/81, 90%, completed all questions), by 96% (78/81) of the physicians (78 completed questionnaires by the physicians for the 81 enrolled patients) and 12% (10/81) of the nurses (10 completed questionnaires by the nurses for the 81 enrolled patients; Figure 3 and Tables 1 and 2). Nurse questionnaires were only completed when they took part in symptom and history taking and when ED workload allowed. For all survey questions, the analysis approach was to report all data with respect to the number of responses to that survey question (ie, to use the pairwise deletion method for handling missing data [29], with the denominator in analyses being lower where there were missing data).

Table 2. Summary of patient, physician, and nurse ratings of the tool for phase 1, phase 2, and the phases combined. Two modified Likert scales were used: a 4-level Likert scale (1=Strongly disagree, 2=Disagree, 3=Agree, and 4=Strongly agree) and a 10-level Likert scale (1=Unlikely to 10=Highly likely). The mean and percentage positive ratings were calculated on the basis of the provided answers for each question. See Table S1 in Multimedia Appendix 2 for detailed data.

Statistic	All patients (N=81)	Patients included in phase 1 (n=45)	Patients included in phase 2 (n=36)
Patient-provided ratings			
1. Was the digital history-taking experience engaging?			
Patients, n (%)	81 (100)	45 (100)	36 (100)
Mean (SD), out of 4	3.4 (0.7)	3.2 (0.7)[a]	3.7 (0.5)[a]
Positive rating proportion[b], n (%)	73 (90)	38 (84[a])	35 (97[a])
2. Could you understand the questions asked by the tool?			
Patients, n (%)	81 (100)	45 (100)	36 (100)
Mean (SD), out of 4	3.4 (0.8)	3.2 (0.9)	3.6 (0.6)
Positive rating proportion[b], n (%)	70 (86)	35 (78)	35 (97)
3. Do you think that the tool could facilitate better treatment in the ED?			
Patients, n (%)	75 (93)	42 (93)	33 (92)
Mean (SD), out of 4	2.9 (1.0)	2.7 (0.9)	3.1 (1.0)
Positive rating proportion[b], n (%)	51[c] (68)	27[d] (64)	24[e] (73)
4. Did you feel better understood when speaking to the physician because they were already aware of your medical problem?			
Patients, n (%)	73 (90)	40 (89)	33 (92)
Mean (SD), out of 4	3.0 (0.9)	2.9 (0.9)	3.1 (1.0)
Positive rating proportion[b], n (%)	55[f] (75)	28[g] (70)	27[e] (82)
5. How do you rate the user experience provided to you in the tool (ie, its usability)?			
Patients, n (%)	79 (98)	43 (96)	36 (100)
Mean (SD), out of 10	7.3 (2.3)	7.4 (2.2)	7.2 (2.5)
Positive rating proportion[b], n (%)	66[h] (84)	36[i] (84)	30 (83)
6. Would you recommend the tool to others?			
Patients, n (%)	77 (95)	41 (91)	36 (100)
Mean (SD), out of 10	7.3 (2.3)	7.1 (2.3)	7.5 (2.4)
Positive rating proportion[b], n (%)	60[j] (78)	31[k] (76)	29 (81)
Physician-provided ratings			
1. Would the tool facilitate rapport with the patient?			
Patients, n (%)	78 (96)	43 (96)	35 (97)
Mean (SD), out of 3	2.8 (0.8)	2.8 (0.9)	2.9 (0.7)
Positive rating proportion[b], n (%)	57[l] (73)	29[i] (67)	28[m] (80)
2. Would the tool provide medically helpful information?			
Patients, n (%)	78 (96)	43 (96)	35 (97)
Mean (SD), out of 4	2.6 (0.9)	2.6 (0.9)	2.7 (0.8)
Positive rating proportion[b], n (%)	43[l] (55)	21[i] (49)	22[m] (63)
3. Would the tool (as currently implemented) save time?			
Patients, n (%)	78 (96)	43 (96)	35 (97)
Mean (SD), out of 4	2.2 (0.9)	2.3 (1.0)	2.2 (0.8)
Positive rating proportion[b], n (%)	27[l] (35)	16[i] (37)	11[m] (31)

Statistic	All patients (N=81)	Patients included in phase 1 (n=45)	Patients included in phase 2 (n=36)
4. Would you recommend the tool to colleagues?			
Patients, n (%)	77 (95)	43 (96)	34 (94)
Mean (SD), out of 10	5.7 (1.9)	5.4 (2.2)	6.1 (1.4)
Positive rating proportion[b], n (%)	41[j] (53)	20[i] (47)	21[n] (62)
Nurse-provided ratings			
1. Would the tool facilitate rapport with the patient?			
Patients, n (%)	10 (12)	1 (2)	9 (25)
Mean (SD), out of 4	3.7 (0.5)	4.0[o]	3.7 (0.5)
Positive rating proportion[b], n (%)	10[p] (100)	1[q] (100)	9[r] (100)
2. Would the tool provide medically helpful information?			
Patients, n (%)	10 (12)	1 (2)	9 (25)
Mean (SD), out of 4	3.4 (0.5)	3.0[o]	3.4 (0.5)
Positive rating proportion[b], n (%)	10[o] (100)	1[p] (100)	9[q] (100)
3. Would the tool (as currently implemented) save time?			
Patients, n (%)	10 (12)	1 (2)	9 (25)
Mean (SD), out of 4	3.2 (0.8)	3.0[o]	3.2 (0.8)
Positive rating proportion[b], n (%)	8[p] (80)	1[q] (100)	7[r] (78)
4. Would you recommend the tool to colleagues?			
Patients, n (%)	10 (12)	1 (2)	9 (25)
Mean (SD), out of 8	7.8 (0.8)	8.0[o]	7.8 (0.8)
Positive rating proportion[b], n (%)	10[p] (100)	1[q] (100)	9[r] (100)
Patient self-sufficiency			
Degree of patient self-sufficiency (on the 4-level scale of assistance: 1=high, 2=medium, 3=low, and 4=none).			
Patients, n (%)	80 (99)	44 (98)	36 (100)
Mean (SD), out of 4	3.1 (1.0)	2.8 (1.1)	3.4 (0.7)

Statistic	All patients (N=81)	Patients included in phase 1 (n=45)	Patients included in phase 2 (n=36)
Proportion who received little or no help, n (%)	61[s] (76)	28[t] (64)	33 (92)

[a]Statistically significant difference in Likert scores, according to the Mann–Whitney U test.

[b]Percentage of positive ratings on the 4-level Likert scale, that is, the percentage of *3* and *4* ratings.

[c]n=75.

[d]n=42.

[e]n=33.

[f]n=73.

[g]n=40.

[h]n=79.

[i]n=43.

[j]n=77.

[k]n=41.

[l]n=78.

[m]n=35.

[n]n=34.

[o]SD not defined as the group size is 1.

[p]n=10.

[q]n=1.

[r]n=9.

[s]n=80.

[t]n=44.

Qualitative Learnings

The qualitative learnings are summarized in Table 3.

Table 3. Summary of qualitative insights.

Study phase, theme, and users	Observation
Phase 1	
Usability	
Patients	• Operating the digital tool and finishing the question flow was successful for most patients
Usefulness	
Physicians	• Provision of additional patient information (past medical history, medications, and allergies) was considered very important to the overall utility of the tool
Physicians	• Treating physicians expected the tool to collect patient information beyond medical history and acute complaints, for example, "collecting a social anamnesis for a full picture of patient background"
Physicians	• The ability to read a patient's history and symptoms before consultation was generally described as making the interaction with the patient more pleasant for both patient and physician (rapport) and making treating physicians more prepared

Physicians	• The tool primarily provided clinical value for newly occurring problems. It provided less added value for the following:
	  • patients with severe or visible trauma
	  • patients with complaints resulting from chronic conditions
	  • patients with multiple comorbidities

Physicians	• The highly mixed patient population (age, language, digital literacy, medical complaints, and socioeconomic status) makes the emergency department a challenging setting for the tool
Physicians and nurses	• Greater integration of the tool with the electronic health record systems and clinical workflows is desirable
Comprehension	

Patients	• The language used in the tool was difficult for several observed users to understand because of the following:
	  • limited German language ability (nonnative German speakers)
	  • limited language reading level (some native German speakers)

Patients	• Several observed users had a low ability to articulate complaints in a manner that the tool could process (when using either medical or layperson's terms). This related to the following:
	  • difficulty in localizing pain to provide the tool precise answers to localization questions
	  • difficulty entering multiple symptoms separately
	  • general difficulty in verbalizing symptoms
	  • difficulty finding accurate synonyms in the tool for the feelings they were experiencing

Phase 2 (only new or changed insights recorded)	
Usefulness	
Physicians and nurses	• The V2 tool asked patients to provide information on past medical history, medications, and allergies, thereby providing a fuller picture of the patient's health beyond their acute symptoms and was recognized by physicians and nurses as more beneficial and better supporting the physician–patient consultation
Nurses	• Nurses were enthusiastic regarding time-saving potential in history taking with V2 tool features

Summary of Changes Made Between the V1 and V2 Systems

We made three changes to the V1 system to create the V2 system based on qualitative observations and quantitative findings:

1. The addition of 3 free-text input fields, where the patient can supply initial information on their medical history, current medications, and allergies: This change was requested by ED physicians so that patients would have the opportunity to pass on information in their own words that was not always collected by the tool's question flow. This was implemented in a manner that did not change the core AI-based symptom assessment. This change affected the

information entered into the patient-facing tool and the information presented on the HCP-facing tool.

2. Improvement of the user interface at the transition between patient information and symptom assessment: A minor rephrasing was made to the patient-facing tool in response to a small number of patients who reported that they misunderstood a specific direction of how to proceed from one step to the next in the early phase of the assessment (a clear direction originally developed in an earlier English language prototype had been poorly expressed in German).
3. Fixing a minor bug affecting the HCP handover report: A minor bug was removed that had resulted in a small number of handover reports not accurately displaying the transcript of questions that Ada asked the patient alongside the patient response.

The changes were anticipated to have a minor impact on the evaluation of the tool. The patients' ratings of the tool (Table 2) improved by 9.5% and those of the physicians by 9.1% across all ratings with a significant improvement (P=.003; $a_{corrected}$=0.0083) in the responses to the patient question "Was the digital history-taking experience engaging?" The changes in nurse ratings of the system between phase 1 and phase 2 are reported in Table 2 and Table S1 of Multimedia Appendix 2; however, there were insufficient nurse assessments in phase 1 to allow statistical significance testing.

Patient, Physician, and Nurse Ratings of the Tool

Patients were positive or highly positive (73/81, 90%) about how engaging the tool was to use, its comprehension (70/81, 86%), its usability (68/81, 84%), its ability to facilitate understanding and rapport with the HCPs (61/81, 75%), and about recommending the tool to peers (68/81, 84%; Table 2 and Table S1 of Multimedia Appendix 2). Likewise, 73% (57/78) of the physician ratings were positive or highly positive about the tool's potential for facilitating understanding and rapport with patients.

Nurses were even more positive or highly positive about the tool's potential for facilitating understanding and rapport with patients (10/10, 100% of nurse ratings), about the helpfulness of the medical information provided (10/10, 100% of nurse ratings), about the tool's time-saving potential (8/10, 80% of nurse ratings), and about recommending the tool to peers (10/10, 100% of nurse ratings).

Subanalyses by Medical Specialty

A post hoc subanalysis by medical specialty of ED discharge diagnosis is shown in Table S1 of Multimedia Appendix 2 and described in detail in Multimedia Appendix 2. This subanalysis showed that there was lower comprehension of the tool by patients receiving care from a neurologist and that these patients also gave lower ratings to the conversation facilitation provided by the tool. Physicians rated the level of helpful information, the time-saving utility of the report, and the likelihood of them recommending the system more positively for internal medicine handovers than for handovers for all recruited patients or for other specialties.

Degree of Patient Assistance Provided

Across both phases, most patients were able to use the tool with little or no assistance (62/81, 76%). This measure improved significantly between phase 1 (52/81, 64%) and phase 2 (74/81, 92%; P=.003; $a_{corrected}$=0.05).

Variability of Physicians' Perceptions

Patient symptom and history data handover were evaluated for 96% (78/81) of the patients by 24 different ED physicians. Qualitative interviews with physicians in phase 1 of the study revealed a number of physicians who were exceptionally enthusiastic about the performance of the system, ie, they exhibited an *early adopter* mentality. The quantitative analysis of the distribution of Likert scores (Figure S1 in Multimedia Appendix 2) supported these qualitative findings, in that there is a skewed distribution of physician scores, with 2 physicians (one of whom evaluated handover reports for 6 patients) providing highly positive evaluations of the tool.

Sample Representativeness

The mean age of patients was 38.7 (SD 15.0) years, which is comparable to the mean age of 41.8 (SD 19.3) years reported in a recent cross-sectional study of patients attending German EDs [31]. The study showed that those who refused to participate in the general ED study, which did not involve assessment of a digital health tool, were older (difference in mean ages 3.6 years) than those who agreed to participate. Our study had a moderately lower participant age group than that reported in the study by Scherer et al [31] (difference in mean ages 3.1 years) and this might reflect, to a minor degree, older patients being less willing to try a digital tool. However, 24% (19/81), 7% (6/81), and 4% (3/81) of the patients in our study were aged >50 years, >60 years, and >70 years, respectively. Overall, the patient population closely reflects the proportion of the German ED patient population that agrees to participate in studies and is reflective of the overall self-referred walk-in German ED patient population [31,32].

Larger Study Planning

A 100% (81/81) response rate was achieved for the patient questionnaire completion and a very high response rate also for the physician questionnaire (78/81, 96%). This was made possible through diligent tracking and follow-up of the patients and physicians; a planned and systematic approach was taken because high questionnaire completion was seen as a criterion for judgment of the feasibility of an RCT, after tool optimization, for regulatory approval. The nurse questionnaire completion rate was much lower (10/81, 12%); it had been understood in planning that (1) not every patient has an interaction with nurses that would be meaningful to assess, (2) the complexity and variability of the timings of the nurse interactions with the patient provided no single time point at which the nurse questionnaire could be completed, and (3) the nurses' workload pressures would mean that completion would be an additional burden. For these reasons, completion was optional for the nurses in the study design. We accept that firm conclusions on nurse perception cannot be made on the basis of completed nurse questionnaire for 12% (10/81) of the patients; however, we report these data for completeness.

Discussion

Principal Findings

The performance of the system was evaluated with respect to two hypotheses: the primary hypothesis was that the tool had the potential to facilitate bidirectional patient-clinician conversation and rapport building and the secondary hypothesis was that the tool would have the potential to facilitate documentation and thereby save HCP time.

There was a strongly positive rating by patients, physicians, and nurses of the tool as an aid in patient–clinician conversation, communication, and rapport building. The proportion of physicians who were positive or highly positive about recommending the system to their peers (in 41/77, 53%, of physician ratings) was not as high as it was for patients and nurses (63/81, 78%, of patient ratings and 10/10, 100% of nurse ratings). Overall, physicians had mixed views on the degree of helpful clinical information and time-saving potential of the clinical handover report. Underlying the mixed results in the physician ratings (Table S1 of Multimedia Appendix 2) were a large number of patients for whom the ED physicians described the presence of a *visual diagnosis*, that is, one that was immediately apparent by simply looking at a patient and for whom patient-directed symptom taking, in whatever form it is designed, is unlikely to save history-taking and recording time. For some other patients, the tool did not adequately draw together and summarize a highly complex, sometimes multimorbid medical presentation in a manner that made it useful for physicians. Both qualitative findings and quantitative analysis revealed a subset of highly enthusiastic early-adopter physicians, as has been recognized in other studies [33], who were highly positive about the history and symptom-taking information provided by the tool and for its potential to improve rapport, to provide helpful clinical information, and to save ED time.

The relatively lower comprehension of the tool by patients receiving care from a neurologist and the lower rating of its conversation facilitation may reflect effects of neurological symptoms on their use of the tool. This could be related to inherent challenges of achieving an acceptable user interface for these patients for self-history and symptom recording, or it may be that the tool requires further specific optimization for this patient group—further quantitative and qualitative study of this group in a larger study is required before definitive conclusions can be drawn. Our interpretation is that the positive or highly positive evaluations from physicians for internal medicine handovers is likely due to the system's AI reasoning engine being better at directing the question flow for conditions that have many subtle interlinked clinical symptoms.

An aim for the development of a waiting room patient history–taking and HCP handover tool is that it should be easy to use without assistance. Many patients visit only once; therefore, they would have little opportunity to learn to use the tool over time. In this study, a degree of patient assistance was provided (where needed) by the recruiting clinical research physician (JSB), and in all cases the degree of assistance provided was documented. Most (62/81, 76%) of the patients required little or no help with the use of the tool, and in some subspeciality, patients were highly independent using the tool (76/81, 94%, in internal medicine with no subspeciality), whereas 67% (54/81) of patients receiving care from an orthopedist required little or no help. A priority in the future development of the prototype will be to adapt the reasoning engine and the user interface to minimize the level of help required for all patient groups, and this may involve making the tool available in the patient's primary language. It is known that the degree of patient eHealth literacy can be an enabler or a barrier for patients, especially for older adults [34-36]. Internationally, particularly in low- and middle-income countries, eHealth literacy of HCPs, including medical and paramedical staff, may be a barrier to widespread adoption [37,38]. We propose that the potential limitation to adoption as a result of low eHealth literacy, for both patients and HCPs, is best resolved through user-focused design and development, followed by user-focused mixed methods pilot evaluation (as in this study). After tool refinement, multicenter RCTs and real-world–performance modeling are appropriate and necessary.

The study incorporated 2 study phases, providing an opportunity to rapidly modify a prototype tool's design based on feedback from patient and physician evaluations and to further evaluate the modified tool. A comparison of tool ratings showed statistically significant improvements in performance between the V1 and V2 tools in two evaluation categories: (1) *how engaging or interesting patients found the tool to use* and (2) *the degree of self-sufficiency of the patients in using the tool*. These changed scores reflect improvements made in tool usability and functionality. Many useful insights on tool performance and usability were obtained in the study. Only a relatively small number of optimizations could be executed between the V1 and V2 tools because of the need to prioritize only those in-study changes that could be completed within the possible study recruitment period. The remaining insights, including those regarding the V2 tool, will be used in later prototype development and optimization. The qualitative findings broadly reflect the quantitative findings of the study, particularly with respect to the groups of patients who benefit most from the tool. Importantly, the qualitative findings also provided a means in the study to iterate on the tool design and to gain deeper insights into further improvements that can be made in tool design, particularly with the interface and language level.

Open questions for future research include the potential for handover of home-completed digital symptom and history taking to HCPs for rapport and communication building and time saving and the potential of the tool in a general nonemergency setting. Other open questions relate to the precise contribution that the evaluated tool or technologies related to it could make at times of substantial impact of COVID-19 on health care provision. This study was conducted in a period of low COVID-19 case numbers. It was recognized early in the COVID-19 pandemic that eHealth tools were needed to reduce the likelihood of health care facilities being overwhelmed and to assist in providing health care without face-to-face contact [39,40]. Although eHealth systems contributed substantially

throughout the pandemic, they did not provide a panacea for care delivery [41].

Limitations

This was a single-site study in Germany, and the findings may not be generalizable to other facilities or to other countries. The patient and physician quantitative evaluations were not influenced by selection bias; however, because the nurse evaluations were completed on a voluntary basis, the influence of selection bias cannot be excluded (9/10, 90% of the nurse evaluations related to study phase 2). The study questionnaire was not based on a previously validated instrument; therefore, response bias cannot be excluded. The pairwise deletion approach was used for handling missing data [29], which results in the analysis being based on different sets of data with different sample sizes. The approach of modeling the missing data and estimating missing values is an alternative and is considered the optimal approach by Kang et al [29]. The qualitative methods were applied to a smaller subgroup of users (patients and HCPs) in the first days after tool introduction and cannot be taken to be fully representative of the full study population. This study explored tool use in the German language and in the German setting only, and the sample size was relatively small, given the resource constraints within the ED. The study did not include patients receiving care from a pediatrician. Infection control with patients using a tablet computer in the ED could be challenging for large-scale deployment. This study explored the early phase after prototype system implementation, training, and first experience of its application by users. Although minimally relevant for ED patients, the skill and speed of ED physician and nurse use of the new tool and their perception of its performance are likely to change over time. Nurse questionnaires were completed for 12% (10/81) of the patients; therefore, firm conclusions cannot be drawn about their perceptions of the tool. Despite the limitations, our approach allowed us to investigate implementation at this single site in considerable depth.

Comparison With Prior Work

There are no previous studies describing similar Bayesian digital history-taking and handover systems in the literature. Arora et al [42] explored patient impressions and satisfaction of a self-administered, automated medical history–taking device in the ED and also reported high levels of patient enthusiasm and potential for building rapport. However, handover to HCPs was not carried out in the study; nor did the tool have a system to enable handover. A protocol has been published for a prospective cohort study of a self-reported computerized medical history–taking tool (Clinical Expert Operating System [CLEOS]) [43]. The CLEOS tool has also been evaluated in an observational study that looked at history taking using the tool subsequent to clinician history taking [44,45]. The focus of these studies was on history taking for the sake of record completeness and less so on the potential for building rapport. The CLEOS tool is fundamentally different in its interface with patients and in the underlying logic that drives patient questioning. The CLEOS tool uses a decision tree approach, whereas the tool evaluated in this study (from Ada) uses a Bayesian network that is defined upon a medical knowledge base and on which approximate inference is carried out, followed by information theoretical methods used to decide which questions to ask the user. There is no simple answer as to whether decision trees or Bayesian network approaches are superior for patient symptom and history taking. One argument in support of the Ada approach is that it aims to ask specific questions based on a large array of possible conditions that a patient may have to gather a highly personalized medical history. This would require an unmanageable size of decision trees, leading to inconsistencies, imbalances, and low accuracy [46].

Implications for Clinicians and Policy Makers and Implications for Future Research

It is recognized that empathy and rapport can be lacking in EDs, both of which are important for staff stress levels and satisfaction and for patient outcomes and empowerment [1,6,7,47,48]. Although a digital tool can never be a complete solution to improve human-to-human empathy, this study has nonetheless demonstrated patient and HCP enthusiasm for the rapport building and conversation facilitation enabled through such tools. This study only evaluated a prototype, and further developments are required to facilitate patient–HCP conversations and optimize rapport in a manner that is complete (ie, clinically helpful information for every patient), streamlined, and seamless (ie, no additional tasks or systems for clinicians). This mixed methods pilot study adopted a user-driven approach, and the results show that the tool has large potential for rapport and communication building. This study has provided a foundation for the further development of the tool, which will be followed by an RCT of a completed tool.

Although the results of this study were definitive regarding the potential for rapport and conversation facilitation in the ED, they were equivocal on the potential for the tool, as it is currently implemented, to save clinician time. Clinicians could only identify opportunities for saving time in selected types of patient presentations, namely internal medicine and surgery. Nurses were more positive about the general potential of the tool for clinician time saving. It is recognized that further development of the prototype is required if the aim of the completed approved regulatory product is to deliver definitive physician time savings for all patients. The role of a patient history and symptom-taking tool is not only to save clinician time, but to also contribute to documentation accuracy and completeness. It is known that medical performance reduces with stress and overstretching in the ED and is likely to result in more errors, including in documentation [8,31]. It was not possible to measure the contribution of the tool to documentation completeness and accuracy in this study.

Conclusions

The patient- and HCP-reported data from this mixed methods pilot study supported the primary hypothesis that the tool improved rapport between patient and clinician and improved patient-clinician communication. Mixed methods trials are powerful approaches to gain insights into tool potential and to optimize a tool for a particular clinical setting. However, confirmatory evaluation is needed in RCTs and in different ED settings. Notably, patients felt better understood and the tool had utility in efficiently recording their symptoms. Nurses

perceived the tool as having the potential to save time through workflow efficiency. Some physicians were enthusiastic about the potential to improve patient interaction and about the tool's benefits in symptom and history taking. Results regarding the secondary hypothesis of documentation assistance and time saving in the ED were more equivocal, but there was potential for time saving in some medical subspecialties, for example, internal medicine and surgery. Insights from this study will be used for further prototyping and research to extend the range of patients for whom the tool can provide support. The tool is based upon an existing and regularly maintained medical reasoning engine and therefore is a sustainable technological approach for history and symptom taking. The tool is readily adaptable to other related settings in which patient self-symptom and history taking and conversation support are relevant, including at home and at primary care and specialist clinics (eg, specialist rare diseases clinics).

Acknowledgments

This study was funded by Ada Health GmbH. The grant award number is not applicable. This was an investigator-initiated study, and as such the principal investigator was TS, assisted by JSB, and the sponsor was the Katharinenhospital, Klinikum Stuttgart (represented by TS). The initial concept for the study originated entirely with the clinicians in the Stuttgart emergency department (principally JSB). The funder (Ada Health GmbH) made nonbinding suggestions on the design of the study but had no role in the conduct of the study, in data collection, or in data management. Overall responsibility for data analysis, data interpretation, manuscript preparation, review, approval of the manuscript, and decision to submit the manuscript for publication was with Klinikum Stuttgart (JSB). The funder (Ada Health GmbH) assisted with data analysis and manuscript preparation but every overall decision regarding, and responsibility for, the collection of data, integrity of the data, and accuracy of the data analysis was undertaken by JSB.

Authors' Contributions

JSB had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. The concept and design were the responsibility of JSB, BB, FC, EdB, M Ondresik, TS, and SG. The coordination of the study in the emergency department was carried out by JSB and TS. Acquisition, analysis, and interpretation of data were the responsibility of JSB, BB, FE, JN, M Ott, GP, TS, AS, PW, and SG. The manuscript was drafted by JSB, FC, JN, M Ondresik, AS, PW, and SG. Statistical analysis was conducted by JSB, TS, and SG. Administrative, technical, and material support were provided by JN and M Ondresik. The study was supervised by JSB, TS, and SG.

Conflicts of Interest

BB, FC, JN, M Ondresik, AS, and SG are employees of Ada Health GmbH, and some of them hold stock options in the company. PW has a consultancy contract with Ada Health GmbH and is an employee of, and owns shares in, Wicks Digital Health. Wicks Digital Health has received funding from Ada Health, AstraZeneca, Baillie Gifford, Bold Health, Camoni, Compass Pathways, Coronna, European Institute of Innovation and Technology, Happify, HealthUnlocked, Inbeeo, Kheiron Medical, Sano Genetics, Self Care Catalysts, The Learning Corp, The Wellcome Trust, VeraSci, and Woebot. PW has received speaker fees from Bayer and honoraria from Roche, ARISLA, AMIA, IMI, PSI, and the BMJ. The Ada Health GmbH research team has received research grant funding from Fondation Botnar and the Bill & Melinda Gates Foundation.

Multimedia Appendix 1
User research guide.
[PDF File (Adobe PDF File), 240 KB - formative_v6i2e28199_app1.pdf]

Multimedia Appendix 2
Results supplement with a subanalyses by medical specialism and the Likert score distribution for the emergency department physician.
[DOCX File , 251 KB - formative_v6i2e28199_app2.docx]

References

1. Pun JK, Matthiessen CM, Murray KA, Slade D. Factors affecting communication in emergency departments: doctors and nurses' perceptions of communication in a trilingual ED in Hong Kong. Int J Emerg Med 2015 Dec;8(1):48 [FREE Full text] [doi: 10.1186/s12245-015-0095-y] [Medline: 26667242]
2. Morley C, Unwin M, Peterson GM, Stankovich J, Kinsman L. Emergency department crowding: a systematic review of causes, consequences and solutions. PLoS One 2018 Aug 30;13(8):e0203316 [FREE Full text] [doi: 10.1371/journal.pone.0203316] [Medline: 30161242]
3. Olsson M, Hansagi H. Repeated use of the emergency department: qualitative study of the patient's perspective. Emerg Med J 2001 Nov;18(6):430-434. [doi: 10.1136/emj.18.6.430] [Medline: 11696488]

4. Edwards M, Moczygemba J. Reducing medical errors through better documentation. Health Care Manag (Frederick) 2004;23(4):329-333. [doi: 10.1097/00126450-200410000-00007] [Medline: 15638340]
5. Lorenzetti DL, Quan H, Lucyk K, Cunningham C, Hennessy D, Jiang J, et al. Strategies for improving physician documentation in the emergency department: a systematic review. BMC Emerg Med 2018 Oct 25;18(1):36 [FREE Full text] [doi: 10.1186/s12873-018-0188-z] [Medline: 30558573]
6. Communicating in Hospital Emergency Departments. Berlin, Heidelberg: Springer; 2015.
7. Slade D, Scheeres H, Manidis M, Iedema R, Dunston R, Stein-Parbury J, et al. Emergency communication: the discursive challenges facing emergency clinicians and patients in hospital emergency departments. Discourse Commun 2008 Aug 01;2(3):271-298. [doi: 10.1177/1750481308091910]
8. Zavala AM, Day GE, Plummer D, Bamford-Wade A. Decision-making under pressure: medical errors in uncertain and dynamic environments. Aust Health Rev 2018 Aug;42(4):395-402. [doi: 10.1071/AH16088] [Medline: 28578757]
9. Miller S, Gilbert S, Virani V, Wicks P. Patients' utilization and perception of an artificial intelligence-based symptom assessment and advice technology in a British primary care waiting room: exploratory pilot study. JMIR Hum Factors 2020 Jul 10;7(3):e19713. [doi: 10.2196/19713] [Medline: 32540836]
10. Montazeri M, Multmeier J, Novorol C, Upadhyay S, Wicks P, Gilbert S. Can symptom and history recording patient apps help optimize patient flow in urgent care centers?: a simulation study. JMIR Form Res 2020 Dec;5(5) [FREE Full text] [doi: 10.2196/26402]
11. Gilbert S, Mehl A, Baluch A, Cawley C, Challiner J, Fraser H, et al. How accurate are digital symptom assessment apps for suggesting conditions and urgency advice? A clinical vignettes comparison to GPs. BMJ Open 2020 Dec 16;10(12):e040269 [FREE Full text] [doi: 10.1136/bmjopen-2020-040269] [Medline: 33328258]
12. Morse KE, Ostberg NP, Jones VG, Chan AS. Use characteristics and triage acuity of a digital symptom checker in a large integrated health system: population-based descriptive study. J Med Internet Res 2020 Nov 30;22(11):e20549 [FREE Full text] [doi: 10.2196/20549] [Medline: 33170799]
13. Chambers D, Cantrell AJ, Johnson M, Preston L, Baxter SK, Booth A, et al. Digital and online symptom checkers and health assessment/triage services for urgent health problems: systematic review. BMJ Open 2019 Aug 01;9(8):e027743 [FREE Full text] [doi: 10.1136/bmjopen-2018-027743] [Medline: 31375610]
14. Verzantvoort NC, Teunis T, Verheij TJ, van der Velden AW. Self-triage for acute primary care via a smartphone application: practical, safe and efficient? PLoS One 2018 Jun 26;13(6):e0199284 [FREE Full text] [doi: 10.1371/journal.pone.0199284] [Medline: 29944708]
15. Chisolm DJ, Purnell TS, Cohen DM, McAlearney AS. Clinician perceptions of an electronic medical record during the first year of implementaton in emergency services. Pediatr Emerg Care 2010 Feb;26(2):107-110 [FREE Full text] [doi: 10.1097/PEC.0b013e3181ce2f99] [Medline: 20093997]
16. Furaijat G, Kleinert E, Simmenroth A, Müller F. Implementing a digital communication assistance tool to collect the medical history of refugee patients: DICTUM Friedland - an action-oriented mixed methods study protocol. BMC Health Serv Res 2019 Feb 06;19(1):103 [FREE Full text] [doi: 10.1186/s12913-019-3928-1] [Medline: 30728030]
17. Kleinert E, Müller F, Kruse S, Furaijat G, Simmenroth A. [Usability and efficiency of a digital communication assistance tool to obtain medical history from non-German-speaking patients]. Gesundheitswesen 2021 Jul;83(7):531-537. [doi: 10.1055/a-1144-2848] [Medline: 32413911]
18. Müller F, Chandra S, Furaijat G, Kruse S, Waligorski A, Simmenroth A, et al. A Digital Communication Assistance Tool (DCAT) to obtain medical history from foreign-language patients: development and pilot testing in a primary health care center for refugees. Int J Environ Res Public Health 2020 Feb 20;17(4):1368 [FREE Full text] [doi: 10.3390/ijerph17041368] [Medline: 32093286]
19. Ogrinc G, Davies L, Goodman D, Batalden P, Davidoff F, Stevens D. Squire 2.0 (Standards for Quality Improvement Reporting Excellence): revised publication guidelines from a detailed consensus process. Am J Crit Care 2015 Nov;24(6):466-473. [doi: 10.4037/ajcc2015455] [Medline: 26523003]
20. Liu X, Rivera SC, Moher D, Calvert MJ, Denniston AK, SPIRIT-AICONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI Extension. BMJ 2020 Sep 09;370:m3164 [FREE Full text] [doi: 10.1136/bmj.m3164] [Medline: 32909959]
21. Emergency Severity Index (ESI): a triage tool for emergency departments. Agency for Healthcare Research and Quality. 2020. URL: https://www.ahrq.gov/patient-safety/settings/emergency-dept/esi.html [accessed 2022-01-11]
22. Brooke J. SUS: a retrospective. J Usab Stud 2013;8:29-40 [FREE Full text]
23. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar 11;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
24. Holtzblatt K, Beyer H. Contextual design: using customer work models to drive systems design. In: Proceedings of the Conference Companion on Human Factors in Computing Systems. 1996 Presented at: Conference Companion on Human Factors in Computing Systems; Apr 13 - 18, 1996; Vancouver British Columbia Canada. [doi: 10.1145/257089.257379]

25. Hogan B, Rasche C, von Reinersdorff AB. The First View Concept: introduction of industrial flow techniques into emergency medicine organization. Eur J Emerg Med 2012 Jun;19(3):136-139. [doi: 10.1097/MEJ.0b013e32834bbd93] [Medline: 22186149]
26. Moore CG, Carter R, Nietert P, Stewart PW. Recommendations for planning pilot studies in clinical and translational research. Clin Transl Sci 2011 Oct;4(5):332-337 [FREE Full text] [doi: 10.1111/j.1752-8062.2011.00347.x] [Medline: 22029804]
27. Thabane L, Ma J, Chu R, Cheng J, Ismaila A, Rios LP, et al. A tutorial on pilot studies: the what, why and how. BMC Med Res Methodol 2010 Jan 06;10:1 [FREE Full text] [doi: 10.1186/1471-2288-10-1] [Medline: 20053272]
28. Altman DG, Gore SM, Gardner MJ, Pocock SJ. Statistical guidelines for contributors to medical journals. Br Med J (Clin Res Ed) 1983 May 07;286(6376):1489-1493 [FREE Full text] [doi: 10.1136/bmj.286.6376.1489] [Medline: 6405856]
29. Kang H. The prevention and handling of the missing data. Korean J Anesthesiol 2013 May;64(5):402-406 [FREE Full text] [doi: 10.4097/kjae.2013.64.5.402] [Medline: 23741561]
30. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
31. Scherer M, Lühmann D, Kazek A, Hansen H, Schäfer I. Patients attending emergency departments. Dtsch Arztebl Int 2017 Sep 29;114(39):645-652 [FREE Full text] [doi: 10.3238/arztebl.2017.0645] [Medline: 29034865]
32. Holzinger F, Oslislo S, Möckel M, Schenk L, Pigorsch M, Heintze C. Self-referred walk-in patients in the emergency department - who and why? Consultation determinants in a multicenter study of respiratory patients in Berlin, Germany. BMC Health Serv Res 2020 Sep 10;20(1):848 [FREE Full text] [doi: 10.1186/s12913-020-05689-2] [Medline: 32912185]
33. 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]
34. Norman CD, Skinner HA. eHealth literacy: essential skills for consumer health in a networked world. J Med Internet Res 2006 Jun 16;8(2):e9 [FREE Full text] [doi: 10.2196/jmir.8.2.e9] [Medline: 16867972]
35. Shiferaw KB, Tilahun BC, Endehabtu BF, Gullslett MK, Mengiste SA. E-health literacy and associated factors among chronic patients in a low-income country: a cross-sectional survey. BMC Med Inform Decis Mak 2020 Aug 06;20(1):181 [FREE Full text] [doi: 10.1186/s12911-020-01202-1] [Medline: 32762745]
36. Terp R, Kayser L, Lindhardt T. Older patients' competence, preferences, and attitudes toward digital technology use: explorative study. JMIR Hum Factors 2021 May 14;8(2):e27005 [FREE Full text] [doi: 10.2196/27005] [Medline: 33988512]
37. Ong SW, Hassali MA, Saleem F. Community pharmacists' perceptions towards online health information in Kuala Lumpur, Malaysia. Pharm Pract (Granada) 2018;16(2):1166 [FREE Full text] [doi: 10.18549/PharmPract.2018.02.1166] [Medline: 30023025]
38. Shiferaw KB, Mehari EA. Internet use and eHealth literacy among health-care professionals in a resource limited setting: a cross-sectional survey. Adv Med Educ Pract 2019;10:563-570 [FREE Full text] [doi: 10.2147/AMEP.S205414] [Medline: 31440113]
39. Greenhalgh T, Wherton J, Shaw S, Morrison C. Video consultations for covid-19. BMJ 2020 Mar 12;368:m998. [doi: 10.1136/bmj.m998] [Medline: 32165352]
40. Heymann DL, Shindo N, WHO ScientificTechnical Advisory Group for Infectious Hazards. COVID-19: what is next for public health? Lancet 2020 Feb 22;395(10224):542-545 [FREE Full text] [doi: 10.1016/S0140-6736(20)30374-3] [Medline: 32061313]
41. Bokolo AJ. Application of telemedicine and eHealth technology for clinical services in response to COVID-19 pandemic. Health Technol (Berl) 2021 Jan 14:1-8 [FREE Full text] [doi: 10.1007/s12553-020-00516-4] [Medline: 33469474]
42. Arora S, Goldberg AD, Menchine M. Patient impression and satisfaction of a self-administered, automated medical history-taking device in the Emergency Department. West J Emerg Med 2014 Feb;15(1):35-40 [FREE Full text] [doi: 10.5811/westjem.2013.2.11498] [Medline: 24695871]
43. Brandberg H, Kahan T, Spaak J, Sundberg K, Koch S, Adeli A, et al. A prospective cohort study of self-reported computerised medical history taking for acute chest pain: protocol of the CLEOS-Chest Pain Danderyd Study (CLEOS-CPDS). BMJ Open 2020 Jan 21;10(1):e031871 [FREE Full text] [doi: 10.1136/bmjopen-2019-031871] [Medline: 31969363]
44. Zakim D. Development and significance of automated history-taking software for clinical medicine, clinical research and basic medical science. J Intern Med 2016 Sep;280(3):287-299 [FREE Full text] [doi: 10.1111/joim.12509] [Medline: 27071980]
45. Zakim D, Braun N, Fritz P, Alscher MD. Underutilization of information and knowledge in everyday medical practice: evaluation of a computer-based solution. BMC Med Inform Decis Mak 2008 Nov 05;8:50 [FREE Full text] [doi: 10.1186/1472-6947-8-50] [Medline: 18983684]
46. Wagholikar KB, Sundararajan V, Deshpande AW. Modeling paradigms for medical diagnostic decision support: a survey and future directions. J Med Syst 2012 Oct;36(5):3029-3049. [doi: 10.1007/s10916-011-9780-4] [Medline: 21964969]
47. Cypress BS. The emergency department: experiences of patients, families, and their nurses. Adv Emerg Nurs J 2014;36(2):164-176. [doi: 10.1097/TME.0000000000000017] [Medline: 24785669]

48. Eisenberg EM, Murphy AG, Sutcliffe K, Wears R, Schenkel S, Perry S, et al. Communication in emergency medicine: implications for patient safety1. Commun Monographs 2005 Dec;72(4):390-413. [doi: 10.1080/03637750500322602]

Abbreviations

AI: artificial intelligence
CLEOS: Clinical Expert Operating System
CONSORT-AI: Consolidated Standards of Reporting Trials–Artificial Intelligence
ED: emergency department
ESI: Emergency Severity Index
HCP: health care provider
RCT: randomized controlled trial
SPIRIT-AI: Standard Protocol Items: Recommendations for Interventional Trials–Artificial Intelligence

Edited by A Mavragani; submitted 24.02.21; peer-reviewed by F Magrabi, A Maier, V Stara, J Benoit; comments to author 06.05.21; revised version received 21.05.21; accepted 14.12.21; published 07.02.22.

Please cite as:
Scheder-Bieschin J, Blümke B, de Buijzer E, Cotte F, Echterdiek F, Nacsa J, Ondresik M, Ott M, Paul G, Schilling T, Schmitt A, Wicks P, Gilbert S
Improving Emergency Department Patient-Physician Conversation Through an Artificial Intelligence Symptom-Taking Tool: Mixed Methods Pilot Observational Study
JMIR Form Res 2022;6(2):e28199
URL: https://formative.jmir.org/2022/2/e28199
doi:10.2196/28199
PMID:35129452

Original Paper

Fertility Workup With Video Consultation During the COVID-19 Pandemic: Pilot Quantitative and Qualitative Study

Hilde Grens[1*], MD; Jan Peter de Bruin[1*], MD, PhD; Aleida Huppelschoten[2*], MD, PhD; Jan A M Kremer[3*], Prof Dr, MD, PhD

[1]Department of Obstetrics and Gynaecology, Jeroen Bosch Hospital, 's-Hertogenbosch, Netherlands

[2]Department of Obstetrics and Gynaecology, Catharina Hospital, Eindhoven, Netherlands

[3]Scientific Institute for Quality in Healthcare, Radboud University Medical Centre, Nijmegen, Netherlands

[*]all authors contributed equally

Corresponding Author:
Hilde Grens, MD
Department of Obstetrics and Gynaecology
Jeroen Bosch Hospital
Henri Dunantstraat 1
's-Hertogenbosch, 5223 GZ
Netherlands
Phone: 31 073 5538660
Email: h.grens@jbz.nl

Abstract

Background: Due to the COVID-19 pandemic, major parts of elective health care in the Netherlands, such as reproductive medicine, were paused. When health care was resumed, video consultation was used as a new solution to continue consultations with the new governmental rules of social distancing. Prior to this COVID-19 situation, video consultation was not used extensively in the Netherlands; therefore, physicians and patients are not familiar with this way of consultation.

Objective: The purpose of this study was to measure the level of patient centeredness and shared decision making in infertile couples who have undergone fertility workup through video consultation.

Methods: This is a questionnaire study with an additional qualitative part for a more in depth understanding. Infertile couples (ie, male and female partners with an unfulfilled wish for a child after 1 year of unprotected intercourse) were referred to a fertility center and underwent fertility workup through video consultation. The fertility workup consisted of 2 separate video consultations, with diagnostic tests according to a protocol. After the last video consultation couples received a digital questionnaire, which consisted of a modified version of the Patient-Centered Questionnaire-Infertility (PCQ-I) and CollaboRATE questionnaire. Fifty-three eligible infertile couples were approached, and of these, 22 participated. Four women were approached for a semistructured interview.

Results: The median score on the modified PCQ-I (scale of 0 to 3) was 2.64. The highest rating was for the subscale communication and information, and the lowest rating was for the subscale organization of care. The median score on the CollaboRATE questionnaire (scale of 1 to 9) was 8 for all 3 subquestions. Patients mentioned privacy, less travel time, and easy use of the program as possible benefits of video consultation. However, patients preferred the first consultation with their physician to be face-to-face consultation as video consultation was considered less personal.

Conclusions: The high levels of patient centeredness and shared decision making show that video consultation is a promising way of providing care remotely, although attention has to be payed to mitigate the more impersonal setting of video consultation when compared with face-to-face consultation.

(JMIR Form Res 2022;6(2):e32000) doi:10.2196/32000

KEYWORDS

COVID-19; patient centeredness; video consultation; fertility care; telemedicine; shared decision making

Introduction

In the initial months of 2020, the COVID-19 pandemic put a stop to most societal activities and major parts of elective health care, such as reproductive medicine. If possible, on-going in vitro fertilization cycles were completed, after which fertility treatments were postponed, based on the advice of both the Dutch and European scientific societies. When elective health care was restarted after a few months, medical departments faced new challenges, having to implement governmental rules on social distancing and minimizing patient traffic in hospitals. One of the solutions for these challenges, in order to continue consultation without physical visits to the hospital, is video consultation.

Video consultation has existed for some time and was mentioned before as a solution to scarcity in health care. Although it was not being used extensively in the Netherlands, it seems to have advantages for patients, such as less travel time and parking costs, and avoidance of waiting in the hospital. Besides, patients are not dependent on others for hospital visits and can experience privacy in the comfort of their own homes. However, technical issues, inferior computer skills, data security issues, and a less personal approach compared with face-to-face consultation are possible pitfalls [1,2]. The less personal approach might be the reason that video consultation has rarely been used during initial consultations. This is supported by the results of a survey by the Dutch Patient Federation and the Amsterdam UMC. Only 6.7% of the respondents opted for video consultation during first contact with a doctor [3].

We have been applying video consulting for first consultations for several years in our project *Fertility Consult*. It involves a secured online platform [4] for infertile couples who are seeking independent advice on fertility treatment. The service includes online questionnaires about prior fertility treatments and medical history, followed by video consultation with a fertility expert. Despite the fact that the video consultation was the first contact couples had with the physician, they were positive about the service. The first results on the patient centeredness of *Fertility Consult* have recently been published [5].

The decision to pause elective health care during the COVID-19 pandemic had a big impact on all patients, with infertility patients being no exception to the rule. It is known that involuntary childlessness is a heavy burden on patients' quality of life. Therefore, infertile patients benefited from quick resumption of care, with a form of care that fitted with the social distancing rules (ie, video consultation). To assess how this care was experienced, a pilot study was set up at the Reproductive Medicine Centre in Jeroen Bosch Hospital, The Netherlands. The aim of this pilot study was to explore patient experiences with a fertility workup through video consultation and extract the possible advantages and disadvantages of video consultation. These results can be used to improve the implementation and quality of video consultation in daily fertility care.

Methods

Overview

We performed a study consisting of both quantitative and qualitative parts to evaluate the experiences of infertile couples with an online fertility workup and obtain more in-depth understanding of their experiences.

Study Design and Recruitment

For the online fertility workup, couples with an unfulfilled wish for a child were referred by their general practitioner after having unprotected intercourse for more than 12 months, just as before the COVID-19 pandemic. Couples completed a questionnaire about their medical situation, after which a video consultation was scheduled instead of a regular appointment in the outpatient clinic. This video consultation was carried out by either a gynecologist or a fertility physician. The program used for the video consultation was called Webcam Consult [6], which works according to the international quality standard ISO 27001 and is linked to the local electronic hospital patient file. Based on the video consultation, the woman was invited to the hospital for a gynecological ultrasound. At this appointment, the woman received laboratory forms for her and her partner for blood and sperm analysis, if necessary.

Consequently, a second video consultation was performed, with the same physician as in the first video consultation, to discuss the results, diagnosis, and possible therapy. After the second video consultation, all couples were asked to participate in the study, which was performed between May and June 2020. Those who did not understand the Dutch language were excluded, as the questionnaire was only available in Dutch. One questionnaire was sent per couple, which could be completed in a secured online environment. Nonresponders received 2 reminders. Four patients were approached, at random, to participate in a semistructured interview. A separate written consent form was collected before the interview.

As the study had minimal impact on the study subjects (ie, filling in a single questionnaire or participating in a single semistructured interview), it was judged that the Dutch Medical Research with Human Subjects Law was not applicable. Therefore, approval of the medical ethics committee was not needed.

Quantitative Part

The quantitative part consisted of a modified version of the validated Patient-Centered Questionnaire-Infertility (PCQ-I) [7] and the CollaboRATE questionnaire [8] measuring patient centeredness and shared decision making, respectively.

The original PCQ-I consists of 7 subscales with 46 questions covering many different subjects of patient-centered infertility care, such as information provision, communication, and continuity of care. The PCQ-I scale has a range from 0 to 3, where a higher score implies a higher level of patient centeredness. Since not all questions were suitable for the specific video consultation setting, we used the modified PCQ-I, as used in the study by Huppelschoten et al [5], which consists of 18 questions. The main modifications were done in the

following categories: accessibility, information, and cooperation within care, since a number of aspects in these questions are not applicable for a setting with a video consultation [5].

The CollaboRATE questionnaire consists of the following 3 questions: How much effort was made to help you understand your health issue? How much effort was made to listen to the things that matter most to you about your health issues? and How much effort was made to include what matters most to you in choosing what to do next? Patients answer these questions on a 10-point Likert scale from 0 (“no effort was made”) to 9 (“every effort was made”).

In addition to the 2 questionnaires, some questions were added about the technical and practical aspects of video consultation.

Qualitative Part

For the qualitative part, several women were approached for semistructured interviews to investigate in more depth the possible advantages and disadvantages of participating in video consultation. A topic list was drafted in advance, based on available literature, and discussed with a small team of experts.

The interviews were held by telephone. The conversation was recorded and transcribed to extract patient comments. Interviews were held until saturation of information was reached. The participants of the interviews also completed the PCQ-I and CollaboRATE questionnaire.

Data Storage and Privacy

The data from the questionnaires were collected from a secure online environment called “Digitale kinderwens poli” (Digital Childwish clinic), which is used in a clinical setting, and exported to Excel. The data from the questionnaire and from the semistructured interviews were handled by a local investigator, who had access to the participant code. For the other members of the research group, the data were anonymous. The data were stored according to the standards of the local facility.

Statistical Analysis

The demographics and background characteristics of the participants and the results of the PCQ-I and CollaboRATE questionnaire were analyzed using descriptive statistics. The normally distributed outcome measures are represented by means of the means with standard deviations, and the nonnormally distributed outcomes are represented by means of the medians. Analyses were performed using SPSS version 22 (IBM Corp).

Results

Quantitative Part

Out of the 53 couples who completed the online fertility workup, 22 responded (response rate 42%). The mean age was 31.7 years (SD 4.1 years). The characteristics of all patients are summarized in Table 1.

Table 1. Baseline characteristics of the participants (N=22).

Characteristic	Value
Age (years), mean (SD)	31.7 (4.1)
Level of education[a], n (%)	
High	14 (64)
Other	8 (36)
Ethnic background, n (%)	
Dutch	19 (85)
European	1 (5)
Asian	1 (5)
Other	1 (5)
Duration of infertility, n (%)	
Less than 2 years	19 (87)
2 to 5 years	3 (13)
Pregnant during the completion of the questionnaire, n (%)	2 (9)

[a]High education indicates higher professional education or university education.

The results of the PCQ-I and CollaboRATE questionnaire are presented in Table 2. The median of the total PCQ-I score was 2.64 (range 0-3). The highest rating was for the subscale communication and information (median 2.83), and the lowest rating was for the subscale organization of care (median 2.33). For the CollaboRATE questionnaire, the median score for all 3 questions was 8 on a scale of 1 to 9.

Table 2. Questionnaire results (N=22).

Outcome values	Value
PCQ-I[a] questionnaire (range 0-3), median (range)	
Total score	2.64 (1.78-2.94)
Communication and information score	2.83 (2.00-3.00)
Respect for patients' values score	2.75 (0.50-3.00)
Staff's competence score	2.70 (1.80-3.00)
Organization of care score	2.33 (1.33-3.00)
CollaboRATE questionnaire (range 0-9), m edian (range)	
Helping to understand health issues score	8 (2-10)
Listen to things that matter most score	8 (5-10)
What to do next score	8 (5-10)
Personal experience	
Overall satisfaction with the Centre for Reproductive Medicine (range 0-10), median (range)	8 (5-10)
Technical problems, n (%)	9 (40)

[a]PCQ-I: Patient-Centered Questionnaire-Infertility.

Qualitative Part

Four interviews were held with female patients. The interviews were analyzed, and quotations were categorized. Acceptance of video consultation was high, given the current COVID-19 situation. Three women expressed that for the first consultation, in a more optimal situation without the COVID-19 pandemic, they would prefer to visit their physician in the hospital.

Example quotations can be found in Textbox 1. Most reported benefits were shorter travel time, a secure feeling concerning privacy, and easy use of the video consultation program.

Textbox 1. Quotations from interviews (N=4).

Information

- *A video consultation is a good replacement for a conversation, a telephonic call is inferior.*
- *It worked perfect to hear the results of the tests per video consultation.*

Privacy

- *I was not worried about the privacy.*
- *I felt very secure in my own environment.*

Communication/respect for patient value

- *I would prefer to talk to my physician face-to-face on the first appointment, because it is possible that video consultation lacks body language.*
- *I found the video consultation a bit impersonal.*

Involvement of partner

- *I could evaluate the appointment with my partner immediately, because we were already home.*

Organization of care

- *I don't like waiting in waiting rooms, and now I was able to wait at home.*
- *We saved a lot of time, without traveling to the hospital.*

Technical aspect

- *Everything went very smooth, no problems.*
- *There was a clear instruction beforehand.*
- *On beforehand I was a bit worried if the connection would be good enough.*
- *We didn't have any sound, so the doctor called us. That made it a bit chaotic.*

Discussion

Principal Findings

This is one of the first studies to explore patient experiences with video consultation in the context of the first consultation. In general, patients perceived a high level of patient centeredness as reflected in a high score on the PCQ-I. Video consultation also scored high in shared decision making on the CollaboRATE questionnaire. During the interviews, patients expressed a relatively high satisfaction rate with video consultation, although they did express that their preference would be to meet their physician face-to-face for the first consultation.

Strengths and Limitations

An important strength of this study is that, to our knowledge, this is one of the first studies to explore patient experiences in their first contact with a new medical department. As both quantitative and qualitative techniques were used, we were able to gather more in-depth information about patient experiences. Besides, we used 2 validated questionnaires, which can be used in the future to compare results.

Some limitations should be mentioned as well. First, since this was a pilot study, the sample size was rather small. The response rate was only 42%, and it would be interesting to know why the other patients did not respond. Second, the relatively high median scores on the PCQ-I and CollaboRATE questionnaire may be affected by the circumstances of the current COVID-19 pandemic [9]. The patients who participated in our study were the first patients to receive fertility consultation after all fertility treatments and appointments had been paused. Therefore, there might have been bias, as the patients were probably relieved to be able to start fertility workup.

Comparison With the Literature

In general, literature on video consultation in health care is limited. Often, only small groups were studied in different settings and almost exclusively using video consultation in the context of a follow-up consultation. Previous Dutch research did study the preferences of a larger group (968 patients), but only 1.7% of the responders had prior experience with video consultation [3]. In another systematic review, physicians seemed to prefer face-to-face consultation, but patients were just as satisfied with video consultation as with face-to-face consultation and preferred it to telephone and email consultations [10].

A British study in primary care showed that video consultation involved less information when compared with face-to-face contact. For example, during video consultation, patients raised fewer issues, the physician provided less information, and the psychosocial context was more often lost [11]. Our study had similar results; in some but not in all cases, patients felt the possibility to express their emotions during the video consultation.

In addition to the content of the video consultation, its implementation is important. In this pilot study, technical problems were recorded in 40% of cases, which is a relatively high percentage. A different British study emphasized this. In the study, clinicians and patients were interviewed after a video consult, and both groups mentioned that technological problems sometimes disrupted the consultation [12]. Both the loss of possible content and the need of higher technical skills entail risks, especially in those patients who already have a psychosocial or socioeconomic disadvantage. To overcome these challenges, the practicing physician must be aware of the possible pitfalls of video consultation and act accordingly. For example, by actively asking patients about the emotional consequences of involuntary childlessness, they are invited to express their emotions.

Conclusions

This pilot study explored infertile patients' experiences using video consultation for their first contact with a gynecologist, after being referred for an unfulfilled wish of having a child. This study showed high scores for patient centeredness and shared decision making. Video consultation is a way of providing care remotely and could potentially be superior to other forms of telecommunication, such as telephone and email contact. Although the acceptance of video consultation might be influenced by COVID-19 circumstances, we do think that the results are promising for the future.

As a next step, we will continue research on video consultation in fertility care. A randomized controlled trial will compare online fertility workup with the standard fertility workup involving face-to-face consults (Dutch trial register number 8554) in terms of patient centeredness and health care costs. Future studies should also focus on how to promote the implementation of video consultation in our current health care setting and minimize technical issues, so that video consultation will be continued after the COVID-19 pandemic.

Conflicts of Interest

None declared.

References

1. Schers H, Buitenhuis E, Besemer Y. Videoconsultatie in de huisartsenpraktijk. Ned Tijdschr Geneeskd. 2014. URL: https://www.ntvg.nl/artikelen/videoconsultatie-de-huisartsenpraktijk [accessed 2022-01-19]
2. Westra I, Niessen FB. Implementing Real-Time Video Consultation in Plastic Surgery. Aesthetic Plast Surg 2015 Oct;39(5):783-790. [doi: 10.1007/s00266-015-0526-4] [Medline: 26169952]

3. Barsom E, Rauwerdink A, Heldoorn M, Schijven M. Videoconsult ten onrechte nog weinig gebruikt. Medisch Contact. 2018. URL: https://www.medischcontact.nl/nieuws/laatste-nieuws/artikel/videoconsult-ten-onrechte-nog-weinig-gebruikt.htm [accessed 2022-01-19]
4. Fertility Consult. URL: https://fertilityconsult.nl/ [accessed 2022-01-20]
5. Huppelschoten AG, de Bruin JP, Kremer JA. Independent and Web-Based Advice for Infertile Patients Using Fertility Consult: Pilot Study. JMIR Form Res 2019 Jun 04;3(2):e13916 [FREE Full text] [doi: 10.2196/13916] [Medline: 31165714]
6. Webcam Consult. URL: https://webcamconsult.com/en/ [accessed 2022-01-20]
7. van Empel IWH, Aarts J, Cohlen B, Huppelschoten DA, Laven JSE, Nelen WLDM, et al. Measuring patient-centredness, the neglected outcome in fertility care: a random multicentre validation study. Hum Reprod 2010 Oct;25(10):2516-2526. [doi: 10.1093/humrep/deq219] [Medline: 20719811]
8. Forcino RC, Barr PJ, O'Malley AJ, Arend R, Castaldo MG, Ozanne EM, et al. Using CollaboRATE, a brief patient-reported measure of shared decision making: Results from three clinical settings in the United States. Health Expect 2017 Jul 05;21(1):82-89. [doi: 10.1111/hex.12588]
9. Gilbert AW, Billany JCT, Adam R, Martin L, Tobin R, Bagdai S, et al. Rapid implementation of virtual clinics due to COVID-19: report and early evaluation of a quality improvement initiative. BMJ Open Qual 2020 May 21;9(2):e000985 [FREE Full text] [doi: 10.1136/bmjoq-2020-000985] [Medline: 32439740]
10. Zandbelt LC, de Kanter FE, Ubbink DT. E-consulting in a medical specialist setting: Medicine of the future? Patient Educ Couns 2016 May;99(5):689-705. [doi: 10.1016/j.pec.2015.11.005] [Medline: 26597384]
11. Hammersley V, Donaghy E, Parker R, McNeilly H, Atherton H, Bikker A, et al. Comparing the content and quality of video, telephone, and face-to-face consultations: a non-randomised, quasi-experimental, exploratory study in UK primary care. Br J Gen Pract 2019 Sep;69(686):e595-e604 [FREE Full text] [doi: 10.3399/bjgp19X704573] [Medline: 31262846]
12. Donaghy E, Atherton H, Hammersley V, McNeilly H, Bikker A, Robbins L, et al. Acceptability, benefits, and challenges of video consulting: a qualitative study in primary care. Br J Gen Pract 2019 Sep;69(686):e586-e594 [FREE Full text] [doi: 10.3399/bjgp19X704141] [Medline: 31160368]

Abbreviations

PCQ-I: Patient-Centered Questionnaire-Infertility

Edited by G Eysenbach; submitted 13.07.21; peer-reviewed by MS Shafi, B Ma; comments to author 23.10.21; revised version received 25.11.21; accepted 27.11.21; published 07.02.22.

Please cite as:
Grens H, de Bruin JP, Huppelschoten A, Kremer JAM
Fertility Workup With Video Consultation During the COVID-19 Pandemic: Pilot Quantitative and Qualitative Study
JMIR Form Res 2022;6(2):e32000
URL: https://formative.jmir.org/2022/2/e32000
doi:10.2196/32000
PMID:34936981

Original Paper

A Conceptual Framework to Design Connected Mental Health Solutions in the United Arab Emirates: Questionnaire Study

Nidal Drissi[1], MSc; Sofia Ouhbi[1], PhD; Leena Amiri[2], PhD; Fadwa Al Mugaddam[2], MPH; Reem K Jan[3], PhD; Minna Isomursu[4], PhD

[1]Department of Computer Science & Software Engineering, College of Information Technology, United Arab Emirates University, Abu Dhabi, United Arab Emirates

[2]Department of Psychiatry and Behavioral Science, College of Medicine and Health Sciences, United Arab Emirates University, Al Ain, United Arab Emirates

[3]College of Medicine, Mohammed Bin Rashid University Of Medicine and Health Sciences, Dubai, United Arab Emirates

[4]Faculty of Information Technology and Electrical Engineering, University of Oulu, Oulu, Finland

Corresponding Author:
Sofia Ouhbi, PhD
Department of Computer Science & Software Engineering
College of Information Technology
United Arab Emirates University
Al Ain
Abu Dhabi, 15551
United Arab Emirates
Phone: 971 37135568
Email: sofia.ouhbi@uaeu.ac.ae

Abstract

Background: Connected mental health (CMH) is a field presenting information and communications technology–based mental care interventions that could help overcome many mental care delivery barriers. Culture and background influence people's attitudes, preferences, and acceptance of such solutions. Therefore, the suitability of CMH solutions to the targeted population is an important factor in their successful adoption.

Objective: The aim of this study is to develop a framework for the design and creation of CMH solutions suitable for the UAE context. The framework is based on investigating enablers and barriers of CMH adoption in the United Arab Emirates, from the mental health professional's (MHP) perspective and from related literature.

Methods: A survey of literature on relevant studies addressing the use of technology for mental care in Arab countries, and a web-based questionnaire-based survey with 17 MHPs practicing in the United Arab Emirates investigating their attitudes and views toward CMH was conducted. Results from the questionnaire and from related studies were analyzed to develop the design framework.

Results: On the basis of findings from the literature survey and analyzing MHP answers to the web-based survey, a framework for the design of CMH solutions for the UAE population was developed. The framework presents four types of recommendation categories: favorable criteria, which included blended care, anonymity, and ease of use; cultural factors including availability in multiple languages, mainly Arabic and English, in addition to religious and cultural considerations; technical considerations, including good-quality communication, availability in formats compatible with mobile phones, and providing technical support; and users' health and data safety considerations, including users' suitability testing, confidentiality, and ensuring MHP integrity.

Conclusions: CMH has the potential to help overcome many mental care barriers in the United Arab Emirates in particular and in the Arab world in general. CMH adoption in the United Arab Emirates has a potential for success. However, many factors should be taken into account, mainly cultural, religious, and linguistic aspects.

(JMIR Form Res 2022;6(2):e27675) doi:10.2196/27675

KEYWORDS

mental health; digital health; eHealth; connected health; mHealth; perceptions; attitudes; framework; design; UAE; mental health care professionals; Arab culture

Introduction

Background

The connected mental health (CMH) field, referring to the use of information and communication technologies (ICTs) for mental health care, has become an established field of research, fueled by the continuous advances in technology and the existing barriers to mental care delivery. CMH includes different mental care solutions including, among others, mobile mental health, e–mental health, digital mental health, and telemental health, which have been increasingly investigated as alternatives or adjuncts to traditional care [1-3]. CMH solutions are available for several mental health conditions, including anxiety disorders [4-6], depression [7], and addiction [8]. CMH solutions are often associated with a number of benefits; they are convenient mental care delivery methods for the patient with regard to time and accessibility to treatment. CMH solutions can also help achieve a more widely spread delivery of mental care, reaching different groups of people. They can also address some of the limitations of the mental care systems, especially regarding costs and mental health professional (MHP) availability, which is currently further challenged by the COVID-19 pandemic. Moreover, CMH solutions can offer the treatment in a discreet and anonymous manner, which may help overcome the stigma barrier.

There is evidence on the effectiveness and acceptability of CMH solutions. Effectiveness has been previously demonstrated for interventions addressing anxiety and depression [9,10], as well as for issues related to exposure to traumatic events such as stress, insomnia, and substance abuse [11]. Acceptance and appropriateness of CMH solutions have been demonstrated globally by studies conducted in a number of countries including Germany [12], Australia [13,14], the Unites States [15], and Canada [16]. In general, positive attitudes around CMH solutions seem to stem from studies testing the outcomes of specific interventions and from participants completing the treatments [17]. However, several studies have revealed low acceptance and negative attitudes toward the use of CMH solutions, which were mainly reported to be related to preference of traditional in-person treatment [18,19], lack of prior awareness or use of CMH solutions [20,21], fear of exposure to false information, and access limitations [22].

MHPs play an important role in the success or failure of the adoption of CMH solutions with patients. With their help and inclusion, effective solutions could be co-designed, and weak aspects of mental care delivery could be identified and addressed. However, MHP attitudes toward CMH solutions are mixed and depend on many factors, including patients' characteristics, types of CMH solutions, and adoption conditions. Several studies have investigated MHP attitudes toward CMH solutions; examples of such studies include, among others, a study conducted in England, which showed overall positive MHP attitudes toward digital solutions. The study showed that younger MHPs are more open to using CMH solutions. However, it also identified limitations to adoption, including that some MHPs reported lack of confidence using technology and expressed concerns regarding security and confidentiality [23]. A study conducted in Australia also showed positive MHP attitudes and reported their concerns regarding patients' accessibility to CMH solutions and their ability to use them [24]. Similar concerns regarding patients were also reported in a study conducted in the United States [15]. In addition, a study conducted in the United Kingdom reported MHP preference of guided use of CMH solutions by the MHP rather than independent use by the patient [25].

Culture and cohort characteristics are also important factors that significantly influence CMH adoption. Although people from different cultures may share some characteristics [26] and preferences regarding certain ICT tools [27], a standard unified approach of ICT intervention design for all cultures might not be successful [27]. Culture is an important factor, as it influences attitudes and preferences [28]. Culture has been found to have an influence on users' preferences and attitudes toward different ICT-based interventions [29-31]. Examples include a study investigating important factors to take into account when designing websites in different cultures (British and Omani), revealing that the users might have similar preferences; however, there were significant differences in the importance and priorities of the preferences. Another study investigated the influence of culture on computer interface acceptance by comparing interface preferences between a group of Australian students and a group of international students, has reported that culture not only influenced the design preferences but also influenced the acceptance, attitudes, and behavior [32].

Objective

There is a scarcity of data on CMH use from the Arab region, even though many of the Arab countries struggle with barriers to mental care delivery [33-35] that could be solved or moderated with the adoption of CMH interventions. To address the paucity of CMH-related research in Arab countries, and to gain insight into the possibility of adoption of CMH in an Arab environment, this study focuses on CMH use in the context of the United Arab Emirates. The Arab world includes the areas known as the Middle East and North Africa. Arab countries can be defined as those where Arabic is the dominant language; they are also religiously and ethnically diverse, with Islam being the dominant religion in most countries [36]. The United Arab Emirates is an Arab country, with distinctive characteristics that include significant numbers of expatriates from different nationalities, which makes it home for different cultures. It must be noted that although different Arab cultures are usually referred to as a unified *Arab culture*, each Arab country has its own unique specifications and characteristics [37]. Therefore, investigating the specific needs and properties of the targeted population is important for the success of any intervention.

This study proposes a framework to design CMH interventions for the UAE population by analyzing related publications and MHP attitudes and views toward the use of CMH interventions. The study identifies facilitators, barriers, and strengths of CMH adoption and extracts necessary elements for the success of CMH adoption in the United Arab Emirates. The framework could also be used in the design of CMH solutions for other Arab countries or populations. Given that design preferences influence the acceptance and behavior [32], the development

of suitable design frameworks for the United Arab Emirates and Arab countries might help increase users' acceptability of this type of mental care intervention and promote CMH adoption.

Methods

Research Design

This study aims to develop a framework to design CMH solutions for the United Arab Emirates. To achieve this objective, a 2-step investigation method was followed. The first part consisted of a literature survey of related studies conducted in Arab countries, to extract relevant findings regarding the advantages and barriers of CMH adoption, in addition to recommendations and guidelines on CMH use. The second part investigated the perceptions and attitudes of MHPs in the United Arab Emirates toward the use of digital technologies for mental health care as well as enablers and barriers of CMH adoption. A survey was carried out using a web-based questionnaire delivered to a group of MHPs practicing in the United Arab Emirates. The questions were designed in collaboration with 2 MHPs and investigated MHP knowledge on CMH, their previous use and willingness to use CMH solutions, best modalities of CMH adoption, as well as their concerns regarding use of CMH. The multiple-choice answers included options that had a significance to the UAE context and culture.

Literature Survey

To construct an understanding of studies addressing the use of technology for mental health in Arab countries, we have conducted a survey of literature following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) protocol [38] in the databases Scopus and Google Scholar using search strings based on the following keywords: *Arab*, *Countries*, *Culture*, *Attitudes*, *Acceptance*, *Use*, *Connected mental health*, *Digital mental health*, *e-mental health*, *Mobile mental health*, and *Tele-mental health*.

The selected studies were investigated and analyzed to extract three main categories of information: advantages of CMH adoption, barriers to CMH adoption, and recommendations for successful CMH adoption.

Web-Based Survey

Data Collection

A self-administered web-based questionnaire, created using Google Forms, was shared with MHPs in the United Arab Emirates via emails. The mailing list to target was provided by MHPs collaborating as coauthors in the study, which contained emails of MHPs from different health institutions in the United Arab Emirates. The questionnaire was pretested by the authors before sending it to the MHPs. The estimated time to complete the questionnaire was 10 minutes, and the answers were collected anonymously, which was stated in the questionnaire. Permission was obtained from the relevant authorities at the UAE University. Information provided about the questionnaire is based on the Checklist for Reporting Results of Internet E-Surveys [39].

Survey Questions

The web-based questionnaire included 18 questions, 13 main questions, and 5 subquestions depending on the answers selected, as presented in Table 1. The first 2 questions were basic questions to provide demographic data (job title and gender). Question Q3 investigated knowledge of MHPs on terms used in the literature to refer to the use of ICT for mental care. Q4-Q9 investigated MHP opinions on different aspects of use of CMH, including cases where CMH could be helpful, whether CMH could assist or replace traditional care, whether CMH could be adopted in the United Arab Emirates, and what elements could promote and improve its adoption. Q10 and Q11 investigated previous MHP use of CMH with their patients, the latest experience with CMH, and MHP willingness to use or reuse CMH with their patients in the future. Q12 and Q13 investigated MHP concerns regarding CMH use and elements that should exist in CMH solutions. Table 1 presents the questions, objectives, and types of answers.

Table 1. Survey questions.

ID and objective	Question	Type of answer
Providing basic information on the participants		
Q1	Job title	Open answer
Q2	Gender	Multiple choice
Investigating MHP[a] knowledge of the terms used to refer to the use of ICT[b] for mental health care, as well as investigating their opinion on different aspects of using digital solutions for mental health care		
Q3	Do you know the following terms? (e-mental health, mobile mental health, connected mental health, digital mental health, telemental health)	Yes or no (for each term)
Q4	In your opinion, in which case could the use of digital technology for mental health care be helpful?	Multiple choice
Q5	Do you think that digital technology for mental health can assist psychiatric therapy?	Yes, no, I don't know or in some cases
Q5.1	If you answered "In some cases," please provide examples of cases	Open answer
Q6	Do you think that therapies delivered via digital technology can replace those delivered face-to-face?	Yes, no, I don't know or in some cases
Q6.1	If you answered "In some cases," please provide examples of cases	Open answer
Q7	In your opinion, what could be the barriers to seeking mental care in the United Arab Emirates that could promote the use of digital solutions?	Multiple choice
Q8	Do you think digital solutions for mental care can be adopted in the United Arab Emirates?	Yes, no, or I don't know
Q8.1	If you answered "No," please explain why	Open answer
Q9	In your opinion, what of the following can improve the adoption of digital solutions for mental health in the UAE culture?	Multiple choice
Investigating MHP previous use of digital solutions with their patients, and their willingness to use them in the future		
Q10	Have you ever used a digital solution with your patients?	Yes or no
Q10.1	If you answered "Yes," how were your patients' attitudes toward it?	Open answer
Q11	Would you be willing to use a digital solution with your patients in the future?	Yes, no, or I don't know
Q11.1	If you answered "No," please explain why	Open answer
Identifying MHP concerns regarding the use of digital solutions for mental health care and their recommended features to be implemented in such solutions		
Q12	What concerns do you have regarding the use of digital solutions for mental care by patients?	Open answer
Q13	In your opinion, what are the critical elements and features that should exist in digital health solutions for mental care?	Open answer

[a]MHP: mental health professional.

[b]ICT: information and communications technology.

Results

Literature Survey Results

Selection Results

A total of 6 relevant studies were selected. The studies have addressed different Arab countries, including Lebanon, Egypt, and Saudi Arabia, and have put forward several significant findings as presented in the following subsections. The studies' selection as well as information of the countries addressed and the aims of the studies are presented in Table 2.

Table 2. Selection results.

Reference	Country or cohort	Aim
Kamel et al [34]	Egypt	Understanding the opinions of psychiatrists on the state of mental health care services in Egypt, their attitudes toward web-based interventions and telemedicine for mental health, and their current knowledge and perceived advantages regarding electronic mental health
Ashfaq et al [40]	Syrian refugees and other vulnerable Arab populations	Evaluating available literature on the use acceptability of mobile mental health in Syrian refugees and other vulnerable Arab populations
Harper Shehadeh et al [41]	Lebanon	Presenting preliminary findings on the feasibility of a minimally guided World Health Organization e-mental health intervention in Lebanon
Abi Ramia et al [33]	Lebanon	Informing the cultural adaptation of an internet-delivered mental health intervention in Lebanon based on a multi-stakeholder perspective
Knaevelsrud et al [42]	War-traumatized Arab patients, focusing on Iraq	Investigating the efficacy of a cognitive behavioral internet-based intervention for war-traumatized Arab patients, with a focus on Iraq
Binhadyan et al [35]	Saudi Arabia	Assisting mental health services in Saudi Arabia by focusing on e-mental health and introducing possibilities and challenges in transforming the e-mental health services of Australia to the Saudi Arabian health care context

Main Findings

Advantages of CMH Adoption

Few studies have investigated CMH adoption in Arab populations, revealing different possible advantages, including reduction of financial, physical, and societal barriers to mental care [33-35], overcoming access limitations related to location and time [35], provision of anonymous access to care [34], reduction of the stigma associated with traditional care [33-35], lowering mental care systems' accessibility threshold, improving therapists' efficacy, provision of psychoeducation [34,35], and wider mental care coverage, as well as reaching vulnerable people [40] and people in remote areas [34].

Barriers to CMH Adoption

Even though CMH might bring solutions to several mental care issues in Arab countries, certain barriers might obviate its adoption. Existing barriers include lack of awareness of the potential seriousness of mental issues [33]; prioritizing other life needs over mental care [33]; high level of illiteracy [33,34,41]; electronic illiteracy, especially among older adults [33,35,40]; concerns regarding confidentiality, privacy, and security [33,34]; resistance to change [34]; lack of content in Arabic [34]; cultural incompatibility [34,40]; technological problems [34]; beliefs that CMH interventions would negatively affect the rapport between patients and MHPs [34]; inadequate technological access, especially for people in conflict regions [40,42]; poor CMH credibility [40]; preference for traditional care [40]; lack of knowledge about CMH [41]; and lack of trust in CMH [42].

Wide use of mobile phones and the wide internet coverage were reported as facilitators to CMH adoption [33,40]. However, several barriers to technology use were reported as well; for example, lack of access and difficulties connecting to the internet [33,40], slow and low quality of internet connection [33], and costs of mobile and internet access [40]. Stigmatization was among the barriers reported to CMH adoption, owing to the stigma of admitting the existence of mental issues and the associated fear of rejection even by family or society [40,42].

Recommendations for Successful CMH Adoption

To overcome some of the barriers to mental care delivery and to CMH adoption in the Arab world, some studies put forward some recommendations. These included ensuring access to good-quality internet connection [41] and making the interventions available in formats compatible with mobile phones [33]. CMH interventions should take into consideration the users' time availability and lifestyle [33] as well as demographics, employment status, family status, previous experiences, and traumas such as war or losing loved ones [33]. When catering to the Arab world, differences between genders in the culture should be considered to make the solutions more relatable; however, unhelpful gender stereotypes should be avoided [33,35]. It is also important to investigate and focus on the mental care needs specific to the targeted population and to provide compatible solutions with those needs [34,40].

To overcome trust and credibility issues as well as the preference of traditional care barrier, blended care could be a helpful approach, in addition to promoting the interventions through trusted parties [33,34]. CMH interventions for Arab populations should incorporate linguistic, cultural, social, and religious considerations [33,35,40].

Web-Based Survey of MHPs

Demographics of the Web-Based Survey Participants

A total of 17 MHPs participated in the survey, and the majority of participants were females (12/17, 71%). The participants occupied different psychiatric posts including consultant psychiatrists, faculty members in psychiatry, medical research specialists, and psychiatrists, who all could be referred to as *experts*, as well as psychiatry residents, who constitute the majority of the respondents (5/17, 29%). Table 3 summarizes the demographic characteristics of the participants.

Table 3. Demographic of participants (N=17).

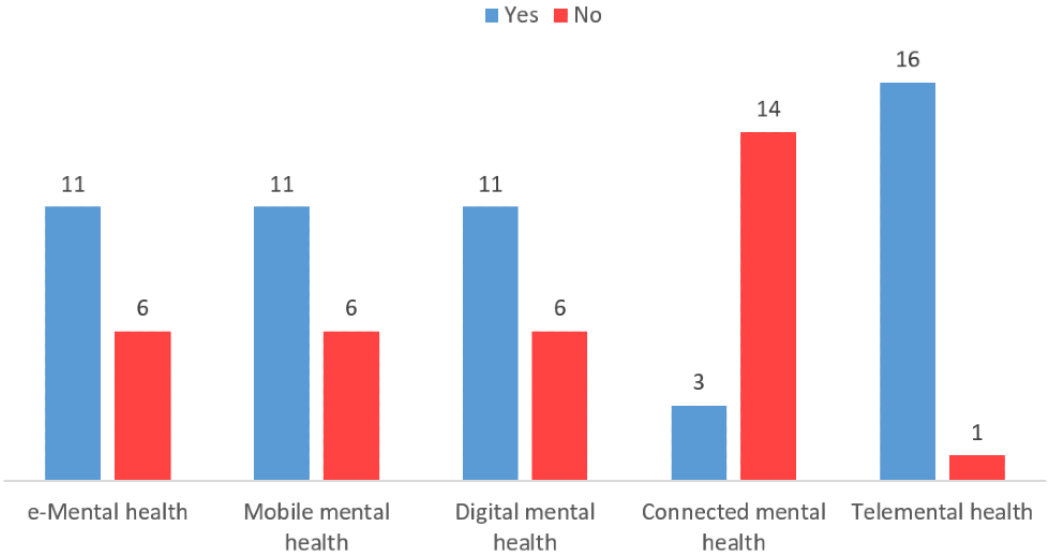
Variables	Participants, n (%)
Gender	
Female	12 (71)
Male	5 (29)
Job titles	
Consultant psychiatrist	5 (29)
Faculty member in psychiatry	3 (18)
Medical research specialist	1 (6)
Psychiatrist	3 (18)
Psychiatry resident	5 (29)

Survey Results

Q3: MHP Awareness of the Terms e-Mental Health, Mobile Mental Health, CMH, Digital Mental Health, and Telemental Health

When investigating knowledge of MHPs on the different terms referring to the use of technology for mental health, *telemental health* was found to be the most known term, as only 6% (1/17) of the participants did not recognize it, whereas *CMH* was the least known term, as 82% (14/17) of the participants did not recognize it. The rest of the terms were found to be equally recognized, with 65% (11/17) of the participants recognizing each term. Figure 1 summarizes the answers to this question.

Figure 1. Mental health care professionals' knowledge on the terms referring to the use of information and communication technologies for mental care.



Q4, Q5, and Q6: Cases and Modalities of CMH Use

Q4 investigated in which cases digital technology could be helpful in mental health care. Answers to this question included common mental problems (depression, stress, anxiety, etc), serious mental problems (schizophrenia, bipolar disorder, dementia, etc), none, and other. Almost all respondents (16/17, 94%) reported that digital solutions could be beneficial in case of common mental disorders, whereas 24% (4/17) of them indicated that it could be beneficial for both common and serious mental disorders, and 6% (1/17) of the participants reported not having an idea of best cases for use.

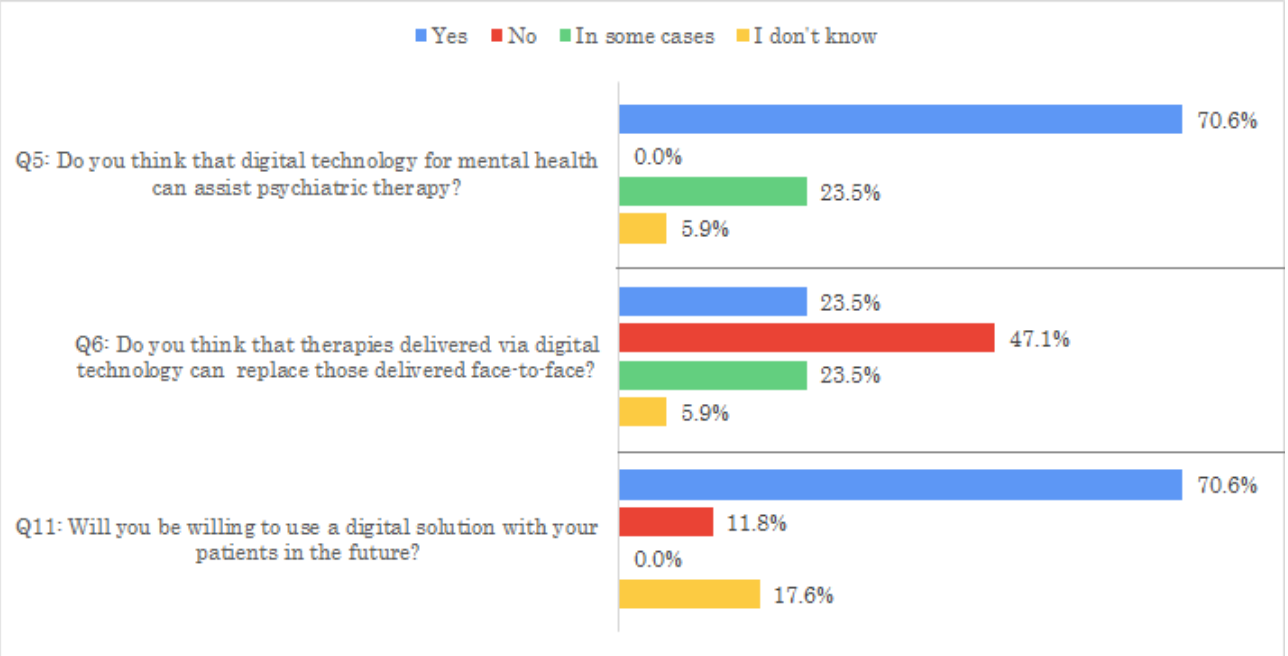
Q5 investigated, based on MHP opinions, whether digital solutions can assist psychiatric therapy. The majority of the participants (12/17, 71%) answered "Yes," 24% (4/17) of the participants answered "in some cases," and 6% (1/17) of the participants answered "I don't know." Figure 2 summarizes the answers to this question. When comparing answers of experts in the sample with those of residents, no major differences were identified for this question, as both reported answers that varied between "Yes" and "In some cases," and none of the participants reported that digital solutions could not assist in psychiatric therapy.

Participants who answered "In some cases" were asked to provide examples of such cases in Q5.1. Their answers are listed in Textbox 1.

Q6 investigated whether digital solutions could replace traditional face-to-face methods of mental care delivery. Overall, 47% (8/17) of the participants answered "No," whereas equal proportions (4/17, 24%) responded "Yes" and "In some cases," as presented in Figure 2. When comparing answers of experts with those of residents for this question, residents seem to be more open to the possibility of CMH replacing traditional therapy, as 58% (7/12) of experts and 20% (1/5) of residents answered "No."

Participants who answered "In some cases" were asked in Q6.1 to provide examples of such cases. Their answers are presented in Textbox 2.

Figure 2. Answers to questions Q5, Q6, and Q11.



Textbox 1. Cases where digital solutions could assist psychiatric therapy.

Participant quotes

- "Some disorders can be managed through digital technology, such as anxiety disorders that does not require medication, application of cognitive behavioral therapy" [P5]
- "Patients who have low risk to harm them or to other, have social support, have normal IQ" [P8]
- "Anxiety; Some cases of depression; insightful schizophrenic patient and bipolar; poor insight schizophrenic and bipolar patients with strong family support; stable dementia patient with family support" [P13]
- "Mild cases with low risk" [P14]

Textbox 2. Cases where digital solutions could replace the traditional face-to-face methods.

Participant quotes

- "I believe most cases requiring cognitive behavioral therapy can be done entirely through digital technology" [P5]
- "Cases with low risk to harm themselves or others; patients with social support; patients that have normal IQ" [P8]
- "Anxiety; some cases of depression; insightful schizophrenic patient and bipolar; poor insight schizophrenic and bipolar patients with strong family support; stable dementia patient with family support" [P13]
- "People with a certain intellectual capacity can engage, however those on the spectrum or intellectually disabled or other neuro-cognitive issues such as dementia or traumatic brain injury (TBI) etc, need in person assistance. Digital health should be an adjunct to human therapy not a replacement" [P17]

Q7: Barriers to Mental Care Seeking in the United Arab Emirates

Q7 was a multiple-choice question. The answers' options for Q7 included cost, stigma, shortage in MHPs, distance from MHPs, lack of knowledge on mental health, and other. The majority of participants (13/17, 76%) reported that the most probable barrier to mental care delivery in the United Arab Emirates was stigma, followed by distance from MHPs, and shortage in MHPs. Table 4 presents the answers to Q7.

Table 4. Barriers to mental care delivery in the United Arab Emirates based on mental health professional (MHP) views (N=17).

Mental care delivery barriers	Participants, n (%)
Stigma	13 (76)
Distance from MHPs	11 (65)
Shortage in MHPs	10 (59)
Cost	8 (47)
Lack of knowledge on mental health	7 (41)

Q8 and Q9: Adoption of CMH Solutions in the United Arab Emirates and Factors That Could Help Their Adoption

Q8 investigated whether digital solutions for mental care can be adopted in the United Arab Emirates based on the MHP opinion. This question can be answered by "Yes," "No," or "I don't know." Of 17 participants, 16 (94%) answered "Yes" to this question, whereas only 1 (6%) participant, who was an expert, answered "No." Participant (P14) who answered "No" justified their choice with the following statement:

> *Digital solutions are unlikely to be able to replace face to face interactions as observation of mental state is such a vital part of our assessment. As mental state is dynamic, digital solutions provide little opportunity to access it repeatedly.*

Q9 investigated which elements can be incorporated in digital solutions for mental care to promote their use in the United Arab Emirates. It is a multiple-choice question with the following answer choices: digital solutions developed in the United Arab Emirates; availability in Arabic and English; religious content, such as Ayat from the Quran or Adkar; or other. The majority reported that availability in Arabic and English (14/17, 82%), as well as being developed in the United Arab Emirates (13/17, 76%), could help promote the use of digital solutions for mental care in the United Arab Emirates. In addition, 1 participant also reported the necessity to consider the different cultures in the UAE population by the following statement: "catering for Urdu/Hindi/Tamal/Arabic/English, at least." Table 5 presents the answers to Q9.

Table 5. Factors that could help the adoption of connected mental health in the United Arab Emirates (N=17).

Factors	Participants, n (%)
Availability in Arabic and English	14 (82)
Digital solutions developed in the United Arab Emirates	13 (76)
Religious content, such as Ayat from the Quran or Adkar	9 (53)

Q10 and Q11: MHP Previous Use of CMH Solutions With Their Patients and Their Willingness to Use Them in the Future

In response to Q10, 53% (9/17) of the participants reported previous use of digital solutions, representing 50% (6/12) of experts and 60% (3/5) of residents. These participants were asked about their patients' attitudes toward the use of digital solutions in their treatment. Their answers are presented in Textbox 3.

Q11 investigated the willingness of MHPs to use digital solutions for mental care in the future. The majority of MHPs (12/17, 71%) answered "Yes," reporting their willingness to use digital solutions in the future, whereas 18% (3/17) of the MHPs answered "I don't know," 2 (67%) of whom were residents and 1 (33%) expert. Figure 2 summarizes the answers to this question.

Only 2 MHP, who were *experts*, answered "No" to Q11, of whom 1 (50%; P14) provided a reason for their answer as follows:

> *Not on a long-term basis. I think it will affect the quality of assessment and also will affect the patient doctor relationship, which is so important in psychiatry.*

Textbox 3. Patients' attitudes toward the use of digital solutions in their treatments.

Participant quotes

- "Generally positive" [P2]
- "Acceptable" [P4 and P6]
- "Happy about it, and some of them were cooperative" [P8]
- "Satisfied" [P12]
- "With serious mental health disorders, it was difficult to reach some poor insight patients, with stable patients and insightful it was welcomed" [P13]
- "Not everyone finds it useful. Some are okay with it as a short-term solution" [P14]
- "Appreciated" [P16]
- "Depends on the patient. Some like it and maximize its benefit, some can barely follow through" [P17]

Q12: MHP Concerns Regarding the Use of Digital Solutions for Mental Care

The participants provided a list of concerns, including, among others, concerns regarding the importance of in-person and physical examinations, security and confidentiality concerns, and patients' suitability concerns. The reported concerns are presented in Textbox 4.

Textbox 4. Reported concerns regarding the use of digital solutions for mental care.

Participant quotes

- "The community response towards change in treatment therapy; Affordability (insurance coverage for some categories of the community); The adaptation of digital solutions into the cultural, behavioral, and language aspects of the country" [P1]
- "Lack of human touch" [P2]
- "Need for physical examination" [P4]
- "Missing high risk patients, inability to reach them in time; Confidentiality; Adherence of patient" [P5]
- "Assessment is jeopardized in many of the psychiatric conditions such as psychosis; Hard to implement it with demented patients unless already diagnosed and the contact is with the family member" [P6]
- "I need face to face assessment as so many symptoms could be reveal only by us" [P7]
- "Need for face to face evaluation and assessment for reaching a plan" [P8]
- "The geriatric patient will need assistance, people might use it to play doctor on others" [P10]
- "Impaired rapport" [P11]
- "Sometimes you need to see the patient face to face to exclude medical causes" [P12]
- "Some Patients are unable to provided the comfortable space that they can find by meeting the doctor face to face, that can easily released their emotions adequately, the signs of relapses can be missed if no strong family support; The compliance to medications and Follow-up (FU), the Electrocardiogram (ECG), lab investigators that done regularly to patients under psychotropic medication" [P13]
- "Lack of frequent observation and proper assessment" [P14]
- "Shouldn't replace the psychiatrist or therapist, medico-legal and ethical concerns" [P15]
- "Technical support regarding opening files, documentation and insurance coverage" [P16]
- "Over utilization and relying solely on digital input, not utilizing or building on coping skills. Social isolation" [P17]

Q13: Critical Elements and Features That Should Exist in Digital Health Solutions for Mental Care

Participants provided a list of critical features that they believe should be incorporated into CMH solutions for mental care. Provided elements included, among others, cultural and linguistic adaptation, patients' suitability testing, ensuring communication between MHPs and patients, and risk or crisis management features. The participants' recommended features are presented in Textbox 5.

Textbox 5. Features and elements that should exist in connected mental health solutions.

Participant quotes

- "Adaptation of language and cultural aspects; Affordability; Good Advocacy and promotion for digital Health Methods" [P1]
- "Digital health solutions should supplement the traditional practice rather than replacing that" [P2]
- "Mix of digital and face to face contact" [P4]
- "Confidentiality, culturally sensitive, uses both English and Arabic, is able to rule out high risk patients" [P5]
- "Camera" [P6]
- "Missing important point in assessment, difficult to reach to some people, at times need to reach to the relative but was difficult" [P8]
- "Easy use and understandable instructions and information for the patients" [P10]
- "Video conferences should be an option" [P11]
- "Video" [P12]
- "Video camera from patients side, the speed of WiFi" [P13]
- "It should be user friendly, dynamic and approachable by vast majority. There should be means of auditing the use of digital health pros and cons. Training for professionals is essential" [P14]
- "Unified regulations" [P15]
- "Confidentiality" [P16]
- "For me the most critical entity that is lacking severely and can be built through Digital health for Mentally ill is a platform to connect patients to resources. In a society with limited resources for the lonely and isolated depressed individuals, these platforms should create network and decrease the isolation" [P17]

Discussion

Principal Findings

Awareness of MHPs in the United Arab Emirates of the terms used to refer to the use of ICT for mental care solutions is aligned with the changes in this research field. Almost all surveyed MHPs had knowledge of the term *telemental health*, which is related to the term *telehealth*, one of the oldest known terms referring to the use of ICT in health, as it has been used in the literature since the 1990s [43,44]. The terms *e-mental health*, *mobile mental health*, and *digital mental health*, are also known terms that have been used in the literature, especially in more recent years [1], which was mirrored in the participants' answers, as the majority reported knowledge of those terms. The majority of MHPs were unaware of the term *connected mental health*; this finding is in line with the changes in the literature, as even though *connected health* is an established field [45], the term *connected mental health* is rarely used in the literature [1].

Surveyed MHPs acknowledged the benefits of CMH, as the majority believed that CMH could assist in psychiatric therapy. However, it must be noted that the majority of participants, especially experts, reported that CMH solutions could not be a replacement for traditional therapy but should be an adjunct to traditional care. Openness of resident participants to the idea of CMH replacing traditional therapy might be explained by them generally being from the young generation, who are familiar with the use of ICT tools, which might have resulted in them having more trust in CMH than experts. Combining traditional and CMH-based therapies or including MHPs in the CMH solutions could be beneficial in promoting and adopting CMH interventions in the United Arab Emirates and in Arab countries in general, as physicians in the Arab culture are much respected and trusted [33,46]. Involvement of MHPs in the CMH solutions may present additional benefits in the UAE context. It may increase the credibility of the treatment and hence adherence to it and may help overcome the barrier of preference of traditional care. Blended care and inclusion of MHPs could also ensure suitability of patients to use CMH solutions. The majority of surveyed MHPs reported that CMH could mainly be used with people with common mental issues, patients with normal IQ and normal cognitive skills, and those who would not present any harm to themselves or to others.

Arab countries face many barriers to mental care seeking and delivery. Such barriers include financial problems, unemployment [33], poorly equipped mental care systems [34], stigma, therapists' availability, as well as cultural barriers such as faith healing beliefs and sex segregation [35]. On the basis of our results, in the case of the United Arab Emirates, stigma followed by MHP availability were the most reported barriers to mental care access. Stigma is one of the barriers most often associated with mental health issues. The suffering of people with mental disorders is not limited to their mental issues' symptoms, which include distress and disability that keep them from living a normal life and achieving their goals; their suffering is extended by the stigma related to the mental issues, causing them to experience social injustice, discrimination, misinterpretation of their mental state, and stereotyping [47,48]. Stigma might be one of the major barriers to mental care seeking in the Arab countries, as in the Arab culture, seeking help from MHPs is viewed as a sign of weakness and a shameful act that impacts not only the individual but also his or her family [37]. The Arab culture is more family or tribe and community oriented and not individual oriented; decisions are usually taken on a

collective level rather than on an individual level to best serve the collective interest [46].

Adoption of CMH interventions could mitigate issues related to stigma and community concerns by providing anonymous and discreet access to care. This, however, does not imply that stigmatization against people with mental issues is normal; it provides a possible solution to overcome it and receive the needed care, as changing stigmatization attitudes have been proven to be a difficult challenge, and implemented strategies to face it were found to be ineffective [47]. On the basis of the results, the availability of MHPs is another barrier to seeking mental care in the United Arab Emirates. However, this barrier could as well be mitigated with the adoption of CMH. CMH interventions could ensure access to care for people in isolated or remote areas and reduce the load on mental care systems, especially considering the wide coverage of internet access and technology use in the United Arab Emirates.

When asked if CMH interventions could be adopted in the United Arab Emirates, the majority of the participants confirmed its possibility and provided several factors that could boost its adoption. The reported factors were mainly related to cultural and linguistic considerations. Lack of consideration of those factors was reported as a barrier to CMH adoption in studies addressing other Arab countries [34,40]. Inclusion of religious content was also believed to be an important factor by more than half of the participants. Religion plays an important role in the Arab culture; it influences the individual's beliefs, life, and behavior [46]. In addition, many Arabs rely on religion for psychological symptom formation, attribution, and management [46]; therefore, its inclusion could make the user more comfortable, accepting, and trusting toward CMH interventions.

As every country or population has its own characteristics, beliefs, and struggles, cultural, economical, and religious considerations may be necessary factors for the successful adoption of CMH in the Arab countries, including the United Arab Emirates. However, it must be noted that in the case of the United Arab Emirates, the majority of the citizens are expatriate residents from different countries with different backgrounds, which should not be neglected when creating interventions for the general population. This insight was also expressed by 1 of the participants who proposed to cater to the different cultures in the United Arab Emirates as a factor that could boost the CMH adoption in the United Arab Emirates.

Only 1 participant expressed unacceptability regarding the CMH adoption in the United Arab Emirates, with concerns mainly around the importance of continuous and face-to-face observation of the mental status of the patients and the possibility that CMH might limit or jeopardize these aspects. Similar findings were reported in a study investigating views of Egyptian psychologists on e-mental health, who reported that it might negatively affect the rapport between patients and MHPs [34]. These concerns further confirm the importance of the inclusion of MHPs in the design and creation of CMH solutions.

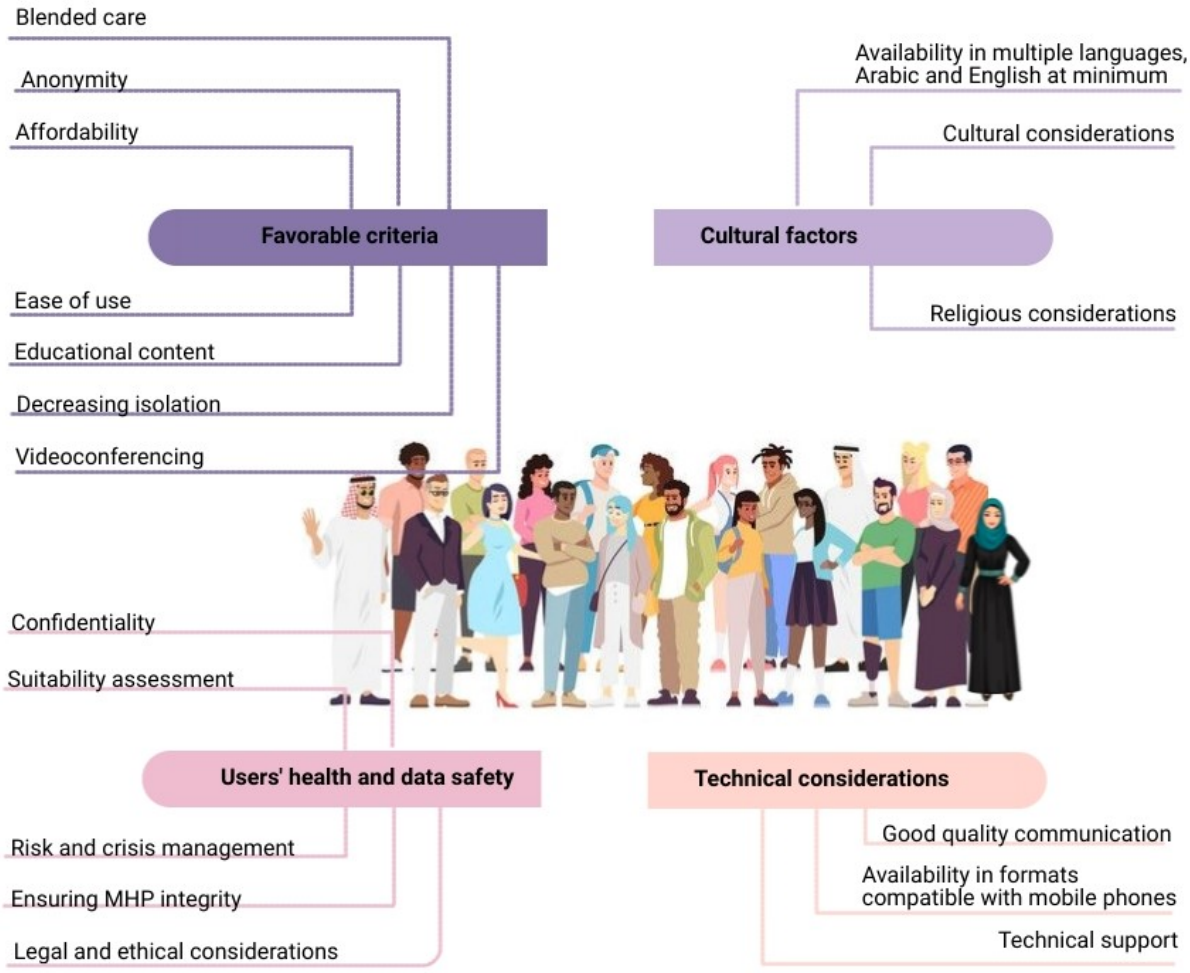
CMH solutions were moderately used by the surveyed MHPs with their patients; approximately 53% (9/17) reported previous use of such solutions, which is in line with the general slow and limited adoption of CMH solutions in the clinical context [49]. Investigated MHPs who have previously used CMH solutions with their patients generally reported positive results. Although some of the participants reported unfavorable outcomes, these were mainly related to patients' characteristics and suitability, which emphasizes the need for users' suitability assessment when using CMH solutions. The majority of participants, including those who have never used CMH interventions with their patients, were open to using them in the future; 2 participants were not open to using CMH, 1 of whom expressed concerns regarding the importance of the physician–patient relationship and the quality of treatment, especially in the long term. Throughout their answers, MHPs repeatedly expressed concerns regarding the importance of MHP input and management of the patients, which again implies that blended care might be the most suitable approach for CMH interventions to be adopted in the United Arab Emirates.

A Conceptual Framework to Design CMH Interventions for the UAE Context

Overview

Surveyed MHPs were asked to provide their concerns regarding the adoption of CMH interventions as well as critical elements and features that they think should exist in CMH solutions. On the basis of the list of concerns and critical elements as well as the answers of participants to the entire survey, a framework for the design of CMH solutions for the UAE population, including Emiratis and expatriates, with a focus on people of an Arab nationality was developed. The framework was also reviewed and validated by 2 MHPs collaborating in the study. The main points of this framework are presented in Figure 3.

Figure 3. Factors to consider when designing connected mental health interventions for the UAE population. MHP: mental health professional.



Favorable Criteria

Blended Care

Blended care may be a trusted approach that has a better chance of success in the United Arab Emirates than independent CMH solutions. Blended care can provide the benefits of technology without losing those of traditional care and of communicating with MHPs. In addition, MHP inclusion and presence in CMH solutions could increase the patients' trust in the solutions and promote their use.

Anonymity

Anonymity could help encourage people to use CMH solutions to seek the needed care, overcoming stigma barriers related to mental health.

Affordability

Costs of professional mental care in the United Arab Emirates are relatively high, which constitutes a barrier to mental care seeking, especially for non-Emirati expatriates. The cost issue is further aggravated by either limited or nonexistent coverage of mental care from insurance companies, as the majority of insurance plans do not include mental care coverage [50,51]. CMH interventions should help overcome the cost barrier by offering free and more affordable solutions. In addition, providing insurance coverage for the use of paid CMH solutions could mitigate the affordability issue.

Ease of Use

CMH solutions should be simple and easy to use and understand for the patients to be comfortable using them.

Providing Educational Content on Mental Health

There is a lack of knowledge on mental health and its importance. CMH solutions should provide educational content that is easy to access and understand, which could help spread awareness on the importance and seriousness of mental health.

Providing Recommendations and Advice That Decrease Isolation

One of the concerns regarding the use of CMH interventions is decreasing the user's social interaction, which is usually already impaired in people with mental issues. Therefore, CMH solutions should include tips and recommendations that would encourage the user to have social interactions with their families, friends, and community, which would decrease their isolation. Examples of such recommendations include proposing outdoor activities in addition to tasks that incorporate family or friends.

Videoconferencing

Videoconferencing is one of the communication modalities that could be integrated in blended care, which could provide face-to-face communication without the need to be in the same location as the MHP. Face-to-face communication could be imperative for certain cases. In addition, MHPs could extract important information from physical observation of the patients, including their gestures, movements, and facial expressions. However, certain points should be respected when adopting blended care in the United Arab Emirates through solutions such as videoconferencing, with the major one being accommodation for Hijabi women. Most Emirati women wear Hijab, which requires them to dress in a certain way in public and when around males who are not related to them, so is the case for many Arab and Muslim women. There are also women who wear Niqab, which requires them to cover their faces. Therefore, when communicating with an MHP especially via videoconferencing, women should be informed beforehand of the gender of the MHP, so they could be prepared and not to cause them any discomfort [52]. In addition, they should preferably have the ability to choose the MHP gender. Both men and women in the Arab culture might not feel comfortable discussing certain sensitive subjects. Therefore, CMH solutions for the United Arab Emirates and for Arab countries in general should provide a gender choice. Not respecting limitations and habits regarding interactions between the genders might present a major barrier to CMH adoption in the United Arab Emirates.

Cultural Factors

Availability in Multiple Languages, At Least in Arabic and English

Providing CMH solutions in the native languages of all residents of the United Arab Emirates might not be feasible. Therefore, CMH interventions should at least be provided in Arabic for Emirati and Arab expatriates and in English for non-Arab expatriates. Arabic and English are the 2 most spoken languages in the United Arab Emirates, and availability in these 2 languages is important to reach a large proportion of the United Arab Emirates's population.

Cultural Considerations

The Arab culture has its own characteristics, including beliefs, traditions, and gender-related specifications, which should be taken into consideration, to offer suitable, relatable solutions and avoid inappropriate content.

Religious Considerations

When creating CMH solutions specifically for Emiratis or Arabs in general, including religious content might offer a sense of trust and comfort to many users. However, as the United Arab Emirates has a mixed population of people with different religious backgrounds, religion should be respected in the sense of avoiding the inclusion of any shapes, images, illustrations, or colors that might have religious meanings and could offend certain users.

Users' Health and Data Safety

Confidentiality

Health information in general is sensitive, especially mental health information. Data privacy and security should be a top priority when designing CMH interventions to protect the users; if patients' data fall into the wrong hands, it could be used to harm them [53]. Patients should be assured that their data would not be hacked or leaked to encourage them to use the CMH interventions and share their data.

Respecting Legal and Ethical Guidelines

Law and ethics for psychological treatment in general and for use of technology differ from one country to another. Existing laws and ethics should be investigated and respected. There are laws in the United Arab Emirates regarding use of technology and data disclosure. For example, Article 379 of the UAE Penal Code prohibits a person, who by means of their profession is entrusted with a *secret*, from disclosing that *secret* or information without consent [54]. In addition, Federal Law No. 5 of 2012 and its amendment Federal Law No. 12 of 2016 prohibit individuals from using any electronic information systems or any information technology tools to offend or invade another person's privacy without authorization [55,56].

Users' Suitability Assessment

CMH suitability testing is imperative to avoid putting any user at risk. Testing could be either conducted by an MHP or based on validated psychological tests such as the Depression Anxiety and Stress Scale [57,58]. Psychological testing would help acquire an overview on the psychological state of the user to determine whether the intervention could be beneficial and appropriate for the specific user.

Risk and Crisis Management Features

Some users might go through crises and situations that may require immediate intervention from MHPs or family members. Features to manage these situations should be considered to ensure the safety of users. Such features include automatic messages to family or caregivers, providing hotlines' contact, and providing directions that could help the user overcome the crisis or risk situation.

Ensuring MHP Integrity and Content Safety

CMH interventions should ensure that any content or input presenting mental care treatments, advice, or practices is sourced from legitimate and licensed MHPs. Illegitimate or harmful input can be a major issue when the CMH solution provides features such as chat between users, web-based communication between the community of the interventions' users, posts, comments, or any features that enable insertion of an input. All these features should imperatively be monitored by an MHP to eliminate harmful content.

Technical Considerations

Good and Clear Communication

Good and clear communication between the patient and the MHP or caregiver is crucial. Mental health is a sensitive subject, and the patient should be put in an environment without

interruptions and difficulties, such as unclear video communication or sound. The United Arab Emirates offers good-quality internet network, with one of the fastest data and download speeds in the world [59,60]. CMH interventions should take advantage of the UAE network quality and ensure that technical issues that would interrupt the communication between the patient and the MHP are avoided.

Providing Technical Support and Technology Education to Both Patients and MHPs

Some MHPs and patients may not be accustomed to or comfortable using technology; therefore, technology education and training may be necessary before use. In addition, technical support should be available in case of need.

Availability of Solutions in Formats Compatible With Mobile Phones

Mobile phones are one of the most used devices. Availability of the CMH solutions in formats that could be used and accessed through mobile phones could help increase the accessibility to CMH solutions and promote their use.

Limitations

This study may have some limitations: (1) the number of MHPs who responded to the study was low, mainly because of the timing of the investigation being during the COVID-19 lockdown, which made reaching a large number of participants difficult, and (2) conducting interviews with the participants would have provided valuable insights, but because of the timing of the study, conducting face-to-face interviews was not possible.

Conclusions

CMH interventions could help overcome many of the existing mental care delivery barriers in the United Arab Emirates, mainly stigma and the availability of MHPs. Surveyed MHPs confirmed the utility and possible adoption of CMH in the United Arab Emirates and provided a number of factors that should be considered when creating CMH solutions for the UAE population. These factors mainly included cultural, religious, and linguistic considerations. However, MHPs also expressed certain concerns around the importance of physician–patient relationship, which could be jeopardized with patients' reliance on CMH for mental care, in addition to patients' suitability concerns, as CMH interventions may not be suitable and beneficial for all patients. On the basis of the MHP input, a blended care approach, combining both traditional treatment and CMH solutions, was concluded to be potentially suitable, beneficial, and successful in the UAE context.

Surveyed MHPs have also reported a set of concerns regarding CMH adoption and a set of critical features that should exist in CMH solutions. On the basis of the input of MHPs, a list of design features to consider when creating CMH interventions for the United Arab Emirates was presented.

Acknowledgments

This work is part of the Abu Dhabi Young Investigator Award 2019 (AYIA19-001) awarded by the Abu Dhabi Research and Development Authority and the Startup project (31T131) funded by the United Arab Emirates University.

Authors' Contributions

All authors contributed to the creation of the manuscript. ND contributed in the following areas: design, conception, acquisition and interpretation of data, analysis of provided answers, drafting of the manuscript, and revision. SO contributed in the following areas: design, conception, statistical support, interpretation of data, drafting of the manuscript, and critical revision. LA and FAM contributed in the following areas: design, distribution of the questionnaire, validation of the framework, and critical revision. RKJ and MI contributed to the critical revision.

Conflicts of Interest

None declared.

References

1. Drissi N, Ouhbi S, Janati Idrissi MA, Fernandez-Luque L, Ghogho M. Connected mental health: systematic mapping study. J Med Internet Res 2020;22(8):e19950 [FREE Full text] [doi: 10.2196/19950] [Medline: 32857055]
2. Olff M. Mobile mental health: a challenging research agenda. Eur J Psychotraumatol 2015;6(1):27882 [FREE Full text] [doi: 10.3402/ejpt.v6.27882] [Medline: 25994025]
3. Langarizadeh M, Tabatabaei M, Tavakol K, Naghipour M, Rostami A, Moghbeli F. Telemental health care, an effective alternative to conventional mental care: a systematic review. Acta Inform Med 2017;25(4):240-246 [FREE Full text] [doi: 10.5455/aim.2017.25.240-246] [Medline: 29284913]
4. Mayo-Wilson E, Montgomery P. Media-delivered cognitive behavioural therapy and behavioural therapy (self-help) for anxiety disorders in adults. Cochrane Database Syst Rev 2013(9):CD005330. [doi: 10.1002/14651858.CD005330.pub4] [Medline: 24018460]
5. Drissi N, Ouhbi S, Janati Idrissi MA, Ghogho M. An analysis on self-management and treatment-related functionality and characteristics of highly rated anxiety apps. Int J Med Inform 2020;141:104243 [FREE Full text] [doi: 10.1016/j.ijmedinf.2020.104243] [Medline: 32768994]

6. Drissi N, Ouhbi S, Idtissi MA, Ghogho M. Mobile apps for post traumatic stress disorder. Annu Int Conf IEEE Eng Med Biol Soc 2019;2019:4279-4282. [doi: 10.1109/EMBC.2019.8857197] [Medline: 31946814]
7. Richards D, Richardson T. Computer-based psychological treatments for depression: a systematic review and meta-analysis. Clin Psychol Rev 2012;32(4):329-342. [doi: 10.1016/j.cpr.2012.02.004] [Medline: 22466510]
8. Riper H, Spek V, Boon B, Conijn B, Kramer J, Martin-Abello K, et al. Effectiveness of E-self-help interventions for curbing adult problem drinking: a meta-analysis. J Med Internet Res 2011;13(2):e42 [FREE Full text] [doi: 10.2196/jmir.1691] [Medline: 21719411]
9. Kaltenthaler E, Parry G, Beverley C, Ferriter M. Computerised cognitive-behavioural therapy for depression: systematic review. Br J Psychiatry 2008;193(3):181-184. [doi: 10.1192/bjp.bp.106.025981] [Medline: 18757972]
10. Van't Hof E, Cuijpers P, Stein DJ. Self-help and Internet-guided interventions in depression and anxiety disorders: a systematic review of meta-analyses. CNS Spectr 2009;14(2 Suppl 3):34-40. [doi: 10.1017/s1092852900027279] [Medline: 19238128]
11. Amstadter AB, Broman-Fulks J, Zinzow H, Ruggiero KJ, Cercone J. Internet-based interventions for traumatic stress-related mental health problems: a review and suggestion for future research. Clin Psychol Rev 2009;29(5):410-420 [FREE Full text] [doi: 10.1016/j.cpr.2009.04.001] [Medline: 19403215]
12. Ebenfeld L, Kleine Stegemann S, Lehr D, Ebert DD, Funk B, Riper H, et al. A mobile application for panic disorder and agoraphobia: Insights from a multi-methods feasibility study. Internet Interv 2020;19:100296 [FREE Full text] [doi: 10.1016/j.invent.2019.100296] [Medline: 31890640]
13. Rodda SN, Merkouris S, Lavis T, Smith D, Lubman DI, Austin D, et al. The therapist experience of internet delivered CBT for problem gambling: Service integration considerations. Internet Interv 2019;18:100264 [FREE Full text] [doi: 10.1016/j.invent.2019.100264] [Medline: 31890617]
14. Strauss P, Morgan H, Wright Toussaint D, Lin A, Winter S, Perry Y. Trans and gender diverse young people's attitudes towards game-based digital mental health interventions: a qualitative investigation. Internet Interv 2019;18:100280 [FREE Full text] [doi: 10.1016/j.invent.2019.100280] [Medline: 31890628]
15. Miller KE, Kuhn E, Yu J, Owen JE, Jaworski BK, Taylor K, et al. Use and perceptions of mobile apps for patients among VA primary care mental and behavioral health providers. Prof Psychol Res Pr 2019;50(3):204-209. [doi: 10.1037/pro0000229]
16. Smail-Crevier R, Powers G, Noel C, Wang J. Health-related internet usage and design feature preference for E-mental health programs among men and women. J Med Internet Res 2019;21(3):e11224 [FREE Full text] [doi: 10.2196/11224] [Medline: 30882361]
17. Kaltenthaler E, Sutcliffe P, Parry G, Beverley C, Rees A, Ferriter M. The acceptability to patients of computerized cognitive behaviour therapy for depression: a systematic review. Psychol Med 2008;38(11):1521-1530. [doi: 10.1017/S0033291707002607] [Medline: 18205964]
18. Apolinário-Hagen J, Hennemann S, Kück C, Wodner A, Geibel D, Riebschläger M, et al. Exploring user-related drivers of the early acceptance of certified digital stress prevention programs in Germany. Health Serv Insights 2020;13:1178632920911061 [FREE Full text] [doi: 10.1177/1178632920911061] [Medline: 32206013]
19. Whealin JM, Seibert-Hatalsky LA, Howell JW, Tsai J. E-mental health preferences of Veterans with and without probable posttraumatic stress disorder. J Rehabil Res Dev 2015;52(6):725-738 [FREE Full text] [doi: 10.1682/JRRD.2014.04.0113] [Medline: 26562090]
20. Hennemann S, Witthöft M, Bethge M, Spanier K, Beutel ME, Zwerenz R. Acceptance and barriers to access of occupational e-mental health: cross-sectional findings from a health-risk population of employees. Int Arch Occup Environ Health 2018;91(3):305-316. [doi: 10.1007/s00420-017-1280-5] [Medline: 29189895]
21. Feng XL, Campbell A. Understanding E-mental health resources: personality, awareness, utilization, and effectiveness of E-mental health resources amongst youth. J Technol Hum Serv 2011;29(2):101-119. [doi: 10.1080/15228835.2011.595276]
22. Keane MC, Roeger LS, Allison S, Reed RL. e-Mental health in South Australia: impact of age, gender and region of residence. Aust J Prim Health 2013;19(4):331-335. [doi: 10.1071/PY13027] [Medline: 23945369]
23. 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;10:344 [FREE Full text] [doi: 10.3389/fpsyt.2019.00344] [Medline: 31133906]
24. Williams A, Fossey E, Farhall J, Foley F, Thomas N. Going online together: the potential for mental health workers to integrate recovery oriented E-mental health resources into their practice. Psychiatry 2018;81(2):116-129. [doi: 10.1080/00332747.2018.1492852] [Medline: 30273103]
25. Breedvelt JJ, Zamperoni V, Kessler D, Riper H, Kleiboer AM, Elliott I, et al. GPs' attitudes towards digital technologies for depression: an online survey in primary care. Br J Gen Pract 2019;69(680):e164-e170 [FREE Full text] [doi: 10.3399/bjgp18X700721] [Medline: 30559111]
26. Weir D, Hutchings K. Cultural embeddedness and contextual constraints: knowledge sharing in Chinese and Arab cultures. Knowl Process Mgmt 2005;12(2):89-98. [doi: 10.1002/kpm.222]
27. Al-Badi AH. Designing usable localised websites: differences between Western and Arab cultures. IBIMA Bus Rev 2009;2:9-19.

28. Papadopoulos I, Koulouglioti C. The influence of culture on attitudes towards humanoid and animal-like robots: an integrative review. J Nurs Scholarsh 2018;50(6):653-665. [doi: 10.1111/jnu.12422] [Medline: 30242796]
29. Van Slyke C, Lou H, Belanger F, Sridhar V. The influence of culture on consumer-oriented electronic commerce adoption. J Electron Commer Res 2010;11(1):30-40.
30. Kaba B, Osei-Bryson KM. Examining influence of national culture on individuals' attitude and use of information and communication technology: assessment of moderating effect of culture through cross countries study. Int J Inf Manage 2013;33(3):441-452. [doi: 10.1016/j.ijinfomgt.2013.01.010]
31. Lopez-Nicolas C, Meroño-Cerdán AL. The impact of organizational culture on the use of ICT for knowledge management. Electron Mark 2009;19(4):211-219. [doi: 10.1007/s12525-009-0020-4]
32. Evers V, Day D. The role of culture in interface acceptance. In: IFIP TC13 International Conference on Human-Computer Interaction. 1997 Presented at: INTERACT '97; July 14-18, 1997; Sydney p. 260-267. [doi: 10.1007/978-0-387-35175-9_44]
33. Abi Ramia J, Harper Shehadeh M, Kheir W, Zoghbi E, Watts S, Heim E, et al. Community cognitive interviewing to inform local adaptations of an e-mental health intervention in Lebanon. Glob Ment Health (Camb) 2018;5:e39 [FREE Full text] [doi: 10.1017/gmh.2018.29] [Medline: 30637112]
34. Kamel MM, Westenberg JN, Choi F, Tabi K, Badawy A, Ramy H, et al. Electronic mental health as an option for Egyptian psychiatry: cross-sectional study. JMIR Ment Health 2020;7(8):e19591 [FREE Full text] [doi: 10.2196/19591] [Medline: 32788155]
35. Binhadyan BS, Peszynski K, Wickramasinghe N. The effect of E-mental health services on Saudi general mental health. In: 28th Bled eConference. 2015 Presented at: BLED eConference; June 7-10, 2015; Slovenia p. 169-179.
36. Office of the Deputy Chief of Staff for Intelligence, U.S. Army Training and Doctrine Command. Arab cultural awareness: 58 factsheets. Fort Eustis: Office of the Deputy Chief of Staff for Intelligence, U.S. Army Training and Doctrine Command; 2006.
37. Sayed MA. Conceptualization of mental illness within Arab cultures: meeting challenges in cross-cultural settings. Soc Behav Pers 2003;31(4):333-342. [doi: 10.2224/sbp.2003.31.4.333]
38. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
39. Eysenbach G. Improving the quality of Web surveys: the checklist for reporting results of internet E-surveys (CHERRIES). J Med Internet Res 2004;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
40. Ashfaq A, Esmaili S, Najjar M, Batool F, Mukatash T, Al-Ani HA, et al. Utilization of mobile mental health services among Syrian refugees and other vulnerable Arab populations-a systematic review. Int J Environ Res Public Health 2020;17(4):1295. [doi: 10.3390/ijerph17041295] [Medline: 32085422]
41. Harper Shehadeh MJ, Abi Ramia J, Cuijpers P, El Chammay R, Heim E, Kheir W, et al. Step-by-step, an E-mental health intervention for depression: a mixed methods pilot study from Lebanon. Front Psychiatry 2019;10:986 [FREE Full text] [doi: 10.3389/fpsyt.2019.00986] [Medline: 32116815]
42. Knaevelsrud C, Brand J, Lange A, Ruwaard J, Wagner B. Web-based psychotherapy for posttraumatic stress disorder in war-traumatized Arab patients: randomized controlled trial. J Med Internet Res 2015;17(3):e71 [FREE Full text] [doi: 10.2196/jmir.3582] [Medline: 25799024]
43. Nickelson DW. Telehealth and the evolving health care system: strategic opportunities for professional psychology. Prof Psychol Res Pr 1998;29(6):527-535. [doi: 10.1037/0735-7028.29.6.527]
44. Simpson J, Doze S, Urness D, Hailey D, Jacobs P. An assessment of routine telepsychiatry services. Edmonton: Alberta Heritage Foundation for Medical Research; 1999.
45. Karampela M, Isomursu M, Porat T, Maramis C, Mountford N, Giunti G, et al. The extent and coverage of current knowledge of connected health: systematic mapping study. J Med Internet Res 2019;21(9):e14394 [FREE Full text] [doi: 10.2196/14394] [Medline: 31573915]
46. Okasha A. The impact of Arab culture on psychiatric ethics. Arab J Psychiatry 2008;19(2):81-99.
47. Corrigan PW, Bink AB. On the stigma of mental illness: practical strategies for research and social change. Washington DC: American Psychological Association; 2005.
48. Brouwers EP. Social stigma is an underestimated contributing factor to unemployment in people with mental illness or mental health issues: position paper and future directions. BMC Psychol 2020;8(1):36 [FREE Full text] [doi: 10.1186/s40359-020-00399-0] [Medline: 32317023]
49. Balcombe L, De Leo D. Digital mental health challenges and the horizon ahead for solutions. JMIR Ment Health 2021;8(3):e26811 [FREE Full text] [doi: 10.2196/26811] [Medline: 33779570]
50. Gokulan D. The cost of keeping the mind healthy in UAE. Khaleej Times. 2019. URL: https://www.khaleejtimes.com/uae/the-cost-of-keeping-the-mind-healthy-in-uae [accessed 2020-10-01]
51. Khalaf H. Mental health in the UAE: limited insurance cover is huge barrier for patients. The National. 2016. URL: https://www.thenationalnews.com/lifestyle/wellbeing/mental-health-in-the-uae-limited-insurance-cover-is-huge-barrier-for-patients-1.219098 [accessed 2020-09-30]

52. Zaman S. Hijabis in UAE share their unique challenge on Zoom, Teams meetings. UAE. 2020. URL: https://gulfnews.com/uae/hijabis-in-uae-share-their-unique-challenge-on-zoom-teams-meetings-1.74853477 [accessed 2020-12-01]
53. Sipilä J. Patients in Finland blackmailed after therapy records were stolen by hackers. CNN Business. 2020. URL: https://edition.cnn.com/2020/10/27/tech/finland-therapy-patients-blackmailed-data-breach-intl/index.html [accessed 2020-11-30]
54. Data protection laws of the world. DLA Piper. URL: https://www.dlapiperdataprotection.com/index.html?t=law&c=AE [accessed 2020-10-01]
55. Data and privacy protection in the UAE. The UAE' Government portal. 2019. URL: https://u.ae/en/about-the-uae/digital-uae/data/data-protection-laws [accessed 2020-10-01]
56. UAE: up to Dh200,000 fine, jail term for accessing data without authorisation. Khaleej Times. 2020. URL: https://www.khaleejtimes.com/uae/uae-up-to-dh200000-fine-jail-term-for-accessing-data-without-authorisation [accessed 2020-10-01]
57. Evans L, Haeberlein K, Chang A, Handal P. Convergent validity and preliminary cut-off scores for the anxiety and depression subscales of the DASS-21 in US adolescents. Child Psychiatry Hum Dev 2021;52(4):579-585. [doi: 10.1007/s10578-020-01050-0] [Medline: 32816139]
58. Lee B, Kim Y. Validity of the depression, anxiety, and stress scale (DASS-21) in a sample of Korean university students. Curr Psychol 2020:1-10 [FREE Full text] [doi: 10.1007/s12144-020-00914-x]
59. Chnadran T. Top 10 countries with the fastest internet speeds. Gulf Business. 2015. URL: https://gulfbusiness.com/top-10-countries-fastest-internet-connection/ [accessed 2020-10-01]
60. Ryan P. UAE has the fastest home internet in the region. The National. 2019. URL: https://www.thenationalnews.com/uae/uae-has-the-fastest-home-internet-in-the-region-1.884109 [accessed 2020-09-30]

Abbreviations

CMH: connected mental health
ICT: information and communication technology
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by A Mavragani; submitted 02.02.21; peer-reviewed by I de la Torre, S Chen, T McCall; comments to author 27.05.21; revised version received 22.07.21; accepted 27.12.21; published 07.02.22.

Please cite as:
Drissi N, Ouhbi S, Amiri L, Al Mugaddam F, Jan RK, Isomursu M
A Conceptual Framework to Design Connected Mental Health Solutions in the United Arab Emirates: Questionnaire Study
JMIR Form Res 2022;6(2):e27675
URL: https://formative.jmir.org/2022/2/e27675
doi:10.2196/27675
PMID:35129458

Original Paper

Home-Based Spirometry Telemonitoring After Allogeneic Hematopoietic Cell Transplantation: Mixed Methods Evaluation of Acceptability and Usability

Ajay Sheshadri[1*], MSc, MD; Sukh Makhnoon[2*], MSc, PhD; Amin M Alousi[3], MD; Lara Bashoura[1], MD; Rene Andrade[2], BS; Christopher J Miller[2], MPH; Karen R Stolar[3], MS; Muhammad Hasan Arain[1], MBBS; Laila Noor[1], MBBS; Amulya Balagani[1], MD; Akash Jain[1], MD; David Blanco[1], RPFT; Abel Ortiz[1], CRT; Michael S Taylor[4], BS; Alex Stenzler[4], MD; Rohtesh Mehta[3], MD; Uday R Popat[3], MD; Chitra Hosing[3], MD; David E Ost[1], MD, MPH; Richard E Champlin[3], MD; Burton F Dickey[1], MD; Susan K Peterson[2], PhD

[1]Department of Pulmonary Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX, United States

[2]Department of Behavioral Sciences, The University of Texas MD Anderson Cancer Center, Houston, TX, United States

[3]Department of Stem Cell Transplantation, The University of Texas MD Anderson Cancer Center, Houston, TX, United States

[4]Monitored Therapeutics, Inc, Dublin, OH, United States

[*]these authors contributed equally

Corresponding Author:
Ajay Sheshadri, MSc, MD
Department of Pulmonary Medicine
The University of Texas MD Anderson Cancer Center
1400 Pressler Unit 1462
Houston, TX, 77030
United States
Phone: 1 7135631987
Email: asheshadri@mdanderson.org

Abstract

Background: Home-based spirometry (HS) allows for the early detection of lung complications in recipients of an allogeneic hematopoietic cell transplant (AHCT). Although the usability and acceptability of HS are critical for adherence, patient-reported outcomes of HS use remain poorly understood in this setting.

Objective: The aim of this study is to design a longitudinal, mixed methods study to understand the usability and acceptability of HS among recipients of AHCT.

Methods: Study participants performed HS using a Bluetooth-capable spirometer that transmitted spirometry data to the study team in real time. In addition, participants completed usability questionnaires and in-depth interviews and reported their experiences with HS. Analysis of interview data was guided by the constructs of performance expectancy, effort expectancy, and social influence from the Unified Theory of Acceptance and Use of Technology model.

Results: Recipients of AHCT found HS to be highly acceptable despite modest technological barriers. On average, participants believed that the HS was helpful in managing symptoms related to AHCT (scores ranging from 2.22 to 2.68 on a scale of 0-4) and for early detection of health-related problems (score range: 2.88-3.12). Participants viewed HS favorably and were generally supportive of continued use. No significant barriers to implementation were identified from the patient's perspective. Age and gender were not associated with the patient perception of HS.

Conclusions: Study participants found HS acceptable and easy to use. Some modifiable technical barriers to performing HS were identified; however, wider implementation of pulmonary screening is feasible from the patient's perspective.

(JMIR Form Res 2022;6(2):e29393) doi:10.2196/29393

KEYWORDS

allogeneic hematopoietic cell transplantation; home-based spirometry; acceptability; usability; mixed methods evaluation; patient perspectives; spirometry; feasibility; mHealth; home-based; remote care; respirology; pulmonary medicine; mobile phone

Introduction

Background

Despite improvements in the recognition and treatment of bronchiolitis obliterans syndrome (BOS) after allogeneic hematopoietic cell transplantation (AHCT), outcomes remain poor [1]. The risk of death after BOS correlates with the severity of airflow impairment at diagnosis [1,2], and recipients of AHCT with severe airflow obstruction at BOS diagnosis have increased nonrelapse mortality compared with recipients of AHCT without BOS [3]. Early BOS presents as lymphocytic bronchiolitis and may be more amenable to treatment [4], whereas late BOS presents as airway fibrosis and is generally treatment refractory [5]. As BOS is a disease with the potential for rapid progression [6], prompt diagnosis may improve therapeutic outcomes.

Home-based spirometry (HS) has been an effective strategy for diagnosing early BOS in recipients of lung allograft and is standard practice for posttransplantation monitoring of lung function. Several studies have documented excellent adherence and proficiency with forced spirometric maneuvers [7-10]. Although fewer studies have been conducted in recipients of AHCT [11-13], the progression of BOS related to graft-versus-host disease (GVHD) may be slowed or halted in recipients of AHCT in whom airflow obstruction is promptly recognized and treated [1]. As a result, despite the lower incidence of BOS in recipients of AHCT than chronic lung allograft rejection [14-16], which occurs in most recipients of a lung transplant, given sufficient time [17], a spirometric telemonitoring program may be valuable for diagnosing BOS in recipients of AHCT at an earlier stage and preventing progression.

Objective

Recipients of AHCT may face barriers to performing HS. First, psychosocial burnout and fatigue are common symptoms among recipients of AHCT and may be more prevalent early in the course after transplantation [18,19]. Therefore, recipients of AHCT may not be willing to perform routine HS measurements during this time, when screening for BOS could be of significant benefit [20]. Second, until recently, real-time data collection was not possible, and HS data were often evaluated by less efficient methods such as by mail or landline phone connections [11,12]. By leveraging the widespread use of smartphones and wireless connectivity [21], we seek to implement an HS pilot program through which spirometry data could be delivered in near real time by wireless networks, allowing for scalable, efficient telemonitoring. We previously reported the technical feasibility of performing home spirometry in recipients of AHCT [22]. Here, we seek to assess the usability and acceptability of HS among recipients of AHCT using mixed methods–patient-reported outcome (PRO) surveys and in-depth interviews.

Methods

Study Overview

This paper describes the results of a pilot feasibility study of HS. The full details of the study protocol have been previously reported [22]. Briefly, at enrollment, participants were given a Bluetooth-compatible home spirometer (GoSpiro, Monitored Therapeutics Inc), which communicates wirelessly with patients' smartphones through an app or custom tablets (GoHome, Monitored Therapeutics Inc; Figure 1). Participants with tablets could additionally take advantage of an automated coaching algorithm which gave instructions to encourage maximal expiratory flow and sufficient expiratory time. Spirometric and questionnaire data were subsequently transmitted to a remote electronic server via an internet-based portal. Participants were instructed to perform 3 maneuvers per session up to 3 times per week. Those who were unable to remain adherent to the study protocol for a sufficient period to generate a baseline forced expiratory volume in 1 second (FEV_1) measurement (ie, they did not perform at least one technically acceptable HS maneuver on 6 separate days within the first 2 weeks) were removed from the study. Participants were instructed to continue measuring until at least 1 year after the transplant or approximately 9 months after enrollment. One week of nonadherence to the study protocol resulted in weekly phone calls or emails to the patient as a reminder to resume measurements at their convenience.

Figure 1. GoSpiro Bluetooth-compatible home spirometer (forefront) and the GoHome wireless tablet (background).

Participants

We consented and enrolled English-speaking adult recipients of AHCT who were seen at an AHCT survivorship clinic at approximately 100 days after transplantation between October 2016 and June 2018. Patients were enrolled in a longitudinal cohort study lasting up to 9 months, involving a baseline visit followed by remote measurement of key clinical and PRO variables. Demographic and clinical data were also collected from electronic health records and institutional databases. To gather information on user experience, participants completed web-based surveys on usability and acceptability at 1-, 3-, 6-, and 9-month follow-ups using REDCap (Research Electronic Data Capture) [23] and completed one-on-one telephone interviews at 1-, 3-, and 6-month follow-ups. The study was approved by the MD Anderson institutional review board (2015-0990).

Quantitative Data Collection

At 1, 3, 6, and 9 months, participants were asked to answer follow-up questions about the HS device, app design, its features, its overall usefulness, satisfaction, their intention to recommend spirometry to other recipients of AHCT, the acceptability of the tablet and mobile app, its impact on health management, and suggestions to improve the spirometer in the future. Usability testing was conducted using a structured questionnaire to collect responses to 14 Likert-style questions scored on a 5-point scale ranging from 0 (not at all) to 4 (extremely), as well as 5 open-ended questions. Example questions include “Overall, how satisfied were you with the GoSpiro Home Spirometer?” and “How useful do you believe the GoHome tablet and GoSpiro app is in helping you to manage your symptoms related to your stem cell transplantation?”

Patient engagement with remote monitoring was measured using the patient activation measure (PAM), which comprises 13 items with strong psychometric properties [24]. Items are focused on constructs of confidence, beliefs, knowledge, and skills about managing one’s health, which respondents can answer with degrees of agreement or disagreement (eg, “I know how to prevent problems with my health” and “I am confident that I can tell a doctor my concerns, even when he or she does not ask”). The measure is scored on a theoretical 0 to 100 scale, with higher scores reflecting a developmental progression from passive receipt of care toward greater activation.

Qualitative Data Collection

We conducted one-on-one telephone interviews with participants at 1, 3, and 6 months of the study period. Interviews followed a semistructured interview guide designed to elicit participants’ views on the usability (functionality, navigation, and interactivity) and acceptability of HS. Open-ended qualitative questions included the following:

- “How has using the HS Spirometer affected your health in general?”
- “Is there anything you do not like about the HS Spirometer or that made you not want to use that device?”
- “How helpful do you think it would be for other patients like you to use the HS spirometer?”

Interviews were audio recorded, professionally transcribed, and redacted before directed qualitative content analyses.

Theoretical Model

The Unified Theory of Acceptance and Use of Technology (UTAUT) model was originally developed as a conceptual framework to explain individuals' intention to adopt and use technological innovations [25]. In this study, we drew from this theory to describe the applicability and likelihood of using the HS device among recipients of AHCT. UTAUT is a widely used model of information technology adoption and has been used to examine the adoption of various technologies in different contexts [26]. The following key theoretical constructs from the UTAUT model guided the analysis of in-depth interview data: (1) performance expectancy, (2) effort expectancy, and (3) facilitating conditions. The construct of social influence from the original model was eliminated as it was not applicable to this study. We also sought to identify potential contextual factors that might affect home spirometer use based on moderating characteristics identified by UTAUT (age, gender, experience, and voluntariness).

Data Analysis

Quantitative Data Analysis

We summarized demographic, clinical, and PROs as mean (SD) or number (percentage). Differences between groups were compared using chi-square analysis for categorical variables or 2-tailed *t* tests for continuous variables. Participant ratings for each usability question were reported as the mean score per item (scale 0-4).

Qualitative Data Analysis

Initial content analysis via line-by-line coding, a form of open coding, was performed by 2 coders independently to identify emergent codes. The coders then met to debrief and discuss emergent codes derived from the process. Codes were discussed and clarified between the 2 analyses, and adjustments were made until intercoder consensus was achieved, which is an approach that ensures rigor. Discrepancies were resolved by discussion, and quotations illustrating the main themes were identified during the coding process. Secondary content analyses were then performed by applying the UTAUT as a semistructured framework. Atlas.ti (version 8.0; ATLAS.ti Scientific Software Development GmbH) qualitative data analysis software was used for the coding and content analysis.

Mixed Methods Integration

The value of mixed methods research lies in the meaningful integration of qualitative and quantitative components. We used merging integration, which comprised comparing qualitative findings with respect to PAM scores and survey responses with usability questions and determining whether the qualitative findings did or did not support, or expand, our understanding of the scores [27].

Results

Study Sample

We screened 207 patients for study participation (Figure 2). Of the 207 patients, 82 (39.6%) completed the baseline assessment and received training to use the spirometer. Of these 82 patients, 51 (62%) performed spirometry and registered baseline values for FEV_1 and were considered to be adequately familiar with HS to complete qualitative and quantitative measures on HS use. Table 1 provides the demographic characteristics of all study participants, whereas Table 2 provides the demographic characteristics of participants who completed the survey and interview at each follow-up time point. Participants who did not remain in the study long enough to register baseline values for FEV_1 (ie, did not perform at least six spirometric maneuvers on separate days within the first month) did not differ significantly from participants included in the study with respect to any demographic characteristics (age, race, ethnicity, and sex). Of the 51 patients, usability surveys were completed by 27 (53%), 22 (43%), 24 (47%), and 15 (29%) participants at the 1-, 3-, 6-, and 9-month follow-ups, respectively. Of the 51 patients approached for recruitment over the course of the study, 35 (69%) recipients of AHCT completed one or more telephone interviews: 3 (9%) completed all 3 interviews, 9 (26%) completed 2 interviews, and 23 (66%) completed 1 interview, with an overall response rate of 69%. Participants who completed open-ended interviews were predominantly older, with a median age of 58 (range 23-75) years, non-Hispanic White, and male. PAM scores (ranging between 80.9 and 86.8) indicated a relatively high level of engagement in the management of self-care while living with a chronic illness (Table 2).

Figure 2. Study enrollment flowchart. AHCT: allogeneic hematopoietic cell transplant; FEV_1: forced expiratory volume in 1 second.

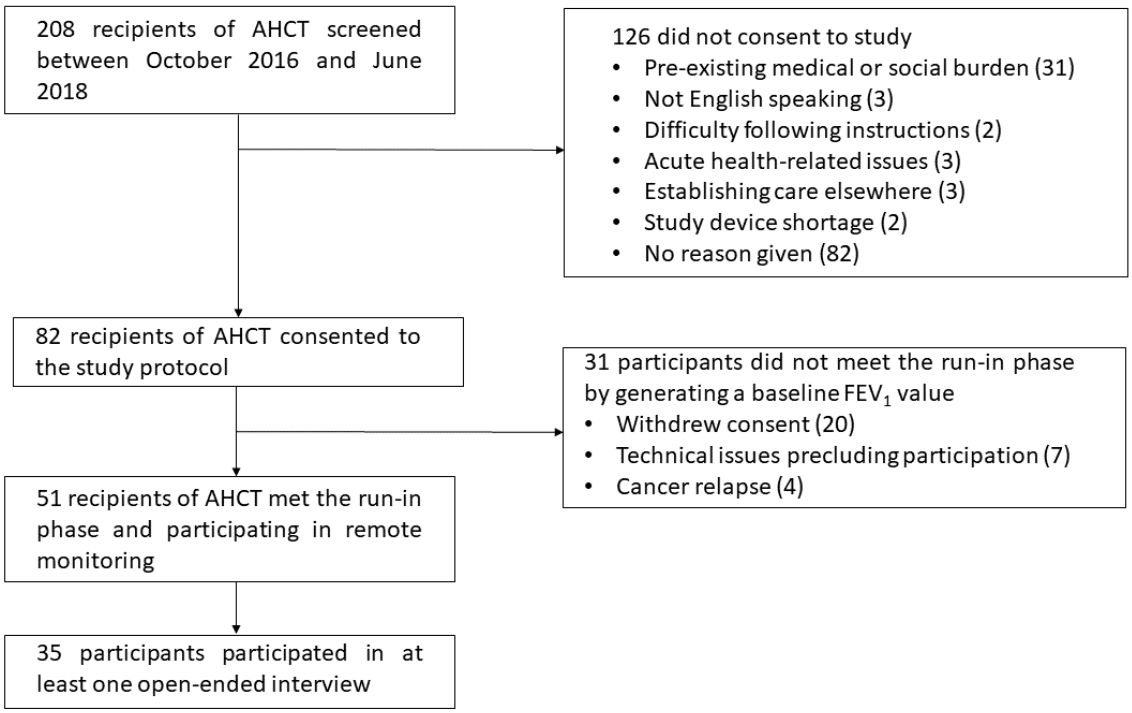


Table 1. Characteristics of the overall study cohort (N=51).

Variable	Values
Age (years), median (IQR)	55 (41-64)
Sex, n (%)	
Male	17 (33)
Female	34 (67)
Race, n (%)	
White	47 (92)
Person of color	4 (8)
Underlying malignancy, n (%)	
Acute myeloid leukemia or myelodysplastic syndrome	25 (49)
Acute lymphoblastic leukemia	4 (8)
Chronic lymphocytic leukemia	4 (8)
Chronic myeloid leukemia	7 (14)
Lymphoma	6 (12)
Myeloma or plasma cell disorder	3 (6)
Myelofibrosis	2 (4)
Acute GVHD[a] before enrollment, n (%)	
Yes	39 (76)
No	12 (24)
Chronic GVHD at enrollment, n (%)	
Yes	18 (35)
No	33 (65)
Baseline PAM[b] score (IQR)	66 (56-80)
Baseline FEV_1[c], % predicted (IQR)	96 (83-106)
Baseline FVC[d], % predicted (IQR)	91 (80-98)
Baseline FEV_1/FVC ratio (IQR)	79 (76-84)

[a]GVHD: graft-versus-host disease.

[b]PAM: patient activation measure.

[c]FEV_1: forced expiratory volume in 1 second.

[d]FVC: forced vital capacity.

Table 2. Characteristics of participants who provided open-ended interviews (N=51).

Variable	1 month (n=25)	3 months (n=22)	6 months (n=24)	9 months (n=15)
Quantitative data (usability survey)				
Age (years), median (range)	59 (32-74)	60 (33-72)	62.5 (40-74)	64 (33-72)
Sex, n (%)				
Male	14 (56)	12 (55)	15 (63)	10 (67)
Female	11 (44)	10 (46)	9 (38)	5 (33)
Race, n (%)				
Asian	0 (0)	1 (5)	0 (0)	0 (0)
Black	1 (4)	0 (0)	0 (0)	0 (0)
White	24 (96)	20 (91)	24 (100)	15 (100)
Other	0 (0)	1 (5)	0 (0)	0 (0)
Ethnicity, n (%)				
Hispanic	1 (4)	2 (9)	3 (13)	1 (7)
Non-Hispanic	23 (92)	19 (86)	21 (88)	14 (93)
Unknown	1 (4)	1 (5)	0 (0)	0 (0)
Patient activation measure, mean (range)	86.4 (67.3-100)	84.3 (63.5-100)	86.8 (73.1-100)	80.9 (57.7-100)
Qualitative data (semistructured interview)				
Completion rate	15 (29.4)	14 (27.5)	21 (41.2)	N/A[a]

[a]N/A: not applicable.

Usability Ratings

Overall, participants rated HS as highly usable, and usability remained high throughout the study period (Table 3). On average, participants believed that HS was helpful in managing symptoms related to AHCT (scores ranging from 2.22 to 2.68 on a scale of 0-4) and for early detection of health-related problems (score range 2.88-3.12). They were also willing to recommend the spirometer to other recipients of AHCT (score range 2.65-3.16) and were satisfied with HS overall (score range 2.31-3.53). The only negatively worded item about their illness interfering with spirometer use received low scores (range 0.62-0.88), indicating that AHCT was not a barrier to using the spirometer. Similarly, patients also rated the GoHome tablet and GoSpiro app as highly usable, satisfactory, and helpful.

Table 3. Usability questionnaire scores (N=51).

Usability items	1 month (n=25)		3 months (n=22)		6 months (n=24)		9 months (n=15)	
	Value, mean (SD)[a]	Positive response, n (%)	Value, mean (SD)	Positive response, n (%)	Value, mean (SD)	Positive response, n (%)	Value, mean (SD)	Positive response, n (%)
GoSpiro home spirometer								
Illness interfered with ability to use the home spirometer	0.88 (1.4)	3 (12)	0.82 (1.05)	1 (5)	0.62 (0.97)	2 (8)	0.79 (1.18)	1 (7)
Ease of use	2.19 (1.52)	12 (48)	2.55 (1.14)	14 (64)	2.67 (1.34)	16 (67)	2.43 (1.28)	9 (60)
Helpful in managing symptoms related to stem cell transplantation	2.22 (1.42)	11 (44)	2.68 (1.17)	14 (64)	2.54 (1.32)	13 (54)	2.33 (1.29)	6 (40)
Helpful for early detection of health-related problems	2.88 (1.22)	16 (64)	3 (0.93)	17 (77)	3.12 (0.95)	20 (83)	3 (1.24)	11 (73)
Belief that GoSpiro home spirometer helps health care providers monitor illness	2.95 (1.31)	17 (68)	3.04 (1.13)	18 (82)	2.92 (0.97)	18 (75)	3.14 (1.17)	11 (73)
Using the GoSpiro home spirometer gives a feeling of security	2.44 (1.42)	14 (56)	2.77 (0.97)	16 (73)	2.75 (1.15)	16 (67)	2.8 (0.94)	9 (60)
Willingness to continue using for up to 2 years	2.35 (1.23)	11 (44)	2.18 (1.3)	10 (45)	2.25 (1.48)	13 (54)	2.26 (1.71)	8 (53)
Willingness to recommend to other patients who receive a stem cell transplantation	2.65 (1.25)	13 (52)	2.9 (1.02)	14 (64)	3.16 (0.96)	19 (79)	3 (1.13)	10 (67)
Overall satisfaction	2.4 (1.39)	13 (52)	2.32 (1.09)	12 (55)	2.79 (1.1)	15 (63)	3.53 (1.3)	9 (60)
GoHome tablet and Go Spiro app								
Ease of use	2.22 (1.55)	13 (52)	2.59 (1.05)	15 (68)	2.92 (1.05)	16 (67)	2.53 (1.25)	9 (60)
Helpful in managing symptoms related to stem cell transplantation	2.34 (1.47)	12 (48)	2.27 (1.16)	11 (50)	2.75 (1.18)	15 (63)	2.33 (1.34)	8 (53)
Willingness to recommend to other patients who receive a stem cell transplantation	2.65 (1.28)	14 (56)	2.72 (0.94)	15 (68)	3.16 (0.96)	19 (79)	2.86 (1.4)	9 (60)
Overall satisfaction	2.26 (1.53)	11 (44)	2.38 (0.97)	13 (59)	2.66 (1.05)	12 (50)	2.4 (1.18)	7 (47)
Automated coaching								
Overall satisfaction	2.59 (1.43)	14 (56)	2.45 (1.36)	12 (55)	2.87 (1.14)	15 (62)	2.93 (1.32)	11 (73)

[a]Scores range from 0=not at all to 4=extremely; scores of 3 or 4 were considered positive responses.

Qualitative Findings

Participant responses around usability and acceptability of HS as well as contextual factors that modify these responses are summarized in Table 4 and discussed below. Qualitative results are organized by the constructs of the UTAUT model, which represent overarching messages emerging out of the cross-sectional and narrative analysis of the interviews over the three time points: 1, 3, and 6 months after transplantation.

Table 4. Example quotes from study participants mapped within constructs of the Unified Theory of Acceptance and Use of Technology construct model.

Unified Theory of Acceptance and Use of Technology construct	Example quotes
Performance expectancy	
Perceived usefulness	• "I think it's a valuable tool and I see benefits for me and I'm sure others would. I think it's easy to operate and I'm pleased with it. I think it's doing me some good." • "I think it would be very helpful, just because they're so many unexpected complications with the stem cell transplant and I think that...it would be very easy to use to get that information." • "I think it's a great product and it'll help a lot of other people. It's also going to help me if there's anything developing that you need to know about in advance." • "Well, I think it's a good system. I think it should be considered as part of the standard treatment for transplant patients. I'm glad to have it and it doesn't take a lot of time, it doesn't take a lot of energy and so it's something that's reassuring, and I don't mind doing it at all"
Outcome expectations	• "If I develop pulmonary GVHD, we're likely to catch it much sooner. It can be interrupted at an earlier stage well before it damages my lungs too much." • "I think it's helped me to try to do better to maintain...better health because I will take my measurements and I always try to like keep higher measurements than before or try to maintain it at the same level." • "From the beginning, I kind of thought that, 'well, it was a good reading,' but yeah what does that mean? That was a little higher—does that mean that's good or bad? I mean, it'd be nice to know that you're still on the right track." • "No. I mean basically, I do the same things that I was doing before, and I haven't changed anything"
Relative advantage	• "Well, I can tell if something's wrong with my lungs easily, but if there's something that I'm not seeing, then maybe [it] makes me feel good to know that they may be able to see a change in pressure...that I can't tell. Identify graft versus host disease in my lungs before I know it" • "It's very easy and you can do it at home, it's a lot easier than going to someplace and get it done professionally...this is reassuring to me that I'm [not] going downhill with GVHD."
Effort expectancy	
Ease of use	• "Yeah, I've got it down where I can do it quickly to the cell phone and the GoSpiro device, they went up quickly on Bluetooth and light turns green two, three times. And using the technique that they showed me and I would say it's very quick, very easy." • "I mean it's pretty quick. I usually do it around dinner time, I'll pull it up, it takes, maybe a minute totally set it up. I take a minute and a half and I'm done with everything." • "Yeah. I've kind of made it into a game to see if I can get up to eight each time and I'm usually right about eight, upper sevens or low eights each time I do it. But yeah, it's almost like making them into a little game, competition, see I can do better than I did the day before. Now there's certainly a technique blowing into it and it took me a while to learn the technique" • "It's part of the routine getting dressed, do breathing exercise three times a week, and it's not a problem whatsoever, it's been easy." • "I think the phone is pretty convenient. I don't really see an issue with that."
Complexity	• "So [charging] is very quirky I find, but, it's not difficult to work with. It's just difficult to get the three tests all in a row...it takes a little time, and it's not something I can do in 10 minutes or so because I have to play with it to get it to work all the time." • "Well, I wanted to use it and it wouldn't work. So we figured it wasn't lying on the charger correctly, it was bent somehow so it would not lie on the charger correctly" • "I continue thinking that the charging port needs to be a little bit longer. The cord is really short, not only that but having, maybe possibly indicator like for charging on the spirometer itself—red while it charged and then it turns green when it's fully charged" • "But I was doing it and apparently it was working, the little Avatar said your measurement has been recorded. So I thought everything was right and then I received a call saying if I wanted to continue with this study. I was like why would you say that and they said because you only have done one measurement during the whole week in a month." • "The communication between the Spiro and the tablet was broken very frequently. I have to repeat several times the process of the measurement, and then I was not able to do it, nothing. The technical things are the problem."
Facilitating conditions	
Portability	• "[I] take it apart and just set it in a suitcase. The battery lasts for the duration of my travel, so I just take the device itself, so it doesn't take up a lot of room and it travels well."

Unified Theory of Acceptance and Use of Technology construct	Example quotes
Reminders and positive reinforcement	• "The bottom line is that you need to have some similar goal system set up, so the participants continue to be interested in doing it. I think it's just natural human nature that if you're being rewarded for whatever you're doing as simple as, doing a great job or hey you're doing great, your scores are continuing to being in pain or whatever, some kind of feedback to encourage you to do at the next time, I think would be helpful."

Performance Expectancy

Perceived expectancy and plans to use the HS device were operationalized through three subconstructs of the UTAUT model: *perceived usefulness*, *outcome expectations*, and *relative advantage*.

> *The perceived usefulness of HS was ascertained from the patients' summative experiences and perceived benefits gained from using the instrument. In agreement with survey responses, most participants felt that HS would be helpful in detecting early complications with stem cell transplantation, including GVHD, and that it should be considered part of the standard treatment for recipients of AHCT. The spirometer was believed to be highly useful and easy to use.*

Outcome expectations are expectations regarding the impact of using HS on health and health behavior. Most participants expected that regular HS use would enable them to better maintain their health by retaining consistent pulmonary function. They reported self-competition and goal setting as factors that motivated them to continue HS use. In addition, the knowledge that a member of the study team was regularly monitoring their lung function and would intervene if needed was reassuring to the study participants. One of the expectations *not met* by the HS device was its inability to provide real-time feedback and interpretation of pulmonary function measurements. None of the participants reported making any health behavior change as a result of using HS. These data are supported by the survey results, where participants expressed a strong belief that HS allowed providers to monitor the illness.

Relative advantage relates to the degree to which an innovation is perceived as being better than using its precursor. In this study, participants believed that the benefits of using HS vastly outweighed the use of clinic-based spirometry, and this belief was a prominent positive influence for their continued completion of spirometric measurements. They were aware of the possible complications of AHCT, including GVHD, and thus, the ability to monitor pulmonary health without having to go to the hospital was deemed immensely advantageous. These interview results were supported by survey findings where use was reported to provide a feeling of security and help manage symptoms related to transplantation.

Effort Expectancy

Effort expectancy is defined as the degree of ease associated with the use of a system with two subconstructs: *ease of use* and *complexity.* Many participants reported that the HS device was easy to use and found the overall HS interface to be intuitive, approachable, and easy to use and synchronize to their smartphone or tablet. Many reported that they could complete the measurements within a few minutes, had turned the measurements into a routine (eg, completing measurements before dinner or after waking up), or had gamified the process to keep themselves motivated. Self-competition was seen as a prominent positive influence on adherence to use. However, others experienced technical difficulties when using the device and continued to experience technical difficulties when using the HS device over the study period. They reported issues with charging the battery, connectivity between devices, and issues with uploading data from the HS device to the server. The consequent failed attempts at recording measurements were frustrating for some participants.

Facilitating Conditions

Most reported that the HS device was easy to maintain and did not report problems with cleaning the mouthpiece and storing the HS device. We specifically asked about the portability of the HS device during interviews, and portability was found to be a facilitating condition for continued use. In fact, participants found the HS device to be portable enough that they reported taking it with them on vacations. On some occasions, participants who were hospitalized or very tired from treatment were not able to complete the HS measurements. This qualitative observation is also reflected in the survey, where participants rarely felt that their illness interfered with HS use.

Participants had suggestions for improvements to facilitate greater use of the device in the future. They suggested adding a trends feature in the app to monitor pulmonary health over time, reminders for weekly use, and positive reinforcements to encourage continued use. Some also suggested that a more comfortable nose clip would be helpful, and reminders to put on the clip before measuring lung function would also be appreciated.

Overall, thematic weaving of results from participants' survey responses and their answers to open-ended interview questions showed a good fit between qualitative and quantitative data and deepened our understanding of the findings. Qualitative findings confirmed the survey responses and provided further insight into the survey responses through specific examples or explanations offered in the interviews. Mixed methods integration showed no apparent relationship between participants' responses and their age or sex. It was not possible to infer a relationship between participants' engagement in their health care and their use and adherence to HS because of the limited variance in PAM scores in our sample.

Discussion

Principal Findings

The aim of this pilot feasibility study was to investigate the acceptability and usability of an HS program among patients who had undergone AHCT using mixed methods. Overall, we found that recipients of AHCT generally found routine, unsupervised, patient-performed HS to be acceptable, despite some technological barriers. Spirometer use was viewed favorably by participants, with enthusiasm for continued use. Although some barriers to performing HS were identified, most were amenable to improvement, and patients were forthcoming with solutions on how to alter their personal use of HS to overcome barriers. Some barriers were unique to this study and the specific devices used, such as periodic difficulty in transmitting data from the HS device to the server. These types of barriers may commonly occur when new technology is introduced in a patient care setting. The results of this study can inform the adoption and implementation of HS on a wider scale and potentially address the need for innovative remote patient monitoring technologies highlighted by the COVID-19 pandemic.

HS is a promising emerging technology that allows for longitudinal monitoring of lung function in high-risk patients. Ideal conditions for HS include screening for diseases that have variable times of onset or monitoring those with variable patterns of progression, particularly when clinic-based monitoring is not practical because of the necessary frequency of testing. This strategy has been successful in recipients of a lung transplant for the monitoring of chronic lung allograft dysfunction, usually in the form of BOS [7-10]. More recently, this strategy has been adopted for patients with idiopathic pulmonary fibrosis, a disease with variable tempos and patterns of progression [28]. Patients with pulmonary fibrosis value home monitoring of symptoms and lung function [29], and these remote monitoring modalities are particularly relevant during the current COVID-19 pandemic, where in-person testing may be infeasible or unsafe [30].

Prompt diagnosis and treatment of BOS are imperative to reduce mortality or permanent loss of lung function [1]. HS is an attractive option for monitoring recipients of AHCT at risk for BOS as BOS often occurs later in the course of a transplant when direct patient contact is less frequent [20,31]; symptoms may often be insidious, even in the setting of significant pulmonary impairment [32]; and the time to onset of BOS and trajectory of impairment after BOS onset is variable [1]. By using an HS program that leveraged cloud-based data delivery to allow participants to perform spirometry anywhere with a cellular connection, we removed major barriers to performing spirometry and interpreting results in a timely fashion [33]. Furthermore, pulmonary impairments could be recognized quickly, allowing for prompt coordination of clinical care. However, none of the patients in our study developed BOS. This may be as we studied recipients of AHCT who were approximately 100 days after hematopoietic cell transplant up to 1 year after hematopoietic cell transplant; BOS more typically occurs in the second year after AHCT, although earlier cases have a poorer prognosis [20]. In addition, we did not take advantage of enriching our cohort with high-risk cases, such as those with high-risk chronic GVHD or prior lung complications [34]. Turner et al [34] found that in an enriched cohort, approximately 25% of patients developed BOS. Therefore, HS initiation would likely be more effective if started later in the course of AHCT in high-risk recipients.

For the most part, participants confirmed that routine HS was feasible and straightforward to perform. This observation was validated by the fact that, on average, participants completed measurements at 69% of possible weeks during the study period, and 94% of weeks with measurements had at least one technically sound loop despite the lack of a face-to-face encounter with a respiratory therapist [22]. As highlighted in our qualitative interviews, participants were highly motivated to participate in our home spirometry program. However, as we noted in our prior work, participants often had urgent health issues that interfered with their participation in home spirometry, as we previously reported [22]. In addition, although we considered missed measurements as nonadherence, it is important to note that some of the observed nonadherence was also because of technical errors with data transmission. The mixed methods design, which combined longitudinal interviews with periodic follow-up contact by study staff, enabled us to identify these technical errors in a timely manner.

Many participants valued knowing their HS measurements as they provided insight into the rapidity of their disease progression. Implementation of immediate feedback could serve those participants who understood how to interpret spirometry data by providing positive feedback for engagement while easing frustrations for those who did not understand how to interpret spirometry data by explaining what these measurements mean in relation to their prior data. Participants who reported not being able to interpret their spirometry data suggested that more immediate feedback on daily measurements would be helpful in understanding progression. Our study did not implement the interpretation of pulmonary function tests in real time for patients, which could improve patients' understanding of their lung function. Interestingly, despite the lack of immediate feedback, our prior work suggested that technical proficiency was maintained or even slightly improved over the duration of the study [22]. Participants used these to set personal FEV_1 goals and to self-monitor. They also felt that extrinsic motivators would improve adherence, for example, feedback from the study team that their lungs were doing well. Including extrinsic patient motivators in the design of HS, and other remote biometric monitoring platforms, may be an important feature for increasing engagement and adherence. Certain factors were found to hinder spirometer use, such as technical issues with charging and synchronization with cloud data. These barriers to use did not dissipate over time. Efforts to address these technical issues should be a priority for investigators and organizations that facilitate the implementation of HS. Negative experiences with new technology may undermine patients' confidence, skill, and willingness to engage with systems such as HS [35].

Participants in our study reported high engagement in managing their own health and health care (as measured by PAM) with little variability, which limited our ability to study the

relationship between patient activation and HS adherence. We found no relationship between age or gender and the use of HS, potentially suggesting that neither of these factors is a barrier to the implementation of HS. We did not directly measure participants' experience with technology, and an additional investigation between comfort with technology use and HS adherence is warranted. However, the high level of acceptability of HS and patients' willingness to overcome usability barriers (barring technical issues such as transmission failure) may reflect participants' self-reported high levels of engagement. Patients with cancer have indicated a willingness to engage in remote monitoring when it potentially adds value to care [36,37].

Including end users in the technology development and modification process can enhance the potential for adoption and successful adherence to remote monitoring platforms. Participants in our study had suggestions for improving the experience of performing HS, including goal setting and feedback to ensure adherence of use. This type of feedback was not present in our study but easily implementable within the GoHome ecosystem in various forms, such as *CareTexts* directly to patients' phones or tablets. Such a strategy may be useful to further reduce the burden of HS on the monitoring team [38], and others have found that implementing SMS text messaging into home spirometry programs may increase adherence [39]. We found that once participants were trained in spirometry, remote data monitoring was not time intensive, requiring approximately 20 minutes/patient every month [40].

Although our work provides evidence that trial participants viewed HS favorably, it remains to be seen whether this would also be true for patients in settings outside of tertiary care cancer centers. Notable differences include the possibility that HS may not be universally covered by payers, which may reduce enthusiasm for monitoring; that access to experts in pulmonary complications of AHCT may not be readily accessible; and that no data exist as to whether prompt interventions after pulmonary decline using real-world HS result in reductions in the incidence of BOS. In particular, the latter highlights the need for a randomized controlled trial comparing early interventions using HS to screen for impairment with the usual standard of care as it is possible that (1) early interventions do not reduce the rates of subsequent BOS and (2) that early interventions expose trial participants to unnecessary procedures, and thereby, unnecessary risks and costs. Therefore, despite positive patient perceptions toward HS, more definitive studies are necessary to prove that HS has a clinical role in detecting and treating BOS after AHCT.

Our study contributes to a growing body of research on remote patient monitoring in oncology. Notably, remote monitoring of electronic PROs, specifically cancer treatment–related symptoms, has been associated with improved quality of life and increased survival when providers intervene in response to worsening symptoms for the prevention of adverse downstream outcomes [41]. However, few studies in oncology have evaluated remote monitoring of biometric outcomes (eg, HS, temperature, blood pressure, weight, saturation of peripheral oxygen, and activity), with or without electronic PROs, using noninvasive digital technology [42]. In contrast, there has been greater progress in evaluating remote monitoring of biometric outcomes and PROs in respiratory, cardiovascular, and metabolic diseases and weight management [43]. Findings from studies in chronic conditions other than cancer are consistent with ours; namely, the data feedback loop wherein remotely captured patient data are enacted upon by providers is critical to affecting patient health outcomes [43]. Patient engagement with remote monitoring, as reflected in adherence to use, may be a critical limitation with regard to the accuracy and fidelity of the collected data [42]. Thus, addressing the technological barriers identified in this study that may affect adherence may optimize the value of remote monitoring.

Strengths and Limitations

A major strength of this study was the use of a longitudinal mixed methods design using the same participants in repeated one-on-one qualitative interviews at multiple time points. This enabled us to understand how the feasibility and acceptability of HS use evolved over time. To the best of our knowledge, this is the first study to evaluate HS using this approach and the first to evaluate HS in patients with AHCT. Another strength is the use of a theoretical model to guide data analysis, thus allowing the sense-making process to be explicit. Finally, to our knowledge, this is the first study to investigate in depth the patient experience of participating in an HS program.

The results of this study should be taken in light of certain limitations. Our study population was predominantly White and English speaking; therefore, our results must be interpreted in the context of the limited cultural and racial diversity of the study cohort and may limit external validity to more diverse patient populations. In particular, although study devices were provided to all participants regardless of background, including providing tablets to participants without a smartphone, it is possible that further barriers to participation existed among participants of color. Participants generally had high PAM scores, which reduced our ability to associate patient activation with other measurements. The higher PAM scores may suggest that patients who were less likely to have higher levels of engagement in self-care declined study participation, although we cannot confirm this. However, we did not interview or measure PAM scores in patients who did not enroll in this study; therefore, we cannot comment on whether our cohort's perspectives are reflective of all recipients of AHCT. We did not directly measure technological or health literacy and, therefore, cannot ascertain whether patients with low technological literacy require more instruction or supervision early in the monitoring period to ensure future adherence. As we did not comprehensively analyze open-ended patient responses in real time, we were unable to suggest modifications to the commercial spirometer used in this study. Finally, study attrition limited our ability to perform longitudinal assessments completely.

Conclusions

In conclusion, we have established that patients found HS acceptable and easy to use. Simple technical and programmatic modifications were identified by the patients, which would improve the quality of HS implementation. Wider implementation of HS would benefit recipients of AHCT as well as other patients at risk for lung disease, particularly during crises such as the COVID-19 pandemic, where access to health

care may be limited for various reasons. Given the limited data on the feasibility and acceptability of remote patient monitoring in oncology, these findings offer support for future research aimed at integrating remote monitoring technology to improve patients' experiences and outcomes during acute cancer care. Future work is necessary to determine the efficacy of HS performed in the real world as a means of detecting and treating BOS and other pulmonary complications of AHCT.

Acknowledgments

This work was supported by the National Institutes of Health/National Institute of Allergy and Infectious Diseases (K23 AI117024 to A Sheshadri) and the National Cancer Institute (Cancer Center Support Grant, P30 CA016672) at the National Institutes of Health.

Conflicts of Interest

A Stenzler is the Chief Science Officer of Monitored Therapeutics, the supplier of the home spirometers.

References

1. Cheng GS, Storer B, Chien JW, Jagasia M, Hubbard JJ, Burns L, et al. Lung function trajectory in bronchiolitis obliterans syndrome after allogeneic hematopoietic cell transplant. Ann Am Thorac Soc 2016;13(11):1932-1939. [doi: 10.1513/annalsats.201604-262oc] [Medline: 27513368]
2. Ahn JH, Jo K, Song JW, Shim TS, Lee SW, Lee JS, et al. Prognostic role of FEV1 for survival in bronchiolitis obliterans syndrome after allogeneic hematopoietic stem cell transplantation. Clin Transplant 2015;29(12):1133-1139. [doi: 10.1111/ctr.12638] [Medline: 26383085]
3. Au BK, Au MA, Chien JW. Bronchiolitis obliterans syndrome epidemiology after allogeneic hematopoietic cell transplantation. Biol Blood Marrow Transplant 2011;17(7):1072-1078. [doi: 10.1016/j.bbmt.2010.11.018] [Medline: 21126596]
4. Holbro A, Lehmann T, Girsberger S, Stern M, Gambazzi F, Lardinois D, et al. Lung histology predicts outcome of bronchiolitis obliterans syndrome after hematopoietic stem cell transplantation. Biol Blood Marrow Transplant 2013;19(6):973-980. [doi: 10.1016/j.bbmt.2013.03.017] [Medline: 23562737]
5. Ryu JH, Myers JL, Swensen SJ. Bronchiolar disorders. Am J Respir Crit Care Med 2003;168(11):1277-1292. [doi: 10.1164/rccm.200301-053SO] [Medline: 14644923]
6. Kwok WC, Liang BM, Lui MM, Tam TC, Sim JP, Tse EW, et al. Rapid versus gradual lung function decline in bronchiolitis obliterans syndrome after haematopoietic stem cell transplantation is associated with survival outcome. Respirology 2019;24(5):459-466. [doi: 10.1111/resp.13472] [Medline: 30663178]
7. Bjørtuft O, Johansen B, Boe J, Foerster A, Holter E, Geiran O. Daily home spirometry facilitates early detection of rejection in single lung transplant recipients with emphysema. Eur Respir J 1993;6(5):705-708. [Medline: 8390943]
8. de Wall C, Sabine D, Gregor W, Mark G, Axel H, Thomas F, et al. Home spirometry as early detector of azithromycin refractory bronchiolitis obliterans syndrome in lung transplant recipients. Respir Med 2014;108(2):405-412. [doi: 10.1016/j.rmed.2013.12.016] [Medline: 24445061]
9. Morlion B, Knoop C, Paiva M, Estenne M. Internet-based home monitoring of pulmonary function after lung transplantation. Am J Respir Crit Care Med 2002;165(5):694-697. [doi: 10.1164/ajrccm.165.5.2107059] [Medline: 11874816]
10. Wang W, Finkelstein SM, Hertz MI. Automatic event detection in lung transplant recipients based on home monitoring of spirometry and symptoms. Telemed J E Health 2013;19(9):658-663. [doi: 10.1089/tmj.2012.0290] [Medline: 23869394]
11. Cheng GS, Campbell AP, Xie H, Stednick Z, Callais C, Leisenring WM, et al. Correlation and agreement of handheld spirometry with laboratory spirometry in allogeneic hematopoietic cell transplant recipients. Biol Blood Marrow Transplant 2016;22(5):925-931. [doi: 10.1016/j.bbmt.2015.12.023] [Medline: 26748162]
12. Guihot A, Becquemin MH, Couderc LJ, Randrianarivelo O, Rivaud E, Philippe B, et al. Telemetric monitoring of pulmonary function after allogeneic hematopoietic stem cell transplantation. Transplantation 2007;83(5):554-560. [doi: 10.1097/01.tp.0000228236.55419.33] [Medline: 17353773]
13. Loiseau C, Lemonnier F, Randrianarivelo O, Itzykson R, Nguyen S, Becquemin MH, et al. Home spirometry in bronchiolitis obliterans after allogeneic haematopoietic cell transplant. Eur Respir J 2018;52(1):1702328. [doi: 10.1183/13993003.02328-2017] [Medline: 29794130]
14. Abedin S, Yanik GA, Braun T, Pawarode A, Magenau J, Goldstein SC, et al. Predictive value of bronchiolitis obliterans syndrome stage 0p in chronic graft-versus-host disease of the lung. Biol Blood Marrow Transplant 2015;21(6):1127-1131. [doi: 10.1016/j.bbmt.2015.02.006] [Medline: 25687798]
15. Chien JW, Martin PJ, Gooley TA, Flowers ME, Heckbert SR, Nichols WG, et al. Airflow obstruction after myeloablative allogeneic hematopoietic stem cell transplantation. Am J Respir Crit Care Med 2003;168(2):208-214. [doi: 10.1164/rccm.200212-1468OC] [Medline: 12649126]

16. Jagasia MH, Greinix HT, Arora M, Williams KM, Wolff D, Cowen EW, et al. National Institutes of Health consensus development project on criteria for clinical trials in chronic graft-versus-host disease: I. The 2014 diagnosis and staging working group report. Biol Blood Marrow Transplant 2015;21(3):389-401.e1 [FREE Full text] [doi: 10.1016/j.bbmt.2014.12.001] [Medline: 25529383]
17. Kulkarni HS, Cherikh WS, Chambers DC, Garcia VC, Hachem RR, Kreisel D, et al. Bronchiolitis obliterans syndrome-free survival after lung transplantation: an International Society for Heart and Lung Transplantation Thoracic Transplant Registry analysis. J Heart Lung Transplant 2019;38(1):5-16 [FREE Full text] [doi: 10.1016/j.healun.2018.09.016] [Medline: 30391193]
18. Andrykowski MA. Psychosocial factors in bone marrow transplantation: a review and recommendations for research. Bone Marrow Transplant 1994;13(4):357-375. [Medline: 8019460]
19. Hacker ED, Fink AM, Peters T, Park C, Fantuzzi G, Rondelli D. Persistent fatigue in hematopoietic stem cell transplantation survivors. Cancer Nurs 2017;40(3):174-183. [doi: 10.1097/ncc.0000000000000405] [Medline: 27333126]
20. Vieira AG, Funke VA, Nunes EC, Frare R, Pasquini R. Bronchiolitis obliterans in patients undergoing allogeneic hematopoietic SCT. Bone Marrow Transplant 2014;49(6):812-817. [doi: 10.1038/bmt.2014.25] [Medline: 24614836]
21. Langford AT, Solid CA, Scott E, Lad M, Maayan E, Williams SK, et al. Mobile phone ownership, health apps, and tablet use in US adults with a self-reported history of hypertension: cross-sectional study. JMIR Mhealth Uhealth 2019;7(1):e12228. [doi: 10.2196/12228] [Medline: 31344667]
22. Sheshadri A, Alousi A, Bashoura L, Stolar K, Baghaie S, Arain MH, et al. Feasibility and reliability of home-based spirometry telemonitoring in allogeneic hematopoietic cell transplant recipients. Ann Am Thorac Soc 2020;17(10):1329-1333. [doi: 10.1513/AnnalsATS.202005-434RL] [Medline: 32614241]
23. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019;95:103208. [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
24. Moljord IE, Lara-Cabrera ML, Perestelo-Pérez L, Rivero-Santana A, Eriksen L, Linaker OM. Psychometric properties of the patient activation measure-13 among out-patients waiting for mental health treatment: a validation study in Norway. Patient Educ Couns 2015;98(11):1410-1417. [doi: 10.1016/j.pec.2015.06.009] [Medline: 26146239]
25. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425-478. [doi: 10.2307/30036540]
26. Williams MS, Rana NP, Dwiwedi YK. The unified theory of acceptance and use of technology (UTAUT): a literature review. J Enterp Inf Manag 2015;28(3):443-488. [doi: 10.1108/JEIM-09-2014-0088]
27. Fetters MD, Curry LA, Creswell JW. Achieving integration in mixed methods designs-principles and practices. Health Serv Res 2013;48(6 Pt 2):2134-2156. [doi: 10.1111/1475-6773.12117] [Medline: 24279835]
28. Russell AM, Adamali H, Molyneaux PL, Lukey PT, Marshall RP, Renzoni EA, et al. Daily home spirometry: an effective tool for detecting progression in idiopathic pulmonary fibrosis. Am J Respir Crit Care Med 2016;194(8):989-997. [doi: 10.1164/rccm.201511-2152OC] [Medline: 27089018]
29. Edwards C, Costello E, Cassidy N, Vick B, Russell AM. Use of the patientMpower app with home-based spirometry to monitor the symptoms and impact of fibrotic lung conditions: longitudinal observational study. JMIR Mhealth Uhealth 2020;8(11):e16158. [doi: 10.2196/16158] [Medline: 33216007]
30. Nakshbandi G, Moor CC, Wijsenbeek MS. Home monitoring for patients with ILD and the COVID-19 pandemic. Lancet Respir Med 2020;8(12):1172-1174. [doi: 10.1016/s2213-2600(20)30452-5] [Medline: 33075296]
31. Bergeron A, Chevret S, Peffault de Latour R, Chagnon K, de Margerie-Mellon C, Rivière F, et al. Noninfectious lung complications after allogeneic haematopoietic stem cell transplantation. Eur Respir J 2018;51(5):1702617. [doi: 10.1183/13993003.02617-2017] [Medline: 29650555]
32. Williams KM. How I treat bronchiolitis obliterans syndrome after hematopoietic stem cell transplantation. Blood 2017;129(4):448-455. [doi: 10.1182/blood-2016-08-693507] [Medline: 27856461]
33. Patrick K, Griswold WG, Raab F, Intille SS. Health and the mobile phone. Am J Prev Med 2008;35(2):177-181. [doi: 10.1016/j.amepre.2008.05.001] [Medline: 18550322]
34. Turner J, He Q, Baker K, Chung L, Lazarevic-Fogelquist A, Bethune D, et al. Home spirometry telemonitoring for early detection of bronchiolitis obliterans syndrome in patients with chronic graft-versus-host disease. Transplantat Cell Ther 2021;27(7):616.e1-616.e6. [doi: 10.1016/j.jtct.2021.03.024] [Medline: 33781975]
35. Lundell S, Modig M, Holmner Å, Wadell K. Perceptions of home telemonitoring use among patients with chronic obstructive pulmonary disease: qualitative study. JMIR Mhealth Uhealth 2020;8(6):e16343. [doi: 10.2196/16343] [Medline: 32490844]
36. Liao Y, Thompson C, Peterson S, Mandrola J, Beg MS. The future of wearable technologies and remote monitoring in health care. Am Soc Clin Oncol Educ Book 2019;39:115-121. [doi: 10.1200/edbk_238919] [Medline: 31099626]
37. Peterson SK, Shinn EH, Basen-Engquist K, Demark-Wahnefried W, Prokhorov AV, Baru C, et al. Identifying early dehydration risk with home-based sensors during radiation treatment: a feasibility study on patients with head and neck cancer. J Natl Cancer Inst Monogr 2013;2013(47):162-168. [doi: 10.1093/jncimonographs/lgt016] [Medline: 24395986]
38. Pangarakis SJ, Harrington K, Lindquist R, Peden-McAlpine C, Finkelstein S. Electronic feedback messages for home spirometry lung transplant recipients. Heart Lung 2008;37(4):299-307. [doi: 10.1016/j.hrtlng.2007.07.001] [Medline: 18620106]

39. Fadaizadeh L, Najafizadeh K, Shajareh E, Shafaghi S, Hosseini M, Heydari G. Home spirometry: assessment of patient compliance and satisfaction and its impact on early diagnosis of pulmonary symptoms in post-lung transplantation patients. J Telemed Telecare 2016;22(2):127-131. [doi: 10.1177/1357633x15587435] [Medline: 26026187]
40. Lindquist R, VanWormer A, Lindgren B, MacMahon K, Robiner W, Finkelstein S. Time-motion analysis of research nurse activities in a lung transplant home monitoring study. Prog Transplant 2011;21(3):190-199 [FREE Full text] [doi: 10.7182/prtr.21.3.u267v51327m276l2] [Medline: 21977879]
41. Basch E, Deal AM, Dueck AC, Scher HI, Kris MG, Hudis C, et al. Overall survival results of a trial assessing patient-reported outcomes for symptom monitoring during routine cancer treatment. JAMA 2017;318(2):197-198. [doi: 10.1001/jama.2017.7156] [Medline: 28586821]
42. McGregor B, Vidal GA, Shah SA, Mitchell JD, Hendifar AE. Remote oncology care: review of current technology and future directions. Cureus 2020;12(8):e10156 [FREE Full text] [doi: 10.7759/cureus.10156] [Medline: 33014652]
43. Vegesna A, Tran M, Angelaccio M, Arcona S. Remote patient monitoring via non-invasive digital technologies: a systematic review. Telemed J E Health 2017;23(1):3-17. [doi: 10.1089/tmj.2016.0051] [Medline: 27116181]

Abbreviations

AHCT: allogeneic hematopoietic cell transplantation
BOS: bronchiolitis obliterans syndrome
FEV_1: forced expiratory volume in 1 second
GVHD: graft-versus-host disease
HS: home-based spirometry
PAM: patient activation measure
PRO: patient-reported outcome
REDCap: Research Electronic Data Capture
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by A Mavragani; submitted 05.04.21; peer-reviewed by P Knapp, P Enright, J Edwards; comments to author 21.05.21; revised version received 06.07.21; accepted 10.12.21; published 07.02.22.

Please cite as:
Sheshadri A, Makhnoon S, Alousi AM, Bashoura L, Andrade R, Miller CJ, Stolar KR, Arain MH, Noor L, Balagani A, Jain A, Blanco D, Ortiz A, Taylor MS, Stenzler A, Mehta R, Popat UR, Hosing C, Ost DE, Champlin RE, Dickey BF, Peterson SK
Home-Based Spirometry Telemonitoring After Allogeneic Hematopoietic Cell Transplantation: Mixed Methods Evaluation of Acceptability and Usability
JMIR Form Res 2022;6(2):e29393
URL: https://formative.jmir.org/2022/2/e29393
doi:10.2196/29393
PMID:35129455

Original Paper

Early Development of a Virtual Coach for Healthy Coping Interventions in Type 2 Diabetes Mellitus: Validation Study

Giulia Bassi[1,2], PhD; Ivan Donadello[3], PhD; Silvia Gabrielli[2], PhD; Silvia Salcuni[1], PhD; Claudio Giuliano[2], PhD; Stefano Forti[2], PhD

[1]Department of Developmental and Socialization Psychology, University of Padova, Padova, Italy

[2]Centre Digital Health & Wellbeing, Fondazione Bruno Kessler, Trento, Italy

[3]KRDB Research Centre, Free University of Bozen-Bolzano, Bolzano, Italy

Corresponding Author:
Giulia Bassi, PhD
Department of Developmental and Socialization Psychology
University of Padova
Via Venezia 8
Padova, 35131
Italy
Phone: 39 3477334405
Email: giulia.bassi@phd.unipd.it

Abstract

Background: Mobile health solutions aimed at monitoring tasks among people with diabetes mellitus (DM) have been broadly applied. However, virtual coaches (VCs), embedded or not in mobile health, are considered valuable means of improving patients' health-related quality of life and ensuring adherence to self-care recommendations in diabetes management. Despite the growing need for effective, healthy coping digital interventions to support patients' self-care and self-management, the design of psychological digital interventions that are acceptable, usable, and engaging for the target users still represents the main challenge, especially from a psychosocial perspective.

Objective: This study primarily aims to test VC interventions based on psychoeducational and counseling approaches to support and promote healthy coping behaviors in adults with DM. As a preliminary study, university students have participated in it and have played the standardized patients' (SPs) role with the aim of improving the quality of the intervention protocol in terms of user acceptability, experience, and engagement. The accuracy of users' role-playing is further analyzed.

Methods: This preliminary study is based on the Obesity-Related Behavioral Intervention Trial model, with a specific focus on its early phases. The healthy coping intervention protocol was initially designed together with a team of psychologists following the main guidelines and recommendations for psychoeducational interventions for healthy coping in the context of DM. The protocol was refined with the support of 3 experts in the design of behavioral intervention technologies for mental health and well-being, who role-played 3 SPs' profiles receiving the virtual coaching intervention in a Wizard of Oz setting via WhatsApp. A refined version of the healthy coping protocol was then iteratively tested with a sample of 18 university students (mean age 23.61, SD 1.975 years) in a slightly different Wizard of Oz evaluation setting. Participants provided quantitative and qualitative postintervention feedback by reporting their experiences with the VC. Clustering techniques on the logged interactions and dialogs between the VC and users were collected and analyzed to identify additional refinements for future VC development.

Results: Both quantitative and qualitative analyses showed that the digital healthy coping intervention was perceived as supportive, motivating, and able to trigger self-reflection on coping strategies. Analyses of the logged dialogs showed that most of the participants accurately played the SPs' profile assigned, confirming the validity and usefulness of this testing approach in preliminary assessments of behavioral digital interventions and protocols.

Conclusions: This study outlined an original approach to the early development and iterative testing of digital healthy coping interventions for type 2 DM. Indeed, the intervention was well-accepted and proved its effectiveness in the definition and refinement of the initial protocol and of the user experience with a VC before directly involving real patients in its subsequent use and testing.

(JMIR Form Res 2022;6(2):e27500) doi:10.2196/27500

KEYWORDS

virtual coach; early development; type 2 diabetes mellitus; healthy coping; Wizard of Oz; ORBIT model; pilot study; mobile phone

Introduction

Background

Technology is revolutionizing both the self-management and self-care of people with diabetes mellitus (DM) [1] by developing digital solutions for the monitoring of blood glucose levels, diet, and physical exercise [2]. However, those based on fostering mental health and emotional symptoms have mainly been developed for the general population [2]. According to the American Association of Diabetes Educators (AADE) guidelines, people with DM can feel depressed, anxious, and stressed, and these emotions can affect the management of their disease [3]. Indeed, AADE guides on 7 areas regarding the healthy management of DM, such as physical activity, monitoring of blood glucose levels, healthy eating, taking prescribed medication, reducing risk behaviors, problem-solving skills, and healthy coping [3]. Altogether, these 7 areas have been included in the National Standards for Diabetes Self-Management Education, referring to the framework of patient-centered diabetes management [4]. The latter area, namely healthy coping, is related to psychological support, in which its key factors are the achievement of healthy coping strategies and the maintenance of a positive attitude toward the management of DM [3]. Therefore, reaching healthy coping skills is fundamental as these allow them to acquire healthy goals. In this regard, diabetes-related emotional distress can influence the motivation of people with DM to self-manage their disease; when this motivation diminishes, the clinical recommendations requested for better self-care are difficult to maintain [3]. Indeed, experiencing diabetes-related emotional distress can inhibit the self-care behaviors of people with DM. For example, studies have found that quality of life shows a significantly positive association with good metabolic control and a negative relationship with diabetes complications [5,6]. Moreover, several psychosocial factors have been identified as risk factors in reducing the capacity of people with DM to maintain metabolic control because of diminished treatment adherence [7]. These psychosocial factors refer to experiencing stressful events, symptoms of depression, family stress, low financial resources, low social support, and the use of maladaptive coping strategies [8,9]. Indeed, adopting and cultivating healthy coping skills allows these people to be more adherent to clinical recommendations, as well as in other areas of life, such as following a healthy diet and engaging in regular physical activity [3].

Virtual Coaches: Related Studies

Virtual coaches (VCs) are generally computer programs that simulate conversations with people by mimicking a human being [2]. More specifically, VCs have mainly been deployed for coaching behavioral strategies. VCs are often developed within messaging apps, websites, or mobile phone apps, and they can interact using various methods, such as SMS text messages, images, audio tracks, and video clips [2]. VCs in the field of DM are mostly focused on the monitoring of blood glucose levels, diet, and physical activity, such as Wellthy, developed within a smartphone-based app [10]. Interestingly, a recent systematic review identified only 2 VCs that were specifically designed for the monitoring of healthy management of DM. One of the VCs was developed for diabetes diagnosis [11], whereas the other VC comprised a variety of chronic diseases, including DM but specifically focused on psoriasis [12]. Monitoring tasks are the most common features of VC approaches for DM [13]. For example, the My Diabetes Coach program is an app-based embodied conversational agent named Laura developed to support diabetes self-management in a home setting over 12 months. This app focuses on several modules, such as the monitoring of blood glucose levels, healthy eating, taking medication, physical activity, and the importance of foot care. This app further includes a scale for assessing the health-related quality of life [14]; however, the app is not addressed to the improvement of this outcome, and the construct of healthy coping has not been considered in the development of Laura. Given this, it is worth noting the relevance of developing a VC to support people with DM in their healthy coping. In this regard, a VC has been established for older adults with type 2 DM (T2DM), using the voice speech of Google Home, taking into account the construct of healthy coping [15]. Individuals can interact with the VC through this voice speech, which is able to initially screen depression symptoms using the Patient Health Questionnaire-9 [16], and then respond to them with personalized suggestions regarding the monitoring and maintenance of blood glucose levels. In general, VC approaches are still unable to handle the emotional needs of individuals with DM [17] as their main goal is to deliver digital interventions without creating a personalized interaction between the VC and the user. Indeed, the only personalized feedback based on people's data is typically related to insulin dosage suggestions [18,19]. Furthermore, people with DM have certain expectations associated with app design, such as being engaging, incorporating activities recommended by the clinical guidelines (eg, AADE), and covering a broad range of content, including psychological and emotional support [20], by taking into consideration their motivational state of change, which is useful in delivering personalized psychological interventions (Transtheoretical Model of Change); therefore, more evidence is required. Most VCs are developed for smartphone-based apps rather than social media communication channels. In this latter case, for example, the use of Telegram, WhatsApp, or SMS text messaging [21] is widely used by the general population, and, indeed, people have already integrated them into their daily lives. On the other hand, mobile health (mHealth) apps present more issues, such as high dropout rates. For example, two-thirds of people who downloaded an mHealth app used it only once [22]. Indeed, the change of habitual behavior requires a person to invest a large amount of time and thus, benefit from this technology for a sufficient period to incorporate them into their daily lives [22]. Overall, the psychological support that VC approaches offer to the user still represents important challenges

for VC solutions. More specifically, the design of psychological digital interventions that are acceptable, usable, and engaging for the target user still represents the main challenge. Finally, the comprehension of emotions, feelings, tone, irony, sarcasm, and therefore the interpretation of the intended user meaning has proved to be a pivotal challenge in the development of a VC [23].

The Design of VCs: The Wizard of Oz Methodology

The development of a robust, natural, and personalized VC is necessary to achieve a wide-ranging, conversion-specific protocol, which reflects the users' needs and preferences in the intervention dialogs with the VC [24-28]. The typical approach is to collect primary stimuli from these resources, such as Wizard of Oz (WOZ), refine them with the users' interactions, validate them with some experts in the field, and then repeat the procedure to reach a protocol of the intervention, which is as acceptable and as effective as possible [24-27,29,30]. Initially, the prototyped intervention should be sufficiently straightforward to ensure an easy evaluation of the user experience (UX) and its refinement to fulfill users' needs [31]. Therefore, the WOZ methodology requires three fundamental elements: first, a script or protocol, which guides what should happen during the interaction; second, the presence of participants who play the role of the final user, thereby using the role-playing technique; and finally, a human operator, called Wizard, who will be the one to perform the virtual coaching tasks in a simulated scenario, whereas at the final stage, these tasks will be automated [32]. The WOZ method is a process that allows users to interact with a counterpart without knowing that the answers are generated by a human being (ie, also called human-controlled) rather than a computer. This deceit aims to let the users experience a greater freedom of expression [32]. A previous study that used the WOZ methodology for the design of a VC for stress detection among older adults showed how this method allows the development of a more natural dialog flow, as well as the importance of the users' needs in designing the intervention [33]. Therefore, one can assume that a VC can offer a convenient, engaging way of providing psychological support to people anytime and anywhere.

The Design of VCs: Standardized Patients Approach

In designing a VC, the WOZ method assumes the presence of people who apply the role-playing technique, which can be combined with the standardized patient (SP) approach. SPs are actors trained to play the role of a patient, thereby simulating a problem in a clinically relevant and realistic way [31,34]. In particular, this approach is useful when an intervention is not yet mature, and thus, testing the protocol in different scenarios of simulation can be more useful and quicker than involving real patients, who are vulnerable user groups. The intention is not to use the representation of an SP to replace an actual encounter with a real patient but to supplement it in an integrative and standardized method [34,35].

To our knowledge, this is the first study that uses the WOZ and SP approaches in the early development of a VC for psychological support among people with DM. In this regard, the digital health technologies available in the literature, whose main aim is to support people in adopting and cultivating positive coping strategies, are still in their infancy. However, the progress of technology has allowed the opportunity for the development of automated support, such as VC approaches.

Objective

The main objective of this study is to develop VC interventions based on psychoeducational and counseling approaches to support and promote healthy coping behaviors in people with T2DM.

In particular, in the early development of VC, the specific objectives are as follows:

1. To develop an initial healthy coping intervention protocol for psychological digital interventions delivered by a VC to people with T2DM
2. To validate the conversational protocol with mental health and behavioral intervention technologies (BITs) experts
3. To test the prototype of VC and thus the intervention protocol with participants who play the role of an individual with T2DM through the use of the WOZ methodology
4. To adapt and refine the VC based on the results that emerge in this study regarding acceptability, usability, and UX
5. To assess user engagement and the accuracy of the users' role-playing by means of clustering techniques applied to the logged conversations between WOZ and participants

The overall data will be useful in future work to convert the WOZ into a conversational agent developed with an automatic VC implemented on neural networks. The VC will be prepared for a proof-of-concept study with patients with DM [36] by making the dialogs as natural and effective as possible to future target users.

Participant Recruitment

Participants in the *Define* phase were selected for the development of the intervention protocol. The only criterion for eligibility was the presence of a team of psychologists.

Participants in the *Refine* phase were 3 experts in mental health and BITs.

Each team works at the Digital Health Lab, the Fondazione Bruno Kessler (Italy), and the Department of Developmental Psychology and Socialization of the University of Padova (Italy).

Participants in the *Refine* phase were 20 psychology students recruited from the University of Padova (Italy). Of the 20 students, 2 (10%) dropped out from the study; thus, the final sample comprised 18 (90%) psychology students (14/20, 78% women), with ages ranging from 19 to 28 (mean age 23.61, SD 1.975) years, who played each 1 out of the 6 SP roles assigned. The participants were randomly recruited by telematics means; that is, by email to the university students' mailing list. The inclusion criteria for the study were (1) being a student in psychology, as they were expected to be better able to empathize with the characteristics, feelings, and needs of the SP profiles assigned and (2) being familiar with smartphones and the WhatsApp messaging channel.

Methods

Study Design

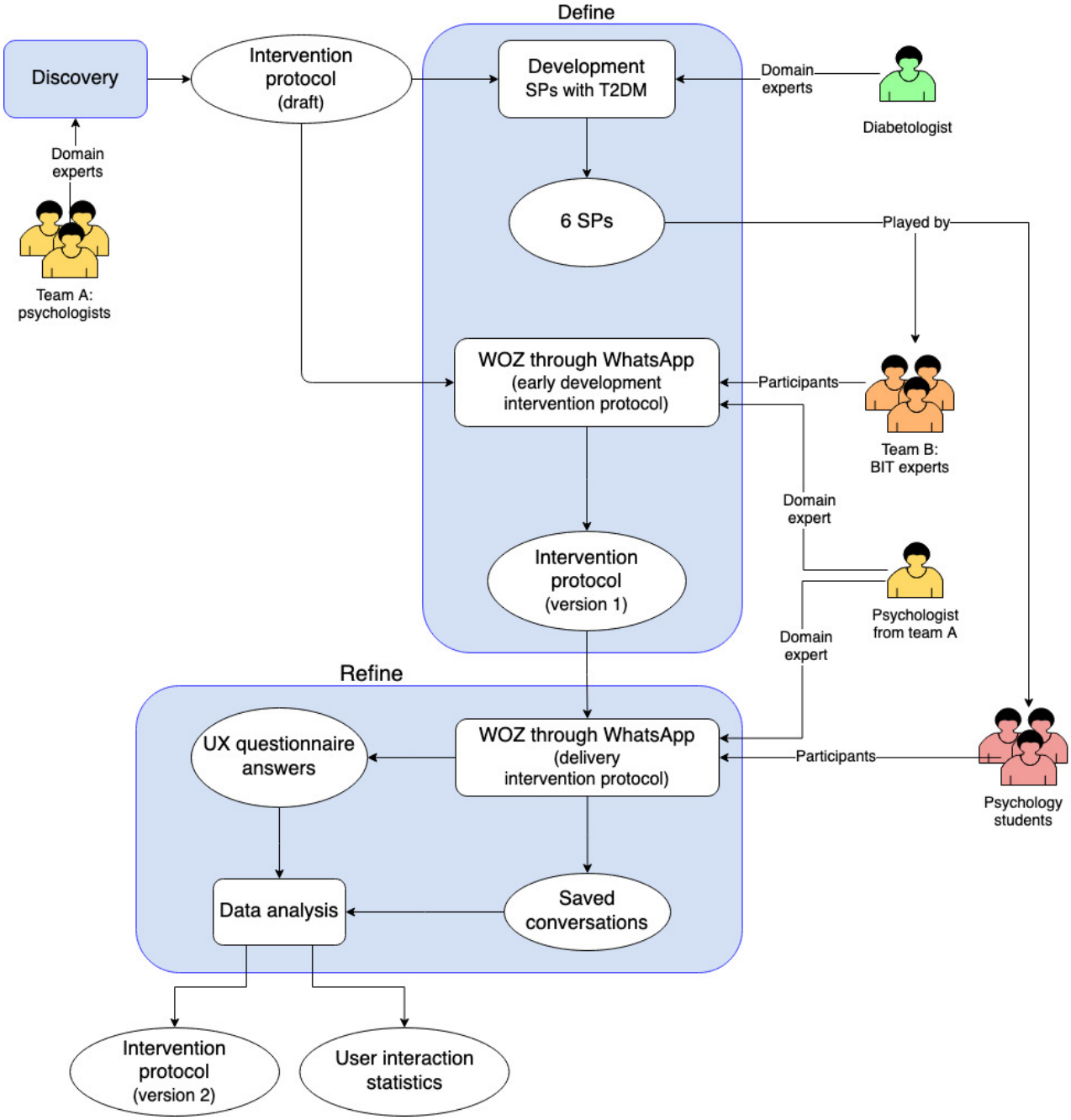
Overview

The design method of this study was 2-fold, and the architecture of the study design is shown in Figure 1.

First, we used the WOZ method, in which the users thought they were interacting with a VC but were interacting with a human being. This deceit allows the participants to express a language that is as natural as possible.

Second, we used the Obesity-Related Behavioral Interventions Trials (ORBIT) model as a guide for the development of this study. The ORBIT model has been developed for behavioral treatment to prevent or manage chronic diseases, such as DM [37,38]. Indeed, for this study, ORBIT allowed for guiding the development of healthy coping interventions by setting clear objectives for each phase of the model, thus enabling the revision of previous phases to improve the intervention in line with the new findings [37]. As an early development and validation study, we followed the first 3 phases of the ORBIT model, which are described in the following sections.

Figure 1. The study design architecture developed in this study. Rounded rectangles represent the architecture components or performed activities, whereas the ellipsis represents the input and output of each activity. SP: standardized patient; T2DM: type 2 diabetes mellitus; WOZ: Wizard of Oz; BIT: behavioral intervention technology; UX: user experience.



Phase 1: Discovery

Overview

As seen in Figure 1, the first phase of our study design was the development of the first draft of the intervention protocol. This process involved background research in behavioral and cognitive psychology, health psychology, the cornerstones of motivational interviewing, and the field of human-computer interaction, with a specific focus on the health care domain. A recent review reported that, in the overall studies, psychological support is underresourced and sometimes inadequate, thereby leading to poor health-related quality of life and decreasing the general well-being of people with DM [39]. Therefore, the main goal of this study was to develop an intervention protocol as part of the implementation of a VC (named Motibot, the abbreviation for Motivational bot) in a messaging system for psychological support to overcome the associated barriers of nonadherence in the context of DM. In this regard, a team of psychologists (team A) developed an intervention protocol based on counseling and psychoeducational approaches focusing on the healthy coping construct. This protocol comprised 2 sessions per week for an overall length of 6 weeks. At the beginning of each session, the VC delivers motivational and behavioral support to improve individuals' motivation to change their behaviors by adopting healthy coping strategies to handle the struggle caused by diabetes-related psychological symptoms. More specifically, motivational interventions are based on dialogs aimed at increasing individuals' awareness of their emotions and on the costs and benefits of adopting a healthy behavior, favoring their psychophysical well-being. Behavioral support is based on video clips or audio tracks inspired by positive psychology interventions and mindfulness practices. Mindfulness exercises are related to mindfulness-based cognitive therapy [40]. In this scenario, the support of VC to the user is not intended as a therapy but rather as psychoeducational support to reach, adopt, and cultivate healthy coping strategies for better diabetes management.

Phase 1a: Define

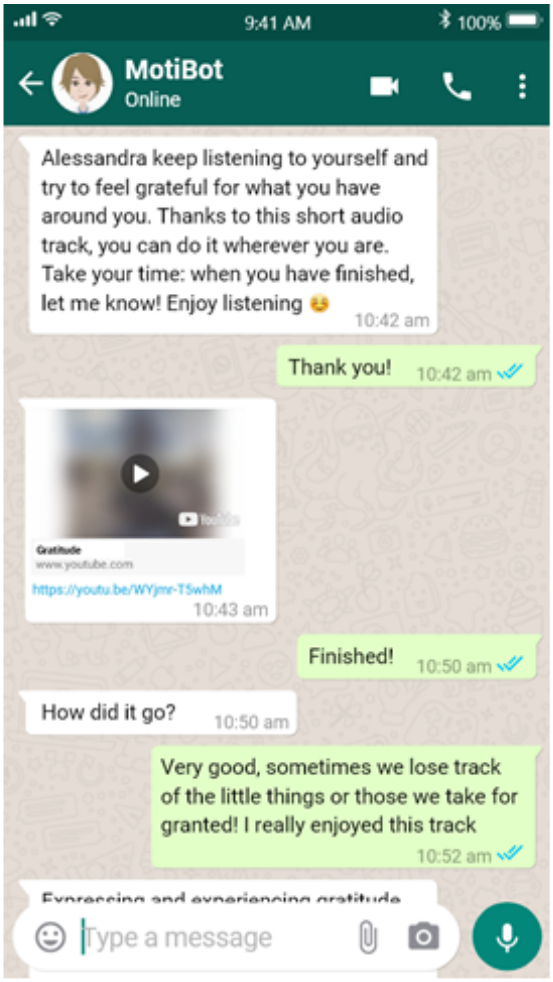
The *Define* phase represents the next step in our study design architecture. Here, the early development of the intervention protocol was performed by involving 3 mental health and BITs experts (team B), who knew who the WOZ was (from team A). These experts interacted with the WOZ by playing the role of 50% (3/6) developed SP profiles using the WhatsApp messaging channel. These 6 profiles were created in collaboration with a diabetologist of the Padova Hospital (Italy), and they represented the profiles of 3 (50%) typical male and 3 (50%) typical female patients with T2DM, which was suitable for every age (ie, youths, adults, and older adults) as they are the target users of the VC. Examples of SPs are shown in the Multimedia Appendix 1. The experts were consulted to define the intervention protocol regarding (1) the psychological approaches targeting DM, (2) the psychoeducational intervention, and (3) the UX assessment tools; indeed, an open-ended question was included to capture additional thematic content related to the UX. The team of experts defined and refined the intervention protocol following an iterative process, including recent scientific evidence based on cognitive behavioral interventions. Therefore, the intervention protocol (version 1) was defined for the next phase.

Phase 1b: Refine

In this phase, as a preliminary test, the WOZ delivered the intervention protocol to psychology students through the WhatsApp messaging channel. The intervention protocol was implemented for 6 weeks, with 2 sessions per week. This phase represents the initial pilot testing of the previously designed protocol. First, participants were asked to interpret and simulate an SP with T2DM by using the role-playing technique during the testing phase. Notably, 6 SP profiles were randomly assigned to the participants. Before starting the intervention, participants were told that they would interact with a VC. In fact, they interacted with a human being called WOZ (from team A). At the end of the second, fourth, and sixth weeks, users filled in the UX questionnaire, in which the items were adapted from the original UX questionnaire [40]. Qualitative data were further collected from the users' answers to the open-ended question, "I'm asking you to express your opinions regarding your experience with me..." At the end of the present phase, the intervention protocol (version 2) was refined based on the users' answers regarding their experience with the VC.

Finally, all the conversations between WOZ and users were saved, and subsequently, the data analysis stage allowed for processing of the intermediate results that emerged in the study (ie, to assess the user interaction statistics). This work will allow the training of a conversational agent to play the role of the simulated VC. An example of the interaction between the WOZ and a user is shown in Figure 2.

Figure 2. Example of the interaction between WOZ and a user within the WhatsApp application.



Measures

To accomplish the purposes of this study, we administered the following tools described in the following paragraphs.

The UXQuestionnaire was developed from the original version of the UX Questionnaire [41] to make the bipolar adjectives in line with the goals of this study. In particular, it includes 26 adjectives, either positive or negative, to assess the experience of interacting with the VC. Each item was rated on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). At the end of the questionnaire, an open-ended qualitative question was added and formulated as follows: “I’m asking you to express your opinions regarding your experience with me...” This latter question allows users to freely express their thoughts about their experiences with the VC. Textbox 1 shows the selection of items for this study.

The Mood Rating Scale was administered to support them in being aware of their emotions using the following question: “What are you feeling at this precise moment?” They were further asked to indicate the intensity of their emotions on a 5-point Likert scale, from 1 (low) to 5 (high).

Data collection and the relative timing during the 6-week intervention are reported in Textbox 2.

Textbox 1. Positive and negative items of the user experience questionnaire.

Positive items

- Pleasant
- Profound
- Cordial
- Comprehensible language
- Empathetic
- Attentive
- Motivating
- Encouraging
- Supportive
- Trustworthy
- Flexible
- Interesting

Negative items

- Annoying
- Not reliable
- Unappealing
- Unclear
- Complicated
- Not efficient
- Too much information
- Dissuading; not stimulating
- Not engaging
- Unpredictable
- Not reflective
- Conventional
- Rigid

Textbox 2. Tools administered and their timing.

User experience questionnaire

- At the end of the second, fourth, and sixth week

Mood Rating Scale

- At the beginning of each dialog session (2 per week) between the Wizard of Oz and the participants

Procedure

Over the 6 weeks of the study, participants were free to interact at any time with the WOZ, who was proactive in starting each session (typically at 10 AM). The UX questionnaire was developed in Google Forms, which was sent via a link through WhatsApp. The study procedure was conducted in compliance with the Declaration of Helsinki (Italian law 196/2003, European Union General Data Protection Regulation 679/2016) and was approved by the ethical committee (3518; April 1, 2020) for psychological research at the University of Padova (Italy). All participants signed and returned the informed consent document before participation and after revealing the WOZ identity at the end of the study.

Data Analyses

Statistical analyses were performed using R (version 64; R Foundation for Statistical Computing) [42], SPSS Statistics (version 24.0; IBM Corp) [43], and Python (version 3.6.6; Python Software Foundation) [44].

We used a mixed methods approach, which combines quantitative and qualitative analyses.

The Shapiro-Wilk test was performed to assess the normal distribution of the answers to the items related to the UX questionnaire.

Bartlett test was further conducted to evaluate homoscedasticity among the participants. The assumption of homoscedasticity, which means the same variance, is the main aspect of linear regression models [45]. Homoscedasticity explains a situation in which the error term (defined as the noise or random interference between the independent variables and the dependent variable) is equal across the overall values of the independent variables [46].

The first 2 objectives of this study (ie, the development of an intervention protocol for healthy coping in DM and its validation with mental health and BITs experts) were performed during the *Discovery* and *Define* phases, respectively. The third objective (ie, intervention protocol testing using the WOZ method among university students), which refers to the *Refine* phase, allowed us to further achieve a fourth objective: the evaluation of user acceptability, usability, and UX. Therefore, the statistical analyses performed to reach this goal were formulated as given in the following paragraphs.

The main *descriptive statistics* (ie, means and SDs) were conducted regarding the following:

1. Descriptive statistics were used for the number of utterances and their length (ie, the number of characters) expressed by both the VC (ie, WOZ) and users. Each utterance is a user's sentence sent via WhatsApp. These utterances are provided as a text file by WhatsApp with the
   Save conversation
   function. Each line of this file contains an utterance, its time stamp, and the name of the relative user. These descriptive statistics are necessary to highlight the differences between the VC and user messages to further understand whether the VC is too informative or too demanding from users. The users' response time (evaluated in minutes) was also assessed to understand the acceptability of the VC intervention.
2. The item response distributions among the 26 items of the UX questionnaire were performed to understand the users' experience with the VC every fortnight (ie, second, fourth, and sixth weeks).

A series of 1-way repeated measuresanalysis of variance was conducted to explore whether there were significant differences in each item's means of the UX questionnaire every fortnight (ie, second, fourth, and sixth weeks). The significance level α=.05 was Bonferroni corrected to α=.0025.

Qualitative analyses were conducted by referring to the analysis of the user's immediate feedback on their experience with the VC, collected at the end of the UX questionnaire. Qualitative contents were analyzed using the thematic analysis method developed by Braun and Clarke [47], thereby following their 6 phases. The approach is inductive and experiential in its orientations and is essential in its theoretical framework. First, the authors read and reread the users' answers to become familiar with the data and thus identify potential themes. Second, 2 authors (ie, GB and SS) analyzed and coded the thematic content, identifying labels for each theme relevant to the aim of this study. Third, the analysis conducted by the aforementioned authors allowed the searching of themes regarding the users' quotes, which was consistent with the scope of this study. Fourth and fifth, the authors (GB and SS) proceeded with the quality checking of themes related to the coded data and the entire data set. Finally, they reviewed the qualitative responses to be included in the final report.

On the other hand, the statistical analyses performed to reach the fifth and, thus, the last objective, namely the assessment of users' engagement and the accuracy of the users' role-playing, were formulated as given in the following paragraphs.

The *Augmented Dickey-Fuller* test was used to check whether the length (ie, the number of characters) of the users' answers or utterances and their response times presented a trend according to the dialog steps. If the trend was not detected, users were reasonably engaged in conversations with the VC. A decreasing trend in the length of the users' answers or an increasing trend in response times can be interpreted as a signal of disengagement.

Clustering techniques (ie, unsupervised learning based on the K-means algorithm) were used to assess the accuracy of the users' role-playing. Here, K-means is applied to vectors of real numbers gathered from the embedding of user utterances. Sentence embeddings are natural language processing techniques that can map natural language sentences or utterances into vectors of real numbers to preserve their semantic distance. Therefore, if 2 user utterances have a similar meaning, the Euclidean distance of the corresponding embedding vectors will be very low. The embeddings are computed with the Italian version of an off-the-shelf trained neural network called the universal sentence encoder [46]. This neural network is state of the art in sentence embedding. The results are based on the Fowlkes-Mallows Index (FMI), in which higher values indicate greater similarity between the computed clusters and the true assignment users—SPs. A great similarity means that users' utterances are in line with the given SP; that is, they correctly performed the role-playing. Therefore, the harvested dialogs can be automatically clustered according to their role, and consequently, they represent a good starting point for training a VC based on neural networks.

Results

Overall, the users' answers showed a nonnormal distribution. There were no missing data, and every participant answered all the questions presented in the UX questionnaire.

Descriptive Analyses

VC and Users: Utterances and Response Time

Table 1 shows the number of VC and user utterances along with the number of characters per utterance and the users' response time for each of the 18 dialogs.

With regard to the number of utterances, the VC expresses twice as many utterances compared with that of the users, as shown

by the means and 50%. In addition, the length of the VC messages (ie, number of characters per utterance) was more than double that of the users. This is aligned with the original intention of co-designing a virtual coaching intervention for healthy coping. In particular, these findings are in line with the role of the VC, which is proactive and guides the user in self-reflecting upon their emotions to learn how to adopt healthy coping strategies in dealing with negative thoughts and feelings.

Regarding the users' response time, half of the users answered on average within 15 minutes, whereas some users only answered within 2.33 minutes and others within 81 minutes. These results indicate a good acceptance of the VC intervention and commitment in the study, which is also confirmed by the low number of dropouts from the study (ie, 2/20, 10% users).

Table 1. Descriptive analyses of the utterances and response times within the dialogs.

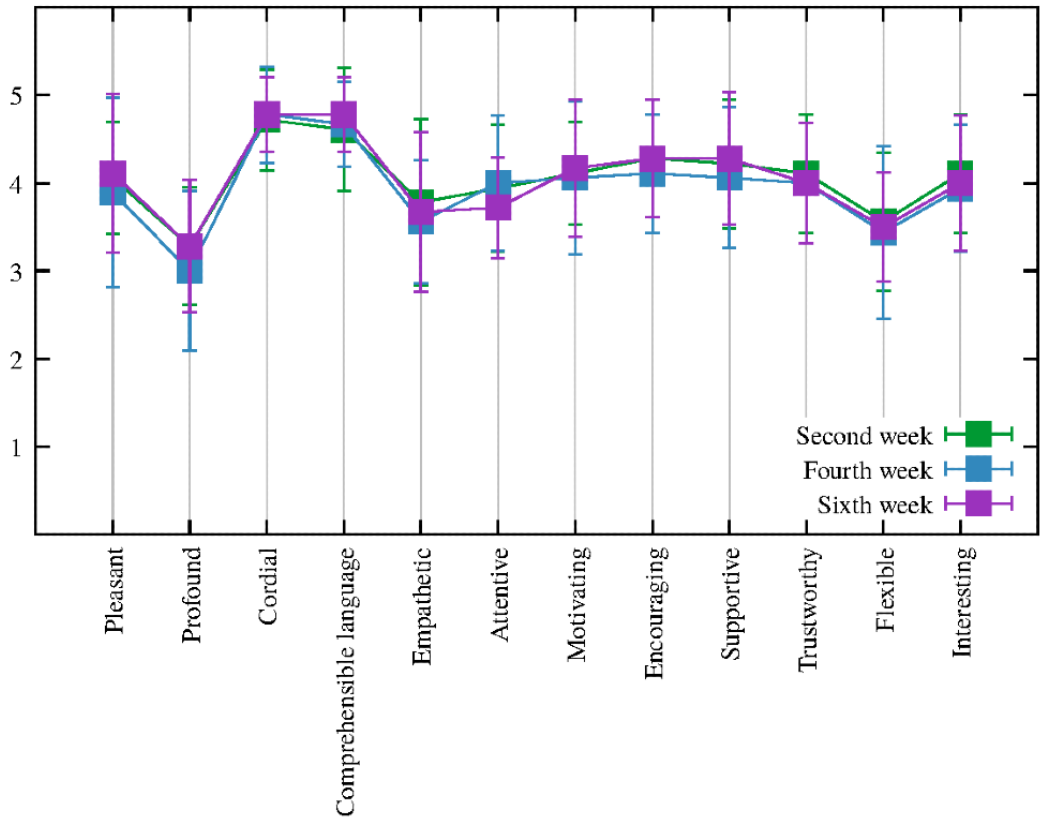
Measure	Number of VC[a] utterances	Number of users' utterances	Number of VC characters per utterance	Number of users' characters per utterance	Users' response time (minutes)
Values, mean (SD)	244.39 (19.05)	122.5 (17.7)	86.23 (3.14)	27.98 (9.7)	27.57 (27.20)
Values, minimum	219	95	79.91	13.06	2.23
25th percentile	231.25	111.5	84.4	18.9	5.24
50th percentile	240	120.5	85.96	30.31	15.45
75th percentile	257.25	130	87.74	35.45	48.28
Values, maximum	283	174	93.61	41.55	81.55

[a]VC: virtual coach.

UX Questionnaire: Positive Item Response Distributions

As displayed in Figure 3, overall, the positive items from the UX assessment showed a mean >3 on a 5-point Likert scale (mean 4.02, SD 0.72). In particular, the items *comprehensible language*, *motivating*, *encouraging*, and *supportive* slightly increased every fortnight week (second, fourth, and sixth week). The precise means and SDs are reported in Multimedia Appendix 2 (Table S1).

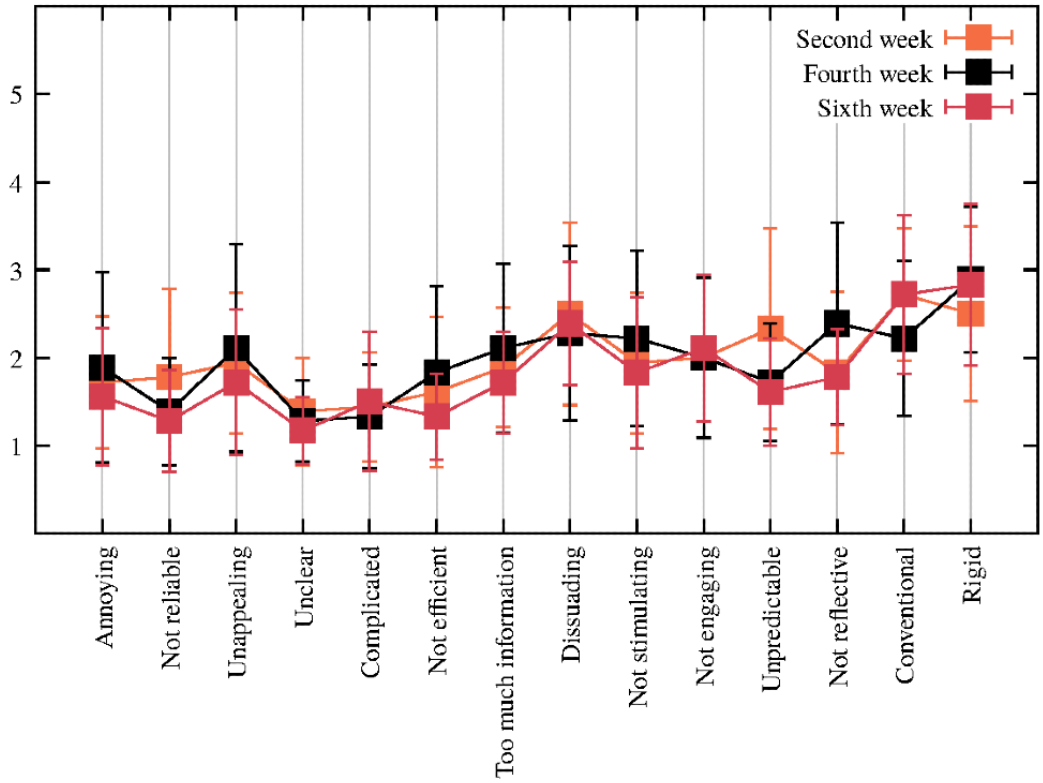
Figure 3. Plot of the positive items every fortnight week (second, fourth, and sixth week). Square dots and error bars correspond to means and SDs, respectively (N=18).



UX Questionnaire: Negative Item Response Distributions

Overall, the mean of the negative items of the UX questionnaire was in the expected direction (mean 1.92, SD 0.81), whereby all the items decreased with the progress of the intervention, as shown in Figure 4. Of particular note is the item *not reflective*, which decreases in week 6 compared with weeks 2 and 4. Moreover, the item *unpredictable* had a decreasing trend over the fortnight weeks, whereas the item *conventional* increased. Finally, the item *rigid* slightly increased in week 4 and slightly decreased in week 6. Again, the precise means and SDs are reported in Multimedia Appendix 2 (Table S2).

Figure 4. Plot of the negative items every fortnight week (second, fourth, and sixth week). Square dots and error bars correspond to means and SDs, respectively (N=18).



Analysis of Variance Repeated Measure: Differences in the UX Answers

A series of 1-way repeated measure analysis of variance with Bonferroni corrections were performed to compare items of the UX questionnaire on the second, fourth, and sixth week. The results showed no significant differences between the variables. The only significant differences were found for 2 negative items, *not efficient* ($F_{1,652}$=3.752; P=.03) and *too much information* ($F_{1,906}$=3.974; P=.03), in which their means decreased during the fortnight week. These differences were detected as significant by post hoc tests.

UX From a Qualitative Perspective

Table 2 shows the qualitative data gathered from participants' free text answers to the UX questionnaire, sent at the end of the second, fourth, and sixth weeks. This process allowed us to better understand their interaction with the VC and thus their UX. Two main categories were identified by two judges (GB and SS), resulting in (1) *insight themes* comprising the subcategories of *pleasant*, *natural*, *supportive*, *stimulating*, and *self-reflective* and (2) *challenge themes* comprising the subcategories of *repetitiveness* and *restrictive*.

Table 2. Users' quotes to the open-ended question provided at the end of the user experience questionnaire for each week (second, fourth, and sixth week).

Weeks and users' quotes	Insight themes	Challenge themes
Week 2		—[a]
"I think it is a pleasant and useful reminder to think about your mood and emotional aspects." [Participant 3]	Self-reflective	
"I find the experience stimulating." [Participant 4]	Stimulating	
"I find this new approach very interesting because without being intrusive, it enables you to have the right support during the day and proposes stimulating activities." [Participant 7]	Supportive	
"The messages stimulate a lot of self-reflection in the present moment. The responses to the messages seem very much in line with what I wrote." [Participant 14]	Natural	
Week 4		
"I am enjoying the interaction. It is not too forced and often the exchange is pleasant." [Participant 7]	Pleasant	—
"I found it to make very stimulating proposals and always try to stimulate a person even when they are not so inclined to undertake a certain activity." [Participant 9]	Stimulating	—
"Interesting but sometimes a little repetitive in the advice and suggestions." [Participant 10]	—	Repetitiveness
"It seems to respond adequately to my answers. When I answer more articulately its answer is often in line with what I wrote." [Participant 14]	Natural	—
"I think this is a good way to support. If you follow the suggestions consistently, I think it can be very helpful." [Participant 15]	Supportive	—
Week 6		
"Motibot offers many interesting and not trivial ideas on how to cultivate your well-being. I think it can be very useful, especially for those who do not know this area. Moreover, beyond the content of the messages, I think that having a regular appointment during which you have to stop for 10/15 minutes and think, can have many positive implications. The only criticism, which comes to my mind, and which is perhaps intrinsically linked to the origin of Motibot, is that its 'sensitivity' to the answers of the writer could be improved because sometimes one gets the impression that it has not 'understood.' Despite this, however, since it is a virtual entity, I believe that it performs its function correctly." [Participant 1]	—	Restrictive
"This path with Motibot has been very positive and useful to take an optimistic, comprehensive, and effective perspective; besides listening to me more, managing my emotions better and taking care of myself as well as others. As a result, I understood how to achieve greater well-being, paying attention to the present moment, and getting in tune with the world." [Participant 6]	Self-reflective	—
"I found the proposed activities very interesting and useful to practice; I think it is good support as it can motivate and encourage. Maybe sometimes I found some proposals and related explanations a bit repetitive among them, but overall, the interaction was very pleasant." [Participant 10]	—	Repetitiveness
"It was very stimulating." [Participant 12]	Stimulating	—
"This bot can help, and support people diagnosed with diabetes. The answers are very much in line with what I wrote. If you write a long and articulate message, the bot will understand if that message is positive or negative and will respond accordingly. Only a few times, it happened that the answers seemed a little out of place. The exercises that are proposed to you are easy to understand and do not require too much time. In my opinion, this allows the person not to interrupt the path, and to follow the advice without too many problems in their working day and not." [Participant 14]	Natural	—
"It was a positive experience, gradually I felt more and more involved and motivated to listen to the tracks and the proactive suggestions." [Participant 17]	Supportive	—

[a]Not available.

The Augmented Dickey-Fuller Test as an Insight for User Engagement

Table 3 displays the results of the Augmented Dickey-Fuller tests for each user. A very low test statistic indicates that the average number of users' characters of their utterances (ie, second column) or their response times (ie, fourth column) does not change significantly according to the dialog steps. Users did not tend to give shorter utterances and their response times did not increase, thereby demonstrating that the engagement of the users remained stable. These findings show that all users, except 1, engaged in interaction with the VC.

Regarding the average number of characters per user utterance, the test presents a low statistic for all users, except for user 14. This later showed a lower critical value of −3.49 at 1%; indeed,

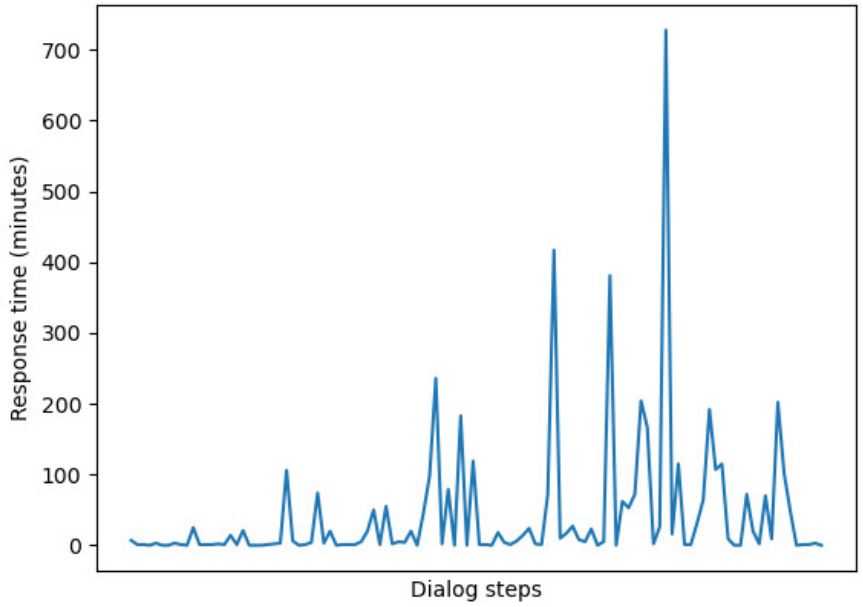
this user wrote longer sentences in the last utterances with the VC, thereby assuming a higher engagement with it.

With regard to the average response time per user, the test has good results for all users, except for users 5, 11, and 15 (critical value of −3.5% at 1% for all). During the last interactions with the VC, users 11 and 15 received longer behavioral interventions (ie, audio tracks and video clips), which required more time to understand the content and, consequently, to react to it. Therefore, even if these cases present a slight increasing trend in response time, they cannot be properly addressed as disengagement. Indeed, in these cases, users may need more time to reflect and answer the VC. In addition, their test statistics were quite close to the critical value, and the *P* value was <.05. At the end of the study, user 5 replied with a longer time, remaining as the one who presented a longer response time. This trend can be easily verified by inspecting the plot of the time series of the response times, as shown in Figure 5. Notably, the increasing trend starts from the central part of the dialog. Therefore, all users, except 1, tended to be engaged in answering with a nonincreasing response time. The intent was not to have a quick response time but to explore whether there was a response time with a nonincreasing trend.

Table 3. The statistical test and corresponding *P* values for the Augmented Dickey-Fuller test regarding the average number of characters and the average response time per user.

User	Number characters, mean		Response time, mean	
	Test statistics	*P* value	Test statistics	*P* value
1	−8.63	5.96×10^{-14}	−9.55	2.65×10^{-16}
2	−9.42	5.57×10^{-16}	−10.22	5.30×10^{-18}
3	−10.12	9.68×10^{-18}	−9.57	2.25×10^{-16}
4	−9.44	4.89×10^{-16}	−11.04	5.50×10^{-20}
5	−8.97	7.85×10^{-15}	−1.94	.31
6	−11.18	2.52×10^{-20}	−10.8	2.01×10^{-19}
7	−11.66	1.89×10^{-21}	−10.77	2.40×10^{-19}
8	−10.22	5.35×10^{-18}	−11.63	2.23E–21
9	−10.31	3.22×10^{-18}	−10.28	3.78E–18
10	−9.81	5.68×10^{-17}	−12.33	6.40E–23
11	−10.6	6.26×10^{-19}	−3.36	.01
12	−6.77	2.71×10^{-09}	−9.49	3.59×10^{-16}
13	−7.12	3.84×10^{-10}	−12.1	1.99×10^{-22}
14	−3.24	.02	−10.88	1.30×10^{-19}
15	−5.76	5.63×10^{-07}	−3.36	.01
16	−5.55	1.64×10^{-06}	−10.07	1.30×10^{-17}
17	−9.82	5.36×10^{-17}	−10.5	1.07×10^{-18}
18	−5.65	9.98×10^{-07}	−9.73	9.18×10^{-17}

Figure 5. The response times for user 5 according to the dialog steps.



Clustering Techniques: Assessing the Accuracy of Users' Role-playing

Table 4 shows the cluster results obtained using the K-means algorithm. The second column shows the true user role ID assignment, hereafter referred to as the gold standard. The third column shows the clustering run by the algorithm. The names of the SP profiles were removed from the dialogs to avoid biases in the neural network performing the embedding. According to Table 4, all users who played role A were clustered, except for user 0. The same holds for role user 4. Users who played role C were split, and user 7 was clustered with user 6, who played a different role. Users who played role E were grouped correctly. FMI is defined as higher values indicating better performance. In this study, FMI=0.32 indicated quite low performance. However, if we remove the users who played the SP of *Mirta* (Mirta is a woman aged 70 years with T2DM; see the whole description in the Multimedia Appendix 1), we thereby obtain an FMI of 0.72, which indicates a high similarity between the cluster and gold standard. These results confirm that the users who role-played Mirta's profile did not play that profile accurately. On the other hand, the remaining users played their roles satisfactorily.

Table 4. The K-mean results (third column) compared with the true role assignment (second column) for each user in the first column.

User ID	Gold standard role ID (name)	Predicted role ID
0	A (Alessandra)	cluster_0
1	A (Alessandra)	cluster_1
2	A (Alessandra)	cluster_1
3	A (Alessandra)	cluster_1
4	B (Alessandra)	cluster_2
5	C (Federico)	cluster_3
6	C (Federico)	cluster_4
7	D (Federico)	cluster_4
13	E (Simona)	cluster_3
14	E (Simona)	cluster_3
15	E (Simona)	cluster_3
16	E (Simona)	cluster_3
17	E (Simona)	cluster_3

Discussion

Overview

To date, many studies have developed behavioral interventions among adults with T2DM, mainly concentrating on the monitoring of their diet, physical activity, and blood glucose levels. In addition, the literature has proliferated, with studies emphasizing the extent to which psychosocial symptoms such as anxiety, stress, depression, and diabetes-related emotional distress directly affect diabetes management and glycemic levels [48]. These studies suggest the importance of including these symptoms in the development of interventions among people with DM [48]. Moreover, technology has opened up the potential for new types of interventions, such as mHealth solutions or VC approaches. Bearing all these aspects in mind, there is a need for more effective and accessible technology-based interventions, such as those delivered through VCs, to improve the self-care and self-management of people with DM. Moreover, the development of a robust, natural, and personalized VC is required to support the psychological needs and coping skills of people with T2DM to ensure adherence to clinical recommendations and motivate them to achieve and adopt healthy coping strategies when they experience symptoms of stress, anxiety, and/or depression.

Early Development, Acceptability, and UX

With regard to the first 4 aims, this study showed how an intervention protocol for healthy coping in the context of T2DM was formalized using both the recommendations of 3 mental health and BITs experts and the results of this study. This pilot study, as a preliminary test, highlighted how the ratio between the number of VC utterances and the number of user utterances provided us important feedback regarding the VC capability of allowing self-reflection. Indeed, a ratio close to one indicates an underinformative VC and, therefore, a low self-reflection role, whereas a high ratio indicates an overinformative VC, resulting in undue stress for the users.

The simulated VC presented a length of dialogs twice as long as the user's answers, which is in line with the design of the proactive VC-driven intervention. Indeed, the VC aims to trigger users' self-reflection on their emotions by teaching them healthy coping strategies, such as mindfulness exercises for better coping with chronic diseases. Furthermore, the VC relies on the Mood Rating Scale, in which users can reflect on and express what they feel every day and, subsequently, be aware of their emotions. On the other hand, the current VC intervention does not require too much text input and information from users, with a low mean of 27.98 (reported in Table 1) regarding the number of characters for user utterances; thus, it makes the VC less demanding but, at the same time, engaging enough for the user. In addition, the low SD in the VC utterances represents useful feedback for its improvement. A more tailored approach might involve different intervention pathways according to the users' characteristics, such as their age, gender, level of education, and habits. With regard to the users' response time, the advantage of the VC is that it gives users the right time to think and reflect on the proposed motivational and behavioral interventions; indeed, some users need more time than others to answer. Furthermore, the average of the positive items was above the median value of 3, and the negative items were under the median of 3. These results suggest that the perception of VC is positive, and it is going in a promising direction. Moreover, of particular interest are the items *motivating*, *encouraging*, and *supportive*, which moderately increased during the 6 weeks, thereby showing a positive and engaging interaction of users with the VC. With regard to the negative items, *not reflective* and *unpredictable* decreased in week 6 compared with weeks 2 and 4, showing that the simulated VC was perceived as thoughtful and predictable as every fortnight week progressed. The item *conventional* increased, indicating that users understood the direction of the dialogs delivered by the VC and felt more confident and secure in interacting with it. Finally, the interaction seemed to become slightly more *rigid*, which may be because of the setup of the intervention protocol. Only the items regarding *not efficient* and *too much information* showed a significant difference, in which their means highlighted a decrease, thereby demonstrating that the VC presents appropriate support to the user without being too intrusive or sending too much information. These latter results are in line with the statistical analyses conducted on the number of utterances and number of characters, which refer to the users' dialogs with the VC.

Users' Quotes

Qualitative analyses well-confirm the aforementioned quantitative results of this study, showing a correspondence between the items of the UX questionnaire and the themes that emerged from the users' responses to the open-ended question, adding other important considerations. There are several lessons learned, starting from the users' open-ended answers to the UX questionnaire to the feedback gathered from the interaction between the VC and users.

The users' responses to UX allowed an in-depth understanding of the users' perception regarding the interaction with the VC. Therefore, the VC was found to be supportive, allowing a certain level of self-reflection on the participants' own emotions, and natural and pleasant in its dialogic interaction. It was also found to be a little restrictive in its interaction and sometimes repetitive in its suggestions but good at stimulating thinking and supporting the use of new coping strategies. On the basis of these results, several refinements of the intervention protocol were conducted. With regard to motivational content, in addition to dialogs to increase motivation to adopt healthy coping strategies, we also included exercises that allow people to adopt a positive attitude toward diabetes management. An example of a simple and short exercise is thinking about a positive episode, describing it, and expressing the associated emotion. The motivational content of the third week was also changed by focusing on mindfulness to aspects related to the quality of sleep among adults with DM from a psychological perspective. Moreover, the frequently asked questions, collected from the users' conversation, and the possible responses were also added to the intervention protocol, such as the following:

- "What did you mean with gratitude?"
- "Aren't you going to send me a video or a recording today?"
- "How can I improve my motivation?"

- "What strategies will you teach me to be more mindful?"

User Engagement and the Accuracy of the Role-playing Technique

With regard to the last aim, which was to evaluate the users' engagement with the VC and the accuracy of their role-playing technique, some interesting results emerged.

All users, except 1, were actively engaged with the VC: they did not shorten their answers to the VC or increase response times as a result of a repetitive or uninteresting interaction. These results confirmed a stable and positive engagement level with the VC intervention. Moreover, the clustering algorithm demonstrated that most users played the assigned role accurately, interpreting the role effectively, even when personalizing it. This is also confirmed by the fact that they did not use the same words for the same role. The clustering analysis showed that the role-playing technique was performed accurately in most cases; indeed, high accuracy emerged in clustering users according to their roles. For instance, Cohen et al [49] found that the SP approach represents the most effective method for primary care physicians and reduces the proportion of patients with uncontrolled asthma by 27%, thereby improving their self-management. Regarding the SP approach, findings that emerged in this study showed that users found difficulties in interpreting Mirta's role. This result may be because of age differences between the SP and users. Indeed, as previously mentioned, Mirta was a woman aged 70 years, and, on the other hand, users presented a mean age of 23.61 years. Therefore, they could have encountered some physiological difficulties in interpreting that role; indeed, none of them played the role accurately. When the dialogs collected are grouped into roles, these dialog clusters further represent an initial data set for training a VC based on neural networks, meaning that they are able to respond to users more adaptively over the intervention. To this extent, the promising results that are presented here regarding the neural networks in the clustering analysis suggest that these algorithms can discriminate between the different SP roles. Therefore, they can potentially respond to personalized interventions according to the specific SP profiles. The results of this study highlighted how the use of the WOZ and SP approaches, combined with the role-playing technique, can prove useful in the early validation of VCs for patients with chronic diseases.

Limitations and Strengths

Pilot testing was performed among university students instead of real patients with DM; however, this choice was considered appropriate. This choice was guided by the need to improve the intervention protocol from the UX perspective without compromising their well-being. The adoption of the WOZ and SP approaches allowed us to iteratively refine our intervention before presenting it to real patients for larger testing in the field. This process allows the collection of several interesting insights from users, which is useful for developing future versions of the VC prototype. This perspective is also in line with the approach recommended by the ORBIT model (in particular for phases 1a and 1b) to implement increasingly more effective interventions with the involvement of relevant stakeholders in the design process. Finally, the use of clustering techniques to assess the accuracy of users' role-playing represents one of the more innovative contributions of this study.

Conclusions

The preliminary results of this study are promising, and the tested VC intervention was well-accepted. Indeed, most users found the VC useful for better coping with emotions and negative thoughts related to the burden of managing DM. Users further showed a good acceptance level for the VC intervention, especially for self-reflection on their emotions. The use of the WOZ and SP approaches allowed effective and rapid refinement of the intervention protocol. The overall results provide the opportunity to train an algorithm to replace the real coach as the training data quality has been checked, which shows that it is possible to capture the mean stream behaviors of each role. This training will refer to phase 2 of the ORBIT model, which is a proof-of-concept study. This study will include real patients with DM to evaluate the efficacy of the psychoeducational intervention, which will be delivered via the Telegram messaging channel [36,37]. The VC will assess symptoms of anxiety, stress, depression, health-related quality of life, and diabetes-related emotional distress at preintervention, postintervention, and follow-up to measure the intervention's efficacy. UX and user engagement will be further assessed. Thus, the ambition of the VC in its new version is to leverage the relevant patients' profiles to personalize content and dialogs delivered and achieve a more effective and higher-quality intervention.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Examples of standardized patient case scenarios for every age (ie, youths, adults, and older adults).
[DOCX File , 29 KB - formative_v6i2e27500_app1.docx]

Multimedia Appendix 2
Means and SDs of the positive and negative items every fortnight week (second, fourth, and sixth week; N=18).
[DOCX File , 22 KB - formative_v6i2e27500_app2.docx]

References

1. Wang Y, Fadhil A, Lange J, Reiterer H. Integrating taxonomies into theory-based digital health interventions for behavior change: a holistic framework. JMIR Res Protoc 2019 Jan 15;8(1):e8055 [FREE Full text] [doi: 10.2196/resprot.8055] [Medline: 30664477]
2. Tudor Car L, Dhinagaran DA, Kyaw BM, Kowatsch T, Joty S, Theng Y, et al. Conversational agents in health care: scoping review and conceptual analysis. J Med Internet Res 2020 Aug 07;22(8):e17158 [FREE Full text] [doi: 10.2196/17158] [Medline: 32763886]
3. Association of Diabetes Care & Education Specialists. URL: https://www.diabeteseducator.org/ [accessed 2022-01-25]
4. Funnell M, Brown T, Childs B, Haas LB, Hosey GM, Jensen B, et al. National standards for diabetes self-management education. Diabetes Care 2010 Jan;33 Suppl 1:S89-S96 [FREE Full text] [doi: 10.2337/dc10-S089] [Medline: 20042780]
5. Testa MA, Simonson DC. Health economic benefits and quality of life during improved glycemic control in patients with type 2 diabetes mellitus: a randomized, controlled, double-blind trial. JAMA 1998 Nov 04;280(17):1490-1496. [doi: 10.1001/jama.280.17.1490] [Medline: 9809729]
6. Bradley C, Speight J. Patient perceptions of diabetes and diabetes therapy: assessing quality of life. Diabetes Metab Res Rev 2002;18 Suppl 3:S64-S69. [doi: 10.1002/dmrr.279] [Medline: 12324988]
7. Kent D, Haas L, Randal D, Lin E, Thorpe CT, Boren SA, et al. Healthy coping: issues and implications in diabetes education and care. Popul Health Manag 2010 Oct;13(5):227-233 [FREE Full text] [doi: 10.1089/pop.2009.0065] [Medline: 20879903]
8. Hill-Briggs F, Cooper DC, Loman K, Brancati FL, Cooper LA. A qualitative study of problem solving and diabetes control in type 2 diabetes self-management. Diabetes Educ 2003;29(6):1018-1028. [doi: 10.1177/014572170302900612] [Medline: 14692375]
9. Hill-Briggs F, Gary TL, Yeh H, Batts-Turner M, Powe NR, Saudek CD, et al. Association of social problem solving with glycemic control in a sample of urban African Americans with type 2 diabetes. J Behav Med 2006 Feb;29(1):69-78. [doi: 10.1007/s10865-005-9037-0] [Medline: 16397820]
10. Sosale AR, Shaikh M, Shah A, Chawla R, Makkar BM, Kesavadev J, et al. Real-world effectiveness of a digital therapeutic in improving glycaemic control in South Asians living with type 2 diabetes. Diabetes 2018 May;67(Supplement 1):866-P. [doi: 10.2337/db18-866-P]
11. Rehman UU, Chang DJ, Jung Y, Akhtar U, Razzaq MA, Lee S. Medical instructed real-time assistant for patient with glaucoma and diabetic conditions. Applied Sciences 2020 Mar 25;10(7):2216. [doi: 10.3390/app10072216]
12. Roca S, Sancho J, García J, Alesanco A. Microservice chatbot architecture for chronic patient support. J Biomed Inform 2020 Feb;102:103305 [FREE Full text] [doi: 10.1016/j.jbi.2019.103305] [Medline: 31622802]
13. Hood M, Wilson R, Corsica J, Bradley L, Chirinos D, Vivo A. What do we know about mobile applications for diabetes self-management? A review of reviews. J Behav Med 2016 Dec;39(6):981-994. [doi: 10.1007/s10865-016-9765-3] [Medline: 27412774]
14. Gong E, Baptista S, Russell A, Scuffham P, Riddell M, Speight J, et al. My diabetes coach, a mobile app-based interactive conversational agent to support type 2 diabetes self-management: randomized effectiveness-implementation trial. J Med Internet Res 2020 Nov 05;22(11):e20322 [FREE Full text] [doi: 10.2196/20322] [Medline: 33151154]
15. Cheng A, Raghavaraju V, Kanugo J, Handrianto Y, Shang Y. Development and evaluation of a healthy coping voice interface application using the Google home for elderly patients with type 2 diabetes. In: Proceedings of the 2018 15th IEEE Annual Consumer Communications & Networking Conference (CCNC). 2018 Presented at: 2018 15th IEEE Annual Consumer Communications & Networking Conference (CCNC); Jan 12-15, 2018; Las Vegas, NV, USA. [doi: 10.1109/ccnc.2018.8319283]
16. 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]
17. Fadhil A, Wang Y, Reiterer H. Assistive conversational agent for health coaching: a validation study. Methods Inf Med 2019 Jun;58(1):9-23. [doi: 10.1055/s-0039-1688757] [Medline: 31117129]
18. Arnhold M, Quade M, Kirch W. Mobile applications for diabetics: a systematic review and expert-based usability evaluation considering the special requirements of diabetes patients age 50 years or older. J Med Internet Res 2014 Apr 09;16(4):e104 [FREE Full text] [doi: 10.2196/jmir.2968] [Medline: 24718852]
19. El-Gayar O, Timsina P, Nawar N, Eid W. Mobile applications for diabetes self-management: status and potential. J Diabetes Sci Technol 2013 Jan 01;7(1):247-262 [FREE Full text] [doi: 10.1177/193229681300700130] [Medline: 23439183]
20. Baptista S, Trawley S, Pouwer F, Oldenburg B, Wadley G, Speight J. What do adults with type 2 diabetes want from the "Perfect" app? Results from the second diabetes MILES: Australia (MILES-2) study. Diabetes Technol Ther 2019 Jul;21(7):393-399. [doi: 10.1089/dia.2019.0086] [Medline: 31166804]
21. Dobson R, Whittaker R, Jiang Y, Maddison R, Shepherd M, McNamara C, et al. Effectiveness of text message based, diabetes self management support programme (SMS4BG): two arm, parallel randomised controlled trial. BMJ 2018 May 17;361:k1959 [FREE Full text] [doi: 10.1136/bmj.k1959] [Medline: 29773539]
22. Lee K, Kwon H, Lee B, Lee G, Lee JH, Park YR, et al. Effect of self-monitoring on long-term patient engagement with mobile health applications. PLoS One 2018;13(7):e0201166 [FREE Full text] [doi: 10.1371/journal.pone.0201166] [Medline: 30048546]
23. Conversational AI. IBM. URL: https://www.ibm.com/cloud/learn/conversational-ai [accessed 2022-01-26]

24. Ogrin R, Viswanathan R, Aylen T, Wallace F, Scott J, Kumar D. Co-design of an evidence-based health education diabetes foot app to prevent serious foot complications: a feasibility study. Pract Diab 2018 Dec 06;35(6):203-29d. [doi: 10.1002/pdi.2197]
25. Dickerson R, Johnsen K, Raij A, Lok B, Hernandez J, Stevens A, et al. Evaluating a script-based approach for simulating patient-doctor interaction. In: Proceedings of SCS 2005 International Conference on Human-Computer Interface Advances for Modeling and Simulating. 2005 Presented at: Proceedings of SCS 2005 International Conference on Human-Computer Interface Advances for Modeling and Simulating; Jan 23-27, 2005; New Orleans, Louisiania.
26. Kenny P, Hartholt A, Gratch J, Swartout W, Traum D, Marsella S, et al. Building interactive virtual humans for training environments. In: Proceedings of the Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC) 2007. 2007 Presented at: Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC); Nov 26-29, 2007; Orlando Florida USA.
27. Kenny P, Parsons T, Gratch J, Rizzo A. Evaluation of Justina: a virtual patient with PTSD. In: Intelligent Virtual Agents. Berlin, Heidelberg: Springer; 2008.
28. Leuski A, Patel R, Traum D, Kennedy B. Building effective question answering characters. In: Proceedings of the 7th SIGdial Workshop on Discourse and Dialogue. 2006 Presented at: Proceedings of the 7th SIGdial Workshop on Discourse and Dialogue; Jul 15 - 16, 2006; Sydney Australia. [doi: 10.3115/1654595.1654600]
29. Reiter E, Sripada SG, Robertson R. Acquiring correct knowledge for natural language generation. J Artificial Intell Res 2003 Jun 01;18:491-516. [doi: 10.1613/jair.1176]
30. Rossen B, Lok B. A crowdsourcing method to develop virtual human conversational agents. Int J Human Comput Stud 2012 Apr;70(4):301-319. [doi: 10.1016/j.ijhcs.2011.11.004]
31. Ruttkay Z, Andre E, Johnson W, Pelachaud C. From Brows to Trust Evaluating Embodied Conversational Agents. Switzerland: Springer Nature; 2004.
32. Dybkjær H, Bernsen N, Dybkjær L. Wizard-of-oz and the trade-off between naturalness and recogniser constraints. In: Proceedings of the Third European Conference on Speech Communication and Technology, EUROSPEECH 1993. 1993 Presented at: Third European Conference on Speech Communication and Technology, EUROSPEECH 1993; Sep 22-25, 1993; Berlin, Germany URL: https://www.researchgate.net/publication/221486754_Wizard-of-oz_and_the_trade-off_between_naturalness_and_recogniser_constraints [doi: 10.1007/978-1-4471-0897-9_5]
33. Ferland L, Sauve J, Lucke M, Nie R, Khadar M, Pakhomov S, et al. Tell me about your day: designing a conversational agent for time and stress management. In: Explainable AI in Healthcare and Medicine. Cham: Springer; 2021.
34. Tamblyn R, Barrows H, Gliva G. An initial evaluation of learning units to facilitate problem solving and self-directed study (portable patient problem pack). Med Educ 1980 Nov;14(6):394-400. [doi: 10.1111/j.1365-2923.1980.tb02390.x] [Medline: 7442578]
35. Becker KL, Rose LE, Berg JB, Park H, Shatzer JH. The teaching effectiveness of standardized patients. J Nurs Educ 2006 Apr;45(4):103-111. [doi: 10.3928/01484834-20060401-03] [Medline: 16629278]
36. Bassi G, Giuliano C, Perinelli A, Forti S, Gabrielli S, Salcuni S. A virtual coach (Motibot) for supporting healthy coping strategies among adults with diabetes: proof-of-concept study. JMIR Hum Factors 2022 Jan 21;9(1):e32211 [FREE Full text] [doi: 10.2196/32211] [Medline: 35060918]
37. Czajkowski SM, Powell LH, Adler N, Naar-King S, Reynolds KD, Hunter CM, et al. From ideas to efficacy: the ORBIT model for developing behavioral treatments for chronic diseases. Health Psychol 2015;34(10):971-982 [FREE Full text] [doi: 10.1037/hea0000161] [Medline: 25642841]
38. Brown JM, Savaglio R, Watson G, Kaplansky A, LeSage A, Hughes J, et al. Optimizing child nutrition education with the Foodbot Factory mobile health app: formative evaluation and analysis. JMIR Form Res 2020;4(4):e15534 [FREE Full text] [doi: 10.2196/15534] [Medline: 32301743]
39. Kalra S, Jena BN, Yeravdekar R. Emotional and psychological needs of people with diabetes. Indian J Endocrinol Metab 2018;22(5):696-704 [FREE Full text] [doi: 10.4103/ijem.IJEM_579_17] [Medline: 30294583]
40. Williams M, Penman D. Mindfulness method. Milan, Italy: Oscar; 2016.
41. Schrepp M. User experience questionnaire handbook. User Experience Questionnaire. 2019. URL: https://www.ueq-online.org/Material/Handbook.pdf [accessed 2022-01-26]
42. R: a language and environment for statistical computing. R Foundation for Statistical Computing. 2019. URL: https://www.Rproject.org/ [accessed 2022-01-26]
43. IBM SPSS Statistics for Windows, Version 24.0. Armonk, NY: IBM Corp; 2016. URL: https://www.ibm.com/it-it
44. Python 3.6.6. Python. URL: https://www.python.org/downloads/release/python-366/ [accessed 2022-01-26]
45. Homoscedasticity. Statistics Solution. URL: https://www.statisticssolutions.com/homoscedasticity/ [accessed 2022-01-26]
46. Cer D, Yang Y, Kong S, Hua N, Limtiaco N, John R, et al. Universal sentence encoder. In: Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations. 2018 Presented at: Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations; Oct 31 – Nov 4, 2018; Brussels, Belgium. [doi: 10.18653/v1/d18-2029]

47. Braun V, Clarke V. Thematic analysis. In: Apa Handbook of Research Methods in Psychology, Vol. 2. Research Designs: Quantitative, Qualitative, Neuropsychological, and Biological. Washington, D.C., United States: American Psychological Association; 2012.
48. The development of a mindfulness-based mobile application to learn emotional self-regulation. CEUR-WS. URL: http://ceur-ws.org/Vol-2524/paper19.pdf [accessed 2022-01-26]
49. Cohen A, Kitai E, David S, Ziv A. Standardized patient-based simulation training as a tool to improve the management of chronic disease. Simul Healthc 2014 Feb;9(1):40-47. [doi: 10.1097/SIH.0000000000000009] [Medline: 24492338]

Abbreviations

AADE: American Association of Diabetes Educators
BIT: behavioral intervention technology
DM: diabetes mellitus
FMI: Fowlkes-Mallows Index
mHealth: mobile health
ORBIT: Obesity-Related Behavioral Interventions Trials
SP: standardized patient
T2DM: type 2 diabetes mellitus
UX: user experience
VC: virtual coach
WOZ: Wizard of Oz

Edited by A Mavragani; submitted 27.01.21; peer-reviewed by MT García-Ordás, G Humphrey; comments to author 17.03.21; revised version received 27.04.21; accepted 20.12.21; published 11.02.22.

Please cite as:
Bassi G, Donadello I, Gabrielli S, Salcuni S, Giuliano C, Forti S
Early Development of a Virtual Coach for Healthy Coping Interventions in Type 2 Diabetes Mellitus: Validation Study
JMIR Form Res 2022;6(2):e27500
URL: https://formative.jmir.org/2022/2/e27500
doi:10.2196/27500
PMID:35147505

Original Paper

Evaluation of mHealth Apps for Diverse, Low-Income Patient Populations: Framework Development and Application Study

Shreya Sharma[1,2], MD; Katherine Gergen Barnett[3,4], MD; John (Jack) Maypole[5,6], MD; Rebecca Grochow Mishuris[7,8], MD, MPH

[1]Boston University School of Medicine, Boston, MA, United States
[2]Beth Israel Deaconess Medical Center, Boston, MA, United States
[3]Department of Family Medicine, Boston Medical Center, Boston, MA, United States
[4]Department of Family Medicine, Boston University School of Medicine, Boston, MA, United States
[5]Department of Pediatrics, Boston Medical Center, Boston, CA, United States
[6]Department of Pediatrics, Boston University School of Medicine, Boston, MA, United States
[7]Department of Medicine, Boston Medical Center, Boston, MA, United States
[8]Department of Medicine, Boston University School of Medicine, Boston, MA, United States

Corresponding Author:
Shreya Sharma, MD
Boston University School of Medicine
72 E Concord Street
Boston, MA, 02118
United States
Phone: 1 (617)358 9600
Email: sshreya@bu.edu

Abstract

Background: The use of mobile technology or smartphones has grown exponentially in the United States, allowing more individuals than ever internet access. This access has been especially critical to households earning less than US $30,000, the majority of whom indicate that smartphones are their main source of internet access. The increasing ubiquity of smartphones and virtual care promises to offset some of the health disparities that cut through the United States. However, disparities cannot be addressed if the medical information offered though smartphones is not accessible or reliable.

Objective: This study seeks to create a framework to review the strengths and weaknesses of mobile Health (mHealth) apps for diverse, low-income populations.

Methods: Focusing on smoking cessation, diabetes management, and medication adherence as models of disease management, we describe the process for selecting, evaluating, and obtaining patient feedback on mHealth apps.

Results: The top 2 scoring apps in each category were QuitNow! and Smoke Free-Quit Smoking Now for smoking cessation, Glucosio and MyNetDiary for diabetes management, and Medisafe and MyMeds for medication adherence.

Conclusions: We believe that this framework will prove useful for future mHealth app development, and clinicians and patient advisory groups in connecting culturally, educationally, and socioeconomically appropriate mHealth apps with low-income, diverse communities and thus work to bridge health disparities.

(JMIR Form Res 2022;6(2):e29922) doi:10.2196/29922

KEYWORDS

mobile health application; apps; mobile health; diverse; low-income; mHealth; framework; chronic disease; condition; smoking cessation; diabetes; medication adherence; safety net hospital; personal; self-management; usability test

Introduction

The use of mobile technology has grown exponentially in the United States, and the COVID-19 pandemic both enforced the need for internet connectivity and laid bare disparities in access. Currently, up to 97% of Americans own a cell phone, including 97% of those with a household income under US $30,000 [1]. However, 27% of households earning under US $30,000 rely on smartphones for internet access, compared to only 6% of households earning greater than US $100,000 [2]. During the

COVID-19 pandemic, 43% of individuals used their cell phones to manage mental health and well-being, 41% to access health care, and 40% to keep fit and exercise [3]. While there still remains a large socioeconomic digital divide, the increasing ownership of ever-evolving smartphones and their functionality may partially offset this disparity [4,5].

The mobile health (mHealth) app market has flourished as consumers, health technology companies, and biomedical researchers have recognized mobile apps as a potential vehicle to lower barriers to accessing preventive medicine and promoting healthy behaviors. This extends to lower-income populations as well [6]. Currently, there are a total of 48,608 mHealth apps available for download in the Apple App Store [7]. The mHealth app industry is forecasted to be worth US $151 billion by 2025 [8,9]. However, little is known about efficacy of these apps, particularly across diverse consumer (and patient) populations.

The COVID-19 pandemic has illuminated the long prevalent health disparities among lower-income populations who experience higher rates of chronic disease such as diabetes and hypertension [10]. Mobile apps could, in theory, address some of these disparities. Indeed, mobile health apps have been proven, in many instances, to impact lifestyle changes and health outcomes positively [11,12]. There are available mobile health apps that promote healthy behaviors (such as exercise or smoking cessation) and support chronic disease management (such as monitoring blood sugar levels and enhancing medication adherence). Patients in diverse, low-income communities have shown more interest than white, high-income communities in the use of mHealth apps, particularly for chronic disease and overall health management [13-15]. This presents an area of opportunity for software developers and health care providers to reduce health inequities. However, the vast majority of mHealth apps do not cater to the needs of lower-income populations, as they have been shown to be difficult to navigate for individuals who may have limited health, digital, or written literacy [11,16-20]. Additionally, there is no universally accepted framework to assess the functionality and usability of mHealth apps, which may further disproportionately impact diverse, low-income populations [6,12].

This study seeks to create a framework for the evaluation of mHealth apps' accessibility for diverse, low-income populations. We developed and tested a rubric (a guide listing specific criteria for grading or scoring) of domains (features of mHealth apps with a common purpose) to measure the functionality and usability of several mHealth apps for patients of an urban safety net institution. Focusing on smoking cessation, diabetes management, and medication adherence as models of disease management, we describe the process of selection of domains of mHealth apps, development of the rubric, and scoring of various mHealth apps. We envision that this framework will prove useful for clinicians, care teams, patient advisory groups, and developers (who seek to design apps with equitable reach and greater impact on health outcomes) in connecting culturally, educationally, and socioeconomically appropriate mHealth apps with communities historically overlooked by this rapidly evolving area of health care.

Methods

Domain Selection

In June 2018, we searched web-based databases (PubMed and Embase) to identify articles related to the evaluation of the usability of mobile apps for health and wellness. Studies from this literature review were assessed to create a list of domains for rating mHealth apps relevant to diverse, low-income populations. Search criteria included "smartphone application underserved community," "usability of commercially available applications for diverse patients," "mHealth apps usability testing underserved," "mheatlh app underserved," "mobile health phone applications for diverse populations," on PubMed and "mobile application'/exp OR 'mobile application' OR 'mobile phone'/exp OR 'mobile phone," "'mHealth under-served' OR (('mhealth'/exp OR mhealth) AND unders-erved)" on Embase.

An extensive literature review indicated important domains of mHealth apps for our target population, including the following: language, literacy, graphics, multimedia, usability, patient-centeredness, data entry mode, data exportability, cost, evidence based content, platform, extent to which the platform was up to date, connectivity, Americans with Disabilities Act (ADA) accessibility, privacy, social network, cultural sensitivity (incorporating diversity in language, graphics, and data), messaging or reminder capability, and benchmarking (comparing a user's performance to the performance of others on the app). These domains were grouped into larger categories, including the following: Usability (graphics, multimedia, usability, and ADA accessibility), Population focus (language, literacy, patient-centeredness, cost, social network, cultural sensitivity, benchmarking, and messaging or reminder capability), Technology (data entry mode, data exportability, platform, connectivity requirement, and privacy), and Clinical Impact (evidence-based content and extent to which the platform was up to date).

Each domain was then weighted in terms of importance to the target population by independent coders (RGM, KGB, and JM). These coders are all primary care physicians at an urban safety net hospital (in General Internal Medicine, Family Medicine, and Pediatrics, each with 10-25 years of experience at this institution) and care for overlapping members of these communities by age. They were asked to rate each domain on a scale of 1 to 5 (1="not important to be included" and 5="must be included"). Then, a reviewer (SS) adjudicated these weights and averaged the "weight" of each domain respectively. Finally, these weights were directionally confirmed with the hospital-based patient and family advocacy committee.

App Selection

To simulate how patients would access recommendations for mHealth apps, we used the most common search engine, Google, to search "Top Ten Mobile Health Apps for Diabetes" and "Top Ten Mobile Health Apps for Smoking Cessation." The first search result referred to articles in a popular health website "Healthline," which provided the top 10 mobile health apps related to each topic [21,22]. The website rated these apps on the basis of quality, reliability, reviews, and community nominations. As no such list existed for medication adherence

focused apps, these apps were chosen from the list recommended by a previous study [23]. We chose apps in these 3 areas because they represent different aspects of medical care delivery: preventative care (smoking cessation), chronic disease management (diabetes management), and general therapeutic intervention (medication adherence).

App Scoring

Ten apps in smoking cessation, diabetes management, and medication adherence were identified and rated by a coder (SS) in each domain (language, literacy, graphics, multimedia, usability, patient-centered, data entry mode, data exportability, cost, evidence-based content, platform, extent the platform was up to date, connectivity, ADA accessibility, privacy, social network, cultural sensitivity, messaging or reminder capability, and benchmarking) with a score of 0 to 3. A score of 0 was assigned if the specific domain was not applicable, and a score of 3 was assigned for the highest applicability (eg, language—available in 3 or more languages; Table 1). Domains were defined on the basis of current research in each respective domain. For example, for "Social Network," higher points were assigned for "competition" than for "social support" because previous studies have shown that social comparison was more important for physical activity [24]. As another example, the definition for "Usability" was based on the design of the app and the ability to navigate it easily. We specifically did not use the industry standard assessment (ie, System Usability Scale) owing to concerns about the applicability for diverse, low–socioeconomic status communities. Lastly, we decided whether apps were "evidence based" if they cited specific research or if the theory underlying their apps had existing evidence (ie, gasification and social networks).

The domain score was multiplied by the domain weight (of importance to the target population) to produce a weighted final score for that domain for the app in question. The weighted scores across all domains were added to assign each app a final overall score of usefulness and applicability for diverse, low-income populations.

Table 1. Domain definitions and scoring.

Domain	Points			
	0	1	2	3
Population focus				
Language	Not applicable/not offered by app/not a function of this app	Only available in English	Available in English + 1 language	Multiple (>2) languages
Literacy	Not applicable/not offered by app/not a function of this app	Complex language, medical jargon	some patient-friendly language, some medical jargon	patient-friendly language at or below 5th grade reading level
Patient-centeredness	Not applicable/not offered by app/not a function of this app	Does not allow patient-entered data/provides no tailored content	Provides moderate tailored content	Provides actionable content based on patient-entered data
Social network	Not applicable/not offered by app/not a function of this app	Provides a forum for people to share information/discuss different topics	Provides a platform for competition between users	Incorporates a forum and also a "competition" between users
Benchmarking	Not applicable/not offered by app/not a function of this app	Provides no sense of patient status among peers	Provides a benchmark against others	Allows patients to define peer groups for benchmarking
Cultural sensitivity	N/A[a]	No attempt at cultural diversity	Some attempt at cultural diversity	Cultural diversity definitely represented
Cost	Not applicable/not offered by app/not a function of this app	Paid app	Can pay for add-on	Free app
Messages or reminders	Not applicable/not offered by app/not a function of this app	Has messages, but not tailored to the user	App has motivational messages tailored to the user/reminders pop-up, but only if the user is on the app	App has option for you to "turn on notification" to receive reminders, messages, and notifications tailored specifically to the user
Usability				
Graphics	Not applicable/not offered by app/not a function of this app	Only text	Text>graphics	Graphics clearly explain the text
Multimedia	Not applicable/not offered by app/not a function of this app	No use of audio or video	Uses audio or video	Uses audio and video
Usability	Not applicable/not offered by app/not a function of this app	Complex app navigation, small icons, multiple screens, not immediately obvious how to use	Small icons, multiple screens, simpler app navigation but still would require instruction	Large icons, few screens, easy to navigate without instruction/mobile literacy
Americans with Disabilities Act accessibility	Not applicable/not offered by app/not a function of this app	No accommodation for visual or hearing impairment	Accommodation for visual or hearing impairment	Accommodation for both visual and hearing impairment
Technology				
Data entry mode	Not applicable/not offered by app/not a function of this app	Manual entry	Manual entry or integrated with information from the phone/other devices	Integrated with external devices or electronic health records
Data exportability	Not applicable/not offered by app/not a function of this app	Does not allow data export	Allows export of data to print or email or message	Integrated with external services or electronic health records
Connectivity requirement	Not applicable/not offered by app/not a function of this app	Requires constant Wi-Fi/cellular connection	Some functions useful without connection	Fully functional without connection/only requires connection to integrate with external devices/system
Platform	Not applicable/not offered by app/not a function of this app	Only iOS	Only android	iOS and android
Privacy	Not applicable/not offered by app/not a function of this app	No privacy statement	N/A	Privacy statement exists
Clinical impact				
Evidence-based content	Not applicable/not offered by app/not a function of this app	No clear identification of evidence for content	Some content has an evidence base	Entire app content is evidence-based

Domain	Points			
	0	1	2	3
Up to date	Not applicable/not offered by app/not a function of this app	No indication of last update date/ updated >18 months ago	Updated within the last 12-18 months	Updated within last 6-12 months

[a]N/A: not applicable.

Patient Advisory Board

In July 2019, after the mHealth apps were chosen and scored, we presented the top 3 apps in each health category to a patient advisory board at a safety net hospital in Boston, Massachusetts. The group consisted of 4 participants on the day of presentation. They were Caucasian women between 35 and 75 years of age from the catchment area of the safety net institution. They were asked to give their overall thoughts on our research idea and the functionality of mHealth apps they considered important to manage their health on a daily basis.

Results

Domain Scoring

Domains of usability and functionality for mHealth apps were assigned weighted scores depending on the importance the clinical experts gave to those categories for diverse, low-income communities. These were directionally confirmed with a patient advisory board. Domains of greatest importance were identified through this process (Table 2). Literacy was found to be the highest scoring domain by clinical experts, with all raters assigning a score of 5 out of 5. Language, usability, cost, evidence based content, and cultural sensitivity had an average weight of 4.5 out of 5. Graphics was rated 4 out of 5. Social media connectivity and timing of the app's latest update were assigned a score of 3.5 out of 5. Multimedia, data exportability, patient centeredness, benchmarking, and messages or reminders were all rated an average of 3 out of 5. ADA accessibility was rated 2.5 out of 5. Data entry and privacy domains were weighted 2 out of 5, being the least important for the specific patient population. Raters were within 1 point of each other on all domains, with 8 out of 19 domain weightings being identical. The higher-level categories into which domains can be grouped also differentiated in terms of importance, with clinical impact and population focus being most highly weighted: Clinical Impact (average weight across component domains 4), Population Focus (average weight across component domains 3.88), Usability (average weight across component domains 3.5), and Technology (average weight across component domains 2.8).

Table 2. Domains of greatest importance.

Domain	Average final weight
Population focus	
Literacy	5
Language	4.5
Cultural sensitivity	4.5
Cost	4.5
Graphics	4
Social network	3.5
Patient-centeredness	3
Benchmarking	3
Usability	
Usability	4.5
Multimedia	3
Messages or reminders	3
Americans with Disabilities Act accessibility	2.5
Technology	
Platform	4
Data exportability	3
Connectivity requirement	3
Data entry mode	2
Privacy	2
Clinical impact	
Evidence-based content	4.5
Up to date	3.5

App Scoring

When using the weighted scoring methodology to rate mHealth apps across smoking cessation, diabetes management, and medication adherence, the framework was able to sufficiently distinguish among 10 apps within each category (Table 3). Weighted scores ranged from 108.5 to 153 for smoking cessation apps, 119 to 147 for diabetes management apps, and 113 to 137.5 for medication adherence apps (Table 4). The top 2 scoring apps in each category were QuitNow! (156/201), Smoke Free-Quit Smoking Now (149.5/201), Glucosio (147/201), MyNetDiary (146/201), Medisafe (137.5/201), and MyMeds (126/201) (Multimedia Appendix 1).

Table 3. Weighted scores for each top mobile app.

	Smoking cessation		Diabetes management		Medication adherence	
	QuitNow!	Smoke Free- Quit Smoking Now	MyNetDiary	Glucosio	Medisafe	My Meds
Rating on app store (out of 5)	4.6	4.8	4.6	N/A[a]	4.7	N/A
Ratings, n	317,000	181,000	107,000	N/A	138,000	N/A
Language (4.5)	13.5	13.5	4.5	13.5	3	1
Literacy (5)	15	15	15	15	15	15
Graphics (4)	8	12	12	12	12	12
Multimedia (3)	3	6	6	3	3	3
Usability (4.5)	13.5	13.5	9	13.5	13.5	13.5
Patient- centeredness (3)	6	9	9	3	9	3
Data entry mode (2)	2	2	4	2	6	2
Data exportability (3)	6	6	9	6	6	3
Cost (4.5)	9	9	6	13.5	9	13.5
Evidence-based content (4.5)	13.5	13.5	13.5	4.5	0	0
Platform (4)	12	12	12	12	12	12
Up to date (3.5)	10.5	10.5	10.5	10.5	10.5	10.5
Connectivity requirement (3)	6	6	6	9	9	9
Americans with Disabilities Act accessibility (2.5)	2.5	2.5	2.5	2.5	2.5	2.5
Privacy (2)	6	6	6	6	6	6
Benchmarking (3)	6	3	3	3	3	3
Social support (3.5)	3.5	0	3.5	0	0	3.5
Cultural sensitivity (4.5)	4.5	4.5	13.5	9	9	4.5
Messages (3)	9	9	9	9	9	9

[a]N/A: not applicable.

Table 4. Weighted scores of top mobile apps.

App name	Cumulative final score	Range of final scores across all 10 apps evaluated
Smoking cessation		108.5-153
Smoke Free-Quit Smoking Now	153	
QuitNow!	146.5	
Diabetes management		122-147
Glucosio	147	
MyNetDiary	146	
Medication adherence		113-137.5
Medisafe	137.5	
My Meds	126	

Discussion

Principal Findings

mHealth apps have the potential to improve individuals' management of chronic diseases and to extend the reach of the health care provider visit. Considering the increasing predominance of smart device use to access the internet among diverse, low-income communities, mHealth apps hold enormous potential for impact on health and could help bridge health equity divides. To date, however, studies have shown that more research is needed to rate the practical functionality of these mobile apps, specifically for this target population [25,26]. By developing a framework to rate the usefulness of mHealth

applications for low income, diverse, patient populations, and showing its effectiveness to differentiate between apps, we aimed to provide a way to curate mHealth apps for better access and engagement among diverse populations.

Existing frameworks such as the Mobile App Rating Scale (MARS) include ratings for domains grouped into categories with broad population applicability, including engagement, functionality, aesthetics, subjective quality, and information. As identified by researchers who created the MARS system, this grading system is agnostic to the needs of specific populations and may not be indicative of the usefulness of apps for specific groups. As a result, it is difficult to apply the grading criteria of MARS to diverse, low-income patient populations [26]. The evaluation mechanism described in this study incorporates domains that are applicable for general consumption as contained in MARS and highlighted by Anderson et al [27] and others as important to engagement ("evidence based content," "privacy," "up to date," "patient centered," "benchmarking," and "social network"), but expands beyond MARS specifically for diverse, low-income patient populations [26-29].

In contrast to the more generalized domains, our research led us to evaluate a greater number of criteria with population-specific focus. The domains expanded upon those in the MARS criteria to include: language, literacy, cultural sensitivity, data entry, data exportability, multimedia, ADA accessibility, cost, platform, and message or reminder function. As identified by prior studies, multimedia availability, such as videos, helped increase mobile app engagement, particularly for individuals with low literacy levels [16,17,19,20,30,31]. Messages or a reminder function helped app users stay motivated in their plan, manage medications, and organize personal health information [29,32-36]. Cost was a domain included in our scoring framework because some past studies showed increased engagement if patients paid for the app. However, there is a question as to whether this would be an undue barrier for patients with low income [27]. Importantly and consistent with our hypothesis, language, literacy, and cultural sensitivity were found to be some of the most important qualities for an app to have to be relevant and useful for diverse, low-income populations, according to both our clinician and patient reviewers [14,16,18-20,29,37-40].

The highest rated mHealth apps in our sample shared notable, similar qualities, regardless of what chronic illness or health state they were developed to manage. Most apps placed importance on offering multiple languages, being written for lower literacy levels, and incorporating graphics; these scored high on usability. However, apps such as MyNetDiary, for diabetes management, which were more difficult to navigate, did offer additional resources to learn about navigating through the app, such as an instructional video. Additionally, privacy, evidence-based content, data entry, data exportability, and patient-centered content were present in most (but not all) of the chosen apps. Benchmarking (comparing a user's performance to the performance of others on the app), social support networks, and multimedia use scores were not robust in any of the chosen apps. ADA accessibility and cultural sensitivity (incorporating diversity in language, graphics, and data) were also lacking. Based on these results, the apps rated most highly in accordance with our framework conveyed information in a manner that a more diverse user base would be able to effectively engage with. However, as a whole, the mobile apps did not focus any specific attention to engaging a diverse population. For example, although many apps were inclusive by providing multiple language options, they lacked content that was tailored to a culturally diverse population. Overall, although these apps were the best of what is currently on the market, they are far from being ideal for our target population.

The patient advisory board was both a versatile and key feature of our research effort providing both a patient perspective on the mHealth apps and on the domains included in the framework we developed. Thus, we gained insight to patients' perspectives and were able to compare their priorities in engaging with these apps to those of health care providers in the same system. Domains including language, literacy, evidence-based content, cultural sensitivity, and providing up-to-date information were ranked highly by both groups. However, some domains such as graphics, social network, and benchmarking were not as important to patients as they were thought to be by providers. In contrast to our qualitative findings, prior studies have shown that apps with simple interfaces that favor graphics over text tend to be more usable with lower literacy populations [19,37]. Additionally, social media functionality has also been found to reduce barriers to sharing information and learning from others [32,41,42]. Therefore, while other studies have shown these domains to be important to engagement, our patients did not rank them highly in terms of importance to usability. These differences could be explained by the small sample size and demographics of our patient advisory group.

Limitations

There were several limitations to this work, which can be used as learning points for future research. First, in selecting which mHealth apps to evaluate we used a common website which had its own subjective way of choosing which apps were the "best." However, this approach to finding the "top 10" apps on a website likely replicates the way in which patients would search for and find mHealth apps to download and use. The mHealth app market is in a constant state of flux and growth, and with it, the "top 10." Second, some of the mHealth apps initially identified for evaluation were not available for download. Churn in the mHealth market may impact engagement with some mHealth apps. However, this better reflected which mHealth apps would be accessible to our patients. Third, as this was a small study with no independent funding, we only had one reviewer grade the mHealth apps across each domain. Though individual user subjectivity is inherently part of the process of evaluating mHealth apps, this process may result in some subjectivity to our analysis of the individual apps. Furthermore, our patient advisory board was not necessarily representative of the diverse patient population of the hospital. While they themselves were patients at the safety net institution, they were all White women aged 35-75 years. As a result, their opinions may not be generalizable and could be subject to conscious and unconscious biases. While this limitation is significant, we also found it important to have some

patient perspective and validation rather than none at all. Finally, it is important to acknowledge our mHealth rubric and research was conducted before the COVID-19 pandemic, and much of mHealth has evolved since then. Our reliance on technology is more than ever, and we may be more dependent on mobile app technologies. However, the nature of this research should withstand the evolution of the mHealth app market, and even seek to improve upon it as it can be applied to make this growing field even more accessible for a diverse user base.

Conclusions

We adapted a framework for evaluating mHealth apps for diverse, low-income populations from the MARS model and created modified domains to characterize critical features as identified by patients and clinical experts who care for these patient populations. Although these app domains are rated separately, it is important to remember that often these different domains work synergistically. Some domains may increase user engagement and retention while others are focused on increasing access, inclusivity, or privacy. Multiple languages remain a worthy goal, but we contend mHealth app developers should also consider incorporating culturally sensitive and specific information such as recipes, videos, and motivational tools.

This novel framework is invaluable as it can be applied to evaluate individual mHealth apps in the context of therapeutic interventions, as well as by app developers to identify those domains important for engagement of diverse, low-income populations. However, right now our app ranks relative performance. To establish specific scoring thresholds and determine cut-off ratings for apps we need to apply this rating framework to a large number of apps. The plan would be to accomplish this during a follow-up study. For now, we believe distinguishing relative performance is important so developer can use an iterative approach to design apps. They can use scores as a performance indicator as they improve or refine existing apps. If developers are able to report app performance, it may be a way to filter apps and create inter-app competition to improve performance.

Future Prospects

In the future, the study team plans to incorporate the selected apps in smoking cessation, diabetes management, and medication adherence into our clinical processes as “prescriptions” for patients. We will be able to track these “prescriptions,” assess patient engagement, and determine impact on health outcomes. In addition, we hope to continue to use our framework to find, scale, and spread mHealth apps that will be most useful for our target population in other areas of health including exercise and mental health. By using this novel framework to identify mHealth apps for recommendation and for future mHealth app development, health care providers, policy makers and developers alike may be able to better incorporate this burgeoning technology into both clinical practice and patient homes for greater impact on health outcomes for all patients, further narrowing digital and health outcome divides.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Domain Definitions and Scoring.
[DOCX File , 37 KB - formative_v6i2e29922_app1.docx]

References

1. Mobile Fact Sheet. Pew Research Center. 2021 Apr 07. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2022-01-26]
2. Vogels EA. Digital divide persists even as Americans with lower incomes make gains in tech adoption. Pew Research Center. 2021 Jun 22. URL: https://www.pewresearch.org/fact-tank/2021/06/22/digital-divide-persists-even-as-americans-with-lower-incomes-make-gains-in-tech-adoption/ [accessed 2022-01-26]
3. Engebreston J. COVID-19 Mobile Usage Increased One Hour a Day; Consumers Value ICT, Report Says. Telecompetitor. 2020 Jun 16. URL: https://www.telecompetitor.com/covid-19-mobile-usage-increased-one-hour-a-day-consumers-value-ict-report-says/ [accessed 2022-01-26]
4. Smith A. African Americans and Technology Use: A Demographic Portrait. Pew Research Center. 2014 Jan 06. URL: https://www.pewresearch.org/internet/2014/01/06/african-americans-and-technology-use/ [accessed 2020-05-20]
5. Smith A. Technology Trends Among People of Color. Pew Research Center. 2010 Sep 17. URL: https://www.pewresearch.org/internet/2010/09/17/technology-trends-among-people-of-color/ [accessed 2020-05-20]
6. Vangeepuram N, Mayer V, Fei K, Hanlen-Rosado E, Andrade C, Wright S, et al. Smartphone ownership and perspectives on health apps among a vulnerable population in East Harlem, New York. Mhealth 2018;4:31 [FREE Full text] [doi: 10.21037/mhealth.2018.07.02] [Medline: 30221166]
7. Number of mHealth apps available in the Apple App Store from 1st quarter 2015 to 1st quarter 2021. Statista. URL: https://www.statista.com/statistics/779910/health-apps-available-ios-worldwide/ [accessed 2022-01-26]
8. The global mHealth market size is expected to reach USD 151.57 billion by 2025, progressing at a CAGR 25.7% over the forecast period. Cision US. URL: https://tinyurl.com/nxr4y3mx [accessed 2022-01-26]

9. MHealth Market Analysis Report By Participants, By Service And Segment Forecasts, 2018 - 2025. ReportLinker. URL: https://www.reportlinker.com/p05741285/mHealth-Market-Analysis-Report-By-Participants-By-Service-And-Segment-Forecasts.html?utm_source=PRN [accessed 2022-01-26]
10. 2016 National Healthcare Quality and Disparities Report. Agency for Healthcare Research and Quality. URL: https://www.ahrq.gov/research/findings/nhqrdr/nhqdr16/index.html [accessed 2020-05-20]
11. Marcolino MS, Oliveira JAQ, D'Agostino M, Ribeiro AL, Alkmim MBM, Novillo-Ortiz D. The Impact of mHealth Interventions: Systematic Review of Systematic Reviews. JMIR Mhealth Uhealth 2018 Jan 17;6(1):e23 [FREE Full text] [doi: 10.2196/mhealth.8873] [Medline: 29343463]
12. Free C, Phillips G, Galli L, Watson L, Felix L, Edwards P, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. PLoS Med 2013;10(1):e1001362 [FREE Full text] [doi: 10.1371/journal.pmed.1001362] [Medline: 23349621]
13. Humble JR, Tolley EA, Krukowski RA, Womack CR, Motley TS, Bailey JE. Use of and interest in mobile health for diabetes self-care in vulnerable populations. J Telemed Telecare 2016 Jan;22(1):32-38. [doi: 10.1177/1357633X15586641] [Medline: 26026179]
14. Ramirez V, Johnson E, Gonzalez C, Ramirez V, Rubino B, Rossetti G. Assessing the Use of Mobile Health Technology by Patients: An Observational Study in Primary Care Clinics. JMIR Mhealth Uhealth 2016 Apr 19;4(2):e41 [FREE Full text] [doi: 10.2196/mhealth.4928] [Medline: 27095507]
15. Purnell JQ, Griffith J, Eddens KS, Kreuter MW. Mobile technology, cancer prevention, and health status among diverse, low-income adults. Am J Health Promot 2014;28(6):397-402 [FREE Full text] [doi: 10.4278/ajhp.120816-ARB-396] [Medline: 24200336]
16. 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]
17. Vu M, Nguyen V, Vishwanath S, Wolfe C, Patel S. Know Your Numbers: Creation and implementation of a novel community health mobile application (app) by student pharmacists. J Am Pharm Assoc (2003) 2018;58(2):191-198.e2. [doi: 10.1016/j.japh.2017.11.003] [Medline: 29249652]
18. Stephan LS, Dytz Almeida E, Guimaraes RB, Ley AG, Mathias RG, Assis MV, et al. Processes and Recommendations for Creating mHealth Apps for Low-Income Populations. JMIR Mhealth Uhealth 2017 Apr 03;5(4):e41 [FREE Full text] [doi: 10.2196/mhealth.6510] [Medline: 28373155]
19. Miller DP, Weaver KE, Case LD, Babcock D, Lawler D, Denizard-Thompson N, et al. Usability of a Novel Mobile Health iPad App by Vulnerable Populations. JMIR Mhealth Uhealth 2017 Apr 11;5(4):e43 [FREE Full text] [doi: 10.2196/mhealth.7268] [Medline: 28400354]
20. Gordon M, Henderson R, Holmes JH, Wolters MK, Bennett IM, SPIRIT (Stress in Pregnancy: Improving Results with Interactive Technology) Group. Participatory design of ehealth solutions for women from vulnerable populations with perinatal depression. J Am Med Inform Assoc 2016 Jan;23(1):105-109 [FREE Full text] [doi: 10.1093/jamia/ocv109] [Medline: 26342219]
21. The Best Diabetes Apps of 2022. Healthline. URL: https://www.healthline.com/health/diabetes/top-iphone-android-apps#mysugr [accessed 2020-05-20]
22. The Best Quit Smoking Apps of 2020. Healthline. URL: https://www.healthline.com/health/quit-smoking/top-iphone-android-apps [accessed 2020-05-20]
23. Santo K, Richtering SS, Chalmers J, Thiagalingam A, Chow CK, Redfern J. Mobile Phone Apps to Improve Medication Adherence: A Systematic Stepwise Process to Identify High-Quality Apps. JMIR Mhealth Uhealth 2016 Dec 02;4(4):e132 [FREE Full text] [doi: 10.2196/mhealth.6742] [Medline: 27913373]
24. Zhang J, Brackbill D, Yang S, Becker J, Herbert N, Centola D. Support or competition? How online social networks increase physical activity: A randomized controlled trial. Prev Med Rep 2016 Dec;4:453-458 [FREE Full text] [doi: 10.1016/j.pmedr.2016.08.008] [Medline: 27617191]
25. Stowell E, Lyson M, Saksono H. Designing and evaluating mHealth interventions for vulnerable populations: A systematic review. 2018 Presented at: CHI '18: CHI Conference on Human Factors in Computing Systems; April 21-26, 2018; Montreal, QC. [doi: 10.1145/3173574.3173589]
26. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar 11;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
27. Anderson K, Burford O, Emmerton L. Mobile Health Apps to Facilitate Self-Care: A Qualitative Study of User Experiences. PLoS One 2016;11(5):e0156164 [FREE Full text] [doi: 10.1371/journal.pone.0156164] [Medline: 27214203]
28. Higgins JP. Smartphone Applications for Patients' Health and Fitness. Am J Med 2016 Jan;129(1):11-19. [doi: 10.1016/j.amjmed.2015.05.038] [Medline: 26091764]
29. Arora S, Peters AL, Agy C, Menchine M. A mobile health intervention for inner city patients with poorly controlled diabetes: proof-of-concept of the TExT-MED program. Diabetes Technol Ther 2012 Jun;14(6):492-496. [doi: 10.1089/dia.2011.0252] [Medline: 22524591]

30. Cordova D, Bauermeister JA, Fessler K, Delva J, Nelson A, Nurenberg R, Youth Leadership Council. A Community-Engaged Approach to Developing an mHealth HIV/STI and Drug Abuse Preventive Intervention for Primary Care: A Qualitative Study. JMIR Mhealth Uhealth 2015 Dec 18;3(4):e106 [FREE Full text] [doi: 10.2196/mhealth.4620] [Medline: 26685288]
31. Stephan LS, Dytz Almeida E, Guimaraes RB, Ley AG, Mathias RG, Assis MV, et al. Processes and Recommendations for Creating mHealth Apps for Low-Income Populations. JMIR Mhealth Uhealth 2017 Apr 03;5(4):e41 [FREE Full text] [doi: 10.2196/mhealth.6510] [Medline: 28373155]
32. Tamony P, Holt R, Barnard K. The Role of Mobile Applications in Improving Alcohol Health Literacy in Young Adults With Type 1 Diabetes: Help or Hindrance? J Diabetes Sci Technol 2015 Aug 06;9(6):1313-1320 [FREE Full text] [doi: 10.1177/1932296815588559] [Medline: 26251369]
33. Buchholz SW, Ingram D, Wilbur J, Pelt P. Using photos to develop text messages to promote walking. J Nurs Scholarsh 2013 Dec;45(4):380-387. [doi: 10.1111/jnu.12043] [Medline: 23870429]
34. Burner ER, Menchine MD, Kubicek K, Robles M, Arora S. Perceptions of successful cues to action and opportunities to augment behavioral triggers in diabetes self-management: qualitative analysis of a mobile intervention for low-income Latinos with diabetes. J Med Internet Res 2014 Jan 29;16(1):e25 [FREE Full text] [doi: 10.2196/jmir.2881] [Medline: 24476784]
35. Connelly K, Siek KA, Chaudry B, Jones J, Astroth K, Welch JL. An offline mobile nutrition monitoring intervention for varying-literacy patients receiving hemodialysis: a pilot study examining usage and usability. J Am Med Inform Assoc 2012;19(5):705-712 [FREE Full text] [doi: 10.1136/amiajnl-2011-000732] [Medline: 22582206]
36. Wen K, Miller SM, Kilby L, Fleisher L, Belton TD, Roy G, et al. Preventing postpartum smoking relapse among inner city women: development of a theory-based and evidence-guided text messaging intervention. JMIR Res Protoc 2014 Apr 03;3(2):e20 [FREE Full text] [doi: 10.2196/resprot.3059] [Medline: 24698804]
37. Chaudry B, Connelly K, Siek K, Welch J. The design of a mobile portion size estimation interface for a low literacy population. 2011 Presented at: 5th International ICST Conference on Pervasive Computing Technologies for Healthcare; May 23-26, 2011; Dublin. [doi: 10.4108/icst.pervasivehealth.2011.246113]
38. Nundy S, Mishra A, Hogan P, Lee SM, Solomon MC, Peek ME. How do mobile phone diabetes programs drive behavior change? Evidence from a mixed methods observational cohort study. Diabetes Educ 2014;40(6):806-819 [FREE Full text] [doi: 10.1177/0145721714551992] [Medline: 25278512]
39. Segall N, Saville J, L'Engle P, Carlson B, Wright MC, Schulman K, et al. Usability evaluation of a personal health record. AMIA Annu Symp Proc 2011;2011:1233-1242 [FREE Full text] [Medline: 22195184]
40. Tirado M. Role of mobile health in the care of culturally and linguistically diverse US populations. Perspect Health Inf Manag 2011 Jan 01;8:1e [FREE Full text] [Medline: 21307988]
41. Schnall R, Bakken S, Rojas M, Travers J, Carballo-Dieguez A. mHealth Technology as a Persuasive Tool for Treatment, Care and Management of Persons Living with HIV. AIDS Behav 2015 Jun;19 Suppl 2:81-89 [FREE Full text] [doi: 10.1007/s10461-014-0984-8] [Medline: 25572830]
42. Grimes A, Bednar M, Bolter J, Grinter R. EatWell: sharing nutrition-related memories in a low-income community. 2008 Presented at: CSCW08: Computer Supported Cooperative Work; November 8-12, 2008; San Diego, CA. [doi: 10.1145/1460563.1460579]

Abbreviations

ADA: Americans with Disabilities Act
mHealth: mobile health

Edited by A Mavragani; submitted 26.04.21; peer-reviewed by B Raiff, H Ranjani; comments to author 26.07.21; revised version received 19.09.21; accepted 16.12.21; published 11.02.22.

Please cite as:
Sharma S, Gergen Barnett K, Maypole J(, Grochow Mishuris R
Evaluation of mHealth Apps for Diverse, Low-Income Patient Populations: Framework Development and Application Study
JMIR Form Res 2022;6(2):e29922
URL: https://formative.jmir.org/2022/2/e29922
doi:10.2196/29922
PMID:35147502

Original Paper

Perceptions on the Use of Wearable Sensors and Continuous Monitoring in Surgical Patients: Interview Study Among Surgical Staff

Meera Joshi[1], PhD; Stephanie Archer[1], PhD; Abigail Morbi[1], MBBS; Hutan Ashrafian[1], PhD; Sonal Arora[2], PhD; Sadia Khan[2], DM; Graham Cooke[3], FRCP; Ara Darzi[1], FRS

[1]Department of Surgery & Cancer, Imperial College London, London, United Kingdom

[2]Chelsea and Westminster Hospital, National Health Service Foundation Trust, London, United Kingdom

[3]Department of Infectious Diseases, Imperial College London, London, United Kingdom

Corresponding Author:
Meera Joshi, PhD
Department of Surgery & Cancer
Imperial College London
Academic Surgical Unit, 10th Floor, QEQM Building, St Mary's Hospital
Praed Street
London, W2 1NY
United Kingdom
Phone: 44 020 3312 6666
Email: meera.joshi03@imperial.ac.uk

Abstract

Background: Continuous vital sign monitoring by using wearable sensors may result in the earlier detection of patient deterioration and sepsis. Few studies have explored the perspectives of surgical team members on the use of such sensors in surgical patients.

Objective: This study aims to understand the views of surgical team members regarding novel wearable sensors for surgical patients.

Methods: Wearable sensors that monitor vital signs (heart rate, respiratory rate, and temperature) continuously were used by acute surgical patients. The opinions of surgical staff who were treating patients with these sensors were collated through in-depth semistructured interviews to thematic saturation. Interviews were audio recorded, transcribed, and analyzed via thematic analysis.

Results: A total of 48 interviews were performed with senior and junior surgeons and senior and junior nurses. The main themes of interest that emerged from the interviews were (1) problems with current monitoring, (2) the anticipated impact of wearables on patient safety, (3) the impact on staff, (4) the impact on patients overall, (5) potential new changes, and (6) the future and views on technology.

Conclusions: Overall, the feedback from staff who were continuously monitoring surgical patients via wearable sensors was positive, and relatively few concerns were raised. Surgical staff members identify problems with current monitoring and anticipate that sensors will both improve patient safety and be the future of monitoring.

(JMIR Form Res 2022;6(2):e27866) doi:10.2196/27866

KEYWORDS

staff feedback; interview; sensors; continuous monitoring; mobile phone

Introduction

The failure to recognize and respond to patient deterioration is a major cause of morbidity and mortality and is predominantly caused by human monitoring factors [1-3]. Patients undergoing major surgery are at risk of life-threatening complications. The earlier recognition of deterioration can improve survival and may also reduce patients' length of hospital stay and the need for higher acuity care [4]. The detection of deterioration occurs by measuring vital signs routinely every 4 to 6 hours or more frequently in patients who are identified as unwell. Until now, continuous monitoring has only been feasible in higher dependency settings via invasive methods. However, the latest

lightweight sensors have the potential to be used for the continuous monitoring of all hospitalized patients.

Previous studies have reviewed the reliability of wearable devices [5,6], but further research is needed to understand the performance of these devices [7] and staff perspectives. There are very few studies that have reviewed staff experiences [8]; a systematic review of patients' and staff's experiences with wearable sensors found a lack of high-quality studies in this subject area [7]. The studies reviewed had a small sample size (most had less than 20 participants), and their reporting of the research process was limited [7]. Further, the end user outcomes were often a secondary aim rather than the primary focus.

In order to assess the acceptability of wearable sensors, the opinions of all who use technology must be reviewed. We anticipated that feedback may be dependent on role and that wearable sensors would have the greatest impact on nurses, as they are the ones who place sensors on patients and are the first to be alerted about deterioration. This study aims to conduct a comprehensive exploration of interdisciplinary staff views on wearable sensor technology for surgical patients.

Methods

Study Design

Semistructured interviews were held for a subgroup of surgical health care staff to explore their experiences with and perceptions of wearable sensors in detail.

Ethical Approval

Ethical approval was granted by Yorkshire & The Humber - Leeds East Research Ethics Committee (reference number: 17/YH/0296).

Participants and Setting

All members of staff had been treating patients wearing a wearable sensor in the wearable patch at West Middlesex University Hospital—a busy hospital located in northwest London that serves an ethnically diverse population. Interviews were conducted starting in March 2018 and were discontinued once preliminary thematic saturation had been attained [9,10].

Sensium Sensor

Acutely unwell patients admitted to surgical wards were offered the Sensium Vitals (The Surgical Company) wearable sensor in addition to standard ward vital sign monitoring by nurses. The sensor is lightweight; measures heart rate, respiratory rate, and temperature every 2 minutes; and has a battery life of 5 days. For longer hospitalization periods, an additional sensor is required. Sensor data flow from the sensor to a web-based server via a bridge before they are sent via Wi-Fi to mobile apps and smartphone devices.

In Figure 1, the sensor placement on a patient's chest can be seen. The sensor was placed by either the trained health care professionals who were looking after the patients or the research team. The patch was attached to the anterior chest wall by using 2 standard disposable electrocardiogram electrodes (Red-Dot2560; 3M Company). Medical tape was used to ensure that the temperature probe was secured in the axilla.

A plastic strip was pulled to activate the sensor. The sensor recorded in a sequential, cyclical, 2-minute fashion. A predictive strategy was used to calculate heart rate based on the RR interval [11]. The RR interval is the time that elapses between 2 successive R waves in a QRS signal on an electrocardiogram. This approach has been described previously [12, 13]. The individual RR intervals from the electrocardiogram strip were rank ordered, and the median value was taken as the average heart rate. Impedance pneumography was the technique used by the Sensium sensor to measure respiratory rate. This is a common technique that is used to measure a person's breathing rate [14]. Impedance was measured through superficial electrodes. The impedance measures both the respiratory volume and the respiratory rate via the relationship between depth and thoracic impedance change [15]. The respiratory rates were derived from changes in the impedance of the thorax due to inhalation and exhalation. A very small current (iK) was injected through the electrocardiogram electrodes. The thoracic impedance changes were detected as variations in the voltage (V) measured at the electrocardiogram electrodes. Inhalation (peak resistance) and exhalation (trough resistance) were detected from a 60-second segment of the impedance pneumography waveform to calculate a median respiratory rate. Temperature was measured by using a calibrated thermistor (ie, a temperature-sensitive resistor; Braun ThermoScan PRO 6000; Welch Allyn Braun), which was placed in patients' axillae. Individual vital sign readings were measured and processed in time order.

Once the sensor captured physiological data, it used an algorithm, which was stored in a microchip, to process the data. This microchip had a built-in processing unit that transmitted the average values for heart rate as beats per minute and those for respiratory rate as breaths per minute to the nearest bridge. These data were then transmitted to the central server [11], thus allowing digital alerts to be sent to health care staff through smartphones or electronic health records (Figure 2).

Figure 1. Properties of the wearable sensor. The picture shows one of the sensors being worn on a patient's chest. The image was reproduced with permission from Sensium (Abingdon, United Kingdom).

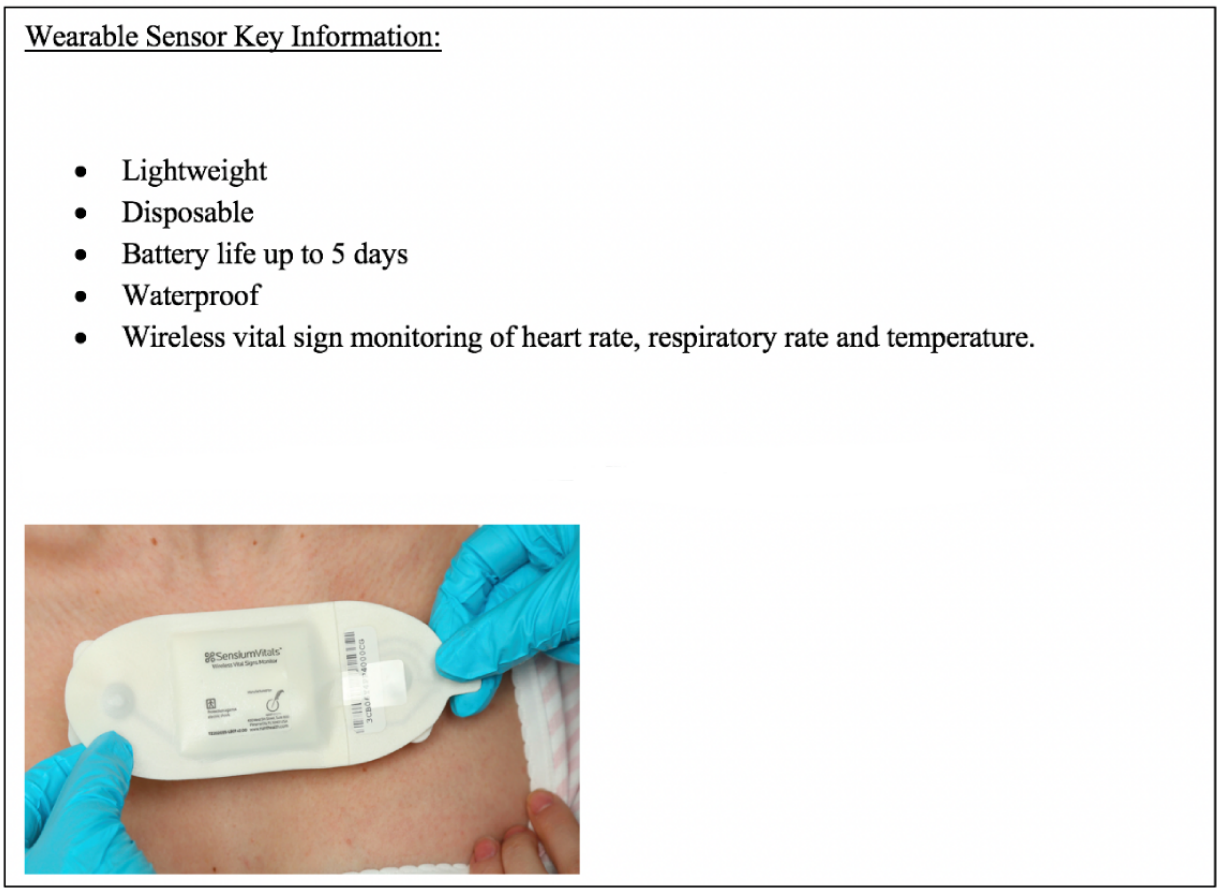
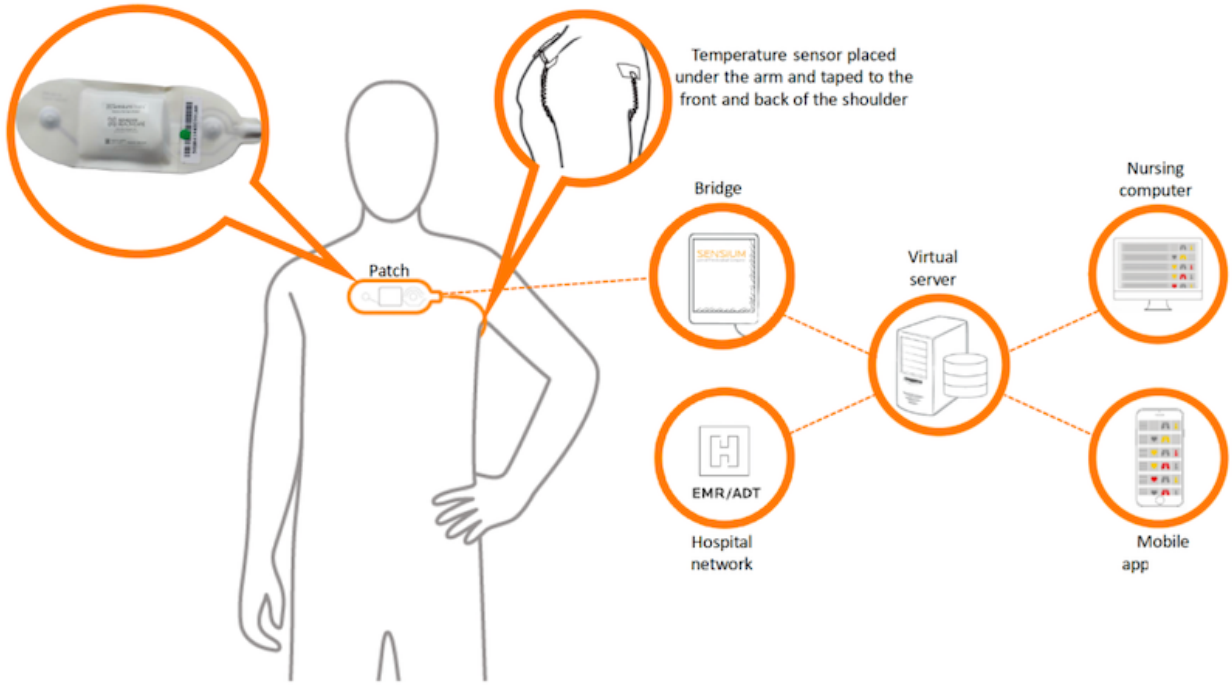


Figure 2. Sensium wearable sensor data transmission to the server and then to mobile apps or computers. The image was reproduced with permission from Sensium (Abingdon, United Kingdom).



Data security is of paramount importance when using wearable technology. The Sensium system is International Organization for Standardization 27001 compliant (information security management), safe, and secure. The Sensium patches are uniquely identified by means of a machine-readable serial number, and these numbers can be matched to a patient ID band

on the Sensium server via a bar code scanner. No patient identifiable information was being communicated from the Sensium patch to the Sensium bridge; just the serial numbers of the devices and values for heart rate, respiratory rate, and temperature were transmitted. Once the information was transferred from the Sensium bridge to the secure Sensium server, only then were the values from the patch put into context with patient identifiable demographic information, usually with the help of a patient administration system. The Sensium patch transmits data to the Sensium bridge every 2 minutes and receives a positive acknowledgement back from the Sensium server when the data have been received. If the patch is out of range of a Sensium bridge or if no acknowledgement is received from the server, the patch continues to attempt communication until it is successful. The Sensium patch stores up to 3 hours' worth of data locally and passes this information to a Sensium bridge once it is back in range.

Once the wearable sensor had been used in the surgical wards for 3 months, semistructured interviews with staff members were conducted. At this stage, the sensor did not alert staff members in real time, and sensor data were only available to the research team.

A wide range of staff participants were sampled, including junior and senior nurses as well as interns/attendings (senior surgeons), to ensure that an interdisciplinary assessment was conducted [16]. All staff members interviewed had applied the wearable sensor, undergone training, or managed a patient that had worn the sensor. The sensor did not alert staff in real time and was not part of day-to-day care; it was only used for research purposes.

Qualitative Data Analyses

Semistructured interviews were conducted to allow for the in-depth exploration of staff perceptions. The interview guide was developed through an extensive literature search and previous piloting. Face-to-face interviews were conducted by the lead researcher (MJ)—a surgical resident with no personal or professional ties to the interviewees.

Interviews were audio recorded and transcribed verbatim. Interview data were analyzed by using thematic analysis [17]. All transcripts were reviewed multiple times and coded by a second independent researcher (AM); any discrepancies in thematic codes were discussed until agreement was reached.

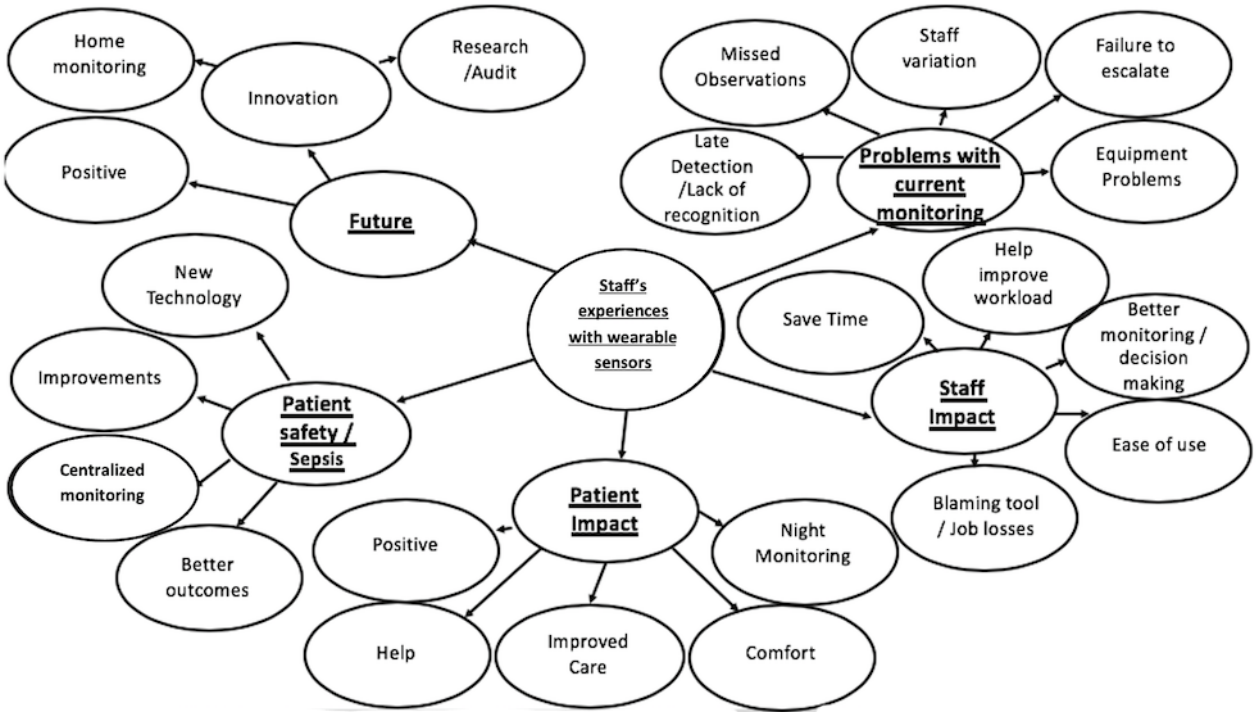
Results

Staff Characteristics

A total of 48 interviews were performed with 12 senior surgeons/attendings (experience: mean 19 years; range 12-44 years), 12 junior physicians (experience: mean 2 years; range 1-5 years), 12 senior nurses (experience: mean 14 years; range 5-20 years), and 12 junior nurses (experience: mean 9 years; range 1-20 years). The staff demographics can be found in Multimedia Appendix 1.

Figure 3 details the themes and subthemes resulting from the thematic analysis; these are expounded in the following sections and illustrated with verbatim quotations.

Figure 3. Themes and subthemes from the thematic analysis.



Problems With Current Monitoring

Staff identified problems with current monitoring; patient deterioration was often identified late. Investigations into adverse outcomes or significant harm in patients found that patient deterioration signs were present several hours before they were detected clinically:

> *Deteriorating patients are often picked late, despite all of the scoring systems that we have. Actually, when you look back at cause analysis and a lot of deterioration, there's always a sign there that was present hours before they were picked up.* [Senior surgeon #11]

Staff suggested that patient deterioration was sometimes overlooked between observations and compounded by missed observations. It was noted that variations in the training and education of junior staff might result in delayed or missed observations. Agency staff were also associated with missed observations:

> *We had a healthcare assistant who was agency, doesn't really work here much, where we thought things were up to date and when we went back to check, an observation hadn't been done, that patient had spiked [temperature], so obviously there's that time period in between we could have been acting sooner.* [Senior nurse #11]

Another concern was a lack of the recognition of an unwell patient. Senior surgeons identified this as being a common problem with inaccurate or incomplete patient assessments and a failure to escalate, even in patients with sepsis:

> *The lack of recognising the development of sepsis and the lack of recognising deteriorating NEWS [national early warning] scores is a relatively common theme amongst serious incident patients.* [Senior surgeon #2]

Staff thought that a lack of confidence might result in a failure to escalate; the greater and sustained education of staff members may combat this issue:

> *Part of the whole sepsis story, it's all about education of staff, being able to have the confidence to institute treatment, I see it so often where the patient's deterioration is accurately monitored but then nothing's done about it.* [Senior surgeon #8]

Practical reasons for failures in monitoring and patient deterioration detection included staff shortages and the resulting additional demands placed on existing staff members. Staff stated that during night shifts, the duties of both nurses and physicians were often stretched. This was particularly true for staff who covered other people's allocated breaks:

> *If on my break or during the night-time it's only one HCA for two bays. It's meaning if I'm on my break definitely nobody is doing [observations] instead of me.* [Junior nurse #12]

Nursing staff also cited faulty equipment and the lack of enough observation machines as compounding problems.

Patient Safety

All staff believed that using wearable sensors and conducting continuous monitoring would improve patient safety. They anticipated that wearable sensors would identify unwell patients earlier and result in a reduction in error:

> *I think it's really good that it's happening a lot quicker and obviously using technology so that the nurses can be alerted if something sort of is going wrong so that even the more junior nurses have that sort of prompt.* [Junior physician #5]

Staff welcomed the use of digital alerts and felt that they would provide encouragement for raising an alarm should a problem arise:

> *It's an element of probably giving people I suppose empowerment is quite a good word, but it gives them more confidence to make decisions.* [Senior surgeon #8]

Staff stated that although observation charts in most hospitals are still kept at the patients' bedsides, centralized monitoring allowed all staff, regardless of their location, to identify unwell patients:

> *Like, if you are not physically there, present, and you can see anywhere in the screen if the patient is unwell or is deterioration in the patient observation.* [Senior nurse #3]

Staff also believed that the abundance of extra data captured allowed for further interpretation through trend analysis rather than just a single set of observations taken at 1 time point:

> *Would be quite interested to see trends in things like their heart rate for example, when they are recovering from an operation, it might give them peace of mind.* [Junior physician #12]

Staff felt that the new technology would be more accurate than their current monitoring technology and could be used to reduce variability when taking observations. In particular, the staff highlighted the sensor's ability to detect respiratory rate as a great strength:

> *Well, we were discussing about the respiratory rate, it would be great to have a way of actually accurately measuring respiratory rate.* [Junior physician #7]

Faster detection, escalation, and treatment were also expected to result in better care and overall reductions in morbidity and mortality. Many staff highlighted the importance of the sensor in the timely identification of sepsis:

> *Now we have a patient with sepsis, and we didn't find out at the beginning. With this machine maybe we should get the result before that, so that's why it's much better than to find that later in every four hours.* [Junior nurse #12]

In addition to the sensor's potential benefits in surgery, staff acknowledged that other high-risk patient groups (eg, older patients) might benefit from wearable sensors and that such sensors improve the overall profile of patient safety, as they can help with making suggestions for a culture change. For instance, senior surgeon #11 said, "It's also a cultural change of actually acting on the abnormalities earlier rather than later."

Impact on Staff

Overall, staff thought that the new technology would positively benefit those working in a surgical setting and expressed their

optimism about the sensor's potential to save time, particularly in settings where the demands are great. Junior physician #1 said, "It also saves doctors time as well in the sense that they know what's appropriate and what they actually need to go." Nursing staff also talked about how the sensors helped them to prioritize patients and allocate care more effectively:

> *It gives us an idea of which patients we need to be looking at, more promptly, who we need to be directing the nurses to, who we need to be sort of escalating more quickly.* [Senior nurse #8]

Another crucial strength of wearable sensors that was identified was the ease of use and interpretation. For example, senior surgeon #8 said, "It's not difficult. Stick it on, plug it in." However, although the impact of wearable sensors on staff workload was generally very positive, some junior physicians raised concerns that the workload of nursing staff may increase if too many alerts are generated:

> *When the nursing staff are busy, and they are having to deal with more alerts and things like that. If it is going off quite a bit, it is going to be difficult to triage, it might end up having to have a protocol for it I think.* [Junior physician #11]

Other junior physicians voiced concern about the use of the new technology in incident investigation and its potential use in litigation:

> *It may become a blaming tool because it sort of you know, registered it as a problem, but then if it's not been acted on, then it becomes somewhere where it's more like a legal sort of thing and people could use it towards that.* [Junior physician #3]

The final concern among junior physicians was that if the technology is too effective, staff who conduct observations regularly may lose their job:

> *There's always a new problem it might actually put some people's job in danger for example Healthcare Assistants, there's no need to do observations on a regular basis because this thing does it for you.* [Junior physician #3]

Impact on Patients

Staff generally believed that wearable sensors would be more comfortable for patients than their current monitoring technology:

> *I never heard any patient complain, if they had any allergy or they found it uncomfortable, I never heard any patient complain about it.* [Junior nurse #3]

Staff also observed that patients could carry out their normal day-to-day tasks, like bathing. This complemented the discreet nature of the patch:

> *It doesn't disturb patients much. Most of the ones that I've seen haven't found that – they've even forgotten that it's on them.* [Senior surgeon #12]

As well as being discrete, the sensor was described as being noninvasive. For example, staff reported that the sensors were particularly beneficial for monitoring patients at night, as they caused less sleep disruption than that caused by their current methods of monitoring.

Changes

A total of 54% (26/48) of participants thought that no changes were needed. Although the size of the patch was not identified as being a problem and was far smaller than the size of the staff's current monitoring technology, all participant groups predicted future miniaturization. Staff participants felt that more parameters should be added in the future to allow for more complete monitoring. Alternative sensors that could be placed on different parts of the body were suggested, such as wrist or necklace sensors. A change to the temperature wire was advised by some of the junior nurses.

The Future

All staff believed that wearable technology was the future and would replace the current monitoring systems. Junior nurse #9 said, "I think the way the future is, everybody will be wearing them." Staff generally thought that wearable sensors were a positive concept and were innovative tools. They welcomed wearable sensors in health care and felt a need for the technology. For instance, junior physician #10 stated, "100%. I think the concept is needed."

Suggestions were made by physicians on the use of wearable sensors in monitoring patients at home following discharge:

> *Also if these could be rolled out to patients who actually go home, so they could monitor themselves at home, rolling it out into maybe a community setting might be something to consider in the future.* [Junior physician #12]

Many participants advocated for the use of monitoring in all patients and felt that this would improve the data quality in clinical and research settings. Further, staff acknowledged that trends in the data, when combined with clinical data, could be used to identify and monitor treatment strategies:

> *So, for instance, if we ever wanted to have a look and see how our treatments are working and what the actual timeline or timeframe is for vital signs to actually normalise, grab all that data and use it for audits as well?* [Junior physician #10]

Discussion

Principal Findings

To our knowledge, this is the first study that provides an interdisciplinary view of wearable sensors for surgical ward patients. All staff groups welcomed the use of wearable sensors and acknowledged deficiencies in current monitoring techniques. Several key themes were identified, including the need for new technology, patient safety, the impact on staff and patients, changes, and future use.

The need for changes to current monitoring methods was clearly identified by delays in identifying unwell patients and a failure to escalate, as reported previously in the literature [2,18-21]. Wider literature shows that clinical deterioration may present several hours prior to an adverse event, and this was identified

by staff members [16]. The nighttime use of wearable sensors in particular was found to be a key strength.

Staffing concerns were raised, including the use of agency staff; variations in junior staff quality; and staff shortages, especially those that occurred overnight. Low numbers of registered nurses and high numbers of admissions per registered nurse have been associated with increased mortality during a hospital admission [22,23]. Problems with staff's current observation technology were also identified; not enough observation machines were available. Staff members highlighted the idea that wearable sensor monitoring may bridge these gaps.

Continuous monitoring was thought to result in the earlier recognition of unwell patients, faster escalation, and faster treatment. Centralized monitoring platforms for the remote monitoring of patients' vital sign information offered staff reassurance. The poor documentation of vital signs has been consistently found for current monitoring techniques, with respiratory rate being the vital sign that is often missed [24]. With current surgical wards being overstretched, any potential assistance from new technologies was embraced and was thought to improve prioritization, multitasking, patient safety culture, and the awareness of unwell patients. The practical applications of small sensors were identified. This was combined with the use of the new technology to allow for easier interpretation, improve documentation, and help with decision-making.

Junior physicians highlighted concerns of increased workload for the nursing staff. Alarm fatigue is a problem in hospitals where many digital alerts are being generated not only from electronic health records but also from other devices [25]. With frequent alarms, there is the chance that staff may become desensitized to them. Sensitive and specific alarms that correctly identify unwell patients must be generated [25]. Novel ways of interpreting such large quantities of data generated from wearable sensors may be required, and one suggestion is to use trend analysis platforms.

All staff felt that wearable technology positively benefits patients and improves care. The sensors were comfortable overall and were well tolerated by patients. In contrast to a previous study [19], staff thought that many patients had forgotten that they were wearing the sensor. However, some improvements were suggested, such as future miniaturization and the addition of other vital sign parameters. These results are similar to those of a previous study on a wrist-worn sensor [26].

Strengths

This is the first study to provide an interdisciplinary view of new continuous remote monitoring technology. Staff participants were recruited from several surgical areas. A clear need for the technology was identified by all staff groups. Staff engagement was high, and feedback has been overwhelmingly positive. Staff also highlighted that patients found that the monitoring technology was comfortable, with many forgetting that they had been wearing the sensor. From our results, we can see that the device was easy to use, and for a continuous monitoring device to be widely adopted and be useful, it must be easy to use [27]. In addition, the modeling of clinical and biological data (partly through the use of wearable sensors) is advancing rapidly and is likely to be a useful adjunct in identifying and predicting physiological changes in critically ill surgical patients [28].

Limitations

This study collected data from a single hospital, and thus the findings may not be generalizable to other hospital settings. Although the Sensium wearable sensor was used in this study, our findings on patient comfort and ease of use for staff may not be transferrable to other sensor technologies. In addition, the wearable patches in this study were being used for research purposes, and real-time alerting was not yet established. It may be appropriate to repeat the interview with staff members in the future after digital alerting is established to determine if their opinions change.

Conclusion

Through our interdisciplinary approach, we found wearable sensors to be a useful tool for identifying unwell patients and improving patient safety largely through the earlier escalation of care and treatment. We hope that this technology will benefit both staff and patients. Although several limitations were identified, wearable sensors have been embraced by staff with the belief that the use of such sensors will be the future of health care.

Acknowledgments

This paper presents independent research funded by the National Institute for Health Research (NIHR) Imperial Biomedical Research Centre, the NIHR Imperial Patient Safety Translational Research Centre, the Royal College of Surgeons England Research Fellowship, and the official charity of Chelsea and Westminster Hospital National Health Service Foundation Trust (CW plus). The views expressed in this publication are those of the authors and not necessarily those of the National Health Service, the NIHR, or the Department of Health and Social Care. GC is supported by an NIHR professorship.

Conflicts of Interest

HA is Chief Scientific Officer, Pre-emptive Medicine and Health Security at Flagship Pioneering.

Multimedia Appendix 1
The demographics of staff interviewed and years of experience.
[DOCX File , 18 KB - formative_v6i2e27866_app1.docx]

References

1. Kause J, Smith G, Prytherch D, Parr M, Flabouris A, Hillman K, Intensive Care Society (UK), Australian and New Zealand Intensive Care Society Clinical Trials Group. A comparison of antecedents to cardiac arrests, deaths and emergency intensive care admissions in Australia and New Zealand, and the United Kingdom--the ACADEMIA study. Resuscitation 2004 Sep;62(3):275-282. [doi: 10.1016/j.resuscitation.2004.05.016] [Medline: 15325446]
2. Centre for Clinical Practice at NICE (UK). Acutely ill patients in hospital: Recognition of and response to acute illness in adults in hospital. National Institute for Health and Clinical Excellence. 2007. URL: https://www.ncbi.nlm.nih.gov/books/NBK45947/pdf/Bookshelf_NBK45947.pdf [accessed 2022-01-04]
3. National Early Warning Score (NEWS) 2: Standardising the assessment of acute-illness severity. Updated report of a working party. Executive summary and recommendations. Royal College of Physicians. 2017 Dec. URL: https://www.rcplondon.ac.uk/projects/outputs/national-early-warning-score-news-2 [accessed 2022-01-04]
4. Vincent JL, Einav S, Pearse R, Jaber S, Kranke P, Overdyk FJ, et al. Improving detection of patient deterioration in the general hospital ward environment. Eur J Anaesthesiol 2018 May;35(5):325-333 [FREE Full text] [doi: 10.1097/EJA.0000000000000798] [Medline: 29474347]
5. 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]
6. Weenk M, van Goor H, Frietman B, Engelen LJ, van Laarhoven CJ, Smit J, et al. Continuous monitoring of vital signs using wearable devices on the general ward: Pilot study. JMIR Mhealth Uhealth 2017 Jul 05;5(7):e91 [FREE Full text] [doi: 10.2196/mhealth.7208] [Medline: 28679490]
7. Bergmann JHM, McGregor AH. Body-worn sensor design: what do patients and clinicians want? Ann Biomed Eng 2011 Sep;39(9):2299-2312. [doi: 10.1007/s10439-011-0339-9] [Medline: 21674260]
8. Prgomet M, Cardona-Morrell M, Nicholson M, Lake R, Long J, Westbrook J, et al. Vital signs monitoring on general wards: clinical staff perceptions of current practices and the planned introduction of continuous monitoring technology. Int J Qual Health Care 2016 Sep;28(4):515-521. [doi: 10.1093/intqhc/mzw062] [Medline: 27317251]
9. Changoor NR, Udyavar NR, Morris MA, Torain M, Mullen JT, Kent TS, et al. Surgeons' perceptions toward providing care for diverse patients: The need for cultural dexterity training. Ann Surg 2019 Feb;269(2):275-282. [doi: 10.1097/SLA.0000000000002560] [Medline: 29095198]
10. Creswell JW. Qualitative Inquiry and Research Design: Choosing Among Five Approaches, Second Edition. Thousand Oaks, California: Sage Publications; 2007.
11. Hernandez-Silveira M, Ahmed K, Ang SS, Zandari F, Mehta T, Weir R, et al. Assessment of the feasibility of an ultra-low power, wireless digital patch for the continuous ambulatory monitoring of vital signs. BMJ Open 2015 May 19;5(5):e006606 [FREE Full text] [doi: 10.1136/bmjopen-2014-006606] [Medline: 25991447]
12. Pan J, Tomkins WJ. A Real-Time QRS Detection Algorithm. IEEE Trans Biomed Eng 1985;32(3):230-236.
13. Hamilton PS, Tomkins WJ. Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database. IEEE Trans Biomed Eng 1986 Dec;33(12):1157-1165.
14. Gupta AK. Respiration rate measurement based on impedance pneumography. Data Acquis Prod. 2011 Feb. URL: https://www.ti.com/lit/an/sbaa181/sbaa181.pdf [accessed 2018-01-23]
15. Pacela AF. Impedance pneumography--a survey of instrumentation techniques. Med Biol Eng 1966 Jan;4(1):1-15. [doi: 10.1007/BF02474783] [Medline: 6003931]
16. Johnston M, Arora S, King D, Stroman L, Darzi A. Escalation of care and failure to rescue: a multicenter, multiprofessional qualitative study. Surgery 2014 Jun;155(6):989-994. [doi: 10.1016/j.surg.2014.01.016] [Medline: 24768480]
17. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
18. An acute problem? National Confidential Enquiry into Patient Outcome and Death. 2005. URL: https://www.ncepod.org.uk/2005report/NCEPOD_Report_2005.pdf [accessed 2022-01-04]
19. Safer care for the acutely ill patient: learning from serious incidents. The fifth report from the Patient Safety Observatory. National Patient Safety Agency. 2007. URL: https://www.baccn.org/static/uploads/resources/NPSA_Report.pdf [accessed 2022-01-04]
20. Recognising and responding appropriately to early signs of deterioration in hospitalised patients. National Patient Safety Agency. 2007 Nov. URL: https://www.patientsafetyoxford.org/wp-content/uploads/2018/03/NPSA-DeteriorPatients.pdf [accessed 2022-01-04]
21. Findlay GP, Shotton H, Kelly K, Mason M. Time to intervene? National Confidential Enquiry into Patient Outcome and Death. 2012. URL: https://www.ncepod.org.uk/2012report1/downloads/CAP_fullreport.pdf [accessed 2022-01-04]
22. Griffiths P, Maruotti A, Saucedo AR, Redfern OC, Ball JE, Briggs J, Missed Care Study Group. Nurse staffing, nursing assistants and hospital mortality: retrospective longitudinal cohort study. BMJ Qual Saf 2019 Aug;28(8):609-617 [FREE Full text] [doi: 10.1136/bmjqs-2018-008043] [Medline: 30514780]

23. Ball JE, Bruyneel L, Aiken LH, Sermeus W, Sloane DM, Rafferty AM, RN4Cast Consortium. Post-operative mortality, missed care and nurse staffing in nine countries: A cross-sectional study. Int J Nurs Stud 2018 Feb;78:10-15 [FREE Full text] [doi: 10.1016/j.ijnurstu.2017.08.004] [Medline: 28844649]
24. Cretikos MA, Bellomo R, Hillman K, Chen J, Finfer S, Flabouris A. Respiratory rate: the neglected vital sign. Med J Aust 2008 Jun 02;188(11):657-659. [doi: 10.5694/j.1326-5377.2008.tb01825.x] [Medline: 18513176]
25. Bonnici T, Orphanidou C, Vallance D, Darrell A, Tarassenko L. Testing of wearable monitors in a real-world hospital environment: What lessons can be learnt? 2012 May 17 Presented at: 2012 Ninth International Conference on Wearable and Implantable Body Sensor Networks; May 9-12, 2012; London, United Kingdom. [doi: 10.1109/bsn.2012.31]
26. Prgomet M, Cardona-Morrell M, Nicholson M, Lake R, Long J, Westbrook J, et al. Vital signs monitoring on general wards: clinical staff perceptions of current practices and the planned introduction of continuous monitoring technology. Int J Qual Health Care 2016 Sep;28(4):515-521. [doi: 10.1093/intqhc/mzw062] [Medline: 27317251]
27. Chan AM, Selvaraj N, Ferdosi N, Narasimhan R. Wireless patch sensor for remote monitoring of heart rate, respiration, activity, and falls. 2013 Sep 26 Presented at: 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); July 3-7, 2013; Osaka, Japan. [doi: 10.1109/embc.2013.6610948]
28. Cohen MJ. Use of models in identification and prediction of physiology in critically ill surgical patients. Br J Surg 2012 Apr;99(4):487-493. [doi: 10.1002/bjs.7798] [Medline: 22287099]

Abbreviations

NIHR: National Institute for Health Research

Edited by G Eysenbach; submitted 05.03.21; peer-reviewed by D Koutsouris, Y Man, M McCorquodale; comments to author 12.08.21; revised version received 08.10.21; accepted 22.11.21; published 11.02.22.

Please cite as:
Joshi M, Archer S, Morbi A, Ashrafian H, Arora S, Khan S, Cooke G, Darzi A
Perceptions on the Use of Wearable Sensors and Continuous Monitoring in Surgical Patients: Interview Study Among Surgical Staff
JMIR Form Res 2022;6(2):e27866
URL: https://formative.jmir.org/2022/2/e27866
doi:10.2196/27866
PMID:35147503

Original Paper

Assessing COVID-19 Health Information on Google Using the Quality Evaluation Scoring Tool (QUEST): Cross-sectional and Readability Analysis

Vismaya S Bachu[1], BS; Heba Mahjoub[1], BS; Albert E Holler[1], BA; Tudor Crihalmeanu[2], MHS; Dheevena M Bachu[3]; Varun Ayyaswami[4], MD; Pearman D Parker[5], MPH, PhD; Arpan V Prabhu[6], MD

[1]School of Medicine, Johns Hopkins University, Baltimore, MD, United States

[2]School of Medicine, West Virginia University, Morgantown, WV, United States

[3]Department of Neuroscience, University of Colorado, Boulder, CO, United States

[4]Department of Internal Medicine, University of Massachusetts Medical School, Worcester, MA, United States

[5]College of Nursing, University of Arkansas for Medical Sciences, Little Rock, AR, United States

[6]Department of Radiation Oncology, University of Arkansas for Medical Sciences, Winthrop P Rockefeller Cancer Institute, Little Rock, AR, United States

Corresponding Author:
Arpan V Prabhu, MD
Department of Radiation Oncology
University of Arkansas for Medical Sciences
Winthrop P Rockefeller Cancer Institute
4130 Shuffield Dr
Little Rock, AR, 72205
United States
Phone: 1 5016644568
Email: avprabhu@uams.edu

Abstract

Background: The COVID-19 pandemic spurred an increase in online information regarding disease spread and symptomatology.

Objective: Our purpose is to systematically assess the quality and readability of articles resulting from frequently Google-searched COVID-19 terms in the United States.

Methods: We used Google Trends to determine the 25 most commonly searched health-related phrases between February 29 and April 30, 2020. The first 30 search results for each term were collected, and articles were analyzed using the Quality Evaluation Scoring Tool (QUEST). Three raters scored each article in authorship, attribution, conflict of interest, currency, complementarity, and tone. A readability analysis was conducted.

Results: Exactly 709 articles were screened, and 195 fulfilled inclusion criteria. The mean article score was 18.4 (SD 2.6) of 28, with 7% (14/189) scoring in the top quartile. National news outlets published the largest share (70/189, 36%) of articles. Peer-reviewed journals attained the highest average QUEST score compared to national/regional news outlets, national/state government sites, and global health organizations (all $P<.05$). The average reading level was 11.7 (SD 1.9, range 5.4-16.9). Only 3 (1.6%) articles were written at the recommended sixth grade level.

Conclusions: COVID-19–related articles are vastly varied in their attributes and levels of bias, and would benefit from revisions for increased readability.

(JMIR Form Res 2022;6(2):e32443) doi:10.2196/32443

KEYWORDS

COVID-19; COVID-19 pandemic; health literacy; readability; QUEST; online health information; cross-sectional; trend; internet; spread; symptom; quality; United States

Introduction

Since the onset of the COVID-19 pandemic, new information is released daily, if not hourly, regarding disease spread, symptomatology, and health and economic consequences. In some cases, news has been rapidly spread only to be contradicted days later. For example, at the beginning of the pandemic, hydroxychloroquine was regularly discussed in lay news and scientific journals alike. Some articles touted a 90% chance of benefit to patients with COVID-19 [1], while peer-reviewed journals soon thereafter released a lack of clinical improvement with use of the drug [2,3]. Given varying accuracy levels of innumerable sources, there is a clear need for standardized quality control of online health information especially in light of current vaccination and other public health campaigns [4].

There was a disjointed public health response, partly due to contradicting information. For example, we now know that universal masking is of the utmost importance in preventing disease transmission, but earlier in the pandemic, it was only recommended for health care professionals [1,3]. These conflicting messages may have left many consumers confused, frustrated, and unsure of what broadcast news channels and online health information to trust. The burden of sorting through the flood of information fell on the consumer and, in many instances, left the consumer feeling paralyzed with information overload and overconcern from frequent use of social media [5]. Furthermore, an analysis of online health information prior to February 6, 2020, showed low quality information relative to several different quality scoring systems, including HONcode, the JAMA benchmark, and the DISCERN instrument [6]. With the prevalence of low-quality information and sudden influx of new conflicting information and associated overwhelming emotion, we felt compelled to analyze the information being consumed by the public.

Google Trends (GT) was used to identify popular COVID-19 search terms and produce a list of related online health articles, after which the Quality Evaluation Scoring Tool (QUEST) was applied to assess validity. QUEST is a verified metric created to assess online health information, or any information available online that patients may read to learn more about their health, in a quantifiable way. It consists of seven questions that numerically measure quality of authorship, attribution, conflict of interest, currency, complementarity, and tone [7]. According to QUEST, a high-quality article is deemed trustworthy and credible, and displays an appropriate level of tone for the reader. We opted for this tool, as opposed to another scale such as DISCERN, because it provides clear guidelines on scoring, with example statements clarifying which articles should receive a score of 0 to 3. Furthermore, the scores are weighted, emphasizing the importance of attribution, conflict of interest, and tone in assessing quality. Though there are many unique tools to analyze online health information, we valued QUEST for its unambiguous scoring and similarity to the US National Library of Medicine's "Medline Plus Guide" in individually judging legitimacy of online health material [8].

In addition to systematically assessing the quality of articles using QUEST, we sought to evaluate the readability of articles resulting from the most frequently Google-searched health-related COVID-19 terms in the United States. Because it is additionally important to recognize the varying degrees of literacy within the public, a readability analysis was performed on each article to compare against the recommended sixth grade reading level for patient health communication materials [9]. Although the production of accurate health information for patients to consume is important, it is equally important for the information to be presented in an understandable manner [10]. We hypothesized that the reading levels of popularly searched health phrases would be too difficult for the average American to understand and that the public was consuming low quality online information regarding COVID-19.

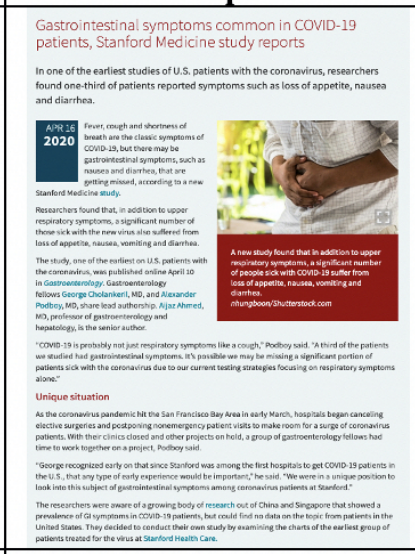
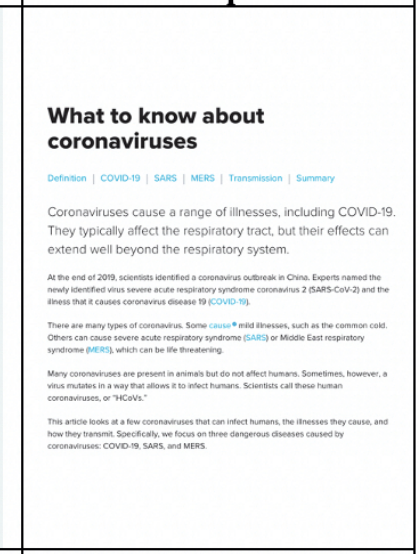
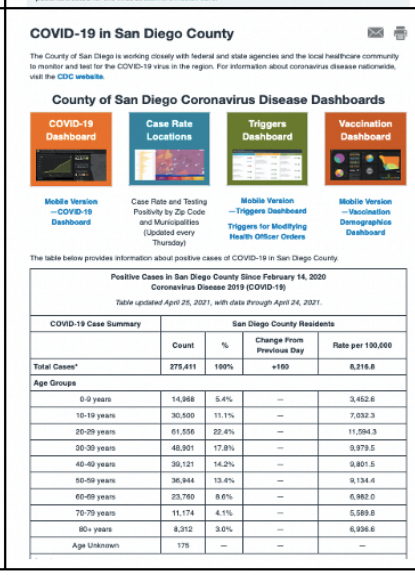
Methods

Article Selection

Institutional review board approval was not required for this study since all information was freely available online. For the purposes of this study, we defined an "article" as being any piece of published writing excluding personal blogs, editorials, and commentaries (Figure 1).

GT has been widely used to capture the most popular queries searched by the public, providing important information regarding emerging patterns. Prior studies have supported the use of GT to monitor COVID-19 incidence and public attention, especially within countries lacking proper diagnostic tools [11,12]. To prevent location bias, online articles were collected using a location-disabled search on Google.com/ncr on April 30, 2020. Using GT, we used the "Explore" option and applied the parameters "United States," "Custom date range 2/29/20 to 4/30/20," "Health," and "Web search." A start date of February 29, 2020, was chosen because this was associated with the first Centers for Disease Control and Prevention–reported death from COVID-19 in the United States [13] and marks an increase in Google searches for the word "coronavirus." From here, we sorted the search queries by "Top" and then collected the top five search queries that had an increased level of Google search frequency (Figure 2.1). These were "coronavirus," "corona," "corona virus," "symptoms,"and "coronavirus update." We then used each of these terms and searched them on GT using the aforementioned methodology. Along with the original five terms, we collected the top five related search queries (including the initially searched word) and excluded any repeat search queries. This resulted in 25 unique health-related search phrases (Figure 2.2). Of note, multi-word keywords were searched using quotations marks, and a comparative analysis of search volume without quotes was not performed.

Figure 1. Description and examples of basic inclusion (article) vs exclusion (nonarticle) criteria. Article, Example 1: “Gastrointestinal Symptoms Common in COVID-19 Patients, Stanford Medicine Study Reports.” Article, Example 2: “What to Know about Coronaviruses.” Nonarticle, Example 1: “COVID-19 in San Diego County.” Nonarticle, Example 2: “Covid-19 Coronavirus Pandemic."

	Description	Example 1	Example 2
Article	We defined “article” as a search result that included a cohesive section of writing that constituted its own separate section on a published page (as opposed to subtitle, caption, etc).	 Gastrointestinal symptoms common in COVID-19 patients, Stanford Medicine study reports In one of the earliest studies of U.S. patients with the coronavirus, researchers found one-third of patients reported symptoms such as loss of appetite, nausea and diarrhea. APR 16 2020 Fever, cough and shortness of breath are the classic symptoms of COVID-19, but there may be gastrointestinal symptoms, such as nausea and diarrhea, that are getting missed, according to a new Stanford Medicine study. Researchers found that, in addition to upper respiratory symptoms, a significant number of those sick with the new virus also suffered from loss of appetite, nausea, vomiting and diarrhea. The study, one of the earliest on U.S. patients with the coronavirus, was published online April 10 in Gastroenterology. Gastroenterology fellows George Cholankeril, MD, and Alexander Podboy, MD, share lead authorship. Aijaz Ahmed, MD, professor of gastroenterology and hepatology, is the senior author. A new study found that in addition to upper respiratory symptoms, a significant number of people sick with COVID-19 suffer from loss of appetite, nausea, vomiting and diarrhea. nhungboon/Shutterstock.com “COVID-19 is probably not just respiratory symptoms like a cough,” Podboy said. “A third of the patients we studied had gastrointestinal symptoms. It’s possible we may be missing a significant portion of patients sick with the coronavirus due to our current testing strategies focusing on respiratory symptoms alone.” Unique situation As the coronavirus pandemic hit the San Francisco Bay Area in early March, hospitals began canceling elective surgeries and postponing nonemergency patient visits to make room for a surge of coronavirus patients. With their clinics closed and other projects on hold, a group of gastroenterology fellows had time to work together on a project, Podboy said. “George recognized early on that since Stanford was among the first hospitals to get COVID-19 patients in the U.S., that any type of early experience would be important,” he said. “We were in a unique position to look into this subject of gastrointestinal symptoms among coronavirus patients at Stanford.” The researchers were aware of a growing body of research out of China and Singapore that showed a prevalence of GI symptoms in COVID-19 patients, but could find no data on the topic from patients in the United States. They decided to conduct their own study by examining the charts of the earliest group of patients treated for the virus at Stanford Health Care.	 What to know about coronaviruses Definition | COVID-19 | SARS | MERS | Transmission | Summary Coronaviruses cause a range of illnesses, including COVID-19. They typically affect the respiratory tract, but their effects can extend well beyond the respiratory system. At the end of 2019, scientists identified a coronavirus outbreak in China. Experts named the newly identified virus severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and the illness that it causes coronavirus disease 19 (COVID-19). There are many types of coronavirus. Some cause mild illnesses, such as the common cold. Others can cause severe acute respiratory syndrome (SARS) or Middle East respiratory syndrome (MERS), which can be life threatening. Many coronaviruses are present in animals but do not affect humans. Sometimes, however, a virus mutates in a way that allows it to infect humans. Scientists call these human coronaviruses, or “HCoVs.” This article looks at a few coronaviruses that can infect humans, the illnesses they cause, and how they transmit. Specifically, we focus on three dangerous diseases caused by coronaviruses: COVID-19, SARS, and MERS.
Non article	We defined “non article” as a search result which did not include a section of cohesive writing. Most non articles were either information dashboards, numerical dashboards, or main menu navigation pages.	 COVID-19 in San Diego County The County of San Diego is working closely with federal and state agencies and the local healthcare community to monitor and test for the COVID-19 virus in the region. For information about coronavirus disease nationwide, visit the CDC website. County of San Diego Coronavirus Disease Dashboards COVID-19 Dashboard Case Rate Locations Triggers Dashboard Vaccination Dashboard Mobile Version —COVID-19 Dashboard Case Rate and Testing Positivity by Zip Code and Municipalities (Updated every Thursday) Mobile Version —Triggers Dashboard Triggers for Modifying Health Officer Orders Mobile Version —Vaccination Demographics Dashboard The table below provides information about positive cases of COVID-19 in San Diego County. Positive Cases in San Diego County Since February 14, 2020 Coronavirus Disease 2019 (COVID-19) Table updated April 25, 2021, with data through April 24, 2021. COVID-19 Case Summary | San Diego County Residents Count | % | Change From Previous Day | Rate per 100,000 Total Cases* 275,411 100% +160 8,216.8 Age Groups 0-9 years 14,968 5.4% — 3,452.6 10-19 years 30,500 11.1% — 7,032.3 20-29 years 61,556 22.4% — 11,594.3 30-39 years 48,901 17.8% — 9,979.5 40-49 years 39,121 14.2% — 9,801.5 50-59 years 36,944 13.4% — 9,134.4 60-69 years 23,760 8.6% — 6,982.0 70-79 years 11,174 4.1% — 5,569.8 80+ years 8,312 3.0% — 6,936.6 Age Unknown 175 — — —	 worldometer Coronavirus Population COVID-19 CORONAVIRUS PANDEMIC Coronavirus Cases: 147,810,736 Deaths: 3,123,003 Recovered: 125,375,246 ACTIVE CASES CLOSED CASES Daily New Cases

Next, between April 30 and May 2, 2020, we searched each keyword phrase and collected all articles (writing that includes more than 100 words) from the first 3 pages of the Google search; this resulted in approximately 10 articles per page (Figure 3). Prior research has shown that internet users tend not to view past these first 3 pages on Google [14]. When queries yielded articles that overlapped, we excluded the repeated articles from analysis. For each article, we collected the Google page number, order on the page, article link, website name, category of website, article title, author, date of publication, and number of references.

Figure 2. (1) Google search words/phrases used to collect most viewed articles. (2) Increase in Google search popularity of the five most commonly searched health-related phrases in the United States between February 29 to April 30, 2020.

1

Health-Related Search Word or Phrase	
Coronavirus	Covid symptoms
Coronavirus cases	Symptoms covid 19
Coronavirus update	Flu symptoms
Coronavirus symptoms	Symptoms of coronavirus
Coronavirus map	Corona virus US
US coronavirus	Corona virus map
Corona	Coronavirus updates
Corona virus	Coronavirus update USA
The corona virus	Coronavirus map update
Corona virus cases	Coronavirus cases update
Corona update	Update on coronavirus
Corona virus update	US coronavirus update
Symptoms	

2

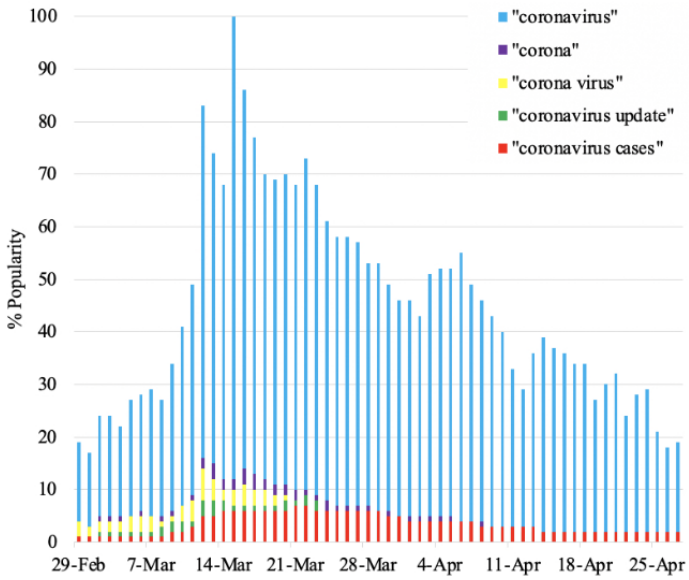
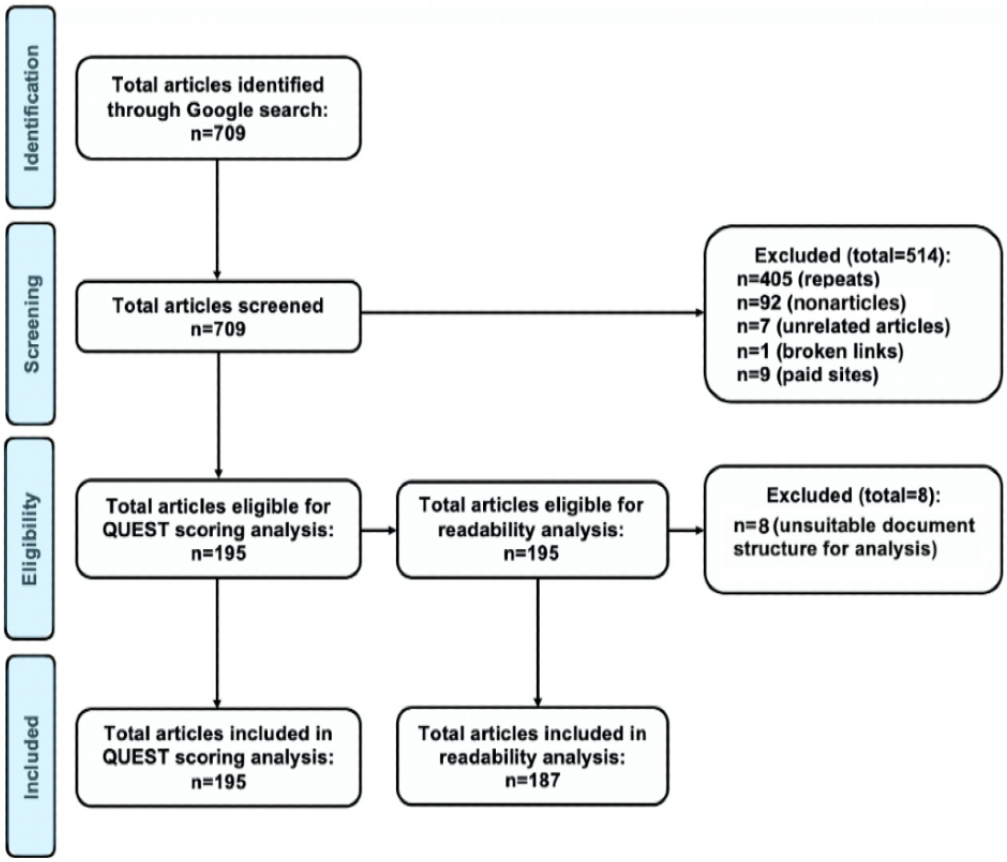


Figure 3. CONSORT (Consolidated Standards of Reporting Trials) diagram depicting article flow and eventual sample size. QUEST: Quality Evaluation Scoring Tool.



QUEST Scoring

Three separate authors scored all articles individually using each of the 7 QUEST questions and associated point values (Textbox 1). Each article's individual sections were then combined into a score between 0 and 28, where 28 was the highest quality article possible. The final score for each article was an average of the three independent scorers' analyses. Interrater consensus was determined using Fleiss's kappa metric.

Textbox 1. The Quality Evaluation Scoring Tool scoring rubric.

Authorship (score x1)

- 0—No indication of authorship or username
- 1—All other indications of authorship
- 2—Author's name and qualification clearly stated

Attribution (score x3)

- 0—No sources
- 1—Mention of expert source, research findings (though with insufficient information to identify the specific studies), links to various sites, advocacy body, or other
- 2—Reference to at least one identifiable scientific study, regardless of format (eg, information in text or reference list)
- 3—Reference to mainly identifiable scientific studies, regardless of format (in >50% of claims)
    - For all articles scoring 2 or 3 on attribution: type of study (score x 1): 0—in vitro, animal models, or editorials; 1—all observational work; 2—meta-analyses, randomized controlled trials, clinical studies

Conflict of interest (score x3)

- 0—Endorsement or promotion of intervention designed to prevent or treat condition (eg, supplements, brain training games, or foods) within the article
- 1—Endorsement or promotion of educational products and services (eg, books or care home services)
- 2—Unbiased information

Currency (score x1)

- 0—No date present
- 1—Article is dated but 5 years or older
- 2—Article is dated within the last 5 years

Complementarity (score x1)

- 0—No support of the patient-physician relationship
- 1—Support of the patient-physician relationship

Tone (includes title; score x3)

- 0—Fully supported (authors fully and unequivocally support the claims, strong vocabulary; eg, "cure," "guarantee," and "easy"; mostly use of nonconditional verb tenses ["can," "will"], no discussion of limitations)
- 1—Mainly supported (authors mainly support their claims but with more cautious vocabulary; eg, "can reduce your risk" or "may help prevent"; no discussion of limitations
- 2—Balanced/cautious support (authors' claims are balanced by caution, includes statements limitations or contrasting findings)

Statistical Analysis

Excel (Microsoft Corporation) was used to conduct the statistical analysis and generate figures for this study. In addition to determining general descriptive metrics (mean, median, etc), we coined the metric search order product to examine if there was any statistical difference in quality between articles toward the beginning and end of the results. Equal to the results page number on Google multiplied by the order of article on that page, the search order product encompasses the article's hierarchy in search results. For example, if an article was second on the third page of the results on Google, its search order product would be 6. A low search order product indicates an earlier appearing article once its associated phrase is searched (meaning increased public exposure), while a high search order product is characteristic of a later-appearing article (decreased public exposure). Comparative *t* tests and Pearson correlation analyses were conducted to stratify scores by variables. The Benjamini-Hochberg false detection rate correction for multiplicity was applied to appropriately adjust P values.

Readability Analysis

Readability analysis was performed using Readability Studio Professional Edition Version 2015 (Oleander Software, Ltd), applying nine validated formulas to quantify article readability: Coleman-Liau Index [15], Flesch-Kincaid Grade Level [16], FORCAST formula [17], Fry graph [18], Gunning Fog Index [19], New Dale-Chall [20], New Fog Count [16], Raygor Reading Estimate [21], and SMOG (Simple Measure of Gobbledygook) [22]. Nine different readability scales were used

to minimize the bias that comes with using only one scale. We subsequently calculated the reading level for each article by averaging estimates derived from all nine scales. These were then compared to the American Medical Association–recommended reading level of sixth grade for health education materials [9]. A 10th readability formula, the Flesch Reading Ease (FRE) [23], was applied separately as it calculates reading level on a different scale. FRE scores of 0 to 30 indicate very difficult, 30 to 50 difficult, 50 to 60 fairly difficult, 60 to 70 standard, 70 to 80 fairly easy, 80 to 90 easy, and 90 to 100 very easy.

Results

QUEST Analysis

A total of 709 Google results listings were initially examined. After excluding repeated articles (n=405), nonarticles (n=92), unrelated articles (n=7), broken links (n=1), and paid sites (n=9) from the analysis, 195 individual articles were scored using QUEST (Figure 3).

The mean article score was 18.4 (SD 2.6) of 28, with only 7% (14/189) of articles in the top score quartile and 89% (173/189) of articles in the top half of scores. National news outlets published the largest share (70/189, 36%) of the analyzed articles, followed by private health-focused entities (45/189, 23%) and regional news outlets (29/189, 15%; Table 1).

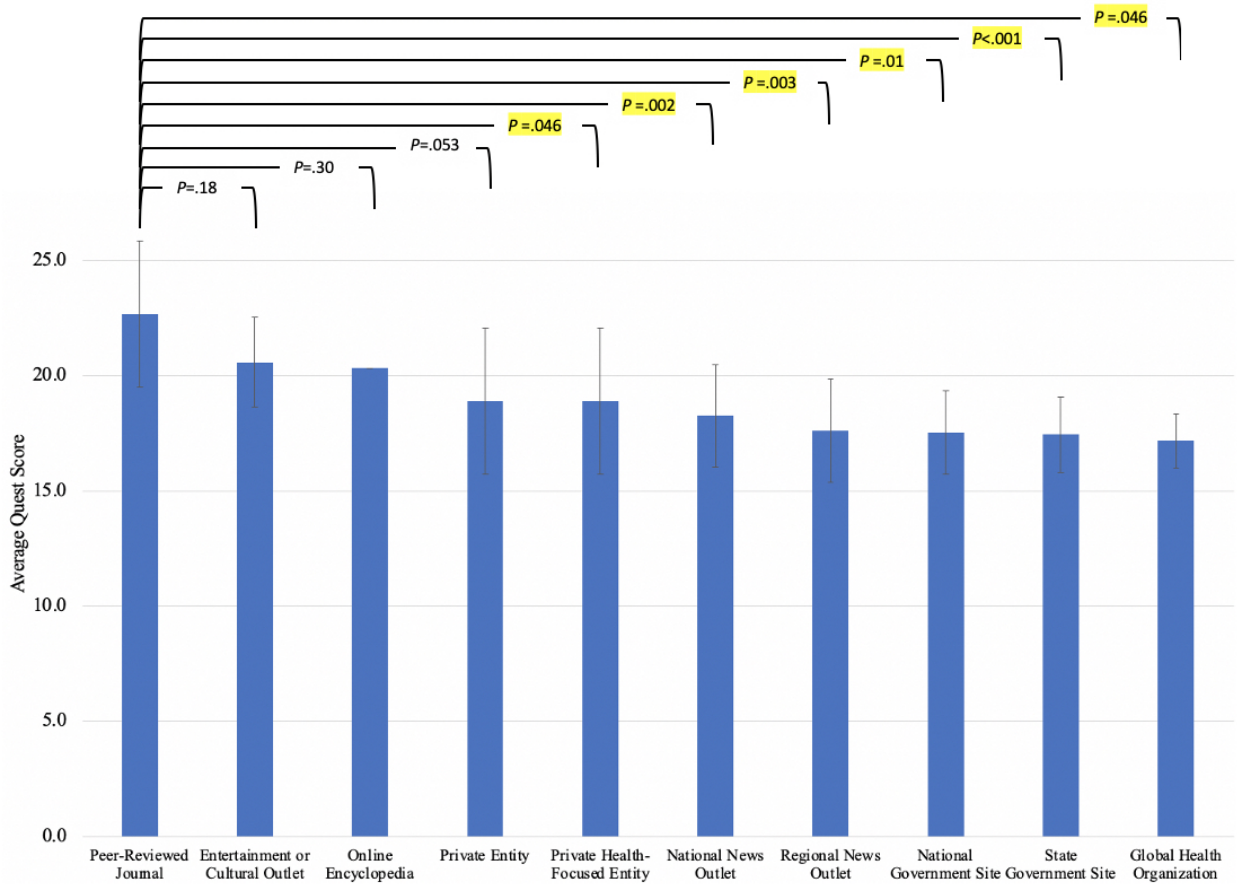
Categorically, global health organization sites had the lowest average score (mean 17.2, SD 1.2) and least dispersive data set (σ^2=1.39; Figure 4). Peer-reviewed journals had the most dispersive data set (σ^2=10.1) and the highest average QUEST score (mean 22.7, SD 3.18), with significantly higher quality averages as compared to national news outlets (mean 18.3, SD 2.20; *P*=.002), regional news outlets (mean 17.6, SD 2.24; *P*=.003), national government sites (mean 17.5, SD 1.81; *P*=.001), state government sites (mean 17.4, SD 1.65, *P*=.002), and global health organizations (mean 17.2, SD 1.18; *P*=.046). In addition, entertainment and cultural outlets (mean 20.6, SD 1.95) also had a significantly higher quality score than regional news outlets (mean 17.6, SD 2.24; *P*=.009) and state government sites (mean 17.4, SD 1.65; *P*=.002).

Table 1. Basic descriptive statistics regarding analyzed articles.

Descriptor	Articles, n (%)
QUEST[a] scoring fraction	
0%-25%	0 (0)
25%-50%	22 (11)
50%-75%	159 (82)
75%-100%	14 (7)
Category	
National news outlet	70 (36)
Private health-focused entity	45 (23)
Regional news outlet	29 (15)
Private entity	18 (9)
State government site	16 (8)
National government site	7 (4)
Peer-reviewed journal	3 (2)
Entertainment or cultural outlet	4 (2)
Global health organization	2 (1)
Online encyclopedia	1 (<1)
Google page	
1	47 (24)
2	76 (39)
3	72 (37)
Order on page	
0-5	106 (54)
6-10	89 (46)

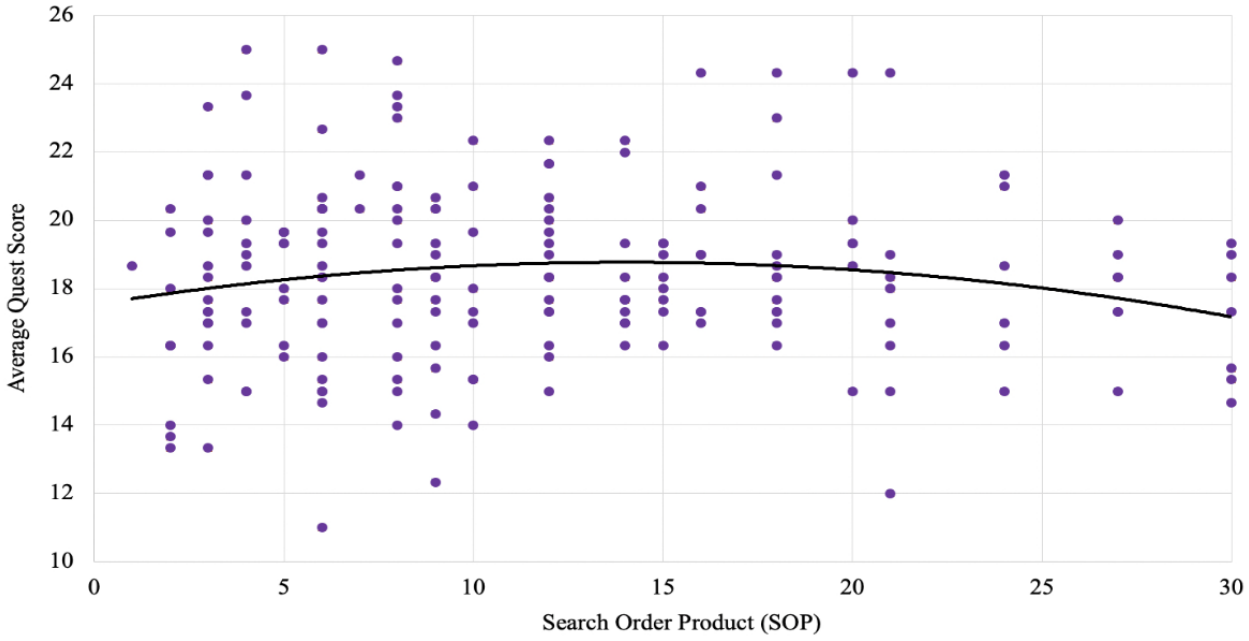
[a]QUEST: Quality Evaluation Scoring Tool.

Figure 4. QUEST score by categorization of 195 articles into the article type. Peer-reviewed journal (PRJ; mean 22.7, SD 2.20); entertainment or cultural outlet (mean 20.6, SD 1.95); online encyclopedia (mean 20.3, SD not applicable); private health-focused entity (mean 18.9, SD 3.17); private entity (mean 18.9, SD 3.17); national news outlet (NNO; mean 18.3, SD 2.2); regional news outlet (RNO; mean 17.6, SD 2.24); national government site (NGS; mean 17.5, SD 1.81); state government site (SGS; mean 17.4, SD 1.65); global health organization (GHO; mean 17.2, SD 1.18). PRJ quality score was significantly higher than NNO (*P*=.002), RNO (*P*=.003), NGS (*P*=.001), SGS (*P*<.001), and GHO (*P*=.046) scores. QUEST: Quality Evaluation Scoring Tool.



Analysis of the QUEST scores by the search order product showed no significant trends (*R*=–0.16; *P*=.75; Figure 5), discrediting any hierarchy by listing order within the sample set. A significant positive correlation (*R*=0.25; *P*<.001) existed between the number of references (≥1) in an article and the QUEST diagnostic score. Due to QUEST already allocating points in a binary fashion for containing references, only articles with ≥1 reference were considered.

Figure 5. QUEST score stratified by search order product (R=–0.16; P>.05); there is no hierarchy in score based on the order of search results. QUEST: Quality Evaluation Scoring Tool.

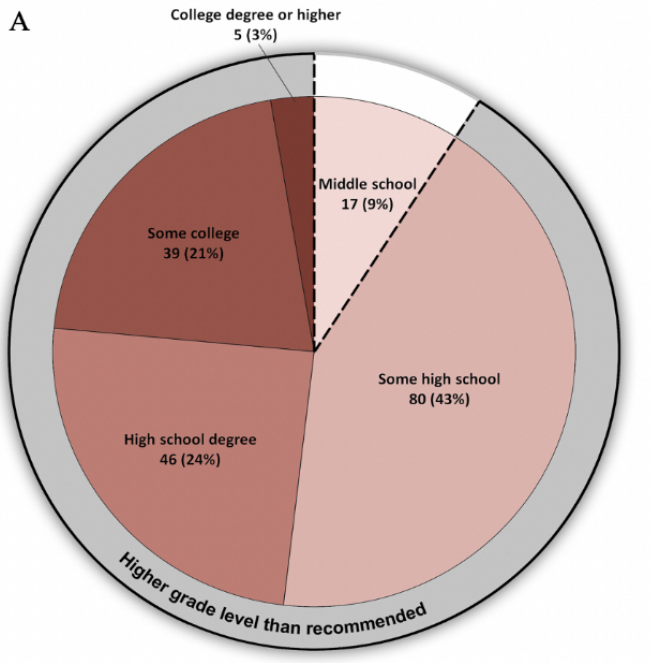


Readability Analysis

The readability levels for 187 of 195 articles were collected; 8 articles containing document structure unsuitable for analysis were excluded. The average reading level across all 187 articles was 11.7 (SD 1.9), ranging from 5.4 (fifth grade reading level) to 16.9 (undergraduate senior reading level). There was variability among the readability scales, with New Fog Count scoring the overall lowest mean readability (mean 9.3, SD 2.7) and Fry scoring the overall highest mean readability (mean 14, SD 2.9). FRE scored an overall average of 47.2 (SD 11.4), corresponding to difficult and representative of college-level reading. Based on the averages of the nine readability scales for each article, only 3 (1.6%) articles were written at the recommended sixth grade levels [9], with 44 (23.5%) written beyond a high school level (Figure 6).

Figure 6. Number of articles (n=187) stratified by educational stages; 187 COVID-19–related articles' readability score breakdown: middle school, <9th grade; some high school, 9th to 11th grade; high school degree, 12th grade; some college, 13th to 15th grade; college degree or higher, ≥16th grade.

A



B

Reading level (Corresponding readability score)	Number of articles
elementary school level (<6)	1
middle school level (6-8.9)	16
high school level (9-12.9)	126
undergraduate level (13-16.9)	44
postgraduate level (>17)	0

Discussion

Analysis of Results

This study systematically assessed and determined that articles resulting from the most frequently Google-searched health-related COVID-19 terms in the United States were of higher quality and readability than hypothesized. QUEST proved versatile in synthesizing aggregate data regarding different aspects of literature including authorship, attribution, conflict of interest, currency, complementarity, and tone. Despite the high prevalence of misinformation on the internet, analysis of our data set revealed that 89% of articles scored in the upper quartiles, suggesting that online information in the United States regarding COVID-19 was of a higher quality than anticipated.

With the uptick in news dissemination by national media after the onset of the pandemic, national news outlets were unsurprisingly the largest source of our sample set followed by private health-focused entities. Interestingly, there was less output from global organizations such as the World Health Organization and United Nations; these organizations only had 1% of total article output, contradicting their organizational goals of far-reaching public health campaigns and initiatives [24]. This discrepancy in expected and observed output could be due to organizational choices to frequently update a centralized information page as opposed to generating new articles that would show up as separate listings. Additionally, different categorical sources allocate varying levels of resources and personnel to public-facing operations that would result in searchable online information [25].

The lack of significant trends associated with an increasing search order product value implied a qualitatively homogenous sampling of articles by exposure in our analysis, validating the decision to analyze only the first three pages of each term's Google search (Figure 5). Using the product of search metrics instead of a direct numerical order of results allowed us to place increased weight on articles listed toward the top of later result pages. Among the sample set, the average quality score of the first articles on the first page (mean 18.7, SD 0.0) was not significantly different from the average quality score of the 10th article on the third page (mean 17.1, SD 1.88) of Google search results.

Stratification of the QUEST scores by article categorization revealed source-based qualitative differences, in part due to different data-gathering and publishing processes. Output from peer-reviewed journals had the highest average score (mean 22.7, SD 3.18), likely due to their rigorous scientific vetting process prior to publication [26]. Increased variation in this category (σ^2=10.1) may be a reflection of the varying requirements set forth by journals and a smaller sample size included in our analysis. Peer-reviewed journals had a significantly increased average QUEST score as compared to national news outlets (mean 18.3, SD 2.20; *P*=.002), regional news outlets (mean 17.6, SD 2.24; *P*=.003), national government sites (mean 17.5, SD 1.81; *P*=.001), state government sites (mean 17.4, SD 1.65; *P*=.002), and global health organizations (mean 17.2, SD 1.18; *P*=.046). Unfortunately, the latter categories were marked by the lowest average information quality, though still mostly in the second (7.5-15) and third (14-22.5) score quartiles. Additionally, entertainment and cultural outlets (mean 20.6, SD 1.95) were characterized by significantly higher quality information compared to regional news outlets and state government sites, perhaps reflecting their tendency to target wider audiences [27].

The majority of articles (170/187, 91%) were written well above American Medical Association–recommended reading levels [21] (Figure 6). Results from the readability analysis substantiated our hypothesis that most COVID-19 articles would be too difficult for the average American to read, in line with the results of prior smaller studies [28]. With the overall FRE score representative of college-level reading and categorically falling under "Difficult," the ability of these most-searched articles to convey accurate information is automatically diminished. Though not directly contributing to misinformation, mismatched comprehension levels lead to knowledge gaps; this may push the public to turn toward other less reliable modalities to stay informed [29]. Ahmed Siddiqui et al [29] specifically discussed the pervasiveness of nonevidence-based medical advice on social media as a "hidden epidemic" considering its ability to transcend geographic and cultural boundaries. Less readable online sources may indirectly facilitate the spread of misinformation regardless of high article quality. This has been affirmed by studies using different criteria from our own including the DISCERN scoring system [6], JAMA benchmarks [30], and even the HONcode system [31]. Additionally, with massive public health awareness efforts underway to encourage COVID-19 vaccination and safe social practices, accurate online media has become increasingly important as a direct source of information for all demographics [32].

Moving forward, publishing sources may benefit from using resources to optimize communication of health information. The Agency for Healthcare Research and Quality updated the second edition of their Health Literacy Universal Precautions Toolkit (HLUPT) recently in September 2020. The document outlines strategies to enhance overall patient understanding, and even contains a section focused on written communication [33]. In a 2015 study, Brega et al [34] determined that applying the sections of the HLUPT pertaining to written materials led to better readability of revised documents. Subjecting patient-facing articles to rigorous quality and literacy guidelines will aide in improving both publishing standards and consumer understanding, both of which are required to best communicate vital information.

Conclusion

The COVID-19 pandemic in the United States was accompanied by an influx of online health information. To investigate the need for quality control of this information, we assessed articles resulting from the most-searched health-related terms in the United States using the QUEST rubric and readability software. Despite the high prevalence and transmission of misinformation during the COVID-19 pandemic, the most frequently searched Google articles had good information quality. Still, the majority of these articles were written above the recommended reading level for the public, diminishing their ability to counteract the spread of misinformation.

Limitations

The limitations of this study include the small sample size, use of only three raters, and lack of individual comparative analysis when determining search keywords. Although GT was used to identify popular keywords, the search volumes of multi-word keywords were not compared against their results without use of quotations. This likely resulted in some bias of listing order because it excluded results only found if only one of the multi-word terms were searched. Furthermore, Google's newer quality rating guidelines adopted in 2019 have resulted in increased automatic filtering criteria, likely resulting in higher quality and personalized results than would have otherwise been listed [35].

Only the QUEST scale was used to measure article validity, and there are a range of other evaluation tools in the literature that may provide differing or complementary data. Additionally, the Fleiss's kappa value of our study was 0.0095, indicating a slight agreement between raters when it came to absolute scores. This may have been due to the subjective nature of the QUEST rubric especially in areas such as attribution, tone, and conflict of interest. Even still, author relative rankings of articles correlated to a greater extent than the absolute score values, indicating a shifted but similarly trending rating among raters. Additionally, facets of QUEST, such as authorship and currency, allocate points for characteristics that do not directly correlate with information accuracy, explaining the lower scores of government and global health organizations due to inherent formatting preferences (ie, omitting authors). The QUEST scale does not address every aspect of misinformation, although it does focus on some aspects such as attribution that is seen in Table 1. Further studies on the spread of misinformation, especially against the backdrop of the COVID-19 pandemic, would benefit from examining media outside articles such as radio, social media, and television.

Acknowledgments

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. PDP is currently supported by the University of Arkansas for Medical Sciences Translational Research Institute grants KL2TR003108 and UL1TR003107 through the National Center for Advancing Translational Sciences of the National Institutes of Health and Arkansas Breast Cancer Research Program.

A supplementary data file (Multimedia Appendix 1) has been provided with the raw data online.

Authors' Contributions

All authors were involved in the design and conception of this manuscript. HM performed the literature search. VSB, AEH, and DMB collected the data. VSB and TC analyzed the articles. VSB, HM, TC, PDP, and AVP wrote the primary manuscript. All authors critically revised the manuscript. AVP supervised the study. All authors have approved the manuscript as it is written.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary data file.
[XLSX File (Microsoft Excel File), 75 KB - formative_v6i2e32443_app1.xlsx]

References

1. Hydroxychloroquine has about 90 percent chance of helping COVID-19 patients. Association of American Physicians and Surgeons. URL: https://aapsonline.org/hcq-90-percent-chance/ [accessed 2020-09-06]
2. Boulware DR, Pullen MF, Bangdiwala AS, Pastick KA, Lofgren SM, Okafor EC, et al. A randomized trial of hydroxychloroquine as postexposure prophylaxis for Covid-19. N Engl J Med 2020 Aug 06;383(6):517-525 [FREE Full text] [doi: 10.1056/NEJMoa2016638] [Medline: 32492293]
3. Skipper CP, Pastick KA, Engen NW, Bangdiwala AS, Abassi M, Lofgren SM, et al. Hydroxychloroquine in nonhospitalized adults with early COVID-19: a randomized trial. Ann Intern Med 2020 Oct 20;173(8):623-631 [FREE Full text] [doi: 10.7326/M20-4207] [Medline: 32673060]
4. Ruiz JB, Bell RA. Understanding vaccination resistance: vaccine search term selection bias and the valence of retrieved information. Vaccine 2014 Oct 07;32(44):5776-5780. [doi: 10.1016/j.vaccine.2014.08.042] [Medline: 25176640]
5. Farooq A, Laato S, Islam AKMN. Impact of online information on self-isolation intention during the COVID-19 pandemic: cross-sectional study. J Med Internet Res 2020 May 06;22(5):e19128 [FREE Full text] [doi: 10.2196/19128] [Medline: 32330115]
6. 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]

7. Robillard JM, Jun JH, Lai J, Feng TL. The QUEST for quality online health information: validation of a short quantitative tool. BMC Med Inform Decis Mak 2018 Oct 19;18(1):87 [FREE Full text] [doi: 10.1186/s12911-018-0668-9] [Medline: 30340488]
8. Evaluating internet health information: a tutorial from the National Library of Medicine. MedlinePlus. URL: https://medlineplus.gov/webeval/webeval.html [accessed 2020-09-06]
9. Weiss BD, Schwartzberg JG, Davis TC, Parker RM, Williams MV, Wang CC. Health Literacy: A Manual for Clinicians. Chicago, IL: American Medical Association Foundation and American Medical Association; 2003.
10. Daraz L, Morrow AS, Ponce OJ, Farah W, Katabi A, Majzoub A, et al. Readability of online health information: a meta-narrative systematic review. Am J Med Qual 2018;33(5):487-492. [doi: 10.1177/1062860617751639] [Medline: 29345143]
11. Ortiz-Martínez Y, Garcia-Robledo JE, Vásquez-Castañeda DL, Bonilla-Aldana DK, Rodriguez-Morales AJ. Can Google® trends predict COVID-19 incidence and help preparedness? The situation in Colombia. Travel Med Infect Dis 2020;37:101703 [FREE Full text] [doi: 10.1016/j.tmaid.2020.101703] [Medline: 32360323]
12. Hu D, Lou X, Xu Z, Meng N, Xie Q, Zhang M, et al. More effective strategies are required to strengthen public awareness of COVID-19: evidence from Google Trends. J Glob Health 2020 Jul;10(1):011003 [FREE Full text] [doi: 10.7189/jogh.10.011003] [Medline: 32373339]
13. CDC, Washington State report first COVID-19 death. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/media/releases/2020/s0229-COVID-19-first-death.html [accessed 2020-09-06]
14. Beitzel SM, Jensen EC, Chowdhury A, Frieder O, Grossman D. Temporal analysis of a very large topically categorized Web query log. J Am Soc Inf Sci Technol 2006 Jan 15;58(2):166-178. [doi: 10.1002/asi.20464]
15. Coleman M, Liau T. A computer readability formula designed for machine scoring. J Appl Psychol 1975 Jul;60(2):283-284. [doi: 10.1037/h0076540]
16. Kincaid JP, Fishburne Jr R, Rogers R, Chissom B, Fishburne JR, Rogers R. Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel. Defense Technical Information Center 1975:1.
17. Caylor J, Sticht T, Fox L, Ford J. Methodologies for determining reading requirements military occupational specialties. Hum Resources Res Organ 1973;7:81.
18. Fry E. A readability formula that saves time. J Reading 1968;11:575-578.
19. Gunning R. The Technique of Clear Writing. New York, NY: McGraw-Hill; 1952.
20. Chall J, Dale E. Readability Revisited: The New Dale-Chall Readability Formula. Brookline, MA: Brookline Books; 1995.
21. Raygor A. The Raygor readability estimate: a quick and easy way to determine difficulty. Read Theory Res Pract 1977:259-263.
22. McLaughlin H. SMOG Grading: a new readability formula. J Reading 1969;12:639-646.
23. Flesch R. A new readability yardstick. J Appl Psychol 1948 Jun;32(3):221-233. [doi: 10.1037/h0057532] [Medline: 18867058]
24. WHO global health days. World Health Organization. URL: https://www.who.int/campaigns [accessed 2020-09-18]
25. Das T. "Basically a house of experts: the production of World Health Organization information". Afr Health Sci 2010 Dec;10(4):390-394 [FREE Full text] [Medline: 21416042]
26. Swartz K. Peer-reviewed journals and quality. Inquiry 1999;36(2):119-121 [FREE Full text]
27. Understanding Media and Culture: An Introduction to Mass Communication. Minnesota: University of Minnesota Libraries; 2016.
28. Basch CH, Mohlman J, Hillyer GC, Garcia P. Public health communication in time of crisis: readability of on-line COVID-19 information. Disaster Med Public Health Prep 2020 Oct;14(5):635-637 [FREE Full text] [doi: 10.1017/dmp.2020.151] [Medline: 32389144]
29. Ahmed Siddiqui MY, Mushtaq K, Mohamed M, Al Soub H, Hussein Mohamedali MG, Yousaf Z. "Social Media Misinformation"-an epidemic within the COVID-19 pandemic. Am J Trop Med Hyg 2020 Aug;103(2):920-921 [FREE Full text] [doi: 10.4269/ajtmh.20-0592] [Medline: 32534600]
30. Fan KS, Ghani SA, Machairas N, Lenti L, Fan KH, Richardson D, et al. COVID-19 prevention and treatment information on the internet: a systematic analysis and quality assessment. BMJ Open 2020 Sep 10;10(9):e040487 [FREE Full text] [doi: 10.1136/bmjopen-2020-040487] [Medline: 32912996]
31. Valizadeh-Haghi S, Khazaal Y, Rahmatizadeh S. Health websites on COVID-19: are they readable and credible enough to help public self-care? J Med Libr Assoc 2021 Jan 01;109(1):75-83 [FREE Full text] [doi: 10.5195/jmla.2021.1020] [Medline: 33424467]
32. Naveed MA, Shaukat R. Health literacy predicts Covid-19 awareness and protective behaviours of university students. Health Info Libr J 2021 Oct 01:10.1111/hir.12404 [FREE Full text] [doi: 10.1111/hir.12404] [Medline: 34595814]
33. Health Literacy Universal Precautions Toolkit, 2nd Edition. Agency for Healthcare Research and Quality. URL: https://www.ahrq.gov/health-literacy/improve/precautions/toolkit.html [accessed 2021-03-13]

34. Brega AG, Freedman MAG, LeBlanc WG, Barnard J, Mabachi NM, Cifuentes M, et al. Using the health literacy universal precautions toolkit to improve the quality of patient materials. J Health Commun 2015;20 Suppl 2:69-76 [FREE Full text] [doi: 10.1080/10810730.2015.1081997] [Medline: 26513033]
35. Google. General Guidelines Overview. Search Quality Rating Program. 2021 Oct 19. URL: https://static.googleusercontent.com/media/guidelines.raterhub.com/en//searchqualityevaluatorguidelines.pdf [accessed 2021-12-01]

Abbreviations

FRE: Flesch Reading Ease
GT: Google Trends
HLUPT: Health Literacy Universal Precautions Toolkit
QUEST: Quality Evaluation Scoring Tool
SMOG: Simple Measure of Gobbledygook

Edited by T Leung; submitted 28.07.21; peer-reviewed by A Mavragani, P Ghezzi; comments to author 10.11.21; revised version received 31.12.21; accepted 06.01.22; published 11.02.22.

Please cite as:
Bachu VS, Mahjoub H, Holler AE, Crihalmeanu T, Bachu DM, Ayyaswami V, Parker PD, Prabhu AV
Assessing COVID-19 Health Information on Google Using the Quality Evaluation Scoring Tool (QUEST): Cross-sectional and Readability Analysis
JMIR Form Res 2022;6(2):e32443
URL: https://formative.jmir.org/2022/2/e32443
doi:10.2196/32443
PMID:34995206

Original Paper

Predictors of Online Patient Portal Use Among a Diverse Sample of Emerging Adults: Cross-sectional Survey

Julie A Wright[1], PhD; Julie E Volkman[2,3], PhD; Suzanne G Leveille[4,5], RN, PhD; Daniel J Amante[3], MPH, PhD

[1]Exercise and Health Sciences, University of Massachusetts Boston, Boston, MA, United States

[2]Department of Communication, Bryant University, Smithfield, RI, United States

[3]Department of Population and Quantitative Health Sciences, UMass Chan Medical School, Worcester, MA, United States

[4]College of Nursing and Health Sciences, University of Massachusetts Boston, Boston, MA, United States

[5]Division of General Medicine, Beth Israel Deaconess Medical Center, Boston, MA, United States

Corresponding Author:
Daniel J Amante, MPH, PhD
Department of Population and Quantitative Health Sciences
UMass Chan Medical School
368 Plantation St
Worcester, MA, 01605
United States
Phone: 1 5088568480
Email: daniel.amante@umassmed.edu

Abstract

Background: Health self-management is increasingly being influenced by emerging health information technologies (IT), especially online patient portals. Patient portals provide patients with direct access to their health information, electronic tools to manage their health, and additional opportunities to engage with their care team. Previous studies have found that patient portal use is highest among patients with high eHealth literacy, the ability to find health information from electronic sources and apply the knowledge gained to solve a health problem. The role of eHealth literacy on patient portal use appears to be especially strong among older adults with chronic diseases. The use of patient portals among emerging adults (ages 18-29) is much less understood. Although generally healthy, emerging adults are more regular IT users and just beginning to independently navigate the health care system. A good understanding of how emerging adults are using online patient portals and what factors, including eHealth, impact portal use is lacking.

Objective: The aim of this study is to describe patient portal use and explore the predictors of portal use among a diverse sample of emerging adults.

Methods: A cross-sectional survey study that used convenience sampling was conducted at two universities. Data on demographics, health care encounters, eHealth literacy, patient engagement, and use of patient portal features (administrative and clinical) were obtained via self-report and summarized. Logistic regression models were used to examine factors associated with portal use.

Results: Of the 340 emerging adults, 257 (76%) were female, 223 (65%) White, 156 (47%) low income, and 184 (54%) reported having patient portal access. Of those reporting access, 142 (77%) used at least 1 portal feature and 42 (23%) reported using none. Significant predictors were patient engagement (odds ratio [OR] 1.08, 95% CI 1.04-1.13, *P*=.001) and total encounters (OR 1.23, 95% CI 1.05-1.44, *P*=.009) but not eHealth literacy. Hispanic and Asian emerging adults were more likely to be frequent users of clinical portal features than White emerging adults (Hispanic: OR 2.97, 95% CI 1.03-8.52, *P*=.04; Asian: OR 4.28, 95% CI 1.08-16.89, *P*=.04).

Conclusions: We found that about half of emerging adults had access to a patient portal. Among those with access, a majority reported using at least one portal feature. Factors associated with increased portal use included increased patient engagement and total clinical encounters. Self-reported eHealth literacy was not associated with patient portal use in this diverse sample of emerging adults. This may have been due to high overall eHealth literacy levels in this population of frequent IT users. There may also be racial/ethnic differences that are important to consider, as we found Hispanic and Asian emerging adults reported more frequent portal use than White emerging adults. Interventions to promote patient portal use among emerging adults should include strategies to increase awareness of portal access and engagement among patients with fewer clinical encounters, with a focus on preventative health management.

(JMIR Form Res 2022;6(2):e33356) doi:10.2196/33356

KEYWORDS

internet; patient portal; emerging adults; portal; predictor; prediction; sample; cross-sectional; survey; usage; young adult; eHealth; literacy

Introduction

Patient self-management of their health is being promoted within health care organizations [1]. Health information technologies (IT) can provide opportunities to enhance patients' self-management [2,3]. One IT that is increasingly adopted by health care systems is the online patient portal. The patient portal allows patients to access and manage their personal health information, request prescription refills, schedule appointments, and message with their health care team [4-6]. The increased availability of online patient portals provides patients with greater access to services and tools to engage in their health management [3,7-10]. Patient portals have demonstrated effectiveness, improving self-management of chronic conditions such as diabetes, hypertension, and depression [3,11] and increasing patient satisfaction [3,12]. Patient portal adoption and use has been shown to vary across patient populations, with increased use associated with patients of White race, those who speak English, and those who have private insurance [13,14]. Older patients and those living in rural areas with poor broadband access have also been shown to be less likely to use a patient portal [15]. An additional factor that could impact the use of patient portals is a patient's eHealth literacy [16]. Extant literature suggests that as a patient's eHealth literacy rises, their ability to use online health resources successfully also increases [3,17,18]. Interventions focused on improving patient portal use could include strategies to enhance eHealth literacy, especially among older patients who use technology less frequently and have greater health management demands due to chronic diseases. Whether such approaches would be helpful for younger, healthier patient populations is less clear.

This study is focused on emerging adults—specifically, higher education students aged 18-29 years old—and examines their use of secure online patient portals. Emerging adults represent a unique demographic, one in a state of transition from their dependency on parents/caregivers, yet not fully engaged in the responsibilities of adulthood [19]. For many emerging adults, university may represent their first independent experiences with the health care system [20]. They may have problems moving into the adult setting after parent-guided pediatric experiences; some may stop their health care altogether during this time of transition [20]. Furthermore, higher education students may have low eHealth literacy skills, despite their familiarity in using the internet to find health information [21,22].

Thus, it is important to understand emerging adults' use of patient portals to inform tailoring of future interventions to improve their engagement with their care and health self-management. Understanding factors associated with online patient portal use will help guide development of interventions for this population. Enabling emerging adults to embrace new tools that increase access to health care services can expedite a culture change in which enhanced patient-provider partnerships lead to more effective care.

The purpose of this study is to describe the usage of online patient portal features among emerging adults. We hypothesized that eHealth literacy and patient engagement would be positively associated with reported portal use.

Methods

Recruitment and Ethics Approval

This cross-sectional survey study used nonprobability convenience sampling to recruit higher education students at a large, urban, public university with many minority students and a small, private university, both located in the Northeastern United States, to complete a cross-sectional survey during the months of March-May 2018 (Multimedia Appendix 1). Convenience sampling was conducted based on the norms associated with student recruitment at each university location and included students being invited via email to participate. For the public university setting, an email was sent via mass listserv to all undergraduate and graduate students. Students had the option to reply to the email and participate in the survey. For the private university setting, emails were sent to the faculty to forward to their students (undergraduate and/or graduate) to participate in the survey.

Email invitations contained a brief description of the purpose of the survey and an embedded survey link, using Qualtrics (April 2018) software, to complete on either a mobile device or desktop computer (web-based). Surveys were anonymous (no identifiers were linked to the respondents' data), used no forced-choice questions, and offered voluntary entry into a random drawing for US $100. At the private university, faculty may offer extra credit as an incentive for participation; however, for this study, faculty members were instructed it was not needed due to the raffle incentive. All study procedures were approved by the UMass Boston and Bryant University Institutional Review Boards (2016148).

Measures

Patient Portal Use

Patient portal use was assessed in two steps. First, participants were primed with an image of a typical health portal sign-in page and were asked to think about where they were currently getting their health care (Multimedia Appendix 2). Participants answered if their health care center offers a secure online patient portal (yes, no, and not sure). The second set of questions assessed their use of 8 common patient portal features. The stem was, "Have you ever used a patient portal for the following reasons?" followed by the 8 features, which were categorized as administrative or clinical in nature based on previous research [23]. The three administrative features included (1) check my appointment date or time, (2) make an appointment, and (3)

request a prescription refill from my physician. The five clinical features were (1) check immunization records, (2) check lab results, (3) email provider, (4) read my visit notes (often called "doctor's notes"), and (5) post my own health information on the patient portal. Response options were (0) No, I don't have access, (1) No, I have never used, (2) Yes, I have used, and (3) Yes, I have used more than once. These items were developed specifically for this study and have not been validated against portal tracking information.

Health Care Utilization

Three items based on Lorig's [24] healthcare utilization questionnaire were used to assess frequency of health care visits in the past 6 months. Participants were asked to report the frequency of visits to a physician's office (not including hospitalizations), hospital emergency department, and urgent care center. The original items were modified by providing response options rather than using an open-ended format. Options ranged from 0 to 6 or more visits. Test-retest reliability has been reported for physician item (r=0.76) and emergency room item (r=0.94) [24,25].

eHealth Literacy

An 8-item modified version of the electronic Health Literacy Scale was used to assess participant eHealth literacy [26]. The modified version used items assessing the critical evaluative skills of information and frequency of engaging in these actions (behavioral literacy). The question stem was, "When looking for health information on the Internet, how often do you do the following?" A total of 8 behaviors were listed and participants were asked to respond on a 6-point frequency scale from never (1) to always (6). The eight behaviors were the following: (1) check the ownership of the health website, (2) check the website's sponsor, (3) evaluate whether the health information is credible, (4) evaluate the credentials of the person providing the information on the website, (5) evaluate whether the coverage of the health topic is comprehensive, (6) check whether other print or web resources confirm the health information provided, (7) check whether the health information is up-to-date, and (8) discuss the health information with your health care provider.

Patient Engagement

The Altarum Consumer Engagement (ACE) Measure was used to assess patient engagement [27]. The ACE contains 12 items and 3 subscales with 4 items each: Commitment (eg, "I can stick with plans to exercise and eat a healthy diet"), Informed Choice (eg, "When choosing a new doctor, I look for official ratings based on patient health"), and Navigation (eg, "I have lots of experience using the healthcare system"). Responses were on a 5-point scale ranging from strongly disagree (0) to strongly agree (4) with a neutral midpoint. The scale has good construct validity with good internal reliability (Cronbach α) for the Informed Choice (α=.82) and Commitment (α=.85) subscales and fair reliability for Navigation (α=.66) [27].

Demographic Questions

Questions to characterize the respondents were asked at the end of the survey and included age, gender, race, ethnicity, language spoken at home, and years in the United States. Respondents were also asked to report on their parent or legal guardian's annual income and the respondent's employment and insurance status.

Data Analysis Plan

Descriptive statistics were calculated to describe the sample's characteristic and summary scores of variables of interest. The sample was defined by only those who self-reported 1 or more health care visits on the Lorig healthcare utilization questionnaire. The sample was further categorized by their use of the portal (nonuser versus user). Users were defined as those who reported using at least one of 8 portal features. Pearson chi-square analyses or independent t tests were used to examine differences by user status. A Levene test for equality of variances was run to determine the t statistic used to evaluate significance.

Multivariable logistic regression was used to examine correlates of portal use and to test our hypotheses that eHealth literacy and patient engagement would be positively associated with reported portal use. Covariates were entered into the model in five blocks: (1) demographics; (2) university type, insurance type, and health condition; (3) total encounters with health care system in last 6 months; (4) eHealth literacy score; and (5) patient engagement score. Participants who reported not having access to a health care portal or not using any patient portal feature were excluded from the analysis. Dichotomous variables were created for age (18-23 years old versus 24-29 years old), gender (female versus not female), ethnicity (Hispanic versus not Hispanic), university (public versus private), insurance (private versus not private), and health condition (condition versus no condition).

A second regression analysis was run to explore predictors of using portal features whose purpose is primarily clinical (eg, read visit notes) rather than administrative (eg, make an appointment) [23]. The outcome variable for this second analysis was dichotomized by high use (3 or more clinical features) versus low use (<3 clinical features). All analyses were completed in SPSS Statistics 25 (IBM Corp).

Results

Figure 1 displays a flow chart of how the analytic samples were defined. Email recruitment resulted in 541 students opening the link to the survey. After data cleaning and exclusions, the remaining analytic sample was 340 emerging adults (Figure 1). The majority of variables had 0-4 (1.2%) missing cases, with one variable missing 6 (2.1%). The majority of the 340 emerging adults (n=184, 54%) reported having access to a patient portal, while 11% (n=39) responded they did not and 34% (n=117) were unsure. We further reduced the sample to include only emerging adults who reported a health care visit in the previous 6 months (n=289). Table 1 describes this sample and classifies each participant as either a nonuser, defined as reporting using no features, or a user, defined as reporting the use of one or more portal features. Using this classification system, there were 124 nonusers and 165 users.

Of the 117 reporting they were unsure about having access to a patient portal, 70.1% (n=82) used none of the features while

29.9% (n=35) used 1 or more features—specifically, 5.1% (n=6) used at least 1 patient portal feature; 13.7% (n=16) used 2-3 patient portal features; and 11.1% (n=13) used 4 or more patient portal features. Of the 165 respondents who reported using portal features, the most used features were checking lab results (n=63, 38%), appointment time (n=61, 37%), and immunization records (n=58, 35%).

Figure 1. Sample flow chart.

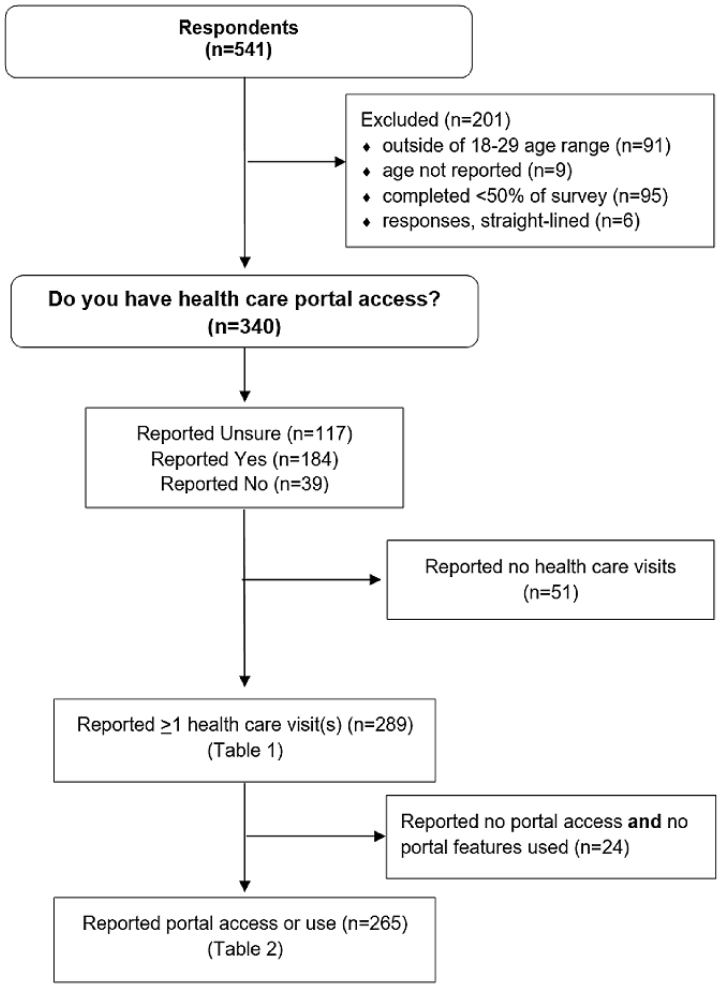


Table 1. Characteristics of students who reported a health care visit in the last 6 months and comparison of nonusers and users of the portal (N=289).

Variable	Nonuser[a] (n=124)	User[b] (n=165)	*P* value[c]
Age group (years), n (%)			*.03*
18-23	107 (86.3)	125 (75.8)	
24-29	17 (13.7)	40 (24.2)	
Gender, n (%)			.30
Female	92 (74.2)	131 (79.4)	
Not female	32 (25.8)	34 (20.6)	
Ethnicity, n (%)			.46
Not Hispanic	109 (87.9)	140 (84.8)	
Hispanic	15 (12.1)	25 (15.2)	
Race, n (%)			.97
White	82 (66.1)	111 (67.3)	
Black	11 (8.9)	15 (9.1)	
Asian	15 (12.1)	17 (10.3)	
Other	16 (12.9)	22 (13.3)	
Years residing in United States, n (%)			.86
0-5 years	9 (7.3)	14 (8.5)	
More than 5 years	113 (91.1)	149 (90.3)	
Prefer not to answer	2 (1.6)	2(1.2)	
Parents' household income (US $), n (%)			.95
<40,000	36 (29.0)	45 (27.3)	
41,000-80,000	39 (31.5)	54 (32.7)	
>80,000	47 (37.9)	62 (37.6)	
Missing	2 (1.6)	4 (2.4)	
University type, n (%)			*.002*
Public	62 (50.0)	112 (67.9)	
Private/not public	62 (50.0)	53 (32.1)	
Insurance type, n (%)			.38
Private	109 (87.9)	139 (84.2)	
Public	15 (12.1)	26 (15.8)	
Do you have a health condition or disease that requires periodic visits or monitoring by a physician? n (%)			.14
No	98 (79.0)	118 (71.5)	
Yes	26 (21.0)	47 (28.5)	
Health care utilization in the previous 6 months, n (%)			
Physician visits (not including emergency department or urgent care)	1.73 (1.16)	2.11 (1.49)	*.02*[d]
Emergency department visits (nonuser n=123, user n=163)	0.22 (0.61)	0.36 (0.82)	.09[d]
Urgent care visits (nonuser n=121, user n=164)	0.48 (0.88)	0.59 (1.00)	.35
Total encounters (sum score)	2.41 (1.80)	3.05 (2.32)	*.009*[d]
e-literacy (sum score, range 8-48), mean (SD)	26.33 (10.45)	27.15 (10.08)	.50
Patient engagement (Altarum Consumer Engagement Measure sum score, range 12-60; nonuser n=123, user n=163), mean (SD)	37.89 (6.79)	41.83 (7.56)	*.001*

[a]Nonusers were defined as those who reported not using any of the 8 portal features.

[b]Users were defined as those who reported using at least one of the 8 portal features.

[c]*P* values for Pearson chi-square test or *t* test. Italicized *P* values indicate significant difference between users and nonusers.

[d]Levene test *P*<.05, equal variances not assumed.

Multivariable logistic regression was used to examine the likelihood of being a user (n=165) versus nonuser (n=100) of the patient portal (Table 2). The respondents (n=265) included only those who reported a health care visit and reported access to a portal or reported they were unsure of access to a portal but reported use of at least one portal feature (Figure 1). Two cases from the sample of 265 were missing patient engagement scores (n=263). The final model included two significant variables, patient engagement (odds ratio [OR] 1.08, 95% CI 1.04-1.13, *P*=.001) and total encounters (OR 1.23, 95% CI 1.05-1.44, *P*=.009).

Table 2. Predictors of using portal features and predictors of high use versus low use of clinical features.

Variables in equation	Users vs nonusers[a] (N=263)		High use vs low use of clinical features[b] (N=263)	
	Adjusted odds ratio[c]	95% CI	Adjusted odds ratio	95% CI
Gender				
Female	1.00		1.00	
Not female	0.89	0.47-1.68	0.840	0.42-1.69
Age (years)				
18-23	1.00		1.00	
24-29	1.43	0.67-3.03	0.63	0.31-1.28
Ethnicity				
Not Hispanic	1.00		1.00	
Hispanic	1.13	0.40-3.20	*2.97*	*1.03-8.52*
Race				
White	1.00		1.00	
Black	0.97	0.37-2.55	2.09	0.678-6.42
Asian	0.75	0.30-1.92	*4.28*	*1.08-16.89*
Other	0.61	0.21-1.79	2.52	0.630-10.05
University type				
Public	1.00		1.00	
Not public	0.625	0.327-2.84	1.65	0.808-3.36
Health insurance type				
Private insurance	1.00		1.00	
Not private	1.20	0.51-2.87	1.03	0.44-2.42
Health condition				
Yes	1.00		1.00	
No	1.08	0.55-2.10	1.37	0.70-2.67
Total health care encounters past 6 months	*1.23*	*1.05-1.44*	*1.16*	*1.01-1.34*
eHealth literacy score	0.99	0.96-1.02	0.97	0.94-1.00
Patient engagement score	*1.08*	*1.04-1.13*	*1.10*	*1.05-1.15*

[a]For this table, nonusers were defined as those who reported not using any of the 8 portal features and users were defined as those who reported using at least one of the 8 portal features.

[b]High users were defined as those who used 3 or more clinical portal features and low users were those who used less than 3 clinical portal features.

[c]Results from multivariable logistic regression models including all variables shown; significant relationships are italicized.

Using the same steps as the first logistic regression models, we examined the outcome variable of high versus low use of clinical portal features. Table 2 displays the results of the regression. Factors associated with high use of clinical portal features included Hispanic versus non-Hispanic ethnicity (OR 2.97, 95% CI 1.03-8.52, *P*=.04), Asian versus non-Asian (OR 4.28, 95% CI 1.08-16.8, *P*=.04), higher scores in patient engagement (OR 1.10, 95% CI 1.05-1.15, *P*=.001), and more health care encounters in the past 6 months (OR 1.16, 95% CI 1.01-1.34, *P*=.04).

Discussion

Principal Results

As hypothesized, patient engagement predicted online patient portal use, supporting previous research [7,9,28]. Findings also support that young adults have an interest in patient portals and engagement with their health care [29]. As they become more engaged, they are likely to find use of patient portals helps with medical decision-making [2].

We did not find that eHealth literacy was associated with patient portal use as we hypothesized. Others who found eHealth literacy was a correlate of patient portal use were reporting on samples with chronic conditions [3,30], compared to relatively healthy emerging adult populations. In our sample, participants had a midrange level of eHealth literacy (see Table 1). Our results echo previous studies' findings that while emerging adults may be familiar with the internet and online information sources, they may not be engaging in health information seeking and thus have average eHealth literacy skills [22]. Interestingly, public university students had higher eHealth literacy scores than private university students, suggesting they can better evaluate eHealth resources than their counterparts. Public university students also had a higher number of health care visits compared to private university students, which may have contributed to greater familiarity with patient portals and influenced eHealth literacy. It will be important to investigate these relationships as emerging adults may engage in more health IT in coming years as portals become even more commonplace and their health management demands increase.

Analyses identified demographic and health factors associated with using clinical features of portals. The findings suggest that non-White emerging adults (specifically, Hispanic and Asian emerging adults), those having more health care encounters, and those with higher levels of patient engagement are using these features more frequently. A study of nearly 50,000 primary care patients showed that patients often look at clinical information through patient portals after a health care visit [23]. In addition, a recent report found patients from vulnerable populations, including those who were less educated, older, and from ethnic racial minority groups, were more likely to report benefits from reading clinic visit notes online compared to less vulnerable patients [31]. It is possible emerging adults from underserved communities may have more to gain from accessing clinical information from their online health records. Accessing patient portals may represent an important first step in taking control of one's health, especially in the emerging adult population.

Limitations

This study had a cross-sectional survey design and used convenience sampling. It is descriptive in nature and cannot determine causality regarding predictors of use. Although a nonprobability convenience sampling method reduces the generalizability of findings, it does allow researchers easier access to study participants and is a common practice in developmental and formative science [32,33]. Due to the nature of sampling, an exact response rate cannot be calculated. The sample size is also relatively small and further research is needed in larger higher education student populations across geographic regions. The study also did not assess if participants are sharing personal health information with others, such as parents or guardians. Furthermore, we did not collect information from faculty about whether extra credit was offered to students for participating, an approach that could bias the sample. Finally, the findings might not be generalizable to other emerging adult populations and health care settings.

Conclusions

Online patient portals represent an important health IT tool offered by many health care organizations. In recent years, available features of patient portals continue to expand, giving patients greater access to their personal health information and tools to engage in self-management. This study adds valuable insight to continuing research into online patient portal use by bringing the emerging adult population into focus. There is still work needed to increase awareness of online patient portals among this age group, but there are also subpopulations that are already frequently using patient portals. Our findings suggest it is important to know more about the health IT available at public and private universities and that there is a need to further investigate eHealth literacy among this population. Interventions should focus on increasing patient portal awareness and engagement in health management in conjunction with improving eHealth literacy skills of the emerging adult patients.

Acknowledgments

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Survey.
[DOCX File , 427 KB - formative_v6i2e33356_app1.docx]

Multimedia Appendix 2
Example of patient portal.
[PNG File , 143 KB - formative_v6i2e33356_app2.png]

References

1. Lorig KR, Holman HR. Self-management education: history, definition, outcomes, and mechanisms. Ann Behav Med 2003 Aug;26(1):1-7. [doi: 10.1207/S15324796ABM2601_01] [Medline: 12867348]
2. Kruse CS, Bolton K, Freriks G. The effect of patient portals on quality outcomes and its implications to meaningful use: a systematic review. J Med Internet Res 2015 Mar 10;17(2):e44 [FREE Full text] [doi: 10.2196/jmir.3171] [Medline: 25669240]
3. Hoogenbosch B, Postma J, de Man-van Ginkel JM, Tiemessen NA, van Delden JJ, van Os-Medendorp H. Use and the Users of a Patient Portal: Cross-Sectional Study. J Med Internet Res 2018 Sep 17;20(9):e262 [FREE Full text] [doi: 10.2196/jmir.9418] [Medline: 30224334]
4. Van den Bulck SA, Hermens R, Slegers K, Vandenberghe B, Goderis G, Vankrunkelsven P. Designing a Patient Portal for Patient-Centered Care: Cross-Sectional Survey. J Med Internet Res 2018 Oct 01;20(10):e269 [FREE Full text] [doi: 10.2196/jmir.9497] [Medline: 30287416]
5. Tang PC, Ash JS, Bates DW, Overhage JM, Sands DZ. Personal Health Records: Definitions, Benefits, and Strategies for Overcoming Barriers to Adoption. Journal of the American Medical Informatics Association 2006 Mar 01;13(2):121-126. [doi: 10.1197/jamia.m2025]
6. The Markle Foundation. The Personal Health Working Group - Final Report. Markle Connecting for Health. 2003 Jul 01. URL: https://www.markle.org/publications/1429-personal-health-working-group-final-report [accessed 2022-02-01]
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. Irizarry T, DeVito Dabbs A, Curran CR. Patient Portals and Patient Engagement: A State of the Science Review. J Med Internet Res 2015 Jul 23;17(6):e148 [FREE Full text] [doi: 10.2196/jmir.4255] [Medline: 26104044]
9. Maurer M, Dardess P, Carman KL, Frazier K, Smeeding L. Guide to Patient and Family Engagement: Enviornmental Scan Report. AHRQ Publication No. 12-0042-EF. 2012 May. URL: https://psnet.ahrq.gov/issue/guide-patient-and-family-engagement-environmental-scan-report [accessed 2022-02-08]
10. Peacock S, Reddy A, Leveille SG, Walker J, Payne TH, Oster NV, et al. Patient portals and personal health information online: perception, access, and use by US adults. J Am Med Inform Assoc 2017 May 01;24(e1):e173-e177 [FREE Full text] [doi: 10.1093/jamia/ocw095] [Medline: 27413120]
11. Davis S. Health literacy, education levels, and patient portal usage during hospitalizations. 2015 Presented at: AMIA Annual Symposium Proceedings; 2015; San Francisco, CA p. 1871.
12. Kruse CS, Argueta DA, Lopez L, Nair A. Patient and provider attitudes toward the use of patient portals for the management of chronic disease: a systematic review. J Med Internet Res 2015 Mar 20;17(2):e40 [FREE Full text] [doi: 10.2196/jmir.3703] [Medline: 25707035]
13. Ancker JS, Barrón Y, Rockoff ML, Hauser D, Pichardo M, Szerencsy A, et al. Use of an electronic patient portal among disadvantaged populations. J Gen Intern Med 2011 Oct 7;26(10):1117-1123 [FREE Full text] [doi: 10.1007/s11606-011-1749-y] [Medline: 21647748]
14. Goel MS, Brown TL, Williams A, Hasnain-Wynia R, Thompson JA, Baker DW. Disparities in enrollment and use of an electronic patient portal. J Gen Intern Med 2011 Oct 3;26(10):1112-1116 [FREE Full text] [doi: 10.1007/s11606-011-1728-3] [Medline: 21538166]
15. Rodriguez JA, Lipsitz SR, Lyles CR, Samal L. Association Between Patient Portal Use and Broadband Access: a National Evaluation. J Gen Intern Med 2020 Dec 10;35(12):3719-3720 [FREE Full text] [doi: 10.1007/s11606-020-05633-4] [Medline: 31925739]
16. Norman CD, Skinner HA. eHealth Literacy: Essential Skills for Consumer Health in a Networked World. J Med Internet Res 2006 Jul 16;8(2):e9 [FREE Full text] [doi: 10.2196/jmir.8.2.e9] [Medline: 16867972]
17. Britt RK, Collins WB, Wilson K, Linnemeier G, Englebert AM. eHealth Literacy and Health Behaviors Affecting Modern College Students: A Pilot Study of Issues Identified by the American College Health Association. J Med Internet Res 2017 Dec 19;19(12):e392 [FREE Full text] [doi: 10.2196/jmir.3100] [Medline: 29258979]
18. Veenhof B, Clermont Y, Sciadas G. Literacy and Digital Technologies: Linkages and Outcomes. Research Paper for Science, Innovation and Electronic Information Division (SIEID). 2005. URL: https://www150.statcan.gc.ca/n1/pub/56f0004m/56f0004m2005012-eng.pdf [accessed 2022-02-08]
19. Arnett JJ. Emerging adulthood: A theory of development from the late teens through the twenties. American Psychologist 2000;55(5):469-480. [doi: 10.1037/0003-066x.55.5.469]

20. Harris MA, Freeman KA, Duke DC. Transitioning From Pediatric to Adult Health Care: Dropping Off the Face of the Earth. American Journal of Lifestyle Medicine 2010 Aug 19;5(1):85-91. [doi: 10.1177/1559827610378343]
21. Stellefson M, Hanik B, Chaney B, Chaney D, Tennant B, Chavarria EA. eHealth literacy among college students: a systematic review with implications for eHealth education. J Med Internet Res 2011 Dec 01;13(4):e102 [FREE Full text] [doi: 10.2196/jmir.1703] [Medline: 22155629]
22. Early J, Stone A, Nieto C, Gonzalez C. eHealth Literacy of 2-Year and 4-Year College Students: Implications for Health Education in a Post-Truth Era. Pedagogy in Health Promotion 2021 Sep 22:23733799211038023. [doi: 10.1177/23733799211038023]
23. Leveille SG, Mejilla R, Ngo L, Fossa A, Elmore JG, Darer J, et al. Do Patients Who Access Clinical Information on Patient Internet Portals Have More Primary Care Visits? Med Care 2016 Jan;54(1):17-23. [doi: 10.1097/MLR.0000000000000442] [Medline: 26565525]
24. Lorig K, Stewart A, Ritter P, Gonzalez V, Lynch J, Laurent D. Outcome measures for health education and other health care interventions. Thousand Oaks, CA: Sage Publications; 1996:53-55.
25. Ritter PL, Stewart AL, Kaymaz H, Sobel DS, Block DA, Lorig KR. Self-reports of health care utilization compared to provider records. Journal of Clinical Epidemiology 2001 Feb;54(2):136-141. [doi: 10.1016/s0895-4356(00)00261-4]
26. Seçkin G, Yeatts D, Hughes S, Hudson C, Bell V. Being an Informed Consumer of Health Information and Assessment of Electronic Health Literacy in a National Sample of Internet Users: Validity and Reliability of the e-HLS Instrument. J Med Internet Res 2016 Jul 11;18(7):e161 [FREE Full text] [doi: 10.2196/jmir.5496] [Medline: 27400726]
27. Duke CC, Lynch WD, Smith B, Winstanley J. Validity of a New Patient Engagement Measure: The Altarum Consumer Engagement (ACE) Measure. Patient 2015 Dec 23;8(6):559-568 [FREE Full text] [doi: 10.1007/s40271-015-0131-2] [Medline: 26097010]
28. Irizarry T, Shoemake J, Nilsen ML, Czaja S, Beach S, DeVito Dabbs A. Patient Portals as a Tool for Health Care Engagement: A Mixed-Method Study of Older Adults With Varying Levels of Health Literacy and Prior Patient Portal Use. J Med Internet Res 2017 Mar 30;19(3):e99 [FREE Full text] [doi: 10.2196/jmir.7099] [Medline: 28360022]
29. Ramsey A, Lanzo E, Huston-Paterson H, Tomaszewski K, Trent M. Increasing Patient Portal Usage: Preliminary Outcomes From the MyChart Genius Project. J Adolesc Health 2018 Jan;62(1):29-35 [FREE Full text] [doi: 10.1016/j.jadohealth.2017.08.029] [Medline: 29169768]
30. Price-Haywood EG, Harden-Barrios J, Ulep R, Luo Q. eHealth Literacy: Patient Engagement in Identifying Strategies to Encourage Use of Patient Portals Among Older Adults. Popul Health Manag 2017 Dec;20(6):486-494. [doi: 10.1089/pop.2016.0164] [Medline: 28384076]
31. Walker J, Leveille S, Bell S, Chimowitz H, Dong Z, Elmore JG, et al. OpenNotes After 7 Years: Patient Experiences With Ongoing Access to Their Clinicians' Outpatient Visit Notes. J Med Internet Res 2019 May 06;21(5):e13876 [FREE Full text] [doi: 10.2196/13876] [Medline: 31066717]
32. Jager J, Putnick DL, Bornstein MH. More than just convenient: The scientific merits of homogeneous convenience samples. Monogr Soc Res Child Dev 2017 Jun 05;82(2):13-30 [FREE Full text] [doi: 10.1111/mono.12296] [Medline: 28475254]
33. Bornstein MH, Jager J, Putnick DL. Sampling in Developmental Science: Situations, Shortcomings, Solutions, and Standards. Dev Rev 2013 Dec;33(4):357-370 [FREE Full text] [doi: 10.1016/j.dr.2013.08.003] [Medline: 25580049]

Abbreviations

ACE: Altarum Consumer Engagement
IT: information technology
OR: odds ratio

Edited by A Mavragani; submitted 13.09.21; peer-reviewed by R Silver, M Bujnowska-Fedak, M Tran; comments to author 03.10.21; revised version received 07.12.21; accepted 30.12.21; published 15.02.22.

Please cite as:
Wright JA, Volkman JE, Leveille SG, Amante DJ
Predictors of Online Patient Portal Use Among a Diverse Sample of Emerging Adults: Cross-sectional Survey
JMIR Form Res 2022;6(2):e33356
URL: https://formative.jmir.org/2022/2/e33356
doi:10.2196/33356
PMID:35166686

Original Paper

Effects of Game Mode in Multiplayer Video Games on Intergenerational Social Interaction: Randomized Field Study

Carmen Zahn[1], PhD; David Leisner[1], MSc; Mario Niederhauser[1], MSc; Anna-Lena Roos[1], PhD; Tabea Iseli[2], MA; Marco Soldati[2], MSc

[1]School of Applied Psychology, University of Applied Sciences and Arts Northwestern Switzerland, Olten, Switzerland

[2]School of Engineering, University of Applied Sciences and Arts Northwestern Switzerland, Windisch, Switzerland

Corresponding Author:
Carmen Zahn, PhD
School of Applied Psychology
University of Applied Sciences and Arts Northwestern Switzerland
Riggenbachstrasse 16
Olten, CH-4600
Switzerland
Phone: 41 (0)629572473
Email: carmen.zahn@fhnw.ch

Abstract

Background: Maintaining social relationships is a basic human need and particularly essential in old age, including when living in a retirement home. Multiplayer video games can promote positive social interactions among players from different generations while playing. Yet, such facilitation of positive social interactions depends on specific game design. To systematically investigate the effects of game design on social interaction between seniors and their coplayers, the game *Myosotis FoodPlanet* was developed in this study, and the impacts of 3 different game modes on social interaction were compared in a controlled field trial.

Objective: This study aims to compare the effects of 3 different game modes (competitive, cooperative, and creative) on social interactions (verbal and nonverbal communication) between seniors and their younger coplayers.

Methods: This study was conducted in a Swiss retirement home as a controlled field trial. Participants were residents of the retirement home (N=10; mean age 84.8 years, SD 5.9 years) and played in pairs with their caregivers. Each pair played 3 game modes in random order. This resulted in 30 game sequences of 20 minutes each. A within-subject design was applied with *game mode* as the within-factor and *social interaction* as the outcome variable. To assess the quality of social interaction, 30 video-recorded game sequences were analyzed based on an event sampling method.

Results: Analysis of variance for repeated measurements revealed significant effects: there was significantly more verbal communication in the creative mode than in the cooperative mode (P=.04) with a strong effect size (Cohen f=0.611). An examination of verbal communication revealed more *game-related* communication in the creative mode than in the cooperative mode (P=.01) and the competitive mode (P=.09) with marginally significant effects and strong effect sizes (Cohen f=0.841). In addition, significantly more biography-related communication occurred in the creative mode than in the cooperative mode (P=.03), with a strong effect size (r=0.707). Regarding nonverbal communication (eg, laughing together), analysis of variance for repeated measurements showed significant differences among the game modes (P=.02) with a strong effect size (Cohen f=0.758). Results showed that there was significantly more laughing together in the competitive mode (competitive>cooperative>creative).

Conclusions: The results show that game mode can be an important factor for shaping the social interactions of players playing together. Compared with other modes, creative game modes can increase verbal communication. In contrast, competitive modes may stimulate more laughing together. This has important implications for game design and the use of computer games to promote social interaction between seniors and their coplayers in practice.

(JMIR Form Res 2022;6(2):e29179) doi:10.2196/29179

KEYWORDS

video games; computer games; older adults; game mode; serious game; social interaction; video analysis methods

Introduction

Playing Digital Games as Social Activity Connecting People From Different Age Groups and Generations

Maintaining social relationships is a basic human need and therefore is highly relevant to psychological well-being [1]. This is particularly true in old age [2-5], including when living in a retirement home. Multiplayer video games can promote well-being and the maintenance of social relationships [6-8] because they can facilitate *positive* social interactions during play, even for players from different generations [9-14]. Imagine, for instance, a visiting grandchild playing a video game with their grandparent in a retirement home, or imagine care persons and residents playing together during activation therapy or recreation in a common room. Playing together is a rewarding experience [15] that connects people through joint action, cooperation, or playful and unthreatening competition. In this way, it can link people from *different* generations, bridging the gap among them. Research indicates that social interaction and entertainment are among the main motivating factors for older adults to play video games [16-18]. Specifically, players enjoy interacting with others, watching others playing, and talking about the game [19-21]. Digital technology, which is ubiquitous in the lives of younger and adult people, has been investigated as a tool to connect generations, including those of advanced age [2]. For instance, intergenerational digital games can enhance social bonding [22] and pave the way for improved communication among players [23]. The term *intergenerational*, thereby does not only include family bonds (eg, grandchildren and grandparents) but can also be considered in a wider sense in terms of age (eg, older and younger people) or community life (eg, youth and older adults) [24].

Taken together, playing video games seems to be a meaningful way to foster positive intergenerational social interaction between younger and older people, which, in turn, is likely to improve psychological well-being in the long run. As emphasized in a review [25] and various studies [26,27], games can have positive effects on the physical, cognitive, social, and emotional states, especially in older adults.

However, there are important research gaps that must be considered concerning the potential of video games in real-life scenarios, especially those concerning specific game design decisions and their impact on social interactions. For example, de la Hera et al [23] state: “The decision to engage players in collaborative, competitive or cooperative competitive games has relevant implications on the effects of these practices.” Previous research on intergenerational digital games describing major research gaps recommend that future empirical studies should directly compare different forms of playing to discover the effects of intergenerational interactions [24]. Likewise, recent research related to digital games in neurorehabilitation has addressed similar research gaps concerning different game modes [28]. Therefore, it is of utmost importance to further contribute to the scientific knowledge on the effects of game design decisions on players’ behaviors via systematic empirical studies. In addition, it has been recommended by previous research on intergenerational games that future research should include more types than only grandparent–grandchildren interactions [23].

In our study, we investigate the possible effects of specific game design decisions regarding game modes on social interactions among players from different age groups in a retirement home. In the following section, we present the rationale behind our game design decisions.

Game Design Decisions and Possible Influences on Players’ Social Interactions

In multiplayer game design [29], *social game mechanics* are used to initialize and increase the social interaction among players within the game. Systematic empirical research is rare in this area; however, isolated evidence has demonstrated that design for social interaction can impact players’ behaviors in old age—not only concerning their verbal behavior but also their nonverbal behavior; a study comparing social interaction design of a pervasive game (pervasive game here interrelates to the close connection between a web-based game and the physical world), as opposed to *no* social interaction design, revealed significant and positive effects on promoting physical activity in older adult players [30].

More specifically, social game mechanics provide different configurations—*game modes* that can influence social interaction among players. The established modes are (1) competitive, if 2 players compete against each other, whereby only one of them can win [31] and (2) cooperative, if 2 players *operate together*, having a dedicated task each, but winning the game together. These contrasting modes are both able to stimulate social interaction and motivation to play [32-34]. They correspond to basic categories of human social interaction behaviors, which are well-known from the long history of social psychology research [35]. In addition, modern games provide a (3) *creative* mode in which the game does not imply any rules and the players are free to explore or modify the game world in a creative way. Such games are referred to as open-ended simulation games or sandbox games [36,37].

Existing research reveals an ongoing debate regarding the influence of game modes on social interaction. Although some [35] assume an increased willingness to communicate in cooperative game structures, others [38] argue for the existence of a correlation between competitive game structures and social interaction. More detailed findings and theoretical considerations on the different game modes and their influence on social interaction are presented in the following section.

Game Modes and Players’ Social Interactions—Empirical Evidence

Competitive Mode and Social Interaction

Research on game gratification shows a positive relationship between social and competitive motives [39,40], and players who engage with games primarily because they seek social interaction are often also competitive gamers [38]. This correlation can be explained by the basic human need for control, according to the fundamental interpersonal relationship orientation theory [41]: competition with others or trying to control each other is an essential part of interpersonal dynamics

[38]. Nevertheless, in contrast to younger players, older players find competition in playing a minor motivator [16,18,22,27], unless there is indirect competition against other teams [42]. In addition, older players have been found to largely reject reflex-oriented games such as fighting or racing games. They experience such games as more difficult, less interesting, and hence less enjoyable to play, owing to their age-related physical condition or disabilities [43].

Cooperative Mode and Social Interaction

Empirical findings suggest that cooperative video games support positive interdependence (eg, for video games) [44,45], meaning players need each other to fulfill a certain task and all members must contribute their knowledge and skills for the group or team to be successful [35]. Positive interdependence plays a crucial role in improving intergenerational social interactions [23]. Thus, cooperative game structures should likely lead to increased willingness to communicate as a team, facilitating the exchange of important information, sharing of ideas, and reacting to the ideas shared by others [35]. However, to the authors' knowledge, there have been no clear empirical results confirming this assumption for older adults or intergenerational games.

Creative Mode and Social Interaction

The creative game experience has been seen to increase when the game does not specify any right or wrong paths [36]. This can foster creativity because players collaboratively generate new ideas beyond what they could have come up with on their own [46-48]. Following this, the given level of freedom in the creative game is likely to initiate and increase communication (about game-related activities and thoughts) and social interaction (joint decision-making) among players. Talking about new ideas or things that are out of the ordinary can lead to discussion of new topics and even more communication. This, in turn, should further increase social interactions. Empirical studies examining this relationship are lacking.

Taken together, these empirical findings suggest that all 3 modes can stimulate social interaction and motivation to play [31,33]. However, empirical studies comparing different modes in relation to social interaction are inconsistent (competitive vs cooperative modes) or lacking (creative mode). As the theoretical considerations described above suggest that there may be differences, a comparative analysis investigating the differential effects of cooperative, competitive, and creative game modes on intergenerational social interaction can fill the research gap as described previously [23] and inform game designers on how to design games that stimulate intergenerational social interaction [22].

Goals of the Field Trial

The goal of this study is to provide original results that contribute to improved scientific knowledge about the effects of social game mechanics (game modes) on social interactions between older players and younger coplayers. To accomplish this, a controlled field trial was presented to investigate these impacts with older participants from a retirement home. The game used for the trial is a serious multiplayer game called *Myosotis FoodPlanet* (see the *Materials and Tools* section and Multimedia Appendix 1) designed specifically as an intergenerational game for use in retirement homes.

This study aims to compare the impacts of 3 different game modes (competitive, cooperative, and creative) on social interactions between seniors and their younger coplayers during game play. With reference to the theory described above and previous research, we assumed that the game modes would differ in the extent to which they influence social interaction. Therefore, we differentiated between verbal social interaction (H1) and nonverbal social interaction (H2) among players. Specifically concerning H1, we expected the creative game mode to stimulate the highest amount of verbal social interaction, as explained earlier. Despite somewhat controversial findings regarding the competitive and cooperative modes, we further assumed that a cooperative mode would stimulate more verbal social interaction, compared with the competitive mode, based on related research [35]. In brief: Creative Mode>Cooperative Mode>Competitive Mode.

Regarding nonverbal social interaction (H2), we hypothesized that there would be differences among game modes; however, owing to the extremely limited research on nonverbal interaction in video games, the assumptions were nondirectional, and our research remains explorative. The study was conducted before the COVID-19 pandemic, when physical interactions (hand shaking and touching another's hands, shoulders, etc) were still not restricted owing to potential health risks.

Methods

Participants and Study Design

In all, 10 older residents from a Swiss retirement home voluntarily participated in a randomized controlled field trial during their leisure time (7 women and 3 men; mean age 84.8 years, SD 5.85 years; range 76-93). The older participants were healthy, with the exception of minor age-related impairments (ie, minor physical limitations and no severe dementia). A within-subject study design was applied with 3 game modes (competitive, cooperative, and creative) administered in a randomized order. This resulted in 30 game sequences. In all, 4 care professionals, 3 activation therapists and 1 nurse (4 women; mean age 44, SD 15.60 years; range 21-55 years) participated as coplayers of a younger generation to ensure the residents' safety at all times during study participation and liability to ethical standards. The study was not part of any therapeutic program and did not include any clinical interventions. Social interaction was the only outcome variable (for details on measures, see the *Measures of Social Interaction* section). No health-related outcomes were addressed in the study. According to the Swiss Federal Human Research Act, the study was not liable for registration.

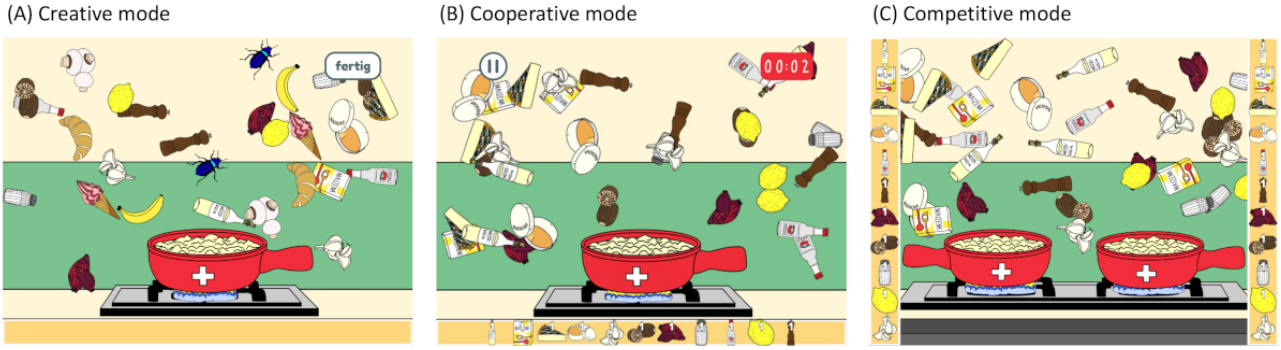
Materials and Tools

The game used in this study is a serious multiplayer game called *Myosotis FoodPlanet*, which was designed specifically as an intergenerational game for use in retirement homes. It is a game involving cooking Swiss cheese fondue together, which is a traditional and well-known dish in Switzerland. From reminiscence therapy [49], it is known that food is an ideal topic

for stimulating social interaction [50], because anybody can be assumed to have an opinion on food, and the interests of different generations can easily be taken into account [11]. *Myosotis FoodPlanet* was developed in the multidisciplinary research project Myosotis Garden [37]. The games developed in this project were designed to provide entertaining positive activity, thereby triggering intergenerational communication and enabling players to find new and exciting access to the memories and biographies of older people [23].

In *Myosotis FoodPlanet*, two players—one older person and one younger care professional in the case of this study—jointly prepare a Swiss cheese fondue on an iPad Pro (32.8 cm) by dragging floating ingredients into a fondue pot (Figure 1). To ensure that older adults, despite possible age-related handicaps, recognize the ingredients, a computer-generated voice announces the name of each ingredient when it is tapped. In addition, the names of the ingredients are presented in a written form. A traditional Swiss folk melody plays in the background, contributing to the creation of a pleasant atmosphere. The game is designed for 2 players sitting in the same room and sharing the touch screen. The entire screen is used by both players, as opposed to a split-screen mode. Although the game is commonly played synchronously, the creative mode also allows a turn-based approach, where one after the other, the players add ingredients to the pot. A total of 3 variants of *Myosotis FoodPlanet*, each offering a different game mode (Figures 1A-C) were used in our field trial.

Figure 1. Game modes: (A) creative, (B) cooperative, and (C) competitive.



In the *competitive mode* (Figure 1A), each player prepares a given fondue recipe as quickly as possible. Each player has their own fondue pot and their own but identical given ingredient list that is displayed on the respective side of the screen closest to the player. The goal of the player is to drag the given ingredients into their own pot faster than the other player. It is also possible to drag the ingredients needed by the opponent into one's own pot. The player who collects all the listed ingredients first wins.

In the *cooperative mode* (Figure 1B), players work together to prepare a fondue by collecting ingredients from a list of typical fondue as quickly as possible into one pot for both. The ingredients are listed in the bar at the bottom of the screen. The order in which the ingredients are collected is irrelevant, and the ingredients can be dragged repeatedly. Once all the listed ingredients have been dragged into the pot, the total time taken and the current high score (best time of all players) are displayed.

In the *creative mode* (Figure 1C), players collaboratively and freely choose ingredients for their cheese fondue. Each ingredient can be dragged into the pot as often as desired, while the bar at the bottom of the screen shows the ingredients already added. Players can choose typical (eg, Gruyère cheese and white wine) as well as atypical (eg, bug and soft ice) fondue ingredients. The latter will likely increase the fun factor and consequently the social interaction. When an atypical ingredient is added, a colored splash appears on the screen, accompanied by a squeaky sound effect. After a short period, the splash disappears, but the cheese fondue now is colored similar to the atypical ingredient (eg, adding a bug turns the fondue blue). The players end the game manually. As a reward, the players obtain the prepared fondue in a recipe form. In addition, the system automatically names the recipe with humorous names (eg, *Dancing Hans* or *Singing Theodora*) to increase the fun factor and social interaction.

Procedure

The study was conducted in the activation room (a room with which the participants were familiar) of the retirement home as a free-choice afternoon leisure activity. Each player participated on 3 different days and played 1 of the 3 game modes in random order. Each game sequence lasted for a maximum of 20 minutes. Participation was voluntary, and it was made clear that participants could withdraw their participation at any time without any consequence. The participants were informed beforehand about the goals of the research, duration and procedure of the study, voluntary nature of participation, and protection of their data. Written informed consent was obtained from all test participants before starting each game sequence, as well as from the institution before starting the trial. The players were seated next to each other so that their dominant hand could easily access the tablet computer (Figure 2).

Figure 2. Study setting: game sequence.

This procedure was performed in accordance with the Helsinki Declaration of 1975, as revised in 2000. As noted, it was not liable to registration according to the Swiss Federal Human Research Act.

Measures of Social Interaction

Social interaction was measured by verbal indicators, such as verbal communication and nonverbal indicators, such as laughing together or eye contact [51]. Thus, data gathering was based on video recordings of verbal and nonverbal communication among players. In all, 30 game sequences were recorded using 2 video cameras. Gameplay recordings, as well as observations, rank among the most frequently used methods to investigate the possible influences of digital game playing on social interaction [23]. Video analysis has proven to be a suitable research tool for observing behavior and social interactions [52]. It enables the storage and reproducibility of complex social interaction data, allowing for multiple observations by different observers at different times, depending on the individual research questions of interest. It was initially intended that self-assessment questionnaires with questions about the players' subjective experiences and well-being would also be administered; however, some older participants had difficulty understanding the questions. (It was unclear whether this was owing to their Swiss language, minor impairments, or a lack of motivation to fill out the questionnaires.) Owing to this, the self-assessment questionnaire component was removed.

Data Analysis

The video data gathered comprised 30 recorded game sequences of approximately 8 hours (7 hours 53 minutes 25 seconds). For video data analysis, a 2-step *coding and counting* approach [53,54] was applied. In the first step, the 30 recorded game sequences were coded using the text analysis software MAXQDA 2018 (VERBI GmbH, Berlin, Germany) based on a preliminary category system (*event sampling* [55]). In the second step, the final category system was developed, including new categories emerging from the data, by rewatching video recordings in an iterative process. As a result, social interaction in terms of *verbal* communication among players was divided into categories: *game-related communication*, *fondue-related biography*, *general biography*, *help seeking*, and *help giving*. Social interaction in terms of nonverbal communication was divided into *laughing together*, *eye contact*, and *body contact*. Table 1 lists the definitions and corresponding anchor examples [56] for each category. The communication behaviors were then systematically coded by 2 trained raters (duration of behavior in minutes and seconds; mm:ss). The raters were trained on examples from the videos; the raters discussed these examples and resolved any uncertainties. They then coded the videos. Intercoder reliability among the raters resulted in substantial agreement (κ =0.69). Finally, the durations of all relevant behavior indicators of a given code were summed up. Verbal communication (eg, game-related communication) and nonverbal communication (eg, eye contact) occurred simultaneously. Owing to overlapping codes caused by this occurrence of simultaneous communication, finally, as an artifact—the sum duration of each observed communication exceeds the effective observation time of total social interaction (Figure 3).

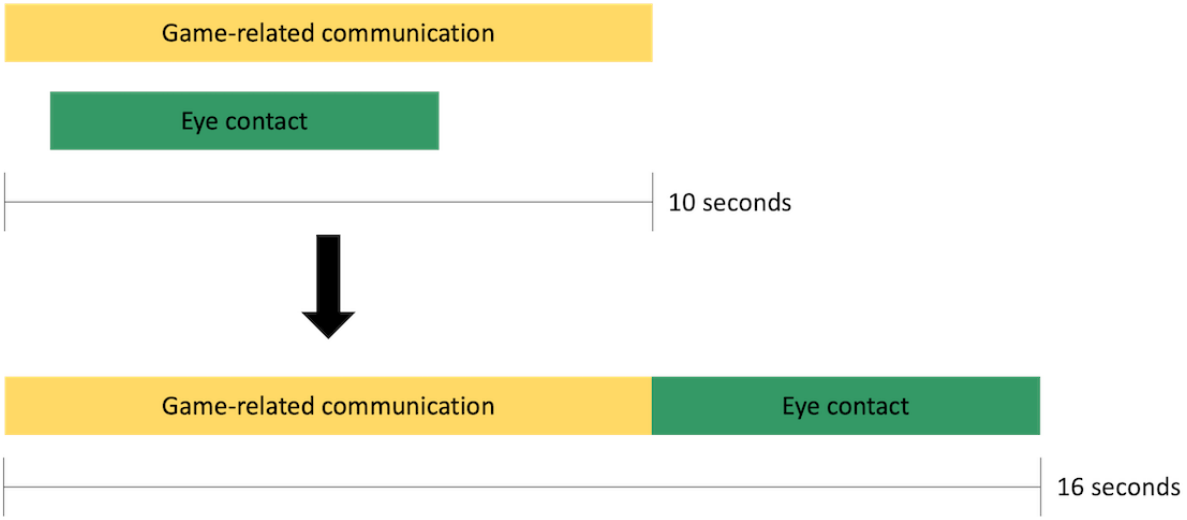
Table 1. Category system for the observation of social interaction behaviors.

Category	Definition	Anchor examples[a]
Verbal communication		
Game-related communication	All statements resulting directly from the game (eg, discussing results, outcomes, or new strategies)	"Oh my God, that's fun" and "You won again, congrats"
Fondue-related biography	When players discuss the method of preparation or consumption of a cheese fondue in the past (ie, anecdotes)	"My son-in-law doesn't like garlic. So, when he came to visit, I always had to make a cheese fondue without garlic" and "Do you like onions in the fondue?"
General biography	When personal biographic information is shared (when players tell or ask about their professional life, family, military, childhood, etc)	"You know, I used to work in a cheese factory and there was this one customer, who..." and "Originally I am from Austria, but after the war my family moved to Switzerland"
Help seeking	When older participants actively ask for help for technical operation of the game or when actively asking about content aspects during the game	"Why can I not grab that cheese over here?" and "Have I collected all the needed ingredients?"
Help giving	When coplayers answer questions regarding technical or content matters. Furthermore, when coplayers help the older adults by giving them hints	"Try using your fingertip instead of your fingernail" and "Look, over there is the onion you still need to collect"
Nonverbal communication		
Laughing together	When players laugh out loud together; also includes when one person laughs and simultaneously talks and the other smiles	N/A[b]
Eye contact	When players look directly into each other's eyes (ie, eyes are meeting)	N/A
Body contact	When one player is patting the other on the back, or when one player is touching the other's hand or arm	N/A

[a]Anchor examples are verbal statements translated from Swiss German.

[b]N/A: not applicable.

Figure 3. Coding example—overlaps.



Results were visualized using MAXQDA 2018, including the overlapping of codes. Thus, it was possible to illustrate which verbal and nonverbal categories tend to occur simultaneously (Table 2). For quantitative analyses, data were transferred to the statistical software SPSS. Owing to the small sample size, the data were checked for normal distribution and variance homogeneity. Statistical analyses compared each category of social interaction by using absolute values (ie, mean duration of social interaction in minutes and seconds; mm:ss), as well as relative values, (ie, percent duration in relation to the total playing time) for the 3 game modes. Hypotheses were tested by analyses of variance with repeated measurements or its nonparametric counterpart (Friedman test).

Table 2. Simultaneous occurrence of verbal and nonverbal codes.

	Laughing together, n	Eye contact, n	Body contact, n
Game-related communication	604	556	83
Fondue-related biography	31	204	1
General biography	12	119	2

Results

Overall Findings

Influence of Game Modes on Total Playing Time

The analysis revealed that the average total playing time per game mode (in mm:ss) varied slightly among different game modes. On average, the creative mode (mean 16:39, SD 6:45) was the longest, followed by the competitive mode (mean 16:01, SD 5:46), and the cooperative mode (mean 14:40, SD 6:32). The playing times did not differ significantly among game modes as shown by the Friedman test (P=.67). Playing time for each mode was restricted by the study instructions to a maximum of 20 minutes.

Influence of Game Modes on Total Time of Social Interactions

The results revealed that for approximately half of the total playing time (in hh:mm:ss), social interactions were observed (3:51:47 with 7570 coded social interactions). It is important to note that different categories of social interaction occurred simultaneously, and the addition of these categories resulted in an *artificially* prolonged total duration (5:10:34, 7570 coded social interactions). Table 3 summarizes the absolute and relative values of verbal and nonverbal social interactions (along with the respective subcategories) for the different game modes. Figures 4-7 show a graphical depiction of these results.

Table 3. Absolute and relative values of social interaction in the different game modes.

Social interaction category	Game mode			Hypothesis
	Creative	Cooperative	Competitive	
Verbal, mean (SD)				H1[a]
Game-related communication				
Absolute values (mm:ss)	7:47 (4:43)	5:05 (3:19)	4:48 (2:20)	Hypothesis confirmed[b]
Relative values (% of total time)	43 (18.7)	33 (15.8)	31.2 (16.4)	Hypothesis confirmed[c]
Fondue-related biography				
Absolute values (mm:ss)	1:15 (1:35)	0:38 (0:47)	0:51 (1:18)	Hypothesis partly confirmed[d]
Relative values (% of total time)	6.4 (7.3)	5.2 (7.8)	4.9 (6.6)	—[e]
General biography				—
Absolute values (mm:ss)	0:21 (0:40)	0:17 (0:30)	0:52 (2:02)	
Relative values (% of total time)	1.8 (3.3)	1.4 (2.3)	4.9 (9.9)	
Help seeking				
Absolute values (mm:ss)	0:09 (0:16)	0:09 (0:11)	0:06 (0:17)	—
Relative values (% of total time)	1 (1.6)	1.6 (2.3)	0.6 (1.4)	Hypothesis partly confirmed[f]
Help giving				—
Absolute values (mm:ss)	0:22 (0:36)	0:25 (0:41)	0:44 (1:28)	
Relative values (% of total time)	2.1 (3.4)	4 (6.2)	4.4 (7.3)	
Total verbal communication				
Absolute values (mm:ss)	9:56 (6:21)	6:36 (4:12)	7:22 (4:47)	Hypothesis partly confirmed[d]
Relative values (% of total time)	54 (25.9)	45 (20.4)	46 (25.3)	—
Nonverbal, mean (SD)				H2[g]
Laughing together				
Absolute values (mm:ss)	1:09 (0:55)	1:25 (1:17)	1:33 (1:22)	—
Relative values (% of total time)	6.1 (4.1)	8 (5.5)	8.8 (6.3)	Hypothesis confirmed[c]
Eye contact				—
Absolute values (mm:ss)	0:32 (0:31)	0:54 (0:49)	1:15 (1:39)	
Relative values (% of total time)	3 (2.4)	5.8 (4.3)	6.8 (7.8)	
Body contact				—
Absolute values (mm:ss)	0:03 (0:08)	0:09 (0:12)	0:04 (0:05)	
Relative values (% of total time)	0.3 (0.7)	1 (1.2)	0.6 (0.8)	
Total nonverbal communication				
Absolute values (mm:ss)	1:45 (1:25)	2:29 (2:07)	2:53 (2:47)	—
Relative values (% of total time)	9.4 (6.2)	14.8 (9.3)	16.2 (12.5)	Hypothesis confirmed[b]
Total social interaction				—
Absolute values (mm:ss)	11:41 (7:11)	9:06 (5:45)	10:16 (6:44)	
Relative values (% of total time)	63.7 (28)	59.9 (23)	62.2 (31.1)	

[a]H1: hypothesis 1.

[b]Significance, *P*=.01.

[c]Significance, *P*=.02.

[d]Marginal significance, *P*=.06.

[e]Hypothesis not confirmed.

[f]Significance, *P*=.04.
[g]H2: hypothesis 2.

Figure 4. Verbal communication (absolute values).

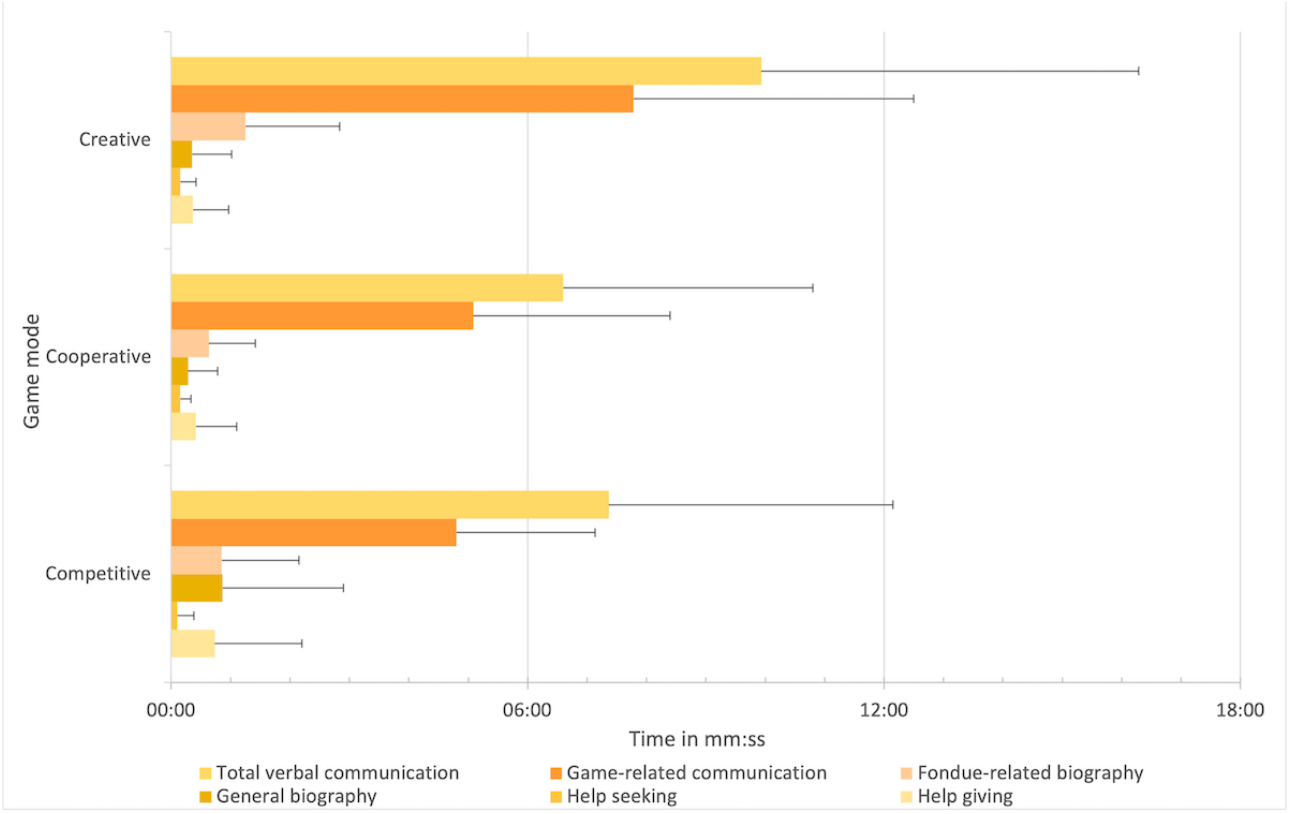


Figure 5. Nonverbal communication (absolute values).

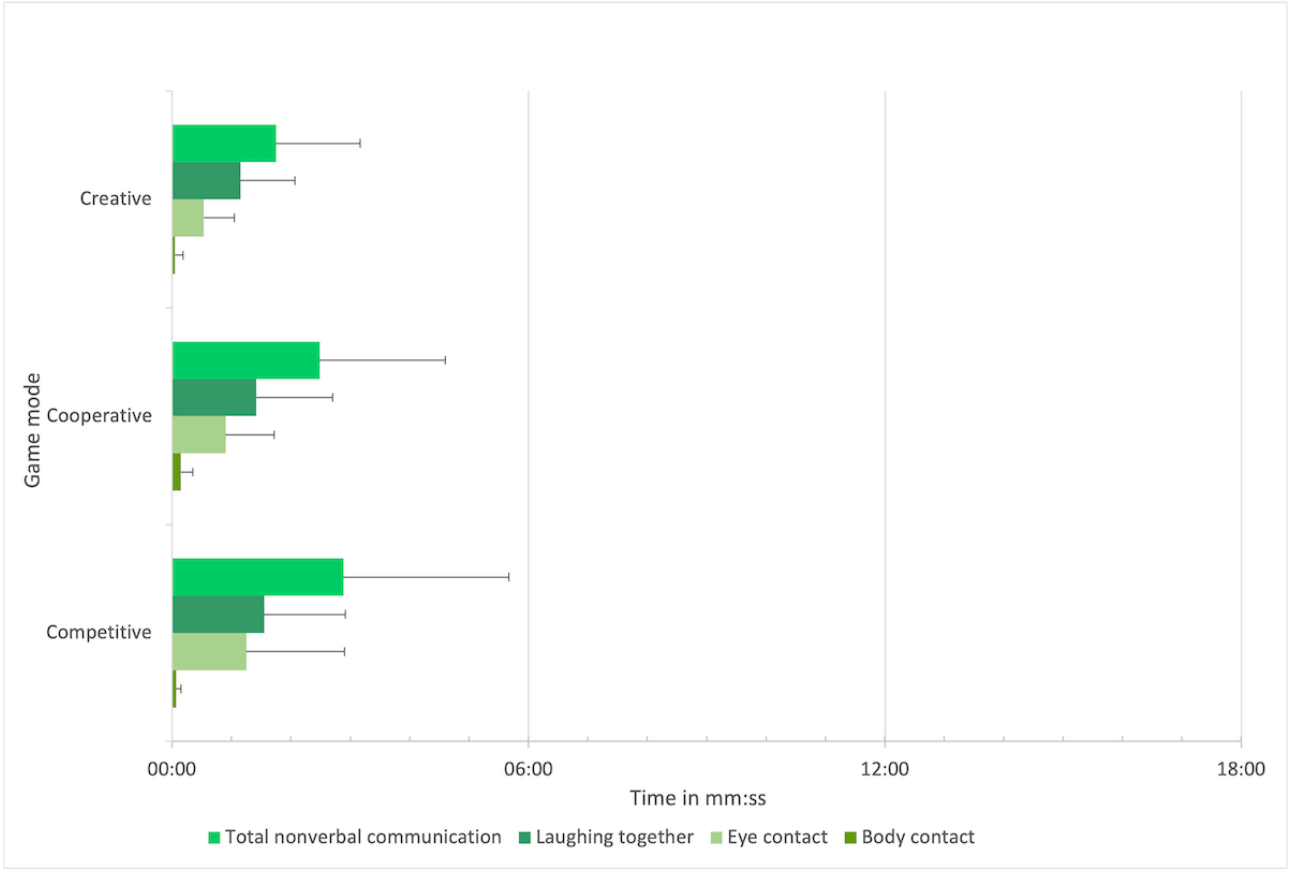


Figure 6. Verbal communication (relative values).

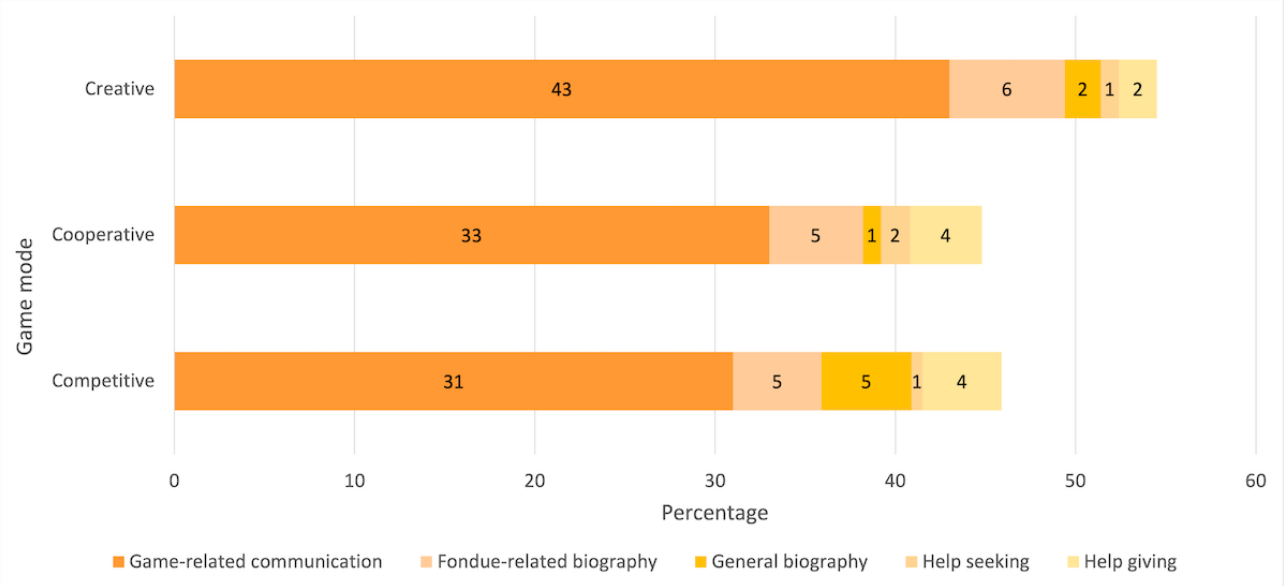
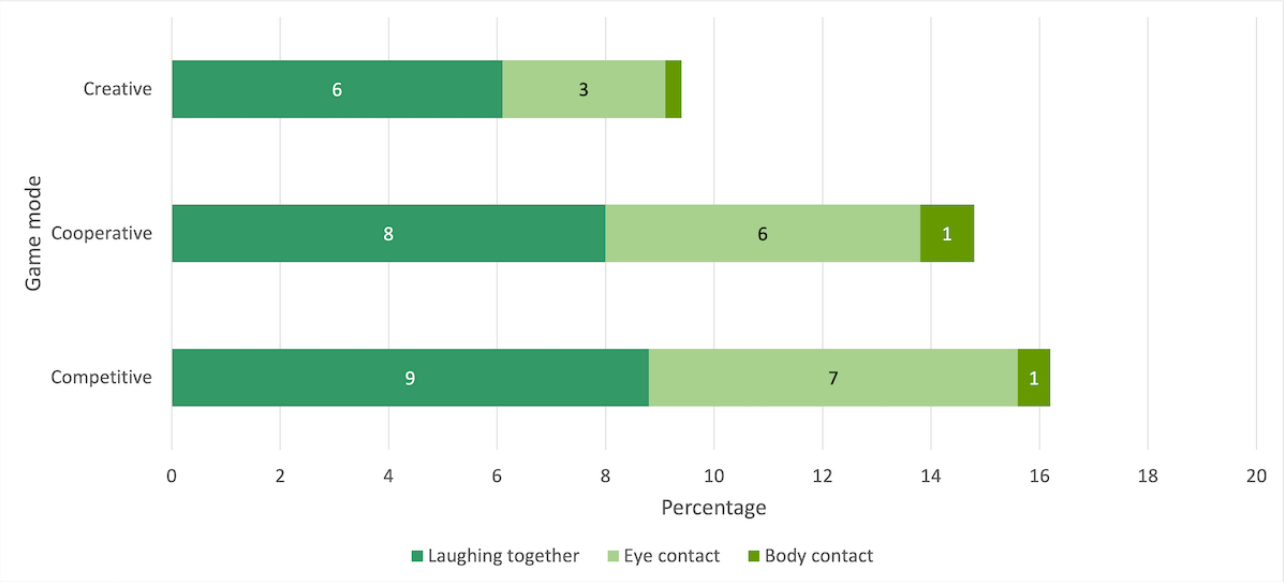


Figure 7. Nonverbal communication (relative values).



As seen in Table 3, on a descriptive level, the *absolute total values* of social interaction were the longest in the creative mode, followed by the competitive and cooperative modes. However, the analysis of variance for repeated measurements showed no significant differences among the game modes (P=.29). The same holds true for the *relative values* (Table 3), that is, percentage of time with observed social interaction in relation to the total playing time (P=.78).

Simultaneous Occurrence of Verbal and Nonverbal Codes

As verbal and nonverbal categories tend to occur simultaneously, overlapping frequencies between verbal and nonverbal codes are presented in Table 2. During game-related communication, players often laugh together and look at each other in their eyes. However, during biographic communication (ie, fondue-related and general biography), players rarely had eye contact or laughed together. Regardless of the content of verbal communication, body contact rarely occurred.

Hypothesis 1: Verbal Communication

Influence of Game Modes on Total Verbal Communication

On a descriptive level, the *absolute total values* of verbal communication were higher in the creative mode, followed by the competitive and cooperative modes (Table 3). Analysis of variance for repeated measures yielded marginal significance ($F_{2,18}$=3.37; P=.06; partial η^2=0.272; N=10). The effect size f, according to Cohen (1988), was 0.611, which corresponds to a strong effect. Bonferroni-corrected pairwise comparisons revealed that there was significantly more verbal communication (P=.04) in the creative mode (mean 9:56, SD 6:21) than in the cooperative mode (mean 6:36, SD 4:12). For the *relative values*, the analyses did not yield statistical significance (P=.11).

Influence of Game Modes on Game-Related Communication

As seen in Table 3, on a descriptive level, the *absolute values* of game-related verbal communication were higher in the creative mode, followed by the cooperative and competitive modes. An analysis of variance for repeated measurements showed that the *absolute values* of game-related communication differed significantly among the game modes ($F_{2,18}$=6.36; *P*=.01; partial η^2=0.414; N=10), with a strong effect (Cohen *f*=0.841). Bonferroni-corrected pairwise comparisons showed that there was significantly more game-related communication (*P*=.01) in creative mode (mean 7:47, SD 4:43) than in cooperative mode (mean 5:05; SD 3:19) and marginally significantly more game-related communication in cooperative mode (*P*=.09) than in competitive mode (mean 4:48, SD 2:20).

Analysis of variance for repeated measurements also showed significant differences among game modes ($F_{2,18}$=4.85; *P*=.02; partial η^2=0.350; N=10), with a strong effect (Cohen *f*=0.734) for the *relative values* of game-related communication. Bonferroni-corrected pairwise comparisons revealed that in the creative mode (mean 43%, SD 18.7%), the percentage of game-related communication was significantly higher (*P*=.01) than in the cooperative mode (mean 32.9%, SD 15.8%).

Influence of Game Modes on Fondue-Related Biography Talk

As seen in Table 3, on a descriptive level, the *absolute values* of fondue-related biography communication were higher in the creative mode, followed by the competitive and cooperative modes. According to relative values, values were higher in the creative mode than in the cooperative mode, followed by the competitive mode. The Friedman test showed marginally significant differences among the game modes for the *absolute values* (χ^2_2=5.56; *P*=.06; N=10), but not for the *relative values* (*P*=.37; Table 3). For the absolute values, post hoc tests (Dunn-Bonferroni tests) revealed significantly more verbal communication regarding fondue-related biography (*z*=2.24; *P*=.03) in the creative mode (mean 1:15, SD 1:35) than in the cooperative mode (mean 0:38, SD 0:47) with a strong effect (*r*=0.707).

Influence of Game Modes on General Biography Talk

As seen in Table 3, on a descriptive level, the *absolute values* of general biography were higher in the competitive mode, followed by the creative and cooperative modes. No significant differences were found among the game modes (*P*=.91). This also holds true for *relative values* (*P*=.91).

Influence of Game Modes on Help Seeking and Help Giving

Table 3 shows on a descriptive level the *absolute values* of help seeking and help giving. No significant differences were found between the game modes for help seeking (*P*=.28) and help giving (*P*=.47). For the relative values, a Friedman test showed a significant difference among the game modes (χ^2_2=6.44; *P*=.04; N=10) for help seeking. Subsequent post hoc tests (Dunn-Bonferroni tests) revealed that help seeking was significantly higher (*z*=2.24; *P*=.03) in the cooperative mode (mean 1.6%, SD 2.3%) than in the competitive mode (mean 0.6%, SD 1.4%) with a strong effect (*r*=0.707). Concerning the absolute and relative values for the category help giving, no significant differences were found among the game modes.

In sum, we expected the creative game mode to stimulate the highest amount of verbal social interaction, because players are given a high degree of freedom, which could stimulate verbal interaction. Despite somewhat controversial findings regarding the competitive and cooperative modes, we assumed in this study that a cooperative mode would stimulate more verbal social interaction than the competitive mode based on related research (H1). The hypothesis could partly be confirmed: for total verbal communication and the subcategories game-related and fondue-related communication, where data indeed show higher values for the creative mode, indicating more time spent on verbal communication here than in the other modes, partly with strong effects. With regard to the effects of cooperative and competitive modes, the picture is less clear. The same is true for the other subcategories (general biography talk, help seeking, and help giving). Considering the rationale behind the assumptions of H1, this pattern of results is interesting and can be explained so that in creative mode, more verbal communication occurs, because players are given a high degree of freedom to exchange and develop new ideas. However, as the data show, this could only hold true for *some* subcategories of communication, namely those that directly relate to playing the game itself (game-related and fondue-related biography). On a more general level, game modes may influence verbal communication in different ways, but only for specific content.

Hypothesis 2: Nonverbal Communication

Influence of Game Modes on Total Nonverbal Communication

The mean *absolute values* of nonverbal communication, as shown in Table 3, were highest in the competitive mode, followed by the cooperative and creative modes. An analysis of variance for repeated measurements (Greenhouse–Geiser correction was applied) indicated marginally significant differences (*P*=.09), but the differences did not reach significance in the subsequent post hoc tests (*P*=.14, *P*=.26, and *P*=.84).

A Friedman test for the *relative* values showed a significant difference among the game modes (χ^2_2=9.80; *P*=.01; N=10). Following post hoc tests (Dunn-Bonferroni tests) revealed that the nonverbal communication was significantly higher (*z*=−2.91; *P*=.004) in the competitive mode (mean 16.2%, SD 12.5%) than in the creative mode (mean 9.4%, SD 6.2%) with a strong effect (*r*=0.919), and significantly higher (*z*=−2.46; *P*=.01) in the cooperative mode (mean 14.8%, SD 9.3%) than in the creative mode with a strong effect (*r*=0.778).

In the following sections, the results will be reported for different categories of nonverbal communication.

Influence of Game Modes on Laughing Together

The mean *absolute values* of laughing together, as shown in Table 3, were highest in the competitive mode, followed by the

cooperative and creative modes. An analysis of variance for repeated measurements with the absolute values for laughing together showed a marginally significant difference among game modes (P=.08) but this difference was not significant in the subsequent post hoc tests (P=.27, P=.26, and P=.99). Regarding the *relative values*, an analysis of variance for repeated measurements showed significant differences among the game modes ($F_{2,18}$=5.18; P=.02; partial η^2=0.365; N=10), with a strong effect (Cohen f=0.758). Bonferroni-corrected pairwise comparisons revealed that there was more laughing together in the competitive mode (mean 8.8%, SD 6.3%) than in the creative mode (mean 6.1%, SD 4.1%; P=.047) and marginally significantly more laughing together (P=.07) in the cooperative mode (mean 8%, SD 5.5%) than in the creative mode.

Influence of Game Modes on Eye Contact

There was no significant difference between the game modes for either the absolute (P=.21) or the relative values of eye contact (P=.12).

Influence of Game Modes on Body Contact

There was no significant difference between the game modes for either the absolute (P=.63) or the relative values of body contact (P=.41).

In sum, regarding nonverbal social interaction, it was expected that there would be differences among game modes (H2), but expectations remained on an explorative level, owing to a lack of a theoretical or empirical research basis for directional assumptions. The results reveal that the total time spent with nonverbal communication and laughing together was highest in the competitive mode and lowest in the creative mode, with significant differences. No significant differences were found between the eye and body contact subcategories.

Discussion

Principal Findings

In the study presented here, the influence of 3 different game modes of the multiplayer video game *Myosotis FoodPlanet* on the social interactions between the older players living in a retirement home and their younger coplayers was investigated. It was expected that different game modes would influence verbal and nonverbal communication among players in different ways (H1 and H2).

First, overall (across all modes), it was demonstrated that the game could successfully be applied and played as intended in the field situation (retirement home) by the residents and their younger coplayers (care professionals). All 3 game modes could *successfully* stimulate positive social interactions, such as talking and laughing together between the older adults and their younger coplayers (albeit in different ways, see below). Simultaneous occurrences (Table 2) of game-related verbal communication and the 2 nonverbal categories of laughing together and eye contact were found. Regardless of the game mode, players tended to stop playing and listen carefully when the other player told a story from the past. Overall, these study results are in line with previous research showing that game-mediated play paves the way for positive social interactions [13,14] and can facilitate communication among players belonging to different age groups or generations [9,10,12]. Thus, important practical goals of the study were met, as it was specifically designed as an intergenerational digital game for use in retirement homes. The game was designed to be fun and entertaining for different generations to support mutual empathy and active listening, and it obviously worked.

Second, partly confirming our assumptions about *the differential* effects of *different* game modes on social interaction, some significant differences were found. To highlight the most important outcomes, it is noted that the creative game mode was significantly more supportive than the other modes with respect to the subcategories of verbal communication, game-related communication, and fondue-related biography talk but not for general biography talk or help giving. In other words, the creative game mode of *Myosotis FoodPlanet* could—better than other game modes—foster social interactions between older adults and their younger coplayers regarding talking about the game and their joint activities during playing. The older adults seemingly related their memories from the past to the present game content (ie, remembering past fondue-events). Furthermore, our results revealed significantly more help seeking communication in the cooperative mode than in the competitive mode. Taken together, these findings support H1. However, with regard to the cooperative and competitive modes, the picture remains fuzzy. The results were inconsistent, mirroring the situation in previous research, where results from empirical studies comparing different modes in relation to social interaction are inconsistent, as described above (in the introduction section on game design decisions and their possible influences on social interaction). Finally, the competitive mode was significantly better than the other modes in promoting the subcategory of laughing together in nonverbal social interaction, whereas no effects were found for eye or body contact. Therefore, H2 was only partly confirmed. This result is in line with related research showing a positive relationship between social and competition motives [39,40]. This seemingly contradicts previous results showing that older adults (in contrast to younger players) would find competition in playing a minor motivator [16,18,22,27]. This can be explained by the unthreatening but still exciting and humorous character of the specific game *Myosotis FoodPlanet* used here.

In summary, the results show how design decisions concerning the choice of game mode can be an important factor in shaping the social interactions of players of different age groups while playing together. In a scientific sense, these original results add to the research gap identified in research on intergenerational games [23], because they provide comparative evidence that can be explained by theory. The creative mode included a high level of freedom for the players and practically demanded communication during the game to *prepare* the players' favorite fondue. It stimulated negotiations about what ingredients to use, resulting in more game-related communication, compared with the other modes. In addition, the creative mode provided an environment in which the players were not under the pressure of wanting to win. Hence, they could talk freely while playing the game, which in turn could have stimulated the cued recall

of associated individual memories concerning fondue-related biographic communication.

Strengths and Limitations

This study has its strengths and limitations. An important strength of this study pertains to the study setting. In comparison to previous studies, which were often conducted in a laboratory situation and therefore might lack ecological validity, our study was conceptualized as a field trial. Thus, we could examine the participants in their natural environment, which is important when looking at aspects such as social interaction and drawing real-world conclusions [57]. However, there are limitations to this approach. In this section, we address these limitations and justify the value of the study despite these limitations.

The small number of participants involved was a serious limitation. This was owing to practical limitations on the side of the retirement homes (eg, their willingness, trust, and available resources to undergo the effort of participating in this field research with researchers coming to their place, bringing in video games for the residents, and placing video cameras to perform recordings). We justify the study despite this limitation by the amount of field data gathered despite the low number of participants. We recorded 30 game sequences, resulting in approximately 8 hours of video material (7 hours 53 minutes 25 seconds). Social interactions were observed, comprising a volume of almost 4 hours of video (3:51:47) with 7570 coded social interactions in the different categories, coded thoroughly according to well-established methods in experimental psychology. Our fine-grained videotaping and coding procedure allowed us to analyze both verbal and especially, nonverbal communication, from which we gained crucial insights into the differences among the game modes. Hence, an important further strength of this study arises from this systematic effort: not only verbal communication but also nonverbal interactions, such as laughing together or eye contact are available and produced interesting new results in this area of research. However, most importantly, significant results were obtained despite the small number of participants (N), and sometimes with effect sizes indicating strong effects, such as the results on the influence of game modes on game-related and fondue-related biography talk, and on total nonverbal communication (see the *Results* section).

We must also note that the younger coplayers in this study were care professionals. Although this fills a gap in research on intergenerational games by looking *further than only grandparents–grandchildren interactions* [23], this may have compromised our study and demands additional caution in interpretation and generalization. Such professionals are trained to read the skill levels of the older adults and adapt their own pace accordingly. Furthermore, older residents may have viewed them as their *caregivers*, in addition to viewing them as *members of the younger generation*. Thus, it remains open at this point how nontrained younger coplayers or family members would behave in such gaming situations and how social interactions would differ from the situation studied here. For instance, it may be that some study outcomes would look different; with family members as younger coplayers, there might be more (joint) general biography-related memories to discuss. The players may invest more time in general biography talk with implications for the effects of the 3 game modes on this subcategory of social interaction. In addition, the patterns of help seeking and help giving types of communication would likely produce different results. The same might be true for body contact. This limitation characterizes our study outcomes as *intergenerational* only in terms of age (younger vs older) but does not necessarily generalize to other characterizations such as family links (eg, grandparents and grandchildren) or organizational membership (eg, juniors and seniors) or other [24]. This may be the subject of future studies.

The game used in this study is a serious multiplayer game called *Myosotis FoodPlanet*, which was designed specifically as a serious intergenerational game for practical use in retirement homes. It was also used for local game events with older adults in different areas of Switzerland. Using this real game supports the ecological validity of the study but somewhat compromises internal validity, as control over all elements and features in the game was not possible. During the game design process, we could not control for all possible confounds across game modes; for example, the creative mode includes some additional elements, such as additional effects and recipes with humorous names. This must be discussed in terms of alternative explanations for the findings showing more verbal communication in the creative mode than in the other modes (confirming H1): Could it be that more talking was stimulated by a more humorous game mode? When viewed together with the results on nonverbal communication, this alternative seems rather unlikely, as significantly more time of *laughing together* was found in the competitive mode than in the creative mode. This suggests that humor was not limited to the creative game mode but a characteristic of the game *Myosotis FoodPlanet* in all modes investigated here. Therefore, we understand our study results still as confirming H1. Nevertheless, this interpretation must remain cautious and initial. Further systematic experimental studies are needed to deepen scientific knowledge on this specific topic.

Associated with the previous information, the game addresses a specific theme (Swiss fondue cooking), represents a specific genre (casual), and has a specific target (intergenerational game playing in retirement homes to improve social interactions). Generalization of the study results to other games and situations is not easily possible. Thus, comparisons among different games (genres, mechanics, and themes) and coplayer's profiles in systematic experimental research could be interesting in future research.

Nevertheless, it is a strength of this study in comparison with previous work that it examines the creative game mode with this specific game in a field trial and that it systematically compares the effects of the 3 game modes on social interactions in a natural setting. Previous studies have primarily addressed the influence of either cooperative or competitive play on social interaction [33,38], whereas a direct comparison of cooperative, competitive, and creative game modes regarding their similarities and differences in promoting different types of social interaction has been lacking so far.

Conclusions

What does the study reveal for the community at the end? In sum, this study highlights the importance of game mode when designing serious multiplayer games for intergenerational game playing and can inform game designers on *how* to design games to stimulate specific types of intergenerational social interaction. Moreover, the results have implications for game-playing situations, such as the one investigated here in a retirement home, when the aim is to promote intergenerational social interaction between care professionals and residents (younger and older). Depending on the type of social interaction (eg, verbal or nonverbal communication) that is intended, a specific game mode or specific elements may be more appropriate than others. When an increase in verbal communication is desired and when older adults should be motivated to talk with younger (care) persons, a creative game mode with no distinct goal and no right or wrong pathway should be considered. On the contrary, when an increase in laughing together is intended (eg, to provide for a positive mood and atmosphere), one might recommend a competitive game mode with an unthreatening, though arousing and stimulating character. These conclusions remain tentative owing to the study limitations and limited research in this area. We hope that our study will stimulate future research.

Acknowledgments

The authors wish to thank Adam Coyle, Maja Riegler, Markus Recher, Michael Kunkel, Shintaro Miyazaki, Thomas Resch, and Tudor Andrievici. The authors thank Anna Caltabiano for proofreading the manuscript. This research was funded by the University of Applied Sciences and Arts of Northwestern Switzerland—Strategic Initiative: “Myosotis Garden.”

Authors' Contributions

CZ, DL, MN, and ALR are psychologists who have conducted the research study, including all analyses and data interpretations. MS and TI created the Myosotis Fondue game and the different game modes needed for this research. They have also provided the background on professional game design. The authors closely collaborated in their work and in the writing of the manuscript.

Conflicts of Interest

MS is the owner of Holunder Games GmbH, a recently funded small game studio, which creates computer games for older adults. Holunder Games GmbH relies on the outcomes of various projects, including the outcome of the Myosotis project.

Editorial notice: This randomized study was not registered. The authors explained that the Swiss Federal Human Research Act agreed that the study was not liable for registration since it addressed no health-related outcomes. The editor granted an exception from ICMJE rules mandating prospective registration of randomized trials, because the risk of bias appears low and the study was considered formative, guiding the development of game design and the use of computer games to shape the social interactions of seniors and their coplayers. However, readers are advised to carefully assess the validity of any potential explicit or implicit claims related to primary outcomes or effectiveness.

Multimedia Appendix 1
Game design of the Myosotis Fondue game.
[MP4 File (MP4 Video), 94173 KB - formative_v6i2e29179_app1.mp4]

Multimedia Appendix 2
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 45926 KB - formative_v6i2e29179_app2.pdf]

References

1. Baumeister RF, Leary MR. The need to belong: desire for interpersonal attachments as a fundamental human motivation. Psychol Bull 1995;117(3):497-529. [doi: 10.1037/0033-2909.117.3.497] [Medline: 7777651]
2. Fang Y, Chau AK, Wong A, Fung HH, Woo J. Information and communicative technology use enhances psychological well-being of older adults: the roles of age, social connectedness, and frailty status. Aging Ment Health 2018;22(11):1516-1524. [doi: 10.1080/13607863.2017.1358354] [Medline: 28777010]
3. Khosravi P, Rezvani A, Wiewiora A. The impact of technology on older adults’ social isolation. Comput Hum Behav 2016;63:594-603. [doi: 10.1016/j.chb.2016.05.092]
4. Leist AK. Social media use of older adults: a mini-review. Gerontology 2013;59(4):378-384. [doi: 10.1159/000346818] [Medline: 23594915]
5. Khoo ET, Merritt T, Cheok AD. Designing physical and social intergenerational family entertainment. Interact Comput 2009;21(1-2):76-87. [doi: 10.1016/j.intcom.2008.10.009]
6. Derboven J, Van Gils M, De Grooff D. Designing for collaboration: a study in intergenerational social game design. Univ Access Inf Soc 2011;11(1):57-65. [doi: 10.1007/s10209-011-0233-0]

7. Al Mahmud A, Mubin O, Shahid S, Martens JB. Designing social games for children and older adults: two related case studies. Entertain Comput 2010;1(3-4):147-156. [doi: 10.1016/j.entcom.2010.09.001]
8. Ijsselsteijn W, Nap HH, de Kort Y, Poels K. Digital game design for elderly users. In: Proceedings of the 2007 Conference on Future Play. New York: Association for Computing Machinery; 2007 Presented at: Future Play '07; November 14-17, 2007; Toronto p. 17-22. [doi: 10.1145/1328202.1328206]
9. Kooiman BJ, Sheehan DP. Intergenerational remote exergaming with family and friends for health and leisure. J Intergener Relatsh 2014;12(4):413-424. [doi: 10.1080/15350770.2014.962442]
10. Chen Y, Wen J, Xie B. "I communicate with my children in the game": mediated intergenerational family relationships through a social networking game. JoCI 2012;8(1):-. [doi: 10.15353/joci.v8i1.3056]
11. Chiong C. The Joan Ganz Cooney Center at Sesame workshop. 2010. URL: https://joanganzcooneycenter.org/publication/can-video-games-promote-intergenerational-play-literacy-learning/ [accessed 2022-01-26]
12. Khoo ET, Cheok AD, Nguyen TH, Pan Z. Age invaders: social and physical inter-generational mixed reality family entertainment. Virtual Real 2008;12(1):3-16. [doi: 10.1007/s10055-008-0083-0]
13. Holladay SJ, Seipke HL. Communication between grandparents and grandchildren in geographically separated relationships. Commun Stud 2007;58(3):281-297. [doi: 10.1080/10510970701518371]
14. Marx MS, Cohen-Mansfield J, Renaudat K, Libin A, Thein K. Technology-mediated versus face-to-face intergenerational programming. J Intergener Relatsh 2005;3(3):101-118. [doi: 10.1300/j194v03n03_07]
15. Othlinghaus J, Gerling KM, Masuch M. Intergenerational play: exploring the needs of children and elderly. In: Eibl M, Ritter M, editors. Workshop-Proceedings der Tagung Mensch & Computer 2011. Chemnitz: Universitätsverlag Chemnitz; 2011:317-322.
16. Salmon JP, Dolan SM, Drake RS, Wilson GC, Klein RM, Eskes GA. A survey of video game preferences in adults: Building better games for older adults. Entertain Comput 2017;21:45-64. [doi: 10.1016/j.entcom.2017.04.006]
17. De Schutter B. Never too old to play: the appeal of digital games to an older audience. Games Cult 2010;6(2):155-170. [doi: 10.1177/1555412010364978]
18. Nap HH, de Kort Y, IJsselsteijn WA. Senior gamers: preferences, motivations and needs. Gerontechnology 2009;8(4):247-262. [doi: 10.4017/gt.2009.08.04.003.00]
19. Ballard M, Visser K, Jocoy K. Social context and video game play: impact on cardiovascular and affective responses. Mass Commun Soc 2012;15(6):875-898. [doi: 10.1080/15205436.2011.632106]
20. Lazzaro N. Why we play games: four keys to more emotion without story. In: Proceedings of the Game Developers Conference 2004. 2004 Presented at: GDC'04; March 22-26, 2004; San Jose.
21. Zagal JP, Nussbaum M, Rosas R. A model to support the design of multiplayer games. Presence (Camb) 2000;9(5):448-462. [doi: 10.1162/105474600566943]
22. Loos E. Designing meaningful intergenerational digital games. In: International Conference on Communication, Media, Technology and Design. 2014 Presented at: ICCMTD'14; April 24-26, 2014; Istanbul p. 46-51.
23. de la Hera T, Loos E, Simons M, Blom J. Benefits and factors influencing the design of intergenerational digital games: a systematic literature review. Societies 2017;7(3):18. [doi: 10.3390/soc7030018]
24. Sánchez-Martínez M, Kaplan M, Bradley L. Using technology to connect generations: some considerations of form and function. Comunicar 2015;23(45):95-104 [FREE Full text] [doi: 10.3916/c45-2015-10]
25. Nguyen TT, Ishmatova D, Tapanainen T, Liukkonen TN, Katajapuu N, Makila T, et al. Impact of serious games on health and well-being of elderly: a systematic review. In: Proceedings of the 50th Hawaii International Conference on System Sciences. 2017 Presented at: HICCS'17; January 4-7, 2017; Hawaii p. 3695-3704. [doi: 10.24251/hicss.2017.447]
26. Martinho D, Carneiro J, Corchado JM, Marreiros G. A systematic review of gamification techniques applied to elderly care. Artif Intell Rev 2020;53(7):4863-4901. [doi: 10.1007/s10462-020-09809-6]
27. Osmanovic S, Pecchioni L. Beyond entertainment: motivations and outcomes of video game playing by older adults and their younger family members. Games Cult 2015;11(1-2):130-149. [doi: 10.1177/1555412015602819]
28. Pereira F, Bermúdez i Badia S, Jorge C, Cameirão MS. The use of game modes to promote engagement and social involvement in multi-user serious games: a within-person randomized trial with stroke survivors. J NeuroEngineering Rehabil 2021;18(1):62 [FREE Full text] [doi: 10.1186/s12984-021-00853-z]
29. Fullerton T. Game design workshop: a playcentric approach to creating innovative games. 4th edition. New York: CRC Press; 2018. [doi: 10.1201/b16671]
30. Santos LH, Okamoto K, Funghetto SS, Cavalli AS, Hiragi S, Yamamoto G, et al. Effects of social interaction mechanics in pervasive games on the physical activity levels of older adults: quasi-experimental study. JMIR Serious Games 2019;7(3):e13962. [doi: 10.2196/13962] [Medline: 31333202]
31. Zagal JP, Rick J, Hsi I. Collaborative games: lessons learned from board games. Simul Gaming 2006;37(1):24-40. [doi: 10.1177/1046878105282279]
32. Heeter C. Play styles and learning. In: Ferdig RE, editor. Handbook of research on effective electronic gaming in education. New York: IGI Global; 2009:826-846.
33. Sweetser P, Wyeth P. GameFlow: a model for evaluating player enjoyment in games. Comput Entertain 2005;3(3):3. [doi: 10.1145/1077246.1077253]

34. Wulf T. Kooperation und Kompetition im Videospiel: Der Einfluss Sozialer Interdependenz auf Stimmung und prosoziales Verhalten. Wiesbaden: Springer; 2017. [doi: 10.1007/978-3-658-16682-3_4]
35. Johnson DW, Johnson RT. Cooperation and competition: theory and research. Minneapolis: Interaction Book Company; 1989.
36. Squire K. Open-ended video games: a model for developing learning for the interactive age. In: Salen Tekinbaş K, editor. The ecology of games: connecting youth, games, and learning. Cambridge: MIT Press; 2007:167-198.
37. Soldati M, Zahn C, Bildibay D, Iseli T, Leisner D, Niederhauser M, et al. Create video games to promote well-being of elderly people - a practice-driven guideline. In: Gao Q, Zhou J, editors. Human aspects of IT for aged population. Healty and active aging. Cham: Springer; 2020:401-418. [doi: 10.1007/978-3-030-50249-2_29]
38. Abeele VV, De Schutter B. Designing intergenerational play via enactive interaction, competition and acceleration. Pers Ubiquitous Comput 2010;14(5):425-433. [doi: 10.1007/s00779-009-0262-3]
39. Jansz J, Tanis M. Appeal of playing online first person shooter games. Cyberpsychol Behav 2007;10(1):133-136. [doi: 10.1089/cpb.2006.9981]
40. Jansz J, Martens L. Gaming at a LAN event: the social context of playing video games. New Media Soc 2005;7(3):333-355. [doi: 10.1177/1461444805052280]
41. Schutz W. FIRO: a three-dimensional theory of interpersonal behavior. New York: Rinehart and Company; 1958.
42. Vasconcelos A, Silva PA, Caseiro J, Nunes F, Teixeira LF. Designing tablet-based games for seniors: the example of CogniPlay, a cognitive gaming platform. In: Proceedings of the 4th International Conference on Fun and Games. 2012 Presented at: FnG '12; September 4-6, 2012; Toulouse p. 1-10. [doi: 10.1145/2367616.2367617]
43. Kaufman D, Sauve L, Renaud L, Sixsmith A, Mortenson B. Older adults' digital gameplay: patterns, benefits, and challenges. Simul Gaming 2016;47(4):465-489. [doi: 10.1177/1046878116645736]
44. de la Hera Conde-Pumpido T, Paz Aléncar A. Collaborative digital games as mediation tool to foster intercultural integration in primary Dutch schools. eLearning Papers 2015;43:13-23.
45. Khoo ET, Merritt T, Cheok A, Lian M, Yeo K. Age invaders: user studies of intergenerational computer entertainment. In: 6th International Conference on Entertainment Computing. Heidelberg: Springer; 2007 Presented at: ICEC '07; September 15-17, 2007; Shanghai p. 231-242. [doi: 10.1007/978-3-540-74873-1_28]
46. Davis MA. Understanding the relationship between mood and creativity: a meta-analysis. Organ Behav Hum Decis Process 2009;108(1):25-38. [doi: 10.1016/j.obhdp.2008.04.001]
47. Stahl G. Meaning and interpretation in collaboration. In: Wasson B, Ludvigsen S, Hoppe U, editors. Designing for change in networked learning environments: proceedings of the International Conference on Computer Support for Collaborative Learning 2003. Dordrecht: Springer; 2003:523-532.
48. Bereiter C. Education and mind in the knowledge age. Mahwah: L. Erlbaum Associates; 2002.
49. Butler RN. The life review: an interpretation of reminiscence in the aged. Psychiatry 1963;26:65-76. [doi: 10.1080/00332747.1963.11023339] [Medline: 14017386]
50. Stinson CK. Structured group reminiscence: an intervention for older adults. J Contin Educ Nurs 2009;40(11):521-528. [doi: 10.3928/00220124-20091023-10] [Medline: 19904866]
51. Argyle M. Non-verbal communication in human social interaction. In: Hinde RA, editor. Non-verbal communication. Cambridge: Cambridge University Press; 1972:243-270.
52. Sternath B, Zweidler R. Einsatz von video in der evaluation: umsetzung von evaluationsstandards und datenschutz. Göttingen: Cuvillier; 2012.
53. Rack O, Zahn C, Mateescu M. Coding and counting: frequency analysis for group interaction research. In: Brauner E, Boos M, Kolbe M, editors. The Cambridge handbook of group interaction analysis. Cambridge: Cambridge University Press; 2018:277-294. [doi: 10.1017/9781316286302.015]
54. Schramm K, Aguado K. Videographie in den Fremdsprachendidaktiken - Ein Überblick. In: Aguado K, Schramm K, Vollmer JH, editors. Fremdsprachliches Handeln beobachten, messen, evaluieren: neue methodische ansätze der kompetenzforschung und der videographie. Frankfurt: Peter Lang; 2010:185-214.
55. Petko D, Waldis M, Pauli C, Reusser K. Methodologische Überlegungen zur videogestützten Forschung in der Mathematikdidaktik. Zentralblatt für Didaktik der Mathematik 2003;35(6):265-280. [doi: 10.1007/bf02656691]
56. Mayring P, Fenzl T. Qualitative Inhaltsanalyse. In: Baur N, Blasius J, editors. Handbuch Methoden der empirischen Sozialforschung. 2nd edition. Wiesbaden: Springer; 2019:633-648. [doi: 10.1007/978-3-658-21308-4_42]
57. Maner JK. Into the wild: field research can increase both replicability and real-world impact. J Exp Soc Psychol 2016;66:100-106. [doi: 10.1016/j.jesp.2015.09.018]

Edited by A Mavragani; submitted 29.03.21; peer-reviewed by Y Fang, M Eckert, R Pine, T Silva; comments to author 11.05.21; revised version received 07.07.21; accepted 17.12.21; published 16.02.22.

Please cite as:
Zahn C, Leisner D, Niederhauser M, Roos AL, Iseli T, Soldati M
Effects of Game Mode in Multiplayer Video Games on Intergenerational Social Interaction: Randomized Field Study
JMIR Form Res 2022;6(2):e29179
URL: https://formative.jmir.org/2022/2/e29179
doi:10.2196/29179
PMID:35171097

Original Paper

Digital Companion Choice to Support Teachers' Stress Self-management: Systematic Approach Through Taxonomy Creation

Julia B Manning[1], BSc; Ann Blandford[2], PhD; Julian Edbrooke-Childs[3], PhD

[1]Interaction Centre, Department of Computer Science, University College London, London, United Kingdom
[2]Institute of Healthcare Engineering, University College London, London, United Kingdom
[3]Evidence-based Practice Unit, University College London and Anna Freud Centre, London, United Kingdom

Corresponding Author:
Julia B Manning, BSc
Interaction Centre
Department of Computer Science
University College London
2nd Floor, 66-72 Gower Street
London, WC1E 6EA
United Kingdom
Phone: 44 07973312358
Email: j.manning.17@ucl.ac.uk

Abstract

Background: There are thousands of digital companions designed for emotional well-being and stress, including websites, wearables, and smartphone apps. Although public evaluation frameworks and ratings exist, they do not facilitate digital companion choice based on contextual or individual information, such as occupation or personal management strategies.

Objective: The aim of this study is to establish a process for creating a taxonomy to support systematic choice of digital companions for teachers' stress self-management.

Methods: We used a 4-step study design. In step 1, we identified the dimension of stress self-management and strategic classifications. In step 2, we identified the dimension of the digital techniques and conceptual descriptions. In step 3, we created 6 criteria for the inclusion of digital companions. In step 4, we used the taxonomy framework created by steps 1 and 2 and populated it with digital companions for stress self-management, as identified in step 3.

Results: First, in the dimension of stress self-management, we identified four classes of strategies: educational, physiological, cognitive, and social. Second, in the digital techniques dimension, we derived four conceptual descriptions for the digital companions' mechanisms of action: fostering reflection, suggesting treatment, peer-to-peer support, and entertainment. Third, we created six criteria for digital companion inclusion in the taxonomy: suitability, availability, evaluation, security, validity, and cost. Using the taxonomy framework and criteria, we populated it with digital companions for stress management ahead of presentation to teachers in a stress study workshop.

Conclusions: The elements of our approach can be generalized as principles for the creation of taxonomies for other occupations or conditions. Taxonomies such as this could be a valuable resource for individuals to understand which digital companion could be of help in their personal context.

(JMIR Form Res 2022;6(2):e32312) doi:10.2196/32312

KEYWORDS

digital technology; digital health; psychological treatment; stress; self-management; mobile phone

Introduction

Background

Self-care digital health smartphone apps, websites, and wearables, referred to collectively in this paper as digital companions, are ubiquitous, but understanding which of these will best support individual needs in a given context is complex. The selection presented to the potential user is immense, with at least 10,000 digital companions targeting behavioral and mental health [1], and the existing approach to digital companion selection is often opportunistic. The availability of mental health apps is hampered by high turnover: 50% of search results change within 4 months, with an app being removed every 2.9 days from web-based platforms [2] and more than 200 health apps being added every day to app stores [3]. Routes to adoption of digital companions for psychological support include recommendations from health professionals [4], although a US study found social media, personal searches, and word of mouth to be more common access routes [5].

User recommendation on app stores is another common route, but it has its limitations, such as including different types and amounts of coverage. In addition, the sources of these reviews are unknown. Taking the reviews at face value, a more detailed exploration of user recommendations of psychological apps has been achieved by machine learning sentiment analysis, revealing the top positive and negative themes for user satisfaction [6]. High cost, app instability, low quality content, and privacy or security concerns were the most common dissatisfaction themes. Tracking, outcome visualization and analytics, and content quality and variety were the most common satisfaction themes. Another study on anxiety apps alone also revealed that price negatively affects adoption, whereas ratings and reviews positively affect downloads, but only up to a point [7]. We also know that app descriptions influence adoption but can be unhelpful. Potentially stigmatizing labeling such as app titles that imply a diagnosis for a mental health condition can constrain access or even cause harm [8]. Some apps use scientific language in their descriptions to verify their clinical validity. However, a study of 73 popular mental and emotional health apps found that although 44% used such language, only 2 apps provided direct scientific evidence associated with app use [9].

More recent studies have begun to elucidate some relevant information on the types of use for technology. One small survey recently showed that although smartphone apps were the most used digital companion to support mental health and well-being, they were often used in conjunction with other tools (eg, social media [10]). Importantly, this study showed a relationship between digital companion medium and purpose: apps are used more for guided activities, relaxation, and tracking; social media is used for sharing experiences and gaining personal understanding; and web-based provision is used for daily stress and anxiety management. This survey did not ask about the use of wearables for stress, but the wearable medical device market continues to grow, with 60% growth predicted between 2019 and 2024 to US $27 billion [11]. Early evidence shows that wearables can accurately capture exposure to psychosocial stress in everyday life [12]. Currently, decisions on wearable choice seem to be guided by perceived value, design, and brand [13] rather than by condition management.

Self-management or treatment techniques are often search terms for digital companions, but critically relevant information such as the suitability of the intervention for an individual's context, occupation, or existing self-management practices are often missing [14]. In meta-analyses of occupational studies where a digital companion had been used to support general well-being [15] or for anxiety, stress, and depression [16], positive effects in these contexts over the short to medium term were noted. However, there is both considerable variation in occupation and little evidence in these studies of any attempt to align an intervention with a particular role or existing individual management strategy. The tendency is simply to trial a digital companion that supports one or more strategies with an occupational cohort, irrespective of the cohort's existing stress management strategies and preferences.

We know that the contexts in which people live and work influence their use of and ability to use health technology [17-20] and previous research has called for tailoring of health care technologies to specific users [21,22]. Contextual or strategic data and insight could logically aid both choice and strategy and, therefore, the potential efficacy of digital companions and user outcomes. As has been noted in the study by de Korte et al [23], research on digital companions designed to have work-related relevance for the mental and physical health of employees is scarce. In this paper, we present the processes of developing both dimensions for a taxonomy and the population criteria that facilitate the selection of contextually appropriate digital support for stress. We chose to work with teachers and focus on their stress self-management because of the very high prevalence of work-related stress, averaging 2100 cases per 100,000 educators in the United Kingdom in 2018 [24]. There are indications that COVID-19 may have exacerbated primary stressors for teachers [25], but we already know that contextual factors such as school organization and culture are critical factors for teachers' experience and management of stress [26-30].

Within the context of schools, individual stress management support can be facilitated by digital companions, particularly if teachers have a taxonomy to inform their choice. This paper, therefore, makes the following contributions:

1. The selection of dimensions within which to classify stress self-management and digital health techniques that could offer support
2. The process applied to develop the taxonomy—one that can potentially be adapted and applied in other contexts where digital support is sought for an individual's health condition to match their practices and values
3. The methodology for populating the taxonomy
4. A populated intervention taxonomy developed for teachers managing stress, with illustrative examples of apps that address teachers' needs, available at the time of writing

Related Work

Overview

We describe here prior work and evidence that fed into our choice of dimensions, classification, and selection. This includes teachers' stress self-management research and previous frameworks and taxonomies on the design and selection of technologies.

Teacher Self-management of Stress

Approaches to aid teachers in stress management have been drawn from the literature on occupational stress and often applied population wide, although not without acknowledgment that "some (strategies) were unnecessary or differentially effective in individual cases" [31]. There is evidence of benefits to teachers from stress awareness education [32] and physiological interventions including adapted mindfulness and relaxation training [33,34] and exercise [35]. Psychological intervention evidence includes, for example, cognitive behavioral therapy (CBT)–based programs [36,37] and mindfulness embedded in psychoeducation with social support adapted for teachers [38]. Reflective supervision and consultation [39] and environmental adjustment or social support [32] have also been shown to be helpful.

Recent systematic reviews have examined teacher stress interventions and found a greater effect size associated with a longer duration of intervention, but most interventions were guided and not self-managed [40,41]. Those interventions that were self-managed demonstrated positive effects, although these varied in size. Such interventions targeted stress or burnout symptom reduction, including positive psychology through gratitude journaling [42] and CBT-based education through bibliotherapy [37].

Digital Companions for Teachers' Stress Management

Delivering stress management interventions digitally can enable uptake. For example, digital delivery could reduce the cost of provision, improving accessibility and reducing risks of stigma [43], which could be highly relevant to teachers. One tailored eHealth (ie, internet or mobile-delivered health care) randomized controlled trial for teachers used an internet-based problem-solving therapy (a form of CBT). Teachers receiving the CBT intervention reported significantly reduced symptoms of depression as well as a reduction in their perceived stress after the trial (7 weeks) and at 3- and 6-month follow-up [44]. Another study examined stress as a contributor to insomnia among teachers, finding that unguided web-based CBT with psychoeducation among mostly female teachers significantly improved sleep [45]. A recent review of the effectiveness of occupational e-mental health interventions identified only one other study that included education sector personnel [46]. This was a self-administered web-based CBT-based intervention, but the participants also received weekly personalized feedback on the modules. The effect on the reduction in perceived stress across all sectors was large [47].

Taxonomy Creation and Digital Technology Selection

Overview

We identified 2 approaches in the literature relevant to our goal of creating and populating a taxonomy. One is the evolution of designer- and researcher-focused frameworks, seeking to improve efficacy and evidence. The other is more focused on clinician and consumer adoption.

Designer and Researcher Frameworks

Frameworks focused on developing and evaluating technologies have led to better formalizing, detailing, and defining of digital companion design. The persuasive design principles discussed by Fogg [48], expanded further by Oinas-Kukkonen and Harjumaa [49] and complemented by a design model by Ritterband et al [50], all informed the development of the behavior intervention technologies model for developers by Mohr et al [51]. This model, along with other theory-based [14,52] and empirically based [20,53] taxonomies and frameworks, has sought to enable both better conceptual design and easier evaluation of digital companions. The Mobile App Rating Scale (MARS) for designers by Stoyanov et al [53], which has been used extensively in the scientific community, was adapted as a consumer assessment version, uMARS [54]. For this study's taxonomy, these models informed our consideration of the digital techniques dimension of the taxonomy.

Clinician and Individual Frameworks

Both the MARS and the uMARS have been used for evaluating apps, with the latter using less technical language for patients to provide feedback on the engagement, functionality, aesthetics, information, and subjective appreciation of quality and impact. The uMARS allows classic human-computer interaction features and elements to be evaluated to assist design iteration, but it was not created to inform final user adoption. Three other relevant *expert review evaluation frameworks* (Reviews) have been created for users.

The ORCHA (Organisation for Review of Care and Health Apps) model, now paywalled, was specifically designed to inform adoption of mostly apps and has some web-based interventions too. Search is based on the condition or digital companion name. Data privacy, user experience, and clinical assurance are each given a score [55,56].

The two other Reviews focus on psychological health: Mindtools and Psyberguide websites [57]. Psyberguide is a public-facing website that enables a search based on conditions or treatment approaches. The approach taken is that the user understands what concepts or treatment they want to choose (eg, tracking or social support), and the focus is on apps. Both websites publish assessment scores on credibility, user experience, and transparency, although Mindtools does not seem to have been updated since 2017. Psyberguide drew on the MARS framework, incorporating additional privacy and security considerations. The American Psychiatric Association app framework [58,59] has also implemented this. Their framework provides a template for user assessment rather than presenting their own assessment scores. It offers both a quick 8-question *screener* and a much more detailed 5-step, 105-question app evaluation process that

allows the end user to judge what is important and a good match. The starting point for this framework is clinical diagnosis, which informs the potential app selection. In theory, their questions could be applied to websites and wearables as well, although this does not appear to have been tested yet.

The main difference between these scales, Reviews, and frameworks, and our intended approach is the starting point. Our goal was to enable digital companion selection to be framed by someone's occupation, condition, and self-management behavior. For this, we required a taxonomy derived for teachers and stress from which they could identify their self-management strategy and supportive technology concept and then identify a digital companion that aligned with these to trial in a future study. To achieve this goal, we first required selecting a logical dimension within which to classify stress self-management. Second, we selected a dimension within which to classify digital techniques that could support these strategies. Finally, we created a rationale for digital companion inclusion and the selection of credible candidates. This outcome is illustrated in Figure 1. This paper describes why we chose the dimensions of self-management strategies and digital companion concepts, how we categorized them, and our approach to identify the potential candidates.

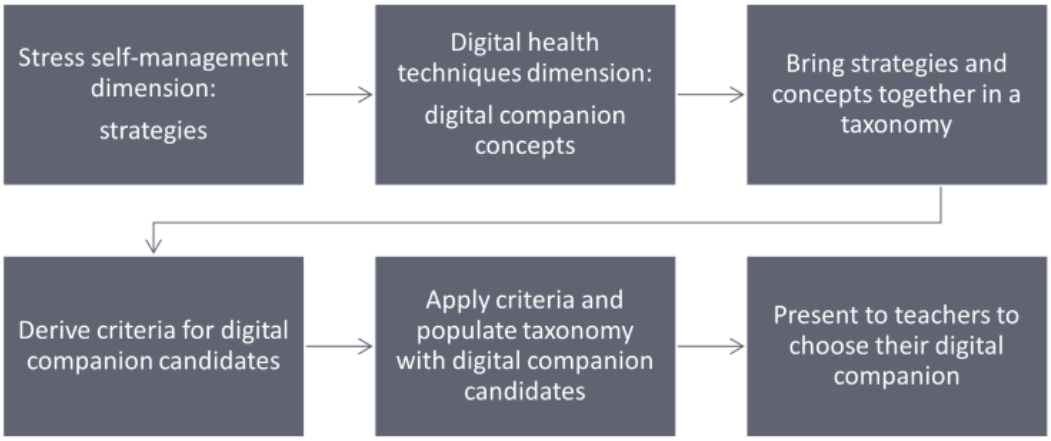
Figure 1. Populated taxonomy with digital stress companion choices for teachers. *Only partial encryption of data **Withdrawn due to lack of updates ***User to provide own device.

Stress self-management strategies → / Digital companion concepts ↓	Educational (eg, knowledge and awareness)	Physiological (eg, relaxation, exercise, reminder)	Cognitive (eg, problem solving, time management)	Social (eg, seeking social support or social skill development)
Fostering reflection via information - health or context	Stress Management at Work Fit2 Teach** TeacherTapp	Withings Steel SR	Daylio Fit2Teach PRO**	Withings Steel SR TeacherTapp
Suggesting treatment and guided self-management	Equoo	Headspace* Calm* Sleepio	SilverCloud Wysa Big White Wall	Big White Wall
Peer-to-peer social support	Fit2Teach PRO**	Withings Steel SR Sleepio	Fit2Teach PRO** Big White Wall	Fit2Teach**
Using entertainment	Equoo	Nintendo/PS4 etc***	SilverCloud	Equoo

Methodology

The study design process is summarized in Figure 2.

Figure 2. Study design process summary.



Study Design

Stress Management Dimension

To choose categories for the stress self-management dimension, we initially extracted descriptions from the qualitative data on the experiences of 14 senior teachers interviewed in a previous study [60]. The participants had provided more than 80 accounts of how they managed their stress. These descriptions were complemented by evidence from systematic reviews of occupational stress [23,40,61].

The interventions informed the *PICO* literature search criteria: patient and problem (eg, teacher and stress), intervention (eg, information, tracking, exercise, or mindfulness), comparison (often none), and outcome (eg, identifying, support, management, and reduction). We adapted the narrative method used in other studies [62,63], including checking references of relevant papers, alerts, and citation tracking along with searches

of academic databases including PsycINFO, Google Scholar, Cochrane, and PubMed. Literature relevant to teachers' self-management of stress was reviewed until repetition of themes revealed no further insight. Quality of papers was determined through their being published in peer-reviewed journals.

Digital Health Techniques Dimension

For the health techniques dimension, we reviewed the literature on persuasive design, digital health taxonomies, and trends in digital health self-care, again using the snowballing method as described above. We were aware of drawing on the different but complementary cultures of human-computer interaction and health and that their definitions of lifecycles, evaluation and implementation differ [64]. Our interest was in producing conceptual descriptions of mechanisms of action that could support the methods of stress management already identified in the literature and those given by teachers in interviews. These concepts would necessarily comprise elements of design, behavior, and theory, and draw on evidenced deployment of a digital companion for health self-management. Our aim was to create a conceptual description of the prevalent overarching technique or action of the digital companion that could be understood without ambiguity or complexity by the end user.

This approach was chosen for several reasons, including the following: (1) many digital companions use multiple techniques, and we wanted to facilitate choice by the primary featured enabled action and (2) other systematic reviews have overlooked or found a paucity in the description of behavior change techniques, which would make categorization of digital companions by such theory harder to achieve [65,66].

Technology Selection

To identify candidate digital companions, we took the following steps to inform our decisions:

1. Suitability: we began with digital interventions used by teachers, as described in a previous qualitative study, followed by a review of the literature for other candidates.
2. Availability: we examined whether the technology is accessible on the 2 main mobile operating systems and had been updated within the last 6 months.
3. Evaluation: we checked whether the technology was ranked positively on 1 of 3 expert review evaluation framework (Review) websites for apps and web-based tools (ORCHA [55] and Mindtools [67]) or apps only (Psyberguide [68]) for credibility and evidence base and for user experience.
4. Security: we reviewed the privacy and security policy to assess whether the technology used encryption for data connection and storage (where relevant).
5. Validity: we searched for significant, published positive clinical trial results.
6. Cost: given that the commercial model for apps that are free means very limited access or a trade in personal data, which we did not want to promote, we set a bar of £50 (US $64) annual fee for smartphone and website apps and £150 (US $202) for a wearable.

Taxonomy Creation

Overview

The process of reviewing the existing literature for the creation of stress management and digital techniques dimensions revealed different approaches to classification. Below, we present our findings and rationale for our choice of classification of strategies and concepts and then share the procedure we followed to enable technology selection.

Stress Self-management Dimension

Overview

We found 3 main approaches to categorize interventions specifically for the support or management of stress experienced by teachers. It is worth emphasizing that the value and goal of this conceptual categorization for our taxonomy was to identify a practical, actionable strategy for the individual [69]. The classification approaches found were as follows: (1) the level targeted by the intervention, (2) the target of intervention, or (3) the intervention strategy. We describe each of these and why we considered the intervention strategy to have the most relevance and explanatory power for the stress dimension.

Level of Intervention

Organizational-, individual-organizational–, or individual-level interventions have been frequently described [61,70-73], with an additional level of a classroom-focused approach being noted more recently [74]. The level of the intervention appears to be a way of describing the agent or group responsible for the stress management strategy. For example, the school leadership team or Multi-Academy Trust directors would be at the organizational level. As our focus was on self-management, this categorization would not provide a practical framework for teachers' own stress management.

Target of Intervention

The primary targets of interventions were the stressors themselves, which could be aspects of the work environment, such as maintaining discipline, time pressures, and workload [75]. The corresponding stress reduction strategies would then seek to reduce the occurrence of occupational stress among employees, such as workload reduction. This primary preventive approach for individuals should be the priority and a normal part of organizational management, as has long been argued in the health care sector [76,77]. Although many targets are well described in the teaching literature, they are beyond the control of the individual.

The secondary targets were the perception or responses of the individual person to the stressor itself, and the interventions were preventive or reactive. By targeting the way someone manages or copes with stress, the aim was to modify the individual's response in a positive way rather than removing the stressor itself. This might include peer support groups or cognitive behavioral techniques.

The tertiary targets of intervention were stress symptoms themselves, such as anxiety, insomnia, or racing heart rate, and the intervention was reactive. The aim of targeting symptoms was to manage or treat the emotional, cognitive, behavioral, or

physical changes brought about by stress. Although identifying secondary and tertiary targets enables a better understanding of stress, they do not indicate a set of potential self-management choices. For instance, if a teacher becomes aware that their response to stress is a behavioral habit (both a response and a symptom), such as to start pacing the floor, this knowledge in itself does not provide any signposting to what action an individual can then take to combat the stress. In addition, stress symptoms, such as nervous tics or fatigue, are not always obvious to the individual. Levels and targets of interventions were used in a prior categorization of occupational stress management from general employee work [78,79]; however, for our study's purposes, this conceptual framework does not always facilitate individual identification of actions that could be taken to self-manage stress.

Intervention Strategies

The third approach we identified was stress management strategies or training approaches [41,80-82]. We identified five overarching, nonmutually exclusive categories that could be supported digitally: (1) educational, (2) physiological, (3) situational, (4) cognitive, and (5) social.

The previous strategies have been described as follows: (1) stress awareness and education, (2) relaxation techniques, (3) cognitive coping, (4) biofeedback, (5) meditation, (6) exercise, (7) lifestyle advice, and (8) interpersonal skills training [81]. We considered that several of these could be grouped together along with more detailed activities simply listed as exemplars. Thus, education, awareness, and lifestyle advice were grouped under education; biofeedback, relaxation, meditation, breathing, aerobic activity, or mindfulness were grouped under physiological; and cognitive coping strategies, such as controlling emotions, problem-solving, or time management, were grouped under cognitive.

Social support was mentioned by the authors but was not listed by them as a category. It goes beyond interpersonal skills training embracing socializing and the therapeutic value of peer support [83] and self-enhancing humor [84]. This social element, along with descriptions of social support, has been described in teachers' stress management research [32,38,41]; hence, we added it as a category. We also noted in the literature some variation in the meaning of mindfulness among educators. This could mean the application of the established 8-week *mindfulness-based stress reduction* program [85,86] or the incorporation as part of a stress reduction program [38] or simply a meditative component of a multi-strategic stress reduction study [29]. Although other authors have used mindfulness-based interventions for categorization [40], the ambiguity in the use of the term meant we decided against using it as a category for strategy.

Digital Techniques Dimension

Overview

Our aim was to create a concise choice architecture that would be meaningful for potential users. This meaning was established through the description of how a digital companion would provide support.

Other condition-specific intervention reviews demonstrate varying approaches to the classification of technologies. Suijkerbuijk et al [87] categorized dementia interventions by purpose, such as support in daily life, safety, meaningful activities, or communication. Singh et al [88] categorized HIV apps and websites by functionality, such as prevention, testing, and management. These approaches sometimes blended the strategy with the mechanism or contained the mechanism within each function and helped us recognize that the primary focus for our categorization should be the broad mechanism of how the technology technique enabled self-care.

Despite an increasing number of studies on the use of digital companions in the workplace for occupational stress [89], reviews often focus on the type of intervention, such as CBT or mindfulness [16], and grouping them as such [40]. Reviews of the mechanism of action or concepts used by these apps are scarce, and others have noted this lack of detail in persuasive technology design [90]. In addition, reviews of wearables mostly seem to have focused on those for physical activity [91], but others have reported on the incorporation of behavior change technique *clusters* [92]. These enabled us to compare and make high-level reconciliation with the motivational affordances described by Orji and Moffatt [93], whose categorization was not always exclusive to one of the condensed descriptions below.

We found that the self-care *opportunities* by Nunes et al [94] essentially conceptualized action-enabled design features which were similar to descriptions given by Klasnja and Pratt [95] for intervention strategies and features. Therefore, we reviewed the descriptions against each other to compare the technique concepts. We then cross-checked them with the descriptions given by Orji and Moffatt [93] to arrive at 5 comprehensive conceptual themes that we now describe as our digital companion concepts.

Fostering Reflection by Making Health and Contextual Information Available

Both Klasnja and Pratt [95] and Nunes et al [94] described the ability to track health data first, and we retained the definition by Nunes et al [94] of "fostering reflection by making health and contextual information available." This data-enabled reflection has been found to be significant for those with severe mental illness [96], bipolar disorder [97], and stress [98], among other psychological conditions.

Suggesting Care Activities or Treatment Adjustments and Guided Self-management

The second description of "suggesting care activities or treatment adjustments" by Nunes et al [94] went beyond the mere "increasing accessibility" of health information described by Klasnja and Pratt [95] to actual adjustments that an individual can make. However, this category also needed to explicitly include delivering guided self-management described in the literature on stress, such as directed breathing or a CBT program. Hence, the second category was adapted to *suggesting care activities or treatment adjustments and guided self-management.*

Peer-to-Peer Social Support

Nunes et al [94] specifically described a trend as "sharing self-care activities and learning from others with the same

chronic condition." The limitation of this for our purposes was the medical emphasis, but we did want to include the significance of peer relationships. Klasnja and Pratt [95] talked about "leveraging social influence," capturing the social-sharing concept, building on the social support principles proposed by Oinas-Kukkonen and Harjumaa [49], so we redefined this category as *peer-to-peer social support*.

Using Entertainment

Klasnja and Pratt [95] also described using entertainment. This went beyond the gamification techniques recognized by Nunes et al [94], which can be used in the technology design of any of their categories. Participating in a purely fun tech-enabled activity not intentionally designed for symptom management has been shown to reduce stress symptoms [99,100].

Involving the Health Care Team

Nunes et al [94] strongly emphasized the patient (not medical) perspective, but 2 of their 5 categories still recognized the shared-care dynamic between patients and their formal and informal carers. Klasnja and Pratt [95] recognized this shared approach but described it under a single form of intervention (involving a health care team), and for our purposes, this sufficed.

For our taxonomy, we did not require the concept of involving the health care team as we were focusing on self-management. Therefore, we brought the 4 digital companion concepts with the 4 stress self-management strategies together in a matrix to give us a taxonomy that could then be the framework for digital companion selection. As a stand-alone taxonomy, this framework provides a structure for anyone seeking to choose a tool to support stress management. Figure 3 depicts this taxonomy.

Figure 3. Taxonomy matrix.

Digital stress support taxonomy matrix

Stress self-management strategies → Digital companion concepts ↓	Educational (eg, knowledge and awareness)	Physiological (eg, relaxation, exercise, reminder)	Cognitive (eg, problem solving, time management)	Social (eg, seeking social support or social skill development)
Fostering reflection via information - health or context				
Suggesting treatment and guided self-management				
Peer-to-peer social support				
Using entertainment				

Taxonomy Population

To populate the taxonomy, we applied the technology selection steps. This selection process was important for ensuring trustworthy digital companion candidates from which teachers in a subsequent study could make informed choices. The process is summarized in Table 1.

Our starting point was suitability and availability, based on a previous qualitative study exploring teachers' familiarity and use of digital tools for stress management [60]. This reflected insight into the influence of context to design as described in both usability study methodologies [101] and the person-based approach [20]. Where that did not provide a candidate, we reviewed the literature, the National Health Service (NHS) App Library, Carlo's behavioral health app review [102], and the scientific literature. Of the 12 apps originally named by teachers, 8 (67%) were available on both iOS and Android platforms (TeacherTapp, Fit2Teach, Headspace, Mindshift, Pacifica [now called Sanvello], Calm, Insight Timer, and Happy not Perfect), but one of these (Fit2Teach) had not been updated for over 2 years. Given that it was uniquely tailored in its approach and that the associated Facebook group had recently been updated, we contacted the developer, but unfortunately, we received no response. Neither Fit2Teach or TeacherTapp had been designed for stress, but both offer education tips and insight, and the opportunity for reflection.

The 2 apps that used diarizing as their prevalent tracking strategy (My Wonderful Days and Now Then Free) were not available on either platform and the other 2 app descriptions were not complete enough for certain identification. The web-based CBT program that had been described by 1 teacher was only available in 1 English county. The wearables used by teachers were Fitbit

(Fitbit, Inc) models (Charge, Alta, and Blaze), Samsung Gear 2 (Samsung Electronics Co, Ltd), Polar M340 (Polar Electro Oy), and Apple Watch (Apple, Inc). No other candidate technologies were identified in the literature on teachers' stress.

We searched for available digital companions within the positive expert review evaluation frameworks (Reviews) but owing to disparities observed between Review assessments [102] and our concern with privacy and safety, we read through all the security and privacy policies. This was also important for all wearables, as none of them were covered in the Reviews. Occasionally, security through encryption was still not evident from the published policy, and in these cases, the developer was emailed for further information.

Table 1. Summary of digital companion population process.

Technology selection steps	Rationale
Suitability: qualitative data from occupation and behavior	We began with digital interventions used by teachers as described in data in a previous qualitative study
Availability: verify whether the technology is accessible on the 2 main mobile operating systems and had been updated within the last 6 months	Ensures the technology is available to a wider audience and supported by the developers
Evaluation: search one or more of the expert review evaluation frameworks (Reviews) to see if the technology is ranked positively	Gives professional or third-party view on the credibility, evidence base, and user experience
Security: review the privacy and security policy	Shows whether the data are stored and transmitted securely with encryption to give an indication of risk
Validity: search for research papers on the technology	Enables any trials with the technology to be considered
Cost: assess cost	Considers whether the technology is in budget

Many digital companions have not been tested through trials, so this step (validity) was not a reason to exclude them, especially wearables where data are sparse. Conversely, some popular apps that did not satisfy the safety inclusion criterion had significant published evidence of their efficacy. For these, we presented this scientific evidence as a reason for inclusion, despite no or partial encryption. Finally, cost was considered.

Our final selection of digital companions for presentation to teachers comprised 4 apps named by teachers in the previous study (Headspace, Calm, TeacherTapp, and Fit2Teach), 4 alternative apps sourced from one or more of the Reviews (Equoo, Sleepio, and Daylio), and 1 app from the scientific literature (Wysa, an artificial intelligence–based chatbot). For websites, 1 was sourced from a Review (Big White Wall, now Togetherall Ltd), 1 from the NHS (Stress Management at Work), and 1 from scientific literature (SliverCloud Health). For wearables, 1 was identified from the scientific literature using medical grade data assurance (Withings Steel HR watch, Withings, Inc).

The stress self-management strategies, digital companion concepts, and selected apps were brought together in the taxonomy matrix shown in the introduction in Figure 1 with caveats shown by asterisks.

Discussion

Principal Findings

This paper describes the process of creating a context-based framework to facilitate the choice of digital companion intervention. Using the dimension of stress self-management, we created classifications of strategies that were derived from empirical research and the literature. Using the dimension of digital techniques, we created conceptual descriptions of the mechanisms of action of digital companions informed by the literature. Bringing these together in a taxonomy gave the framework that we could populate with digital companions for teachers' stress self-management according to availability, evaluation, security, validity, and cost. It is a starting structure for the presentation and selection of contextually appropriate digital companions.

Populating the taxonomy presented some significant challenges. The transience of apps or their ratings (availability and evaluation) meant that by the time we came to present our taxonomy to teachers, 1 peer-to-peer–supported CBT website had been removed. Likewise, a highly rated diarizing app had one of its review ratings plummet during our study, although we found no cause for concern on rechecking the privacy policy. Another CBT course with extensive validation through research publications was included, as it had been commissioned by the local NHS in the areas where the teachers we planned to work with were employed. However, when 1 participant tried to access it, a referral from the general practitioner was required, which precluded pure self-management. Some apps we considered were described as designed for stress but included reference to medical conditions such as psychosis and schizophrenia. We were concerned that their inclusion would imply a medical need or that such a diagnostic association could be too sensitive for a study that was focused on occupational stress.

It became clear as we reviewed candidate smartphone apps that many did not offer comprehensive (if any) encryption of data, even those where the funding model required user payment (thus requiring input of more sensitive data). Our search was not exhaustive: that would have been impossible. To ensure candidates in each category, when we were able to reference scientific studies on app efficacy (eg, headspace and calm), it was decided to include them in the taxonomy with the caveat that although widely used, there was no or only partial encryption of stored and/or transmitted user data.

The sequence of application of our selection criteria was affected for wearables because of their cost. Of the 6 different wearables described in the teachers' study, because of the price, we excluded Samsung Gear 2, Polar 340, and Apple Watch. Obsolescence excluded 2 of the Fitbits (Blaze and Alta), leaving the Fitbit Charge. This failed the encryption requirement being nonspecific and considered external evaluation to be inadequate [103]. Database search, paper retrieval, and website scrutiny enabled us to identify 1 wearable from Withings that satisfied all the set criteria, offered support for 2 of the 4 stress self-management strategies, and fell into the set price bracket.

Importantly, using qualitative field data as a starting point was crucial for identifying digital companions that would not have appeared in a search based on the condition of stress. For example, TeacherTapp was designed as a research tool to voice teachers' opinions. However, its educational content and sense of peer connection were considered valuable for relieving feelings of stress. Likewise, Fit2Teach, although designed for well-being and work-life balance, was listed under *education* and not under *stress* in app stores.

In a world in which automated or unsubstantiated rating systems are prevalent, there is still a need for autonomous, informed, human decision-making that draws on personal knowledge and understanding [104]. Individuals need to be able to confidently identify their personal preferences to improve their chances of adherence [5]. Improving app selection by context-based condition management and conceptual categorization could logically aid both the adoption and potential efficacy of digital health tools and reduce attrition before the desired outcome. However, our findings illustrate that there is no quick route to informed adoption.

The populated taxonomy was presented to 8 high school midmanagement teachers in workshops to enable them to identify how they currently managed their stress and how it could be supported by digital means. Their chosen digital companions were then used during a planned longitudinal study in the school summer term (during partial COVID-19 lockdown) and on into a serendipitous study in the autumn (where teachers were back in hygiene-adjusted school settings). Of the 8 teachers, 6 (75%) still used their digital companion choice 6 months after beginning. The analysis of these findings will be the subject of a subsequent study.

Limitations

Our review of the literature was not exhaustive, and other research may reveal stress management strategies beyond those we identified. In addition, there could be disagreement on the way that we have grouped or limited the explanatory power of digital companion concepts or that they are relevant for conditions other than stress. Further research will be able to substantiate whether these issues are significant.

We have already noted in our process and discussion that the selection of technology can never be complete and is only ever a reflection of what apps and information are available at the time of the search. In addition, our starting point for apps was a previous small study where the participants had self-selected; a different or wider cohort could have produced other findings. There is no circumventing the reality that populating a taxonomy will always have to be revisited at the time of use.

Another limitation of our approach is potentially in embedding the notion that dealing with or coping with workplace stress is just the responsibility of the individual. This individualized approach can place a profound burden on a teacher as it fails to acknowledge the complexity of the origins of stress [105]. It is not our intention to imply that managing stress is only the responsibility of the individual, and through our context-based approach, we acknowledge the structural and environmental influences, in addition to the sociocultural factors within a school.

Conclusions

There is no quick and easy solution to identifying a safe, efficacious, contextually, and individually appropriate app, website, or wearable to support self-management of health, well-being, or a specific health condition. Evaluation frameworks are valuable and evolving but would benefit from complementary information for users to be able to identify their preferences and consider whether the technology on offer fits their current behaviors or contexts.

If an individual can use a taxonomy to identify their preferred management strategy and, from there, make an informed selection of a digital companion for support, the user starts from a strong position. We hope that these procedures can generally inform professionals seeking to facilitate the selection of a digital companion for an individual's self-management of a named health or well-being condition. We also hope that our populated taxonomy can be a specific starting point for teachers' digital companion–supported stress self-management, and one that can be refreshed through repopulation in the future.

Acknowledgments

This research received funding from the Engineering and Physical Sciences Research Council Doctoral Training Partnership awarded to University College London EP/R513143/1.

Authors' Contributions

All authors reviewed and provided feedback on the final manuscript.

Conflicts of Interest

None declared.

References

1. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. Evid Based Ment Health 2018 Aug 05;21(3):116-119. [doi: 10.1136/eb-2018-102891] [Medline: 29871870]
2. Larsen ME, Nicholas J, Christensen H. Quantifying app store dynamics: longitudinal tracking of mental health apps. JMIR Mhealth Uhealth 2016 Aug 09;4(3):e96 [FREE Full text] [doi: 10.2196/mhealth.6020] [Medline: 27507641]
3. Aitken M, Clancy B, Nass D. The growing value of digital health. IQVIA. 2017. URL: https://www.iqvia.com/insights/the-iqvia-institute/reports/the-growing-value-of-digital-health [accessed 2022-01-18]
4. Bennion MR, Hardy G, Moore RK, Millings A. E-therapies in England for stress, anxiety or depression: what is being used in the NHS? A survey of mental health services. BMJ Open 2017 Jan 23;7(1):e014844 [FREE Full text] [doi: 10.1136/bmjopen-2016-014844] [Medline: 28115336]
5. Schueller SM, Neary M, O'Loughlin K, Adkins EC. Discovery of and interest in health apps among those with mental health needs: survey and focus group study. J Med Internet Res 2018 Jun 11;20(6):e10141 [FREE Full text] [doi: 10.2196/10141] [Medline: 29891468]
6. Oyebode O, Alqahtani F, Orji R. Using machine learning and thematic analysis methods to evaluate mental health apps based on user reviews. IEEE Access 2020;8:111141-111158. [doi: 10.1109/access.2020.3002176]
7. Huang H, Bashir M. Users' adoption of mental health apps: examining the impact of information cues. JMIR Mhealth Uhealth 2017 Jun 28;5(6):e83 [FREE Full text] [doi: 10.2196/mhealth.6827] [Medline: 28659256]
8. Bakker D, Kazantzis N, Rickwood D, Rickard N. Mental health smartphone apps: review and evidence-based recommendations for future developments. JMIR Ment Health 2016 Mar;3(1):e7 [FREE Full text] [doi: 10.2196/mental.4984] [Medline: 26932350]
9. 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(1):18 [FREE Full text] [doi: 10.1038/s41746-019-0093-1] [Medline: 31304366]
10. Stawarz K, Preist C, Coyle D. Use of smartphone apps, social media, and web-based resources to support mental health and well-being: online survey. JMIR Ment Health 2019 Jul 12;6(7):e12546 [FREE Full text] [doi: 10.2196/12546] [Medline: 31301126]
11. Wearable medical devices market analysis highlights the impact of COVID-19 (2020-2024) : increasing prevalence of chronic diseases to boost market growth. Business Wire. 2020. URL: https://www.businesswire.com/news/home/20200730005550/en/ [accessed 2020-09-01]
12. Pakhomov SV, Thuras PD, Finzel R, Eppel J, Kotlyar M. Using consumer-wearable technology for remote assessment of physiological response to stress in the naturalistic environment. PLoS One 2020 Mar 25;15(3):e0229942 [FREE Full text] [doi: 10.1371/journal.pone.0229942] [Medline: 32210441]
13. Yang H, Yu J, Zo H, Choi M. User acceptance of wearable devices: an extended perspective of perceived value. Telemat Inform 2016 May;33(2):256-269. [doi: 10.1016/j.tele.2015.08.007]
14. Wang Y, Fadhil A, Lange J, Reiterer H. Integrating taxonomies into theory-based digital health interventions for behavior change: a holistic framework. JMIR Res Protoc 2019 Jan 15;8(1):e8055 [FREE Full text] [doi: 10.2196/resprot.8055] [Medline: 30664477]
15. Carolan S, Harris PR, Cavanagh K. Improving employee well-being and effectiveness: systematic review and meta-analysis of web-based psychological interventions delivered in the workplace. J Med Internet Res 2017 Jul 26;19(7):e271 [FREE Full text] [doi: 10.2196/jmir.7583] [Medline: 28747293]
16. Stratton E, Lampit A, Choi I, Calvo RA, Harvey SB, Glozier N. Effectiveness of eHealth interventions for reducing mental health conditions in employees: a systematic review and meta-analysis. PLoS One 2017 Dec 21;12(12):e0189904 [FREE Full text] [doi: 10.1371/journal.pone.0189904] [Medline: 29267334]
17. Fitzpatrick G. New challenges for Health IT – design fit for life. Design 4 Health. 2011. URL: https://research.shu.ac.uk/design4health/wp-content/uploads/2012/09/D4H2011_proceedings_v5a.pdf [accessed 2018-07-13]
18. O'Kane AA. Individual differences and contextual factors influence the experience and practice of self-management with type 1 diabetes technologies. Doctoral Dissertation, University College London. 2016. URL: https://discovery.ucl.ac.uk/id/eprint/1482175/ [accessed 2019-08-08]
19. Blandford A, Berndt E, Catchpole K, Furniss D, Mayer A, Mentis H, et al. Strategies for conducting situated studies of technology use in hospitals. Cogn Tech Work 2014 Dec 16;17(4):489-502. [doi: 10.1007/s10111-014-0318-7]
20. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. J Med Internet Res 2015 Jan;17(1):e30 [FREE Full text] [doi: 10.2196/jmir.4055] [Medline: 25639757]

21. Doherty G, Coyle D, Sharry J. Engagement with online mental health interventions: an exploratory clinical study of a treatment for depression. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2012 Presented at: CHI '12: CHI Conference on Human Factors in Computing Systems; 2012; Austin Texas USA p. 1421-1430. [doi: 10.1145/2207676.2208602]
22. Rennick-Egglestone S, Knowles S, Toms G, Bee P, Lovell K, Bower P. Health technologies 'In the Wild': experiences of engagement with computerised CBT. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. 2016 Presented at: CHI'16: CHI Conference on Human Factors in Computing Systems; May 7 - 12, 2016; San Jose California USA p. 2124-2135. [doi: 10.1145/2858036.2858128]
23. de Korte E, Wiezer N, Roozeboom M, Vink P, Kraaij W. Behavior change techniques in mhealth apps for the mental and physical health of employees: systematic assessment. JMIR Mhealth Uhealth 2018 Oct 03;6(10):e167 [FREE Full text] [doi: 10.2196/mhealth.6363] [Medline: 30282621]
24. Statistics - working days lost in Great Britain. HSE Statistics. 2018. URL: https://www.hse.gov.uk/statistics/dayslost.htm [accessed 2019-09-18]
25. Johnson A, Phillips D, Horm D, Luk G. Parents, teachers, and distance learning during the COVID-19 pandemic: a snapshot from Tulsa, OK. Tulsa SEED Study. 2020. URL: https://medium.com/@TulsaSEED/parents-teachers-and-distance-learning-during-the-covid-19-pandemic-a-snapshot-from-tulsa-ok-5b5fdb54ea18 [accessed 2022-01-19]
26. Day C, Gu Q. Teacher emotions: well being and effectiveness. In: Schutz P, Zembylas M, editors. Advances in Teacher Emotion Research. Boston, MA: Springer; 2009:15-31.
27. Day C, Gu Q. Resilient Teachers, Resilient Schools. London, UK: Routledge; 2013.
28. Day C, Hong J. Influences on the capacities for emotional resilience of teachers in schools serving disadvantaged urban communities: challenges of living on the edge. Teach Teach Educ 2016 Oct;59:115-125. [doi: 10.1016/j.tate.2016.05.015]
29. Siu OL, Cooper CL, Phillips DR. Intervention studies on enhancing work well-being, reducing burnout, and improving recovery experiences among Hong Kong health care workers and teachers. Int J Stress Manag 2014;21(1):69-84. [doi: 10.1037/a0033291]
30. Skaalvik EM, Skaalvik S. Still motivated to teach? A study of school context variables, stress and job satisfaction among teachers in senior high school. Soc Psychol Educ 2017 Jan 6;20(1):15-37. [doi: 10.1007/s11218-016-9363-9]
31. Bertoch M, Nielsen E, Curley J, Borg W. Reducing teacher stress. J Experim Edu 2014 Apr 16;57(2):117-128 [FREE Full text] [doi: 10.1080/00220973.1989.10806500]
32. Unterbrink T, Pfeifer R, Krippeit L, Zimmermann L, Rose U, Joos A, et al. Burnout and effort-reward imbalance improvement for teachers by a manual-based group program. Int Arch Occup Environ Health 2012 Aug;85(6):667-674. [doi: 10.1007/s00420-011-0712-x] [Medline: 22038086]
33. Flook L, Goldberg SB, Pinger L, Bonus K, Davidson RJ. Mindfulness for teachers: a pilot study to assess effects on stress, burnout and teaching efficacy. Mind Brain Educ 2013 Sep;7(3) [FREE Full text] [doi: 10.1111/mbe.12026] [Medline: 24324528]
34. Kaspereen D. Relaxation intervention for stress reduction among teachers and staff. Int J Stress Manag 2012;19(3):238-250. [doi: 10.1037/a0029195]
35. Austin V, Shah S, Muncer S. Teacher stress and coping strategies used to reduce stress. Occup Ther Int 2005 May;12(2):63-80. [doi: 10.1002/oti.16] [Medline: 16136865]
36. Leung SS, Chiang VC, Chui Y, Mak Y, Wong DF. A brief cognitive-behavioral stress management program for secondary school teachers. J Occup Health 2011 Feb 18;53(1):23-35 [FREE Full text] [doi: 10.1539/joh.l10037] [Medline: 21079374]
37. Jeffcoat T, Hayes SC. A randomized trial of ACT bibliotherapy on the mental health of K-12 teachers and staff. Behav Res Ther 2012 Sep;50(9):571-579. [doi: 10.1016/j.brat.2012.05.008] [Medline: 22750188]
38. Reiser JE, Murphy SL, McCarthy CJ. Stress prevention and mindfulness: a psychoeducational and support group for teachers. J Special Group Work 2016 Mar 30;41(2):117-139. [doi: 10.1080/01933922.2016.1151470]
39. Susman-Stillman A, Lim S, Meuwissen A, Watson C. Reflective supervision/consultation and early childhood professionals' well-being: a qualitative analysis of supervisors' perspectives. Early Edu Dev 2020 Aug 10;31(7):1151-1168. [doi: 10.1080/10409289.2020.1793654]
40. von der Embse N, Ryan SV, Gibbs T, Mankin A. Teacher stress interventions: a systematic review. Psychol Schs 2019 Jul 24:1328-1343. [doi: 10.1002/pits.22279]
41. Iancu AE, Rusu A, Măroiu C, Păcurar R, Maricu oiu LP. The effectiveness of interventions aimed at reducing teacher burnout: a meta-analysis. Educ Psychol Rev 2017 Jul 22;30(2):373-396. [doi: 10.1007/s10648-017-9420-8]
42. Chan DW. Burnout and life satisfaction: does gratitude intervention make a difference among Chinese school teachers in Hong Kong? Edu Psychol 2011 Dec;31(7):809-823. [doi: 10.1080/01443410.2011.608525]
43. Ralston AL, Andrews AR, Hope DA. Fulfilling the promise of mental health technology to reduce public health disparities: review and research agenda. Clin Psychol Sci Pract 2019 Mar;26(1):1-14. [doi: 10.1111/cpsp.12277]
44. Ebert D, Lehr D, Boß L, Riper H, Cuijpers P, Andersson G, et al. Efficacy of an internet-based problem-solving training for teachers: results of a randomized controlled trial. Scand J Work Environ Health 2014 Aug 13;40(6):582-596 [FREE Full text] [doi: 10.5271/sjweh.3449]

45. Thiart H, Lehr D, Ebert DD, Berking M, Riper H. Log in and breathe out: internet-based recovery training for sleepless employees with work-related strain - results of a randomized controlled trial. Scand J Work Environ Health 2015 Mar 15;41(2):164-174 [FREE Full text] [doi: 10.5271/sjweh.3478] [Medline: 25590336]
46. Phillips EA, Gordeev VS, Schreyögg J. Effectiveness of occupational e-mental health interventions: a systematic review and meta-analysis of randomized controlled trials. Scand J Work Environ Health 2019 Nov 01;45(6):560-576 [FREE Full text] [doi: 10.5271/sjweh.3839] [Medline: 31184758]
47. Persson Asplund R, Dagöö J, Fjellström I, Niemi L, Hansson K, Zeraati F, et al. Internet-based stress management for distressed managers: results from a randomised controlled trial. Occup Environ Med 2018 Feb 30;75(2):105-113 [FREE Full text] [doi: 10.1136/oemed-2017-104458] [Medline: 28855344]
48. Fogg BJ. Persuasive Technology: Using Computers to Change What We Think and Do. 1st Edition. Burlington, Massachusetts, United States: Morgan Kaufmann; 2003.
49. Oinas-Kukkonen H, Harjumaa M. Persuasive systems design: key issues, process model, and system features. Commun Assoc Inf Syst 2009;24:485-500. [doi: 10.17705/1CAIS.02428]
50. Ritterband LM, Thorndike FP, Cox DJ, Kovatchev BP, Gonder-Frederick LA. A behavior change model for internet interventions. Ann Behav Med 2009 Aug;38(1):18-27 [FREE Full text] [doi: 10.1007/s12160-009-9133-4] [Medline: 19802647]
51. Mohr DC, Schueller SM, Montague E, Burns MN, Rashidi P. The behavioral intervention technology model: an integrated conceptual and technological framework for eHealth and mHealth interventions. J Med Internet Res 2014 Jun;16(6):e146 [FREE Full text] [doi: 10.2196/jmir.3077] [Medline: 24905070]
52. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol 2008 May;27(3):379-387. [doi: 10.1037/0278-6133.27.3.379] [Medline: 18624603]
53. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
54. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the User Version of the Mobile Application Rating Scale (uMARS). JMIR Mhealth Uhealth 2016 Jun 10;4(2):e72 [FREE Full text] [doi: 10.2196/mhealth.5849] [Medline: 27287964]
55. Organisation for Review of Care and Health Applications (ORCHA). URL: https://orchahealth.com/ [accessed 2022-01-19]
56. Leigh S, Ouyang J, Mimnagh C. Effective? Engaging? Secure? Applying the ORCHA-24 framework to evaluate apps for chronic insomnia disorder. Evid Based Ment Health 2017 Nov 25;20(4):e20. [doi: 10.1136/eb-2017-102751] [Medline: 28947676]
57. Neary M, Schueller SM. State of the Field of Mental Health Apps. Cogn Behav Pract 2018 Nov;25(4):531-537 [FREE Full text] [doi: 10.1016/j.cbpra.2018.01.002] [Medline: 33100810]
58. Torous JB, Chan SR, Gipson SY, 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]
59. Torous J, Luo J, Chan S. Mental health apps: What to tell patients. Curr Psychiatr 2018;17(3):21-25 [FREE Full text]
60. Manning JB, Blandford A, Edbrooke-Childs J, Marshall P. How contextual constraints shape midcareer high school teachers' stress management and use of digital support tools: qualitative study. JMIR Ment Health 2020 Apr 27;7(4):e15416 [FREE Full text] [doi: 10.2196/15416] [Medline: 32338623]
61. Naghieh A, Montgomery P, Bonell CP, Thompson M, Aber JL. Organisational interventions for improving wellbeing and reducing work-related stress in teachers. Cochrane Database Syst Rev 2015 Apr 08(4):CD010306. [doi: 10.1002/14651858.CD010306.pub2] [Medline: 25851427]
62. Debono DS, Greenfield D, Travaglia JF, Long JC, Black D, Johnson J, et al. Nurses' workarounds in acute healthcare settings: a scoping review. BMC Health Serv Res 2013 May 11;13(1):175 [FREE Full text] [doi: 10.1186/1472-6963-13-175] [Medline: 23663305]
63. Short CE, DeSmet A, Woods C, Williams SL, Maher C, Middelweerd A, et al. Measuring engagement in eHealth and mHealth behavior change interventions: viewpoint of methodologies. J Med Internet Res 2018 Nov 16;20(11):e292 [FREE Full text] [doi: 10.2196/jmir.9397] [Medline: 30446482]
64. Blandford A, Gibbs J, Newhouse N, Perski O, Singh A, Murray E. Seven lessons for interdisciplinary research on interactive digital health interventions. Digit Health 2018;4:2055207618770325 [FREE Full text] [doi: 10.1177/2055207618770325] [Medline: 29942629]
65. Buckingham SA, Williams AJ, Morrissey K, Price L, Harrison J. Mobile health interventions to promote physical activity and reduce sedentary behaviour in the workplace: a systematic review. Digit Health 2019 Mar;5:2055207619839883 [FREE Full text] [doi: 10.1177/2055207619839883] [Medline: 30944728]
66. Whitehead L, Seaton P. The effectiveness of self-management mobile phone and tablet apps in long-term condition management: a systematic review. J Med Internet Res 2016 May;18(5):e97 [FREE Full text] [doi: 10.2196/jmir.4883] [Medline: 27185295]
67. MindTools. URL: https://mindtools.io/ [accessed 2022-01-31]

68. One Mind PsyberGuide. URL: https://onemindpsyberguide.org/ [accessed 2022-01-31]
69. Peters D, Calvo RA, Ryan RM. Designing for motivation, engagement and wellbeing in digital experience. Front Psychol 2018 May;9:797. [doi: 10.3389/fpsyg.2018.00797] [Medline: 29892246]
70. Montano D, Hoven H, Siegrist J. Effects of organisational-level interventions at work on employees' health: a systematic review. BMC Public Health 2014 Feb 08;14(1):135 [FREE Full text] [doi: 10.1186/1471-2458-14-135] [Medline: 24507447]
71. Brunton G, Dickson K, Khatwa M, Caird J, Oliver S, Hinds K, et al. Developing evidence-informed, employer-led workplace health. EPPI-Centre. 2016. URL: https://eppi.ioe.ac.uk/CMS/Portals/0/PDF%20reviews%20and%20summaries/Employer-led%20workplace%20health%202016%20Brunton.pdf [accessed 2018-06-27]
72. Greenberg M, Brown J, Abenavoli R. Teacher stress and health: effects on teachers, students and schools. Edna Bennett Pierce Prevention Research Center, Pennsylvania State University. 2016. URL: https://www.prevention.psu.edu/uploads/files/rwjf430428-TeacherStress.pdf [accessed 2019-03-26]
73. DeFrank RS, Cooper CL. Worksite Stress Management Interventions: Their Effectiveness and Conceptualisation. Journal of Managerial Psychology 1987;2(1):4-10. [doi: 10.1108/eb043385]
74. Ouellette RR, Frazier SL, Shernoff ES, Cappella E, Mehta TG, Maríñez-Lora A, et al. Teacher job stress and satisfaction in urban schools: disentangling individual-, classroom-, and organizational-level influences. Behav Ther 2018 Dec;49(4):494-508 [FREE Full text] [doi: 10.1016/j.beth.2017.11.011] [Medline: 29937253]
75. Camacho DA, Vera E, Scardamalia K, Phalen PL. What are urban teachers thinking and feeling? Psychol Schs 2018 Aug 13;55(9):1133-1150. [doi: 10.1002/pits.22176]
76. Firth-Cozens J. Doctors, their wellbeing, and their stress. Br Med J 2003 Mar 29;326(7391):670-671 [FREE Full text] [doi: 10.1136/bmj.326.7391.670] [Medline: 12663377]
77. Bergerman L, Corabian P, Harstall C. Effectiveness of Organizational Interventions for the Prevention of Workplace Stress. Canada: Institute of Health Economic; 2009.
78. Bhui K, Dinos S, Stansfeld S, White P. A synthesis of the evidence for managing stress at work: a review of the reviews reporting on anxiety, depression, and absenteeism. J Environ Public Health 2012;2012:515874 [FREE Full text] [doi: 10.1155/2012/515874] [Medline: 22496705]
79. Dewe P, O?Driscoll M, Cooper C. Stress management interventions. In: Coping with Work Stress: A Review and Critique. Hoboken, New Jersey, United States: John Wiley & Sons; 2010.
80. Brown Z, Uehara D. Coping with teacher stress: a research synthesis for pacific educators. Office of Educational Research and Improvement (ED), Washington, DC. 1999. URL: https://files.eric.ed.gov/fulltext/ED440924.pdf [accessed 2019-06-06]
81. Lundberg U, Cooper C. Preventive strategies. In: The Science of Occupational Health: Stress, Psychobiology, and the New World of Work. Hoboken, New Jersey, United States: Wiley-Blackwell; 2010.
82. Cartwright S, Whatmore L. Stress and individual differences: implications for stress management. In: Antoniou AS, Cooper CL, editors. Research Companion to Organizational Health Psychology. Cheltenham, UK: Edward Elgar Publishing Inc; 2005.
83. Stanford B. Reflections of resilient, persevering urban teachers. Teach Educ Q 2001;28(3):75-87 [FREE Full text]
84. Ho SK. The relationship between teacher stress and burnout in Hong Kong: positive humour and gender as moderators. Edu Psychol 2015 Dec 15;37(3):272-286. [doi: 10.1080/01443410.2015.1120859]
85. Kabat-Zinn J. Mindfulness-based interventions in context: past, present, and future. Clin Psychol Sci Pract 2003;10(2):144-156. [doi: 10.1093/clipsy.bpg016]
86. Gold E, Smith A, Hopper I, Herne D, Tansey G, Hulland C. Mindfulness-Based Stress Reduction (MBSR) for primary school teachers. J Child Fam Stud 2009 Dec 5;19(2):184-189. [doi: 10.1007/s10826-009-9344-0]
87. Suijkerbuijk S, Nap HH, Cornelisse L, IJsselsteijn WA, de Kort YA, Minkman MM. Active involvement of people with dementia: a systematic review of studies developing supportive technologies. J Alzheimer's Dis 2019 Jun 18;69(4):1041-1065. [doi: 10.3233/jad-190050]
88. Singh A, Gibbs J, Estcourt C, Sonnenberg P, Blandford A. Are HIV smartphone apps and online interventions fit for purpose? In: Proceedings of the 2017 International Conference on Digital Health. 2017 Presented at: DH '17: International Conference on Digital Health; July 2 - 5, 2017; London United Kingdom p. 6-15. [doi: 10.1145/3079452.3079469]
89. Kuster A, Dalsbø TK, Thanh BY, Agarwal A, Durand-Moreau QV, Kirkehei I. Computer-based versus in-person interventions for preventing and reducing stress in workers. Cochrane Database Syst Rev 2017 Aug 30;8:CD011899 [FREE Full text] [doi: 10.1002/14651858.CD011899.pub2] [Medline: 28853146]
90. Carolan S, Harris PR, Greenwood K, Cavanagh K. Increasing engagement with an occupational digital stress management program through the use of an online facilitated discussion group: results of a pilot randomised controlled trial. Internet Interv 2017 Dec;10:1-11 [FREE Full text] [doi: 10.1016/j.invent.2017.08.001] [Medline: 30135747]
91. Mercer K, Li M, Giangregorio L, Burns C, Grindrod K. Behavior change techniques present in wearable activity trackers: a critical analysis. JMIR Mhealth Uhealth 2016;4(2):e40 [FREE Full text] [doi: 10.2196/mhealth.4461] [Medline: 27122452]
92. Chia GL, Anderson A, McLean LA. Behavior change techniques incorporated in fitness trackers: content analysis. JMIR Mhealth Uhealth 2019 Jul 23;7(7):e12768 [FREE Full text] [doi: 10.2196/12768] [Medline: 31339101]
93. Orji R, Moffatt K. Persuasive technology for health and wellness: state-of-the-art and emerging trends. Health Informatics J 2018 Mar;24(1):66-91 [FREE Full text] [doi: 10.1177/1460458216650979] [Medline: 27245673]

94. Nunes F, Verdezoto N, Fitzpatrick G, Kyng M, Grönvall E, Storni C. Self-care technologies in HCI. ACM Trans Comput-Hum Interact 2015 Dec 14;22(6):1-45. [doi: 10.1145/2803173]
95. Klasnja P, Pratt W. Healthcare in the pocket: mapping the space of mobile-phone health interventions. J Biomed Inform 2012 Jan 01;45(1):184-198. [doi: 10.1016/j.jbi.2011.09.017]
96. Buitenweg D, van de Mheen D, Grund J, van Oers H, van Nieuwenhuizen C. Visual and personalized quality of life assessment app for people with severe mental health problems: qualitative evaluation. JMIR Ment Health 2020 Dec 03;7(12):e19593 [FREE Full text] [doi: 10.2196/19593] [Medline: 33270036]
97. Morton E, Hole R, Murray G, Buzwell S, Michalak E. Experiences of a web-based quality of life self-monitoring tool for individuals with bipolar disorder: a qualitative exploration. JMIR Ment Health 2019 Dec 04;6(12):e16121 [FREE Full text] [doi: 10.2196/16121] [Medline: 31799936]
98. Ptakauskaite N, Cox AL, Berthouze NL. Knowing what you're doing or knowing what to do: how stress management apps support reflection and behaviour change. In: Proceedings of the CHI Conference on Human Factors in Computing Systems. 2018 Presented at: CHI '18: CHI Conference on Human Factors in Computing Systems; April 21 - 26, 2018; Montreal QC Canada p. 1-6. [doi: 10.1145/3170427.3188648]
99. Babbage C, Jackson GM, Nixon E. Desired features of a digital technology tool for self-management of well-being in a nonclinical sample of young people: qualitative study. JMIR Ment Health 2018 Dec 18;5(4):e10067 [FREE Full text] [doi: 10.2196/10067] [Medline: 30563820]
100. Tucker SJ, Lanningham-Foster LM, Murphy JN, Thompson WG, Weymiller AJ, Lohse C, et al. Effects of a worksite physical activity intervention for hospital nurses who are working mothers. AAOHN J 2011 Sep;59(9):377-386. [doi: 10.3928/08910162-20110825-01] [Medline: 21877670]
101. Yen P, Bakken S. Review of health information technology usability study methodologies. J Am Med Inform Assoc 2012 May;19(3):413-422 [FREE Full text] [doi: 10.1136/amiajnl-2010-000020] [Medline: 21828224]
102. Carlo AD, Ghomi RH, Renn BN, Areán PA. By the numbers: ratings and utilization of behavioral health mobile applications. NPJ Digit Med 2019 Jun 17;2(1):54 [FREE Full text] [doi: 10.1038/s41746-019-0129-6] [Medline: 31304400]
103. Orlosky J, Ezenwoye O, Yates H, Besenyi G. A look at the security and privacy of Fitbit as a health activity tracker. In: Proceedings of the 2019 ACM Southeast Conference. 2019 Presented at: ACM SE '19: 2019 ACM Southeast Conference; April 18 - 20, 2019; Kennesaw GA USA p. 241-244. [doi: 10.1145/3299815.3314468]
104. Calvo R, Peters D. Positive Computing: Technology for Wellbeing and Human Potential. Cambridge, MA: MIT Press; 2014.
105. Taylor-Rodgers E, Batterham PJ. Evaluation of an online psychoeducation intervention to promote mental health help seeking attitudes and intentions among young adults: randomised controlled trial. J Affect Disord 2014 Oct 15;168:65-71. [doi: 10.1016/j.jad.2014.06.047] [Medline: 25038293]

Abbreviations

CBT: cognitive behavioral therapy
MARS: Mobile App Rating Scale
NHS: National Health Service
ORCHA: Organisation for Review of Care and Health Apps

Edited by J Torous; submitted 22.07.21; peer-reviewed by N Mohd Aszemi, H Fulford; comments to author 20.10.21; revised version received 02.11.21; accepted 02.11.21; published 16.02.22.

Please cite as:
Manning JB, Blandford A, Edbrooke-Childs J
Digital Companion Choice to Support Teachers' Stress Self-management: Systematic Approach Through Taxonomy Creation
JMIR Form Res 2022;6(2):e32312
URL: https://formative.jmir.org/2022/2/e32312
doi:10.2196/32312
PMID:35171106

Original Paper

Online Cognitive Behavioral Therapy (CBT) Life Skills Program for Depression: Pilot Randomized Controlled Trial

Christopher Williams[1*], BSc, MD, MBPsS, MMedSc, MBChB; Carrie-Anne McClay[1*], BA (Hons), PhD; Rebeca Martinez[2], LMS, Dr med, MRCGP; Jill Morrison[1], MBChB, MSc, PhD; Caroline Haig[1], BSc, MSc, PhD; Ray Jones[3], BSc, MSc, PhD; Paul Farrand[4], BA (Hons), PG Dip, PhD

[1]Institute of Health and Wellbeing, University of Glasgow, Glasgow, United Kingdom

[2]Mersey Care National Health Service Foundation Trust, Norris Green Community Hub, Liverpool, United Kingdom

[3]Faculty of Health, University of Plymouth, Plymouth, United Kingdom

[4]Sir Henry Wellcome Building for Mood Disorders Research, University of Exeter, Exeter, United Kingdom

[*]these authors contributed equally

Corresponding Author:
Christopher Williams, BSc, MD, MBPsS, MMedSc, MBChB
Institute of Health and Wellbeing
University of Glasgow
Mental Health and Wellbeing, Gartnavel Royal Hospital, Administration building
1st floor, General Office
Glasgow, G12 0XH
United Kingdom
Phone: 44 141 330 3299
Email: chris.williams@glasgow.ac.uk

Abstract

Background: Depression is a common mental health problem with significant personal and social consequences. Studies have suggested that cognitive behavioral therapy (CBT) is an effective treatment for depression and anxiety when delivered one-to-one by an expert practitioner, but access to this talking therapy is often limited, and waiting lists can be long. However, a range of low-intensity interventions that can increase access to services are available including guided CBT self-help materials delivered via books, classes, and online packages.

Objective: This project aimed to pilot a randomized controlled trial (RCT) of an online CBT-based life skills course with community-based individuals experiencing depression.

Methods: Individuals with symptoms of depression were recruited directly from the community via newspaper advertisements. Participants were remotely randomized to receive either immediate access (IA) or delayed access (DA) to a research version of the Living Life to the Full online CBT-based life skills package (3rd edition) with telephone support provided by nonspecialist, charity-based workers while they used the online intervention. The primary end point was at 3 months postrandomization, at which point, the DA group were offered the intervention. Levels of depression, anxiety, social functioning, and satisfaction were assessed.

Results: There were effective recruitment, randomization, and uptake, with 19 IA and 17 DA control participants entering the pilot study via newspaper advertisements and 13 of the 19 participants taking up the intervention. Overall, 72% (26/36) were not currently under the care of their general practitioner. The online package was acceptable to participants; the mean satisfaction score on the Client Satisfaction Questionnaire was 21 out of 32 (SD 8.89). At 3 months, data collection was achieved from 78% (28/36) of the participants. The efficacy and retention data were used for a power calculation indicating that 72 participants in total will be required for a future substantive RCT.

Conclusions: The research design successfully tested the recruitment, data collection, and intervention delivery. The pilot study has provided data for the required sample size for the full RCT.

Trial Registration: ISRCTN registry ISRCTN12890709; https://doi.org/10.1186/ISRCTN12890709

International Registered Report Identifier (IRRID): RR2-10.1186/s13063-016-1336-y

(JMIR Form Res 2022;6(2):e30489) doi:10.2196/30489

KEYWORDS

pilot study; depression; low mood; iCBT; guided self-help; online; psychotherapy; LLTTF; RCT; treatment gap; bibliotherapy; life skills; anxiety

Introduction

Background

Depression is characterized by persistent low mood and loss of pleasure or interest in things. Additional symptoms include sleep problems, loss of energy, changes in appetite, trouble concentrating, feeling like a failure, and thoughts of self-harm or suicide [1]. These symptoms can have a significant impact on social functioning, physical health, work, and relationships and can result in suicide. The most recent UK Wellbeing Survey found that 19.1% of people aged 16 years and older in the United Kingdom show evidence of anxiety and depression [2].

In general practice, symptoms of depression are usually assessed using the Patient Health Questionnaire (PHQ) [3] or similar self-report scale. Following this, there are various treatments that can be recommended. The National Institute for Health and Clinical Excellence (NICE) [4] guidelines state that treatments for depression should follow the stepped care model, taking into account severity and chronicity of symptoms and response to treatment. Treatment options range from active monitoring and psychoeducation at step 1 to psychosocial treatment options such as low-intensity interventions including bibliotherapy (cognitive behavioral therapy [CBT]–informed self-help books), internet-based cognitive behavioral therapy (iCBT), and group CBT within step 2; Step 3 involves face-to-face delivery of therapy such as specialist-delivered CBT, interpersonal therapy, and behavioral activation; and the use of crisis and inpatient services represent steps 4 and 5.

However, the majority of individuals with depression worldwide do not receive appropriate treatment for their depression within health services [5]. This treatment gap may be dependent upon a number of factors, including nonrecognition of the problem (by the individual or health care workers), reluctance to seek treatment from statutory mental or health service provision, limited knowledge about treatment options, or a lack of mental health resources resulting in long waiting lists. A key aim of the World Health Organization mental health Gap Action Programme (mhGAP) is to increase access to psychological therapies via nonexpert support or health care workers. This could be achieved through the application of low-intensity CBT-guided self-help interventions, delivered via bibliotherapy or online in the form of iCBT. These could be delivered, as recommended, in Step 2 of the stepped care model and through enhancing “informal health care” service provision through trained members of community or voluntary sector organizations [6] to improve access for those reluctant to present to health services.

Various iCBT packages are available, including MoodGYM [7], Beating the Blues [8], and Living Life to the Full (LLTTF) [9]. There is evidence that online CBT can be an acceptable, impactful, and cost-effective intervention for depression [10,11] Some studies have found iCBT to be as effective as face-to-face CBT [12,13].

A review carried out by Richards and Richardson [14] found that outcomes were best when iCBT interventions were delivered with some form of support rather than unguided. This review highlighted that drop-out rates are high in studies of iCBT (57% across the 40 studies). The issues of support, guidance, and retention are key in interpreting the “REEACT” study of iCBT for depression by Gilbody et al [15]. This study found no significant effect of free (MoodGYM) or commercial (Beating the Blues) iCBT packages when used with purely technical support (as opposed to clinical support) at 4 months compared with usual general practitioner (GP) care. However, significant methodological concerns have been highlighted concerning the support offered, with caution expressed regarding the conclusion reached [16]. Poor rates of uptake were experienced alongside nearly one-quarter (24%) of participants dropping out by 4 months. Furthermore, those engaging with the intervention only completed 1 to 2 modules with just over 6 minutes in total of technical telephone support received on average. Such small amounts of support are below that recommended for so-called minimal contact support [17]. Furthermore, the support was described as technical support alone, which means that the results need to be interpreted with caution. In contrast, support for iCBT interventions is usually provided by a therapist or trained facilitator and improves outcomes [14]. Such content-focused support usually involves motivating users, goal setting, and review of homework tasks, with the main focus being on working through the online package and applying what is learned.

This study is a pilot study of an online educational life skills program designed for low mood and anxiety based on CBT principles. The LLTTF approach puts bibliotherapy and life skills training at its core and adopts a blended learning approach [18] where the user engages with the course through a variety of media (reading, audiovisual modules, and video). Content addresses key aspects of low mood and stress and is available in different formats to suit the user’s preferred learning style. This includes live classes [19-21], printed books [19], and electronic books in addition to the online course. The online version of the LLTTF intervention adheres to the key NICE recommendations regarding iCBT packages in terms of content, number of modules, and the support element provided [4]. However the person chooses to learn, it is recommended that a clear and structured support protocol is offered to enhance package engagement and application. A nonrandomized feasibility study [22] demonstrated equivalent outcomes for the online course [9] compared with the Beating the Blues online resource [8] and a book-based bibliotherapy intervention [23,24]. However, as there are currently no adequately powered randomized controlled trials (RCTs) of the LLTTF package,

further research is needed to examine the efficacy of the package in its online format.

Research Questions

Primary Question Examining Methodological Uncertainties

The primary research question was: "Is the study design feasible—is it possible to recruit from the community, randomize participants, deliver the online intervention with structured telephone support, and collect data at baseline and 3 months post randomization?"

Secondary Questions Examining Adherence and Acceptability

The secondary research questions were: (1) "To what extent will participants adhere to the online intervention?" (2) "Is the Living Life to the Full online package acceptable to participants?" and (3) "What sample size will be required in the future substantive study?"

Methods

Overview

The study was a parallel, 2-arm, pilot RCT with a 50:50 allocation ratio across the 2 groups with the primary outcome point at 3 months postrandomization. Participants were remotely randomized to receive the online intervention immediately (IA group) or following a 3-month delay after randomization (DA group). The DA group served as the control at the primary outcome point.

Ethics Approval

Ethical approval was granted by the College of Medical, Veterinary and Life Sciences Ethics Committee for Non-Clinical Research Involving Human Subjects, University of Glasgow (Reference number 200140159).

Procedures

Recruitment and Participants

The aim was to recruit 30 to 50 participants, a sample size widely accepted as sufficient for pilot studies to identify problems with recruitment, delivery of the intervention, and evaluation measures [25]. People with symptoms of depression were recruited using a community-based approach using advertisements in the Metro free newspaper in Central Scotland. No participants were recruited directly via the National Health System (NHS). Informed consent was collected from all potential participants prior to entry into the study.

Inclusion Criteria

Participants had to meet the following criteria: age ≥18 years; living in the United Kingdom; able to understand written and spoken English; regular access to a computer, smartphone, and tablet with audio and broadband connections; and PHQ-9 score ≥10 [3].

Exclusion Criteria

Participants were excluded for any of the following reasons: high rating for suicidality (ie, scoring 2 or 3 on PHQ-9 item 9), currently receiving any psychological intervention such as counselling or psychotherapy, new or altered dose of antidepressant in the last month, or taking part in any other research projects.

Randomization

Participants completed initial online questionnaires, and those that were eligible and consented were then remotely randomized to either the IA or DA group. Participants' ID numbers were passed to a separate researcher who used the randomization function in Excel to remotely assign participants to the immediate IA or DA group. As it was a pilot study, we did not stratify for any variables during randomization. After randomization, those in the IA group were given a free access code and link to the LLTTF website, including instructions for how to use the resources. Those in the DA group were told that they would be contacted again in 3 months; the free access code to the LLTTF website was sent once follow-up data had been received.

Intervention

The free-access, online LLTTF course consists of 8 modules that teach a range of CBT-based life skills.

Session 1 involves an introduction to the CBT model and the "vicious cycle" of low mood that aims to help participants understand why they feel the way they do and how their thoughts, feelings, physical symptoms, and behavior are linked.

The session 2 module focuses on the impact of reduced activity. Users are encouraged to consider the things that they have stopped or reduced doing as a result of their low mood and make a plan to re-establish these activities in order to improve their mood.

Session 3 teaches skills to help tackle upsetting thinking, including how to label unhelpful thoughts, identify negative thinking patterns, and turn these thoughts around to create more helpful ones.

Session 4 teaches how self-confidence is developed and teaches confidence-building techniques.

Session 5 involves a problem-solving approach using an "easy 4-step plan" to help break down problems into smaller parts in order to overcome challenges.

Session 6 addresses unhelpful behaviors that may be worsening mood. Participants are encouraged to recognize problem behaviors such as isolating themselves or drinking or smoking too much and then create a plan for reducing them.

In session 7, participants learn to recognize the things that cause them to feel irritable or angry and the early warning signs they experience when they start to feel angry. They then learn techniques for better managing their anger to help them react differently to challenging people and situations.

Session 8, which is the final module, teaches key lifestyle choices that can improve mood, including healthy eating, exercise, and closeness with others.

Additional course worksheets, video, and audio (relaxation) resources supplement the main modules. These are available on the intervention website.

Support

The NICE guidelines [4] for the treatment of adults with depression advise that self-help packages should have support that includes monitoring and reviewing progress. This study included automated weekly email reminders plus 6 weekly support sessions provided over the telephone with a support worker. The support workers were volunteers from the Scottish depression charity Action on Depression who had completed at least one day of training with the Five Areas team (developers of LLTTF) and had experience with delivering LLTTF content and support. Support was protocol-driven and focused on encouraging use and application of the intervention [18]. It was designed to encourage an individualized plan to be made at the end of each session, completing Planner and Review worksheets informed by a "Plan, Do, Review" structure. Within this structure, participants apply the skills and techniques they have learned to make changes in their lives, using the materials to support them in doing this. The amount of support time allocated varied to reflect the complexity of the needs of the person being supported; session durations were not recorded in this pilot study although the recommended duration was 15 minutes to 20 minutes per session.

Data Collection

Baseline measures included age, gender, ethnicity, mood, social function, and past and current sources of support. After 3 months, all randomized participants were invited to repeat mood and social functioning questionnaires. Use of the intervention and satisfaction with the online package in the IA group were assessed. The collection of data for an economic analysis was piloted at this time point. This involved completion of the Client Service Receipt Inventory (CSRI) [26] and EQ5D [27], which assess access to health care services and health-related quality of life. These data were not analyzed, and the purpose was to test the ability to collect the data.

Outcome Measures

The key outcomes of the study were the ability to recruit and retain participants, to test the acceptability of the intervention as measured by the Client Satisfaction Questionnaire-8 (CSQ-8) [28], and the ability to gather data.

The measures used to assess the impact on mood and social functioning were the PHQ-9 [3], the Generalized Anxiety Disorder (GAD-7) questionnaire [29], and the Work and Social Adjustment Scale (WSAS) [30]. These were sent via a SurveyMonkey [31] link at each time point.

Statistical Analyses

Recruitment, uptake, and adherence to the intervention are summarized. Baseline characteristics are summarized for all participants as well as for the IA and DA groups separately. Two-sample *t* tests or Mann-Whitney *U* tests, depending on distributions, and Fisher exact tests were used to compare the 2 treatment groups and confirm that participants had been effectively randomized.

To practice the analysis for efficacy, appropriate linear models were applied using an intention-to-treat approach. Models were adjusted using baseline scores as covariates. These models were used to estimate and test the statistical significance of the within-group change between baseline and 3 months and the between-group differences at each time point.

Results

Recruitment

The community-based recruitment method we employed was successful as our sample size target to recruit 30 to 50 participants was met within the proposed recruitment period (August 2015 to May 2016; see Table 1). Recruitment involved 3 distinct recruitment phases of advertising, each including 1 Metro newspaper advertisement appearance per week for 4 weeks. Therefore, the total number of Metro advertisements placed was 12.

Of the 70 potential participants who were assessed for eligibility, 36 individuals were randomized by the end of the recruitment period. As shown in the CONSORT flow diagram (Figure 1), 5 people were excluded due to thoughts of self-harm or suicide. These individuals were advised to contact their GP regarding these feelings and were also given a list of additional sources of support that they could access.

As outlined in Table 2, in the randomized sample, the mean chronicity of depressive symptoms was 9.20 (SD 8.70) years. Symptoms had persisted for at least 2 years for 78% (28/36) of participants and for at last 5 years for 56% (20/36), reflecting a chronic group. Despite this, 26 of the 36 participants (72%) were not currently under the care of their GP, and 20 of the 36 participants (56%) were getting no support at all for their low mood.

Table 1. Recruitment phases.

Recruitment phase	Dates	Number of advertisements
1	August 2015 to October 2015	4 (1 a week for 4 weeks)
2	January 2016 to March 2016	4 (1 a week for 4 weeks)
3	March 2016 to May 2016	4 (1 a week for 4 weeks)

Figure 1. Study flowchart. PHQ-9: Patient Health Questionnaire-9.

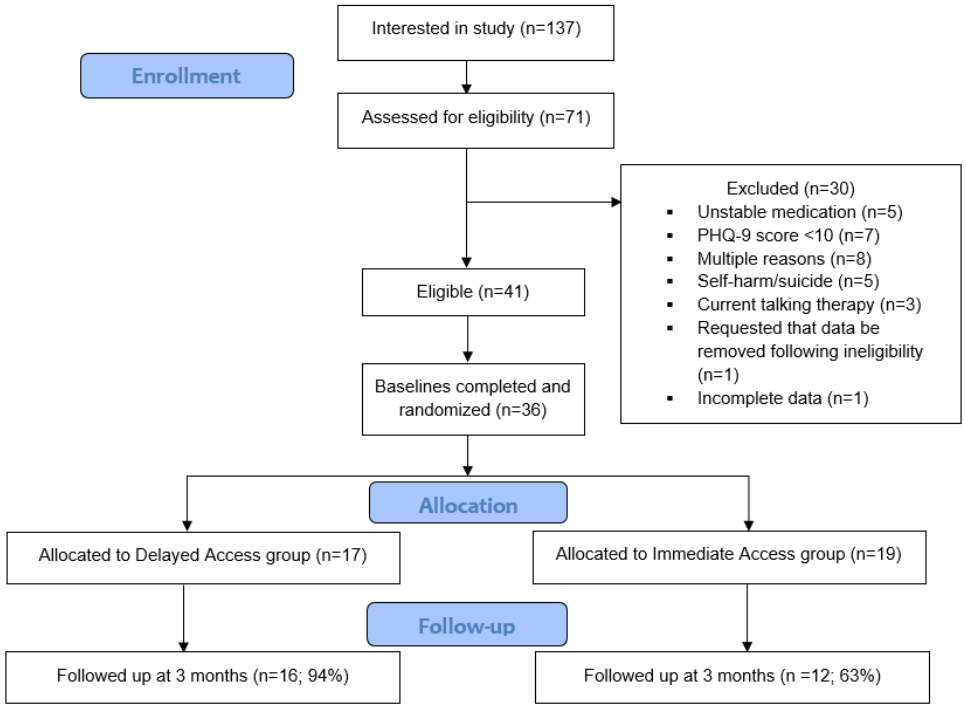


Table 2. Demographic data.

Characteristic	Overall (n=36)	Control (n=17)	Intervention (n=19)
Age (years)			
Mean (SD)	42.33 (13.54)	41.82 (15.38)	42.79 (12.09)
Minimum to maximum	20.00 to 69.00	23.00 to 69.00	20.00 to 63.00
Sex, n (%)			
Female	24 (67)	11 (65)	13 (68)
Male	11 (31)	5 (29)	6 (32)
Transgender	1 (3)	1 (6)	0 (0)
Chronicity (years)[a]			
Mean (SD)	9.20 (8.70)	7.85 (6.54)	10.48 (10.39)
Minimum to maximum	0.17 to 30.00	0.33 to 20.00	0.17 to 30.00
Marital status, n (%)			
Married/cohabiting	17 (47)	8 (47)	9 (47)
Unmarried/separated	19 (53)	9 (53)	10 (53)
Antidepressant use, n (%)			
Yes	10 (28)	3 (18)	7 (37)
No	26 (72)	14 (82)	12 (63)
Ethnicity, n (%)			
White	36 (100)	17 (100)	19 (100)
Non-White	0 (0)	0 (0)	0 (0)

[a]Because chronicity was not available for 1 person in the intervention group, the sample size for the overall group was 35, and the sample size for the intervention group was 18.

Uptake and Adherence

These findings relate to the IA group participants only as the DA had not accessed the intervention at 3 months. Of the 19 IA participants, 13 provided data concerning the number of sessions they completed, and 10 (10/13, 77%) of the responding IA participants completed at least 1 module of the intervention. The average number of the 4 core modules completed was 2

(SD 1.58; minimum 0, maximum 4). One-third (4/13, 31%) of participants completed all 4 core modules. The mean number of optional modules completed was 1 (SD 1.41).

Ability to Collect Data and Questionnaires

At the 3-month follow-up, primary outcome data were collected from 78% (28/36) of the participants. A Fisher exact test showed a significant between-group difference in follow-up rates (P=.03), with 94% (16/17) of the DA participants completing the measures at 3 months, compared with 63% (12/19) of the IA participants. Economic analysis data were collected from 25% (7/28) of responding participants at 3 months.

Participant Satisfaction

Participant satisfaction with the LLTTF online package was measured using the CSQ-8 for the intervention group. Of the 19 IA participants, 10 completed this questionnaire. A mean satisfaction score of 21 (SD 8.89), from a possible score of 32, indicated that participants were reasonably satisfied with the online intervention and support they received.

In relation to side effects of the intervention, very few were reported. Specifically, 1 participant said that they felt bad about their ability to put what they learned in the course into action, 4 people felt that they were letting the support worker down during the intervention, and 1 person said that the course content made them feel worse.

Initial Assessments of Efficacy

Effect on Depression (PHQ-9) Scores

The IA and DA groups improved to a similar level between baseline and 3 months. The mean reduction in PHQ-9 score in the IA group was 6.08 (SD 7.49); the mean difference in the DA group was 5.19 (SD 6.00). As expected in this small unpowered pilot study, the between-group difference was not statistically significant (P=.62; 95% CI –6.10 to 3.69).

Effect on Anxiety and Social Functioning

At 3 months, the IA group had improved by 4.60 (SD 6.53) points on the GAD-7 from baseline, while the average improvement in the DA group was 2.80 (SD 4.97) points. Again, no significant treatment effect was found for anxiety (P=.20; 95% CI –7.76 to 1.70). Similarly, there was no significant effect of treatment on social functioning (WSAS scores; P=.99, 95% CI –8.16 to 8.02).

Discussion

Principal Findings

This pilot study aimed to test key elements of a pilot RCT of the LLTTF online package. It demonstrated that a group of self-referring participants can be successfully recruited from the community and engaged in the RCT, with voluntary sector delivery. Several key methodological uncertainties in relation to study design and intervention delivery have been addressed, thus providing useful information for a future fully powered trial.

Recruitment and Sample

First, recruitment was feasible, with 36 eligible individuals recruited within 3 phases of advertising in the Metro newspaper. The resources were limited in the current study, but it seems likely that, in a future funded RCT with an appropriate budget for UK-wide advertising, the required sample size (see the following section) could be achieved using a community-based recruitment approach. The community-based approach employed successfully identified those living with depression. The majority of participants had at least moderate symptoms of depression, yet less than one-third was currently under the care of their GP. This is in line with reports that there is a significant treatment gap with regard to depression [5]. Online interventions could serve as a potential treatment avenue to bridge the gap between onset of symptoms and access to psychological therapy.

Intervention Delivery and Scalability

The completion and satisfaction rates in this pilot study demonstrate that it is feasible and acceptable to deliver the intervention in a community setting with voluntary sector support. Working alongside an established voluntary sector organization was a strength of our intervention, and the full RCT aims to broaden the number of organizations involved, including providing access more widely across the United Kingdom. This will test the ability to scale up and deliver to a higher volume of participants.

Data Collection and Follow-up

It proved possible to record data, with a 78% follow-up rate at 3 months. However, only one-quarter of responding participants completed the economic analysis questionnaires at 3 months, indicating that the format (emailed rather than completed via SurveyMonkey) and length may need to be reconsidered, as it appears that this outcome measure was too burdensome.

Power Calculation

Mood measures were collected at baseline and 3 months and used along with rates of follow-up to calculate the sample size for a future definitive RCT.

The primary analysis in the future substantive RCT will compare changes in PHQ-9 scores at 6 months between intervention groups. We will power the study to be able to detect a between-group difference of 5 points on the PHQ-9 score. A difference of 5 points is used to reflect a category change (from moderate to mild depression, for example) on the PHQ-9 and is clinically important.

Based on a 2-sample t test, a sample size of 36 participants per arm would be required to have 95% power. In the pilot, follow-up data were available for 78% at 3 months. The proposed study would have 1 follow-up assessment at 6 months, so we have accounted for a 25% drop-out rate for the full RCT.

Limitations

A limitation of this study is that, although we determined that participants had significantly high mood scores upon entry into the study, a diagnostic interview was not used in the screening process. This approach was applied because some people prefer voluntary sector delivery so they can seek help without a

diagnosis being made or creating a formal NHS record. However, it is acknowledged that defining the participants in terms of diagnosis would be useful for a research analysis. Therefore, in order to address this limitation while promoting participant choice, the future RCT could include an optional psychiatric interview or diagnostic assessment at baseline. This method was applied in a previous study in which 65% of participants agreed to take part in the diagnostic interview [19,21]. Negative impacts will be assessed in the full RCT using qualitative interviews including with those who drop out of the study.

Conclusions

This pilot study, designed to test the trial design and the delivery of the online life skills LLTTF course, has helped inform the sample size calculation for a future RCT and determined that such a trial is feasible in relation to recruitment, delivery, and retention. Our results have addressed several methodological uncertainties and demonstrated the feasibility to conduct a definitive RCT within a community-based setting, delivered with support from nonspecialist voluntary sector support workers. The power calculation suggests a sample size of 72 is required for the full RCT. This pilot study strongly supports the potential to successfully recruit this required number of participants using a community-based recruitment strategy.

Delivery of the intervention by charity support workers was associated with reasonable participant satisfaction. However, efforts will need to be taken to improve completion rates in a future definitive RCT. Additionally, it proved problematic to conduct an economic analysis via emailed documents. It seems a telephone or other live interview might work more effectively for any future definitive RCT. Finally, in the future study, there would be advantages to collecting data at 6 months and possibly at 1 year to record long-term outcomes and whether any changes to mood and social function are sustainable.

Acknowledgments

The project was sponsored by the University of Glasgow, with partial funding by NHS Greater Glasgow and Clyde and was carried out within the Institute of Health and Wellbeing. We acknowledge the assistance of the Action on Depression with the study design, recruitment, and support delivery. Thank you also to Michael Helley (volunteer research assistant) for his contribution to the project.

Authors' Contributions

CW is the chief investigator of the study, contributed to the writing of the study protocol, supervised the support workers, and contributed to this outcome paper. C-AM contributed to the writing of the study protocol; carried out recruitment and screening; assigned participants to groups; analyzed the recruitment rate, adherence, and satisfaction data; liaised with the Action on Depression support workers; facilitated access to the intervention; and collected and interpreted follow-up data as well as contributing to writing this outcome paper. CH is the statistician for the study and analyzed the efficacy data, carried out the power calculation for the project, assisted with study design decisions, and provided statistical support in the analysis of the final outcome data. PF contributed to the writing of the study protocol and outcome paper and contributed trial expertise. RJ assisted with the design of the study, shared his trial expertise during the study, and contributed to the interpretation and writing of the findings. JM contributed to the conceptualization of the study, shared her depression research and RCT (randomized controlled trial) expertise during the study, and contributed to the interpretation and writing of the findings. RM contributed to the development of the research design and was available to share her clinical and research experience to assist with the trial and the interpretation of the findings. All authors contributed to the interpretation and reporting of the findings and were involved in the preparation of this manuscript.

Conflicts of Interest

CW is an author of a variety of written and iCBT (internet-based CBT) resources. He is Director and shareholder of Five Areas Ltd, a company that delivers free and licensed online life skills resources including the LLTTF (Living Life to the Full) course.

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 15144 KB - formative_v6i2e30489_app1.pdf]

References

1. Diagnostic and Statistical Manual of Mental Disorders, 5th ed. Washington, DC: American Psychiatric Association Publishing; 2013.
2. Measures of National Well-being. Office for National Statistics. 2019 Oct 23. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/wellbeing/articles/measuresofnationalwellbeingdashboard/2018-04-25 [accessed 2021-12-11]
3. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. JAMA 1999 Nov 10;282(18):1737-1744. [doi: 10.1001/jama.282.18.1737] [Medline: 10568646]

4. Depression in adults: recognition and management. National Institute for Clinical Excellence. 2009 Oct 28. URL: https://www.nice.org.uk/guidance/cg90 [accessed 2021-12-11]
5. Depression. World Health Organization. URL: https://www.who.int/en/news-room/fact-sheets/detail/depression [accessed 2021-12-11]
6. Shidhaye R, Lund C, Chisholm D. Closing the treatment gap for mental, neurological and substance use disorders by strengthening existing health care platforms: strategies for delivery and integration of evidence-based interventions. Int J Ment Health Syst 2015;9:40 [FREE Full text] [doi: 10.1186/s13033-015-0031-9] [Medline: 26719762]
7. moodgym. URL: https://moodgym.anu.edu.au/ [accessed 2021-12-11]
8. Beating the Blues. URL: http://www.beatingtheblues.co.uk/ [accessed 2021-12-11]
9. Living Life to the Full website. URL: https://www.lltf.com [accessed 2022-01-12]
10. Andersson G, Rozental A, Shafran R, Carlbring P. Long-term effects of internet-supported cognitive behaviour therapy. Expert Rev Neurother 2018 Jan;18(1):21-28. [doi: 10.1080/14737175.2018.1400381] [Medline: 29094622]
11. Andersson G, Titov N, Dear BF, Rozental A, Carlbring P. Internet-delivered psychological treatments: from innovation to implementation. World Psychiatry 2019 Feb 02;18(1):20-28 [FREE Full text] [doi: 10.1002/wps.20610] [Medline: 30600624]
12. 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]
13. Karyotaki E, Ebert DD, Donkin L, Riper H, Twisk J, Burger S, et al. Do guided internet-based interventions result in clinically relevant changes for patients with depression? An individual participant data meta-analysis. Clin Psychol Rev 2018 Jul;63:80-92. [doi: 10.1016/j.cpr.2018.06.007] [Medline: 29940401]
14. 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]
15. Gilbody S, Littlewood E, Hewitt C, Brierley G, Tharmanathan P, Araya R, REEACT Team. Computerised cognitive behaviour therapy (cCBT) as treatment for depression in primary care (REEACT trial): large scale pragmatic randomised controlled trial. BMJ 2015 Nov 11;351:h5627 [FREE Full text] [doi: 10.1136/bmj.h5627] [Medline: 26559241]
16. Jones M, Ebert D, Jacobi C, Beintner I, Berger T, Görlich D, ICare Investigators. Rapid Response: Why didn't patients use it? Engagement is the real story in Gilbody et al. (2015), not effectiveness. BMJ 2015:351 [FREE Full text]
17. Glasgow RE, Rosen GM. Behavioral bibliotherapy: a review of self-help behavior therapy manuals. Psychol Bull 1978 Jan;85(1):1-23. [Medline: 343144]
18. Williams C, Morrison J. A new language for CBT: new ways of working require new thinking, as well as new words. In: Bennett-Levy J, Richards D, Farrand P, Christensen H, Griffiths K, Kavanagh D, et al, editors. Oxford Guide to Low Intensity CBT Interventions. Oxford, England: Oxford University Press; 2010:69.
19. McClay C, Collins K, Matthews L, Haig C, McConnachie A, Morrison J, et al. A community-based pilot randomised controlled study of life skills classes for individuals with low mood and depression. BMC Psychiatry 2015 Feb 06;15(1):17 [FREE Full text] [doi: 10.1186/s12888-015-0384-2] [Medline: 25884922]
20. Williams C, McClay C, Matthews L, McConnachie A, Haig C, Walker A, et al. Community-based group guided self-help intervention for low mood and stress: randomised controlled trial. Br J Psychiatry 2018 Feb 01;212(2):88-95 [FREE Full text] [doi: 10.1192/bjp.2017.18] [Medline: 29436324]
21. McClay C, Morrison J, McConnachie A, Williams C. A community-based group-guided self-help intervention for low mood and stress: study protocol for a randomized controlled trial. Trials 2013 Nov 19;14:392 [FREE Full text] [doi: 10.1186/1745-6215-14-392] [Medline: 24252475]
22. Williams C, Wilson P, Morrison J, McMahon A, Walker A, Allan L, et al. Guided self-help cognitive behavioural therapy for depression in primary care: a randomised controlled trial. PLoS One 2013 Jan 11;8(1):e52735 [FREE Full text] [doi: 10.1371/journal.pone.0052735] [Medline: 23326352]
23. Pittaway S, Cupitt C, Palmer D, Arowobusoye N, Milne R, Holttum S, et al. Comparative, clinical feasibility study of three tools for delivery of cognitive behavioural therapy for mild to moderate depression and anxiety provided on a self-help basis. Ment Health Fam Med 2009 Sep;6(3):145-154 [FREE Full text] [Medline: 22477905]
24. Williams C. Overcoming Anxiety, Stress and Panic: A Five Areas Approach. Boca Raton, FL: CRC Press; 2003.
25. Williams C. Overcoming Depression and Low Mood: a Five Areas Approach. London, England: Hodder Education Publishers; 2006.
26. Beecham J, Knapp M. Costing psychiatric interventions. In: Thornicroft G, Brewin C, Wing J, editors. Measuring mental health needs. London, England: Gaskell/Royal College of Psychiatrists; 1992:163-183.
27. EuroQol Group. EuroQol–a new facility for the measurement of health-related quality of life. Health Policy 1990 Dec;16(3):199-208. [doi: 10.1016/0168-8510(90)90421-9] [Medline: 10109801]
28. Nguyen TD, Attkisson C, Stegner BL. Assessment of patient satisfaction: Development and refinement of a Service Evaluation Questionnaire. Evaluation and Program Planning 1983 Jan;6(3-4):299-313. [doi: 10.1016/0149-7189(83)90010-1]
29. 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]

30. Mundt JC, Marks IM, Shear MK, Greist JM. The Work and Social Adjustment Scale: a simple measure of impairment in functioning. Br J Psychiatry 2002 May 2;180(05):461-464. [doi: 10.1192/bjp.180.5.461] [Medline: 11983645]
31. SurveyMonkey. URL: https://www.surveymonkey.com/ [accessed 2021-12-11]

Abbreviations

CBT: cognitive behavioral therapy
CSQ-8: Client Satisfaction Questionnaire-8
CSRI: Client Service Receipt Inventory
DA: delayed access
GAD-7: Generalized Anxiety Disorder-7
GP: general practitioner
IA: immediate access
iCBT: internet-based CBT
LLTTF: Living Life to the Full
mhGAP: mental health Gap Action Programme
NHS: National Health System
NICE: National Institute for Health and Clinical Excellence
PHQ-9: Patient Health Questionnaire-9
RCT: randomized controlled trial
WSAS: Work and Social Adjustment Scale

Edited by G Eysenbach; submitted 17.05.21; peer-reviewed by N Keränen, T Gargot; comments to author 30.07.21; revised version received 21.11.21; accepted 22.11.21; published 17.02.22.

Please cite as:
Williams C, McClay CA, Martinez R, Morrison J, Haig C, Jones R, Farrand P
Online Cognitive Behavioral Therapy (CBT) Life Skills Program for Depression: Pilot Randomized Controlled Trial
JMIR Form Res 2022;6(2):e30489
URL: https://formative.jmir.org/2022/2/e30489
doi:10.2196/30489
PMID:35175203

Original Paper

Think-Aloud Testing of a Novel Safer Drinking App for College Students During COVID-19: Usability Study

Jessica Gomez Smith[1], BA; Nour Sami Alamiri[2], MPH; Grace Biegger[1], BA; Christina Frederick[1], BSc; Jennifer P Halbert[2], MSPH; Karen S Ingersoll[1], PhD

[1]Center for Behavioral Health & Technology, Department of Psychiatry and Neurobehavioral Sciences, University of Virginia School of Medicine, Charlottesville, VA, United States

[2]Center for Addiction and Prevention Research, Department of Public Health Sciences, University of Virginia School of Medicine, Charlottesville, VA, United States

Corresponding Author:
Karen S Ingersoll, PhD
Center for Behavioral Health & Technology
Department of Psychiatry and Neurobehavioral Sciences
University of Virginia School of Medicine
560 Ray C Hunt Dr
Charlottesville, VA, 22903
United States
Phone: 1 434 243 0581
Email: KES7A@hscmail.mcc.virginia.edu

Abstract

Background: Hazardous alcohol consumption, and binge drinking in particular, continues to be common among college students, posing the greatest risk for their health and safety. Despite widespread exposure to evidence-based preventive interventions among US undergraduates, only modest and temporary effects on risky drinking occur. Formative studies have demonstrated that students want a more engaging intervention tool for risky drinking that can be used *just in time*.

Objective: The purpose of this study is to test the appeal, relevance, and perceived utility of a draft mobile app for safer student drinking at a public university in Virginia.

Methods: Undergraduate student participants tested the draft mobile app via a web-based prototype that tailors to individual feedback with hot spots that responded to their taps to mimic app functionality. They narrated their impressions, navigation, and comments in a standardized think-aloud procedure. After each round of think-aloud interviews, researchers debriefed the investigators and developers to discuss findings and brainstorm app modifications.

Results: Minor changes to the functionality and aesthetics would improve usability of the app (eg, option for light mode in app settings). Student testers recommended tailoring the app to the needs of college students and to aspects of the local university's drinking culture.

Conclusions: Findings from this study will be synthesized with information gained from other formative work to determine the final app features. We will test the app in a pilot randomized trial to assess app use and the impact of the app on college student drinking behavior over several months.

(JMIR Form Res 2022;6(2):e32716) doi:10.2196/32716

KEYWORDS

app development; college binge drinking; think aloud testing; formative research; mobile phone

Introduction

Binge Drinking

Binge drinking, defined as consuming 5 (for men) or 4 (for women) standard drinks within approximately 2 hours on a single occasion [1], continues to be a public health concern, particularly among college students [1]. In this millennium, there has been an increase of over 200% in alcohol-related deaths and a 25% increase in alcohol overdose–related hospitalizations in the United States [2]. Binge drinking can result in short-term effects such as poor academic performance, but more seriously, it can lead to detrimental health outcomes

such as alcohol use disorder and other health risks [3]. College binge drinking habits can predict the course of alcohol dependence and problems over a lifetime [4,5]. Developing effective interventions to reduce binge drinking among college students has the potential to improve lifetime drinking outcomes and prevent alcohol-related health and psychosocial problems among college students across the United States.

Alcohol Consumption Among Undergraduates

Undergraduate students are commonly exposed to evidence-based binge drinking preventive and clinical interventions that aim to promote drinking safety [6-8]. An example of a common intervention across universities in the United States is the Brief Alcohol Screening and Intervention for College Students program, which provides a harm reduction intervention that uses personalized normative feedback and structured prompts to guide behavior change for students who have risky drinking habits or alcohol-related consequences [9]. Despite universities continually improving alcohol programming, these efforts yield minimal change and rates of dangerous alcohol consumption have not changed significantly; worse, the effects of these programs diminish over time [10-13]. Therefore, novel interventions that engage students and promote lower risk drinking behavior by college students are still needed. A tailored mobile app for students could provide personalized alcohol safety information in real time, either before or during a drinking episode.

This Study

With extensive student involvement, and based on findings from other formative work [14-16], we designed a draft mobile app for safer student drinking. The aim of this study is to determine college students' perspectives on the appeal, relevance, and perceived utility of the draft app.

Methods

Basis of App Features and Functions

The setting was a public university in Virginia with extensive alcohol programming, alcohol interventions, and student involvement in governance, which has experienced several high-profile alcohol-related injuries and deaths of students in the past decades. We sought to identify new ways to reduce hazardous drinking among college students. We conducted 2 studies before this one. First, we conducted in-depth interviews and focus groups with students and university stakeholders to understand the culture of the university, the role of drinking in student and alumni life, the history and impact of various alcohol interventions, the continuing ubiquity of binge drinking despite deployment of numerous interventions, and students' impressions of the impact of alcohol interventions to which they had been exposed (Alamiri, N, unpublished data, January 2022). Second, we conducted a review of public social media posts to understand how alcohol use and its consequences at the university were documented on the web by students and bystanders [17]. These formative studies demonstrated that (1) risky drinking is a prominent feature of student life, (2) most students have participated in risky drinking behavior, (3) students report minimal change in their own drinking in response to existing alcohol programming, (4) mobile app use by students is ubiquitous, and (5) students wanted flexible intervention tools that are accessible during drinking episodes (their own or when witnessing excessive drinking by their peers). Students requested a novel intervention with evidence-based and student-preferred features and functions. An estimated blood alcohol content (eBAC) calculator (to estimate the percentage of alcohol in a person's bloodstream [15]) was the feature students were most interested in when the concept was presented during the focus groups and surveys. Overall, the formative findings indicated that novel interventions must be highly relevant to drinking behaviors and be usable during drinking.

App Prototype

Study team members created a draft prototype app based on prior experience in mobile app development and prior research on alcohol intervention development. Previous research indicates the benefits of self-monitoring for behavior change [18,19]. Logging risky behaviors, such as drinking, helps reduce those behaviors [14], leading us to include a drink tracker. To boost engagement, the app included personalized encouragement [20,21] to users who tracked several days in a row (streak). We used push notifications [22] to provide safety messages during heavy drinking episodes to alert the user of potential danger and provide recommended actions. In-app messaging, called the *message center* [22], provided *testimonials* from real students about drinking situations. To appeal to college student users' needs, we added a secondary behavior (hydration or sleep) to track that could prompt users to engage with the app even when not drinking alcohol to keep users engaged [23]. The settings feature within the app allowed users to turn off push notifications and adjust their weight to reflect a more accurate eBAC. Once the prototype app was developed, 2 classes of undergraduate and graduate students in a mobile engineering class provided initial feedback on the draft app, and some features were slightly redesigned. Thus, the final prototyped app contained eight features and a home screen (Figure 1): alcoholic drink tracker (Figure 2); additional behavior tracker (hydration or sleep); self-monitoring feedback of days tracked; push messages about drinking and safety; eBAC calculator that provided an eBAC based on weight, sex, and timing or specific drinks logged; learning topics (Figure 3); a message center containing student testimonials (Figure 4); and settings.

Figure 1. Home screen of the intervention (bhoos app). BAC: blood alcohol content.

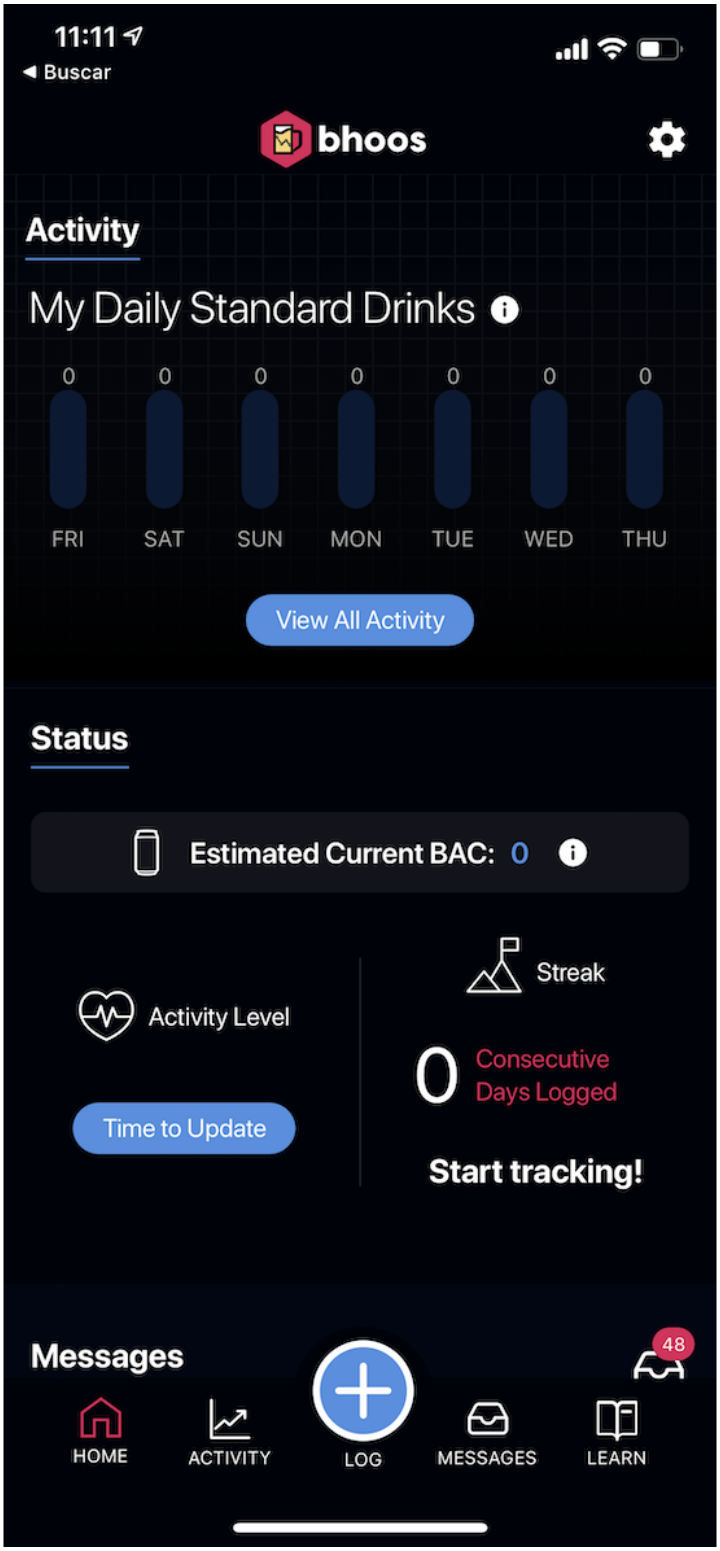


Figure 2. Drink tracker of the intervention (bhoos app).

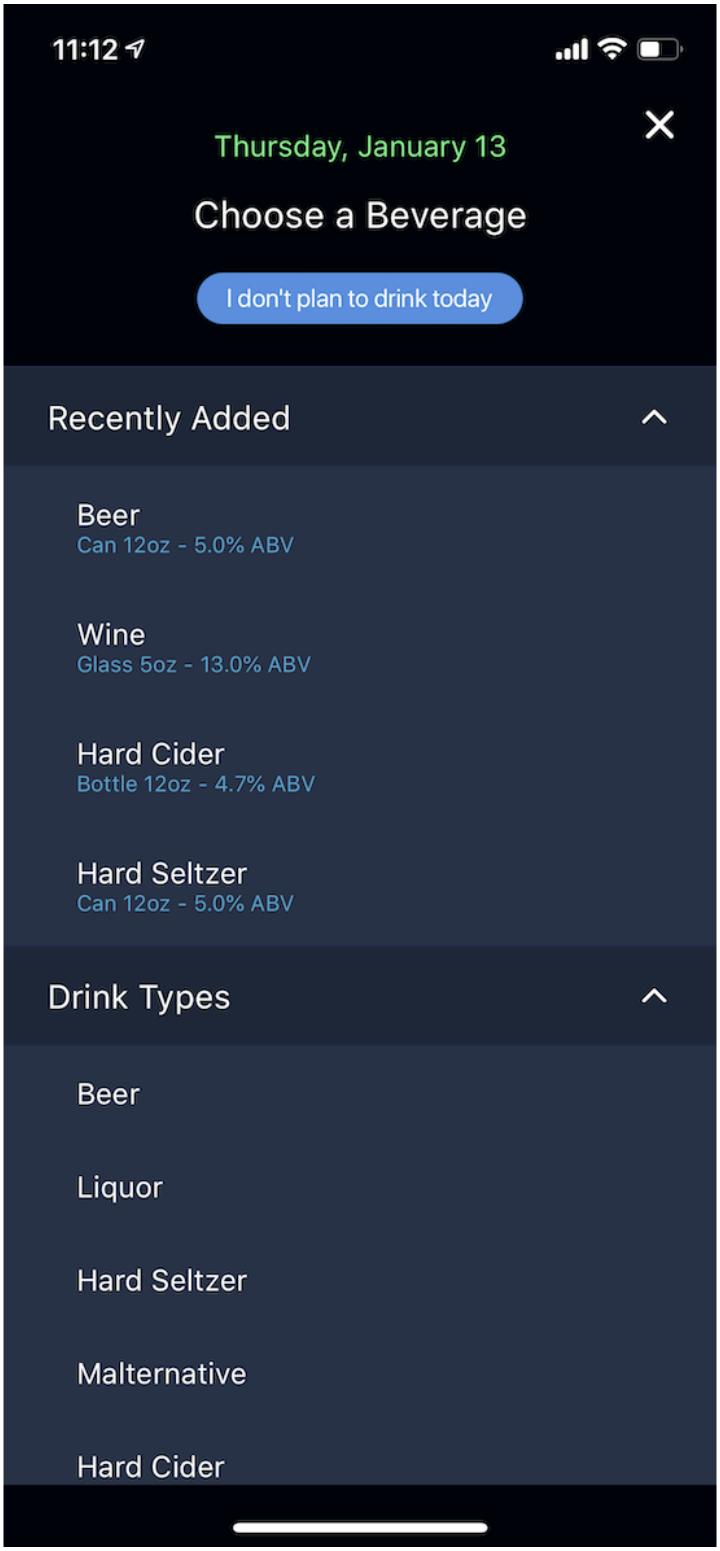


Figure 3. Learn topics of the intervention (bhoos app). eBAC: estimated blood alcohol content.

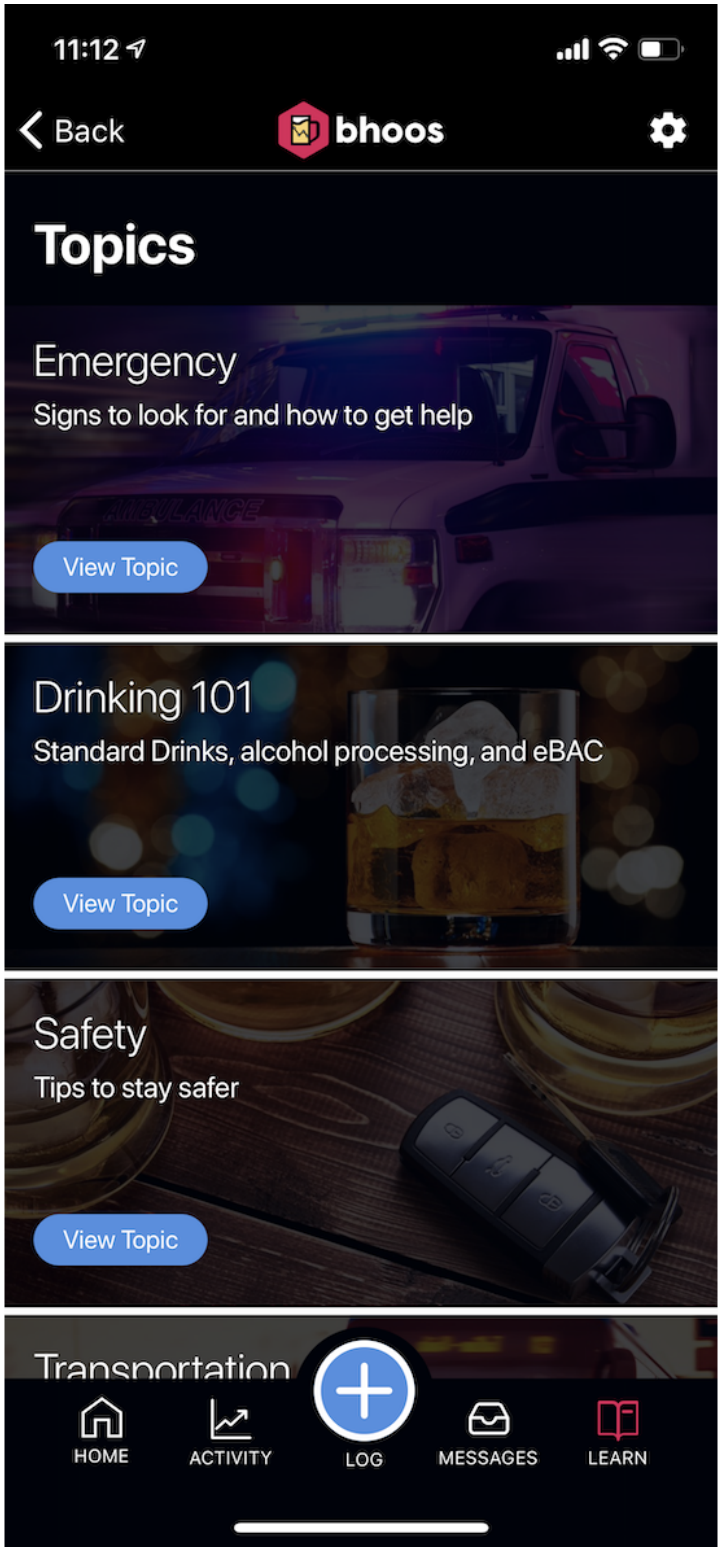
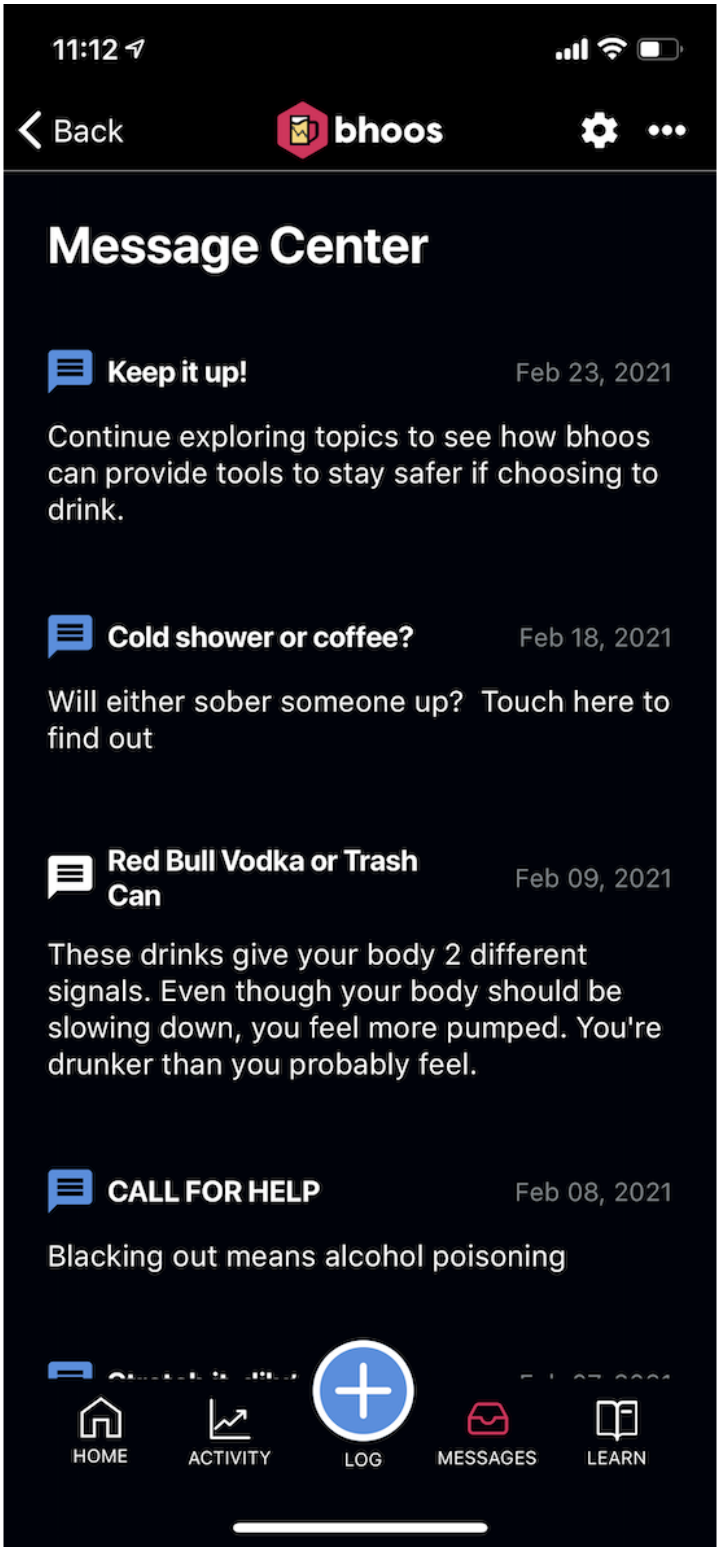


Figure 4. Message center containing student testimonials of the intervention (bhoos app).



Study Design

The study included a single 1-hour qualitative interview of each undergraduate student participant using standardized think-aloud methods [24], described below. This study was approved by the Social and Behavioral Science Institutional Review Board of the University of Virginia (protocol 3282).

Eligibility Criteria

To be eligible, a study applicant had to be (1) a student at the university, (2) able to read and speak English, and (3) aged ≥18 years. Participants reviewed an information form before completing any study activities and provided verbal consent that was audio recorded for documentation. The first round of interviews consisted of students who were not members of Greek organizations or student athletes. Participants in the second round included students who were fraternity and sorority

members and student athletes. These groups tested the prototype app separately because previous research has shown that student athletes or who are in Greek organizations are at a higher risk for dangerous drinking episodes than their non-Greek peers [25-29], and we sought to gather feedback from groups who may have different alcohol use patterns.

Procedures

Recruitment

Research coordinators and undergraduate research assistants (RAs) at the Center for Behavioral Health and Technology and the Center for Addiction and Prevention Research conducted the study. Participants were 10 university students aged 18 to 30 years. The sample was small because of disruption in recruitment as a result of the beginning of the COVID-19 pandemic. Participants were recruited from undergraduate and graduate classes, via flyers posted on bulletin boards in academic buildings around campus, via digital flyers posted on social media (eg, *GroupMe* and *Facebook* class pages), and by word-of-mouth referral from students working at the Center for Behavioral Health and Technology. The flyers contained a scannable QR code that linked directly to a Qualtrics interest form. The interest form collected applicant contact information, year in university, age, availability, and willingness to be contacted.

Think-Aloud Procedure

The *think-aloud* process is widely used for app development, as a common usability tool [24,30]. However, unlike a formal qualitative coding process that involves a thematic analysis to discover main findings [31], we used the think-aloud process to gather feedback on the app. Researchers asked participants to use a prototype mobile app, prompting them to use specific features while continuously speaking their thoughts and actions out loud [24]. Participants (N_{total}=10, $n_{in\text{-}person}$=5, n_{zoom}=5) reviewed an web-based mock-up of the app with hot spots that responded to taps to mimic app functionality. Participants used a QR code they scanned on their personal smartphones to access the app mock-up on *InVision*, a collaborative digital design platform [32]. Participants navigated through the prototyped app, speaking aloud what they were doing and noticing, and providing feedback about app features, navigation, and perceived utility as they used them. When necessary, researchers reminded study participants to maintain a running commentary of their thoughts and noted the student's navigation process as they used all aspects of the app. Researchers audio recorded the interviews and took thorough notes on a Word document while the participant was providing feedback. After completing the think-aloud procedure, participants received a US $10 Amazon gift card.

Modifications to the Protocol Due to COVID-19

The study was conducted in 2 rounds between March 3 and April 17, 2020. The first round of interviews occurred in person from March 3 to 6, 2020, that is, in early 2020, before the emerging COVID-19 pandemic required pausing in-person research. For in-person think-alouds, RAs watched the participants use the prototype while standing nearby and recorded their observations of student interactions and spoken thoughts of specific app pages on a note-taking sheet populated with the app screenshots. We developed a note-taking framework with screenshots of each page of the app grouped by app location (home, activity, log, message, and learn). Each screenshot provided space for think-aloud notes. RAs audio recorded the session using a handheld audio recorder or phone to supplement their written notes; when possible, >1 person in the room took notes or listened to the audio tape and took notes [30]. RAs reviewed the recordings after the trials and added further notes that were missing from the original notes to the summary document, which was finalized for review.

Between Rounds: App Changes Made in Response to Feedback

After the first round, the researchers debriefed the developers and investigators. Feedback showed that some labeling was unclear and users wanted more options for personalization and wanted colors and iconography to make data entry easier and more intuitive. Investigators and developers discussed which changes were needed immediately and prioritized changes (eg, adding and removing testimonials from the message center was prioritized lower than improving the labels and usability of the drink tracker). The development team fixed the functionality and clarity concerns reported in the round 1 trials.

Changed features included the eBAC, alternate behavior tracking, and calendar day view on the activity home screen. The eBAC was perceived as too small in the original mock-ups and was made more prominent on the dashboard. In addition, its colors were changed to correspond to which *zone* of drinking risk the user is in, based on the eBAC (eg, *golden zone*) [33]. Hydration was intended to be an alternative or secondary behavior that students could track, so they could track something other than alcoholic beverages. However, testing showed that the concept of hydration was confusing to users because they did not know if they should enter it once at the end of the day or update it multiple times throughout the day. The team switched to an alternate behavior that can be rated based on the previous full day, such as sleep quality. In addition, the calendar day view became more and more informative. For example, in the first version of the app, the day view showed only the number of drinks, but developers added the ability to record each drink, to compute a standard drink total, and to duplicate or mark drinks as a favorite and added the alternative behavior tracking feature to the calendar day view. Within the weekly and monthly views of behavioral metrics calculated by the app, graphs evolved to show metrics for the alternate behavior over those periods.

Students were sent home from the university in mid-March 2020, and in-person research was temporarily halted. To adapt to the situation, the second round of interviews between April 9 and 17, 2020, were conducted via Zoom video calls and used the second-round version of the app. The interviewer shared the screen, showing each of the app features one at a time. The participants installed the *InVision* mock-up on their personal device to enable navigation as if on an app and navigated simultaneously as the RA showed each screen. The note-taking and recording procedures were unchanged except for conducting them remotely. Information saturation was reached during the

second round of interviews. Second-round participants' comments and ideas were redundant with data from the first round. Major themes emerged naturally as the researchers summarized their notes. The research team determined that coding using qualitative software was not necessary.

Data Review and Analysis

Once all of the notes were captured, the RAs deleted the audio recording. Researchers reviewed the notes and presented the major findings to the app development team for discussion and identification of features and functions to be modified. Reporting was in aggregate only (ie, "5 of the 20 participants said x, while 12 said y and 3 said z"). The app development team continuously iterated and updated the look and feel of the mock-up based on the feedback. Data included audio recordings of the think-aloud protocol process, notes taken by researchers observing participants, and usability questionnaire responses. All participants were assigned a unique ID unrelated to their identity. Stored data were linked only to this unique ID. All data were stored behind the Health Insurance Portability and Accountability Act firewall of an institutional review board-approved School of Medicine highly secure data drive.

Results

Overview

Demographic characteristics of the overall study sample and by condition are presented in Table 1. Participants were mostly women (7/10, 70%) and aged 20 years (5/10, 50%), and 80% (8/10) were enrolled in the most populous school within the university, containing departments in the Arts and Sciences.

Table 1. Sociodemographic characteristics of participants.

Characteristics	Total (N=10), n (%)	Round 1 students (n=5), n (%)	Round 2 students (n=5), n (%)
Gender			
Male	3 (30)	2 (40)	1 (20)
Female	7 (70)	3 (60)	4 (80)
Age (years)			
18	1 (10)	1 (20)	0 (0)
19	2 (20)	0 (0)	2 (40)
20	5 (50)	3 (60)	2 (40)
21	1 (10)	0 (0)	1 (20)
22	1 (10)	1 (20)	0 (0)
Year in university			
1	3 (30)	1 (20)	2 (40)
2	3 (30)	1 (20)	2 (40)
3	3 (30)	2 (40)	1 (20)
4	1 (10)	1 (20)	0 (0)
School			
Arts and Sciences	8 (80)	4 (80)	4 (80)
Engineering	1 (10)	1 (20)	0 (0)
Undergraduate Business	1 (10)	0 (0)	1 (20)

Round 1

The feedback provided by the participants *in person* during round 1 is presented in Table 2. These students provided feedback on the eight essential features of the app: drink tracker, alternate behavior tracker (hydration or sleep), days tracked, messages, settings, eBAC calculator, learn topics, and a message center containing student testimonials. During the first round, RAs asked participants to focus on usability, rather than aesthetics, of the app. Participants expressed confusion about the purpose of the features and provided useful feedback about how they were displayed. For example, the drink tracker had drink choices listed that the students wanted to rename (eg, they recommended renaming *seltzer* or *cider*, to *hard seltzer* and *hard cider*, respectively, to ensure that future users would log alcoholic versions of these drinks) and indicated that such changes would increase utility and relevance. Participants reported that they would prefer having brand name examples (eg, *Corona*, *Coors Lite*, and *White Claw*) for each of the choices (ie, beer, liquor, seltzer, mixed drink, cider, and wine) to help clarify the different types of alcoholic drinks, again increasing utility and relevance as well as appeal. They reported that this would help future student users accurately track their drinks and that they believed many students do not know the alcohol by volume percentages for typical drinks.

Table 2. Round 1: summary of responses about features in prototype app (n=5).

Feature and feedback provided	Value, n (%)
Drink tracker	
Seltzer and *cider* renamed as *hard seltzer* and *hard cider*, respectively, to convey that these are alcoholic drinks	3 (60)
Want examples of each alcoholic drink (eg, for beer, to have options such as Corona, Modelo, and PBR[a]) not only for clarification but to select since not everyone will know the alcohol by volume percentage	3 (60)
Confusion around the green fill bars on the dashboard or home screen; unclear whether this was displaying a proportion of drinks consumed to a limit that was set	2 (40)
Confusion as to whether there's an option to track zero drinks	2 (40)
Option to log a *recent drink* for ease of use rather than having to tediously enter the information for the same drink again	2 (40)
Hydration tracker	
Thinks the metric is too subjective; the scale is too ambiguous (eg, not sure what the difference, for example, in ounces of water, that would be categorized as a *1* or *2*)	5 (100)
Feature not necessarily useful	2 (40)
Does not understand the purpose of this feature	1 (20)
Days tracked	
Metric unclear for measuring days tracked (eg, 10 days tracked since not tracking or 10 days since downloading the app?)	2 (40)
Display next to the drink tracker, rather than the hydration tracker	2 (40)
Add an option to share this with other app users to compare with friends	1 (20)
Messages	
Confusion around the *thumbs up or down* buttons in the header	4 (80)
Confused as to what this feature is entirely	4 (80)
Preference for having messages be set at the dock along the bottom with the home, activity, and learn buttons	1 (20)
Settings	
Profile data (ie, biological sex, age, and weight) accessible here	4 (80)
Option to decide types of messages are received here	3 (60)
Default setting have push notifications set to *off*	2 (40)
App FAQs[b] would fit better if relocated here	1 (20)
eBAC[c] calculator	
Display statistics of *current BAC*[d] and *maximum BAC level*	4 (80)
Remove *Time until legal limit (0.08%)* in case it might be misleading or encourage driving while intoxicated	2 (40)
Option to see an *Average BAC* or drinking trends over time	2 (40)
Should be the first thing you see when you open the app	2 (40)
Learn topics	
Information should be condensed so no scrolling is required	3 (60)
Confusion on the difference between this and other features in the app	3 (60)
Do not think that this will be a used	2 (60)
Message center with testimonials	
Pushing too much content with this feature; this is no different from previous features	3 (60)
Option to control the frequency of messages and notifications	3 (60)
Would like the *thumbs up or down* button moved here so they could indicate preference on types of messages and rate the content for future app updates	2 (40)

[a]PBR: Pabst Blue Ribbon.

[b]FAQ: frequently asked question.

[c]eBAC: estimated blood alcohol content.

[d]BAC: blood alcohol content.

Participants also provided feedback on whether they were interested in and would use certain features. For example, students were not interested in a hydration tracker feature and considered its utility and appeal low. The hydration tracker asks students to rate how hydrated they are on a scale of 1 to 5. Participants expressed confusion about the hydration tracking as it was not as specific as the (alcoholic) drink tracker. Students questioned how the hydration tracker related to the app, failing to see any connection between hydration and drinking alcohol. Thus, the hydration tracker showed low relevance.

Round 2

Participants in round 2 met with the RAs on the web using video meeting technology, but other than this, all other procedures remained similar. In the second round of think-aloud testing, RAs guided participants to focus on the look and feel (eg, aesthetics and design) more than the functionality of the app. For example, the app's default setting is in *dark mode*, with a black background; a participant expressed preference for a *light mode* option to increase its appeal. In addition, participants expressed a preference for symbols for sex rather than the terms *male* and *female* for personal data entry screens. Participant feedback in round 2 is presented in Table 3. Overall, participants expressed interest in several of the features of the app, but not in others. The look and feel of the app were reported to be important to the students. Student navigation across app features indicated that the improved look and feel helped guide them to the features they were most interested in using.

Table 3. Round 2: look and feel and features in prototype app (n=5).

Feature and feedback provided	Value, n (%)
Drink tracker	
Likes the log button and is easily accessible	3 (60)
Likes that there is a *recently added* shortcut for logging recent drinks	2 (40)
Does not think that students will use the feature; thinks that students would forget to log drinks on a night out	1 (20)
Sleep tracker	
Does not understand the connection or link between sleep and drinking	5 (100)
Does not believe the feature be used	2 (40)
Days tracked	
Do not believe that this feature is necessary	2 (40)
Rather than days tracked, would prefer a tracking streak of days not drinking	1 (20)
Messages	
Preference for having messages be set at the dock along the bottom with the home, activity, and learn buttons	2 (40)
Likes the *unread* symbol that alerts the user to open messages	1 (20)
Settings	
Option for *light mode* or a white interface	1 (20)
Option for specific *push* notification	1 (20)
Option to share *data* with friends to boost engagement	1 (20)
eBAC[a] calculator	
Increase in font size for the BAC[b] statistics	2 (40)
Prefer the sex symbols to change to words on home screen	1 (20)
Learn topics	
Believed this section to be visually appealing	2 (40)
Add images to help visualize the concept of *standard drinks*	1 (20)
Message center with testimonials	
Seems to be helpful	3 (60)
Would prefer graduating class over age (eg, Jessica, Class of 2021, UVA[c] activities, instead of—Jessica, 21, UVA activities)	2 (40)
Testimonials are too long and too dense; would prefer bullet points over long sentences	2 (40)

[a]eBAC: estimated blood alcohol content.

[b]BAC: blood alcohol content.

[c]UVA: University of Virginia.

Discussion

Principal Findings

The objective of this study is to assess the appeal, relevance, and perceived utility of a draft mobile app for college students that could intervene at or around the time of alcohol use. We sought to determine if students would want and use the app and if it would be perceived as useful and engaging. Participants reported that minor changes to the functionality of the initial versions of the app would improve the usability of the app (eg, having examples of each alcoholic drink for clarification and stating whether they believe certain features would be used or not). In addition, participants reported that minor changes to the aesthetics would increase the usability (eg, incorporating the logo colors throughout the app to make it more vibrant and having an option to turn off push notifications). The changes were made between rounds of think-aloud user testing to enact modifications based on user feedback.

This process was an iterative step within a series of formative studies to create and improve an app to reduce college binge drinking. Feedback provided by college students assisted the investigators and development team in designing features that were useful and appealing. The purpose was to test usability, and student feedback helped the team modify the app so that it would perform as students hoped in *real-life* scenarios to address binge drinking and alcohol consumption. This process was designed as a feedback loop beginning with the *science* that drafted features about learning what a standard drink is,

self-monitoring tools with immediate and longer-term feedback about risks when drinking, options to get help in an emergency, and vetted information about sex differences in alcohol processing in the body. The features were conceptualized and then mocked up by developer team members with expertise in user experience and user interface design). Finally, the team obtained qualitative feedback on relevance and usability from students regarding the aesthetics and appeal of the app and its ability to engage them.

Limitations: Modifications Due to COVID-19

Although the practical insights from students about the app proved valuable and guided us to make important app modifications, there are several limitations to be considered. Because of COVID-19, students had to participate in the second round of think-aloud trials virtually over Zoom using an adapted protocol. Although participants navigated through the app with ease, the researchers could not always see the participants' phone screen and sometimes had difficulty determining which features the participants were commenting on. RAs asked what feature the participant was viewing and where in the app they were navigating to and prompted the participant to turn their phone screen toward the camera. These moments of clarification seemed to interrupt the natural rhythm and flow of the think-aloud format and may have resulted in reduced information. In person, the researcher could easily see the participant's phone screen without prompting, which allowed for the conversation to flow uninterrupted while the RA could take notes more inconspicuously. During round 1, the participants' comments were prominent in the conversation, whereas researchers listened and asked fewer questions. Conversely, researchers asked more questions during round 2, whereas participants provided less feedback, and their comments were either positive or more superficial in nature.

Previous studies have used apps to help reduce dangerous drinking episodes among college students. However, these apps did not have or include features that tailor to individual personalized feedback such as logging drinks and providing an eBAC [34-36]. One of these previous interventions has justified the use of a digital platform as users did indicate that it offered help and was easy to use [36]. Other alcohol app interventions have reported to significantly reduce drinking [34,35].

Future Research

Data and conclusions from this study will be synthesized with information gained from other qualitative and quantitative formative work, to determine the final features to be tested. We plan a pilot randomized trial to assess the use and impact of the app on drinking behavior over several months; we plan to conduct a prospective brief observational and feasibility study with about 255 students to (1) measure student demographics, attitudes, experiences, and behaviors related to drinking; (2) evaluate the acceptability, feasibility, relevance, and actual use of the app over the intervention period among students; (3) measure student attitudes, experiences, and behaviors related to drinking in a final survey; (4) determine the usability of the app and desired improvements using short interviews after about 4 weeks of app use; and (5) test the impact of incentives on app use to provide crucial information for implementation of a larger subsequent trial. If the pilot of this mobile app intervention shows acceptable use and reporting of frequency and volume of binge drinking, we will consider testing it further for impact on drinking behaviors.

Conclusions

Student feedback clearly indicated preferences for the eBAC calculator and days tracked features. In contrast, student feedback showed a dislike of or lower interest in using the water or hydration tracker, learn topics, and messages or testimonials. On the basis of student ratings of appeal, relevance, and utility, we revised the app, and the finalized app included the drink tracker, days tracked, eBAC calculator with an eBAC level, sleep quality tracker, learn topics, message center, and testimonials. By asking student participants to navigate through, and comment on, the app's features using standard think-aloud procedures, the researchers gained real time perspective on the appeal, relevance, and utility of the app that could be translated into a refined version of the app.

Acknowledgments

The authors thank the Board of Visitors and the President of the University of Virginia for funding this project through the Strategic Investment Fund. The authors also appreciate the support and guidance of the Executive Committee for the project "Reducing the Burden of Addictions in Virginia": Dr Robert Klesges, Dr Lee Ritterband, Dr Nassima Ait-Daoud Tiouririne, Dr Karen Ingersoll, Dr Christopher Holstege, and Dr Margaret Shupnik. Finally, the authors thank the undergraduate research assistants for this study: Grace Biegger, Aasritha Natarajan, David Vann, Jessica G Smith, and Shruti Vachhani. We would also like to thank our development team, Gabe Heath, Nicole Le, and Brendan Murphey, for the design and development of our InVision mock-ups.

Conflicts of Interest

None declared.

References

1. College drinking. National Institute on Alcohol Abuse and Alcoholism. URL: https://www.niaaa.nih.gov/publications/brochures-and-fact-sheets/college-drinking [accessed 2022-02-05]

2. Hingson R, Zha W, Smyth D. Magnitude and trends in heavy episodic drinking, alcohol-impaired driving, and alcohol-related mortality and overdose hospitalizations among emerging adults of college ages 18-24 in the United States, 1998-2014. J Stud Alcohol Drugs 2017 Jul;78(4):540-548 [FREE Full text] [doi: 10.15288/jsad.2017.78.540] [Medline: 28728636]
3. Underage drinking. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/alcohol/fact-sheets/underage-drinking.htm [accessed 2022-02-05]
4. Hasin DS, Stinson FS, Ogburn E, Grant BF. Prevalence, correlates, disability, and comorbidity of DSM-IV alcohol abuse and dependence in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. Arch Gen Psychiatry 2007 Jul 01;64(7):830-842. [doi: 10.1001/archpsyc.64.7.830] [Medline: 17606817]
5. Grant BF, Dawson DA. Age at onset of alcohol use and its association with DSM-IV alcohol abuse and dependence: results from the national longitudinal alcohol epidemiologic survey. J Substance Abuse 1997 Jan;9:103-110. [doi: 10.1016/s0899-3289(97)90009-2]
6. Dejong W, Larimer ME, Wood MD, Hartman R. NIAAA's rapid response to college drinking problems initiative: reinforcing the use of evidence-based approaches in college alcohol prevention. J Stud Alcohol Drugs Suppl 2009 Jul(16):5-11 [FREE Full text] [doi: 10.15288/jsads.2009.s16.5] [Medline: 19538907]
7. Brief Alcohol Screening and Intervention of College Students (BASICS). Youth.gov. URL: https://youth.gov/content/brief-alcohol-screening-and-intervention-college-students-basics [accessed 2022-02-05]
8. College drinking, changing the culture homepage. College drinking, changing the culture. URL: https://www.collegedrinkingprevention.gov/ [accessed 2022-02-05]
9. Baer JS, Kivlahan DR, Blume AW, McKnight P, Marlatt GA. Brief intervention for heavy-drinking college students: 4-year follow-up and natural history. Am J Public Health 2001 Aug;91(8):1310-1316. [doi: 10.2105/ajph.91.8.1310] [Medline: 11499124]
10. Cole HA, Prassel HB, Carlson CR. A meta-analysis of computer-delivered drinking interventions for college students: a comprehensive review of studies from 2010 to 2016. J Stud Alcohol Drugs 2018 Sep;79(5):686-696. [doi: 10.15288/jsad.2018.79.686]
11. Huh D, Mun E, Larimer ME, White HR, Ray AE, Rhew IC, et al. Brief motivational interventions for college student drinking may not be as powerful as we think: an individual participant-level data meta-analysis. Alcohol Clin Exp Res 2015 May;39(5):919-931 [FREE Full text] [doi: 10.1111/acer.12714] [Medline: 25872599]
12. Krieger H, Young CM, Anthenien AM, Neighbors C. The epidemiology of binge drinking among college-age individuals in the United States. Alcohol Res 2018;39(1):23-30 [FREE Full text] [Medline: 30557145]
13. Mitka M. College binge drinking still on the rise. JAMA 2009 Aug 26;302(8):836-837. [doi: 10.1001/jama.2009.1154] [Medline: 19706853]
14. Milward J, Deluca P, Drummond C, Watson R, Dunne J, Kimergård A. Usability testing of the BRANCH smartphone app designed to reduce harmful drinking in young adults. JMIR Mhealth Uhealth 2017 Aug 08;5(8):e109 [FREE Full text] [doi: 10.2196/mhealth.7836] [Medline: 28790022]
15. Carrà G, Crocamo C, Bartoli F, Carretta D, Schivalocchi A, Bebbington PE, et al. Impact of a mobile e-health intervention on binge drinking in young people: the digital-alcohol risk alertness notifying network for adolescents and young adults project. J Adolesc Health 2016 May;58(5):520-526. [doi: 10.1016/j.jadohealth.2016.01.008] [Medline: 27107907]
16. Haug S, Schaub MP, Venzin V, Meyer C, John U, Gmel G. A pre-post study on the appropriateness and effectiveness of a web- and text messaging-based intervention to reduce problem drinking in emerging adults. J Med Internet Res 2013 Sep 02;15(9):e196 [FREE Full text] [doi: 10.2196/jmir.2755] [Medline: 23999406]
17. Cirillo M, Halbert J, Smith JG, Alamiri NS, Ingersoll K. #BingeDrinking: using social media to understand college binge drinking. JMIR 2022 (forthcoming) [FREE Full text] [doi: 10.2196/preprints.36239]
18. Cho H. Influences of self-monitoring and goal-setting on drinking refusal self-efficacy and drinking behavior. Alcoholism Treatment Q 2008 Oct 02;25(3):53-65. [doi: 10.1300/j020v25n03_05]
19. Crane D, Garnett C, Michie S, West R, Brown J. A smartphone app to reduce excessive alcohol consumption: identifying the effectiveness of intervention components in a factorial randomised control trial. Sci Rep 2018 Mar 12;8(1):4384 [FREE Full text] [doi: 10.1038/s41598-018-22420-8] [Medline: 29531280]
20. How to rapidly grow engagement with your new app. App Samurai. URL: https://appsamurai.com/how-to-rapidly-grow-engagement-with-your-new-app/ [accessed 2022-02-05]
21. Covetus. URL: https://www.covetus.com/blog/5-practical-ways-to-boost-mobile-app-engagement-user-retention-rate [accessed 2022-02-05]
22. Clearbridge Mobile. URL: https://clearbridgemobile.com/5-methods-for-increasing-app-engagement-user-retention/ [accessed 2022-01-07]
23. Molina MD, Sundar SS. Can mobile apps motivate fitness tracking? A study of technological affordances and workout behaviors. Health Commun 2020 Jan;35(1):65-74. [doi: 10.1080/10410236.2018.1536961] [Medline: 30358424]
24. Thinking aloud: the #1 usability tool. Nielsen Norman Group. URL: https://www.nngroup.com/articles/thinking-aloud-the-1-usability-tool/ [accessed 2022-02-05]
25. Study explores the relationship between drinking and fraternities. American Psychological Association. URL: https://www.apa.org/monitor/jun01/fraternities [accessed 2022-02-05]

26. Borsari B, Hustad J, Capone C. Alcohol use in the Greek system, 1999-2009: a decade of progress. Curr Drug Abuse Rev 2009 Sep;2(3):216-255 [FREE Full text] [doi: 10.2174/1874473710902030216] [Medline: 20443768]
27. Huchting K, Lac A, Hummer J, LaBrie J. Comparing Greek-affiliated students and student athletes: an examination of the behavior-intention link, reasons for drinking, and alcohol-related consequences. J Alcohol Drug Educ 2011 Dec 01;55(3):61-81 [FREE Full text] [Medline: 25698846]
28. Miley B, Chalk H. Fitting the mold: alcohol use and body image disturbances in athletes and greek-affiliated undergraduates. Mod Psychol Stud 2015;21(1):40-49 [FREE Full text]
29. Relationship between high-risk drinking and team cohesion among college athletes. ProQuest. URL: https://www.proquest.com/docview/1696935533?pq-origsite=gscholar&fromopenview=true [accessed 2022-02-05]
30. Group notetaking for user research. Nielsen Norman Group. URL: https://www.nngroup.com/articles/group-notetaking/ [accessed 2022-02-05]
31. Thematic analysis of qualitative user research data. Nielsen Norman Group. URL: https://www.nngroup.com/videos/thematic-analysis-qualitative-user-research-data/ [accessed 2022-02-05]
32. Transformative collaboration for all the work you do. InVision. URL: https://www.invisionapp.com [accessed 2022-02-05]
33. Office of Health Promotion /Department of Student Health /University of Virginia. BAC to basics. Stall Seat J 2019;18(3) [FREE Full text]
34. Tofighi B, Chemi C, Ruiz-Valcarcel J, Hein P, Hu L. Smartphone apps targeting alcohol and illicit substance use: systematic search in in commercial app stores and critical content analysis. JMIR Mhealth Uhealth 2019 Apr 22;7(4):e11831 [FREE Full text] [doi: 10.2196/11831] [Medline: 31008713]
35. Fucito LM, DeMartini KS, Hanrahan TH, Yaggi HK, Heffern C, Redeker NS. Using sleep interventions to engage and treat heavy-drinking college students: a randomized pilot study. Alcohol Clin Exp Res 2017 Apr;41(4):798-809 [FREE Full text] [doi: 10.1111/acer.13342] [Medline: 28118486]
36. Crane D, Garnett C, Brown J, West R, Michie S. Factors influencing usability of a smartphone app to reduce excessive alcohol consumption: think aloud and interview studies. Front Public Health 2017;5:39 [FREE Full text] [doi: 10.3389/fpubh.2017.00039] [Medline: 28421175]

Abbreviations

eBAC: estimated blood alcohol content
RA: research assistant

Edited by T Leung; submitted 06.08.21; peer-reviewed by V Buss, K Schroeder; comments to author 05.01.22; revised version received 13.01.22; accepted 14.01.22; published 17.02.22.

Please cite as:
Smith JG, Alamiri NS, Biegger G, Frederick C, Halbert JP, Ingersoll KS
Think-Aloud Testing of a Novel Safer Drinking App for College Students During COVID-19: Usability Study
JMIR Form Res 2022;6(2):e32716
URL: https://formative.jmir.org/2022/2/e32716
doi:10.2196/32716
PMID:35175210

Original Paper

The Relation of Attitude Toward Technology and Mastery Experience After an App-Guided Physical Exercise Intervention: Randomized Crossover Trial

Kai Sassenberg[1,2], Dr rer nat; Inka Roesel[3,4], MSc; Gorden Sudeck[5,6], Dr Phil; Katharina Bernecker[7], Dr Phil; Jennifer Durst[3,6], MA; Inga Krauss[3,6], Dr rer soc

[1]Social Processes Lab, Leibniz-Institut für Wissensmedien, Tübingen, Germany

[2]School of Science, University of Tübingen, Tübingen, Germany

[3]Department of Sports Medicine, University Hospital Tübingen, University of Tübingen, Tübingen, Germany

[4]Institute for Clinical Epidemiology and Applied Biostatistics, University of Tübingen, Tübingen, Germany

[5]Institute of Sports Science, University of Tübingen, Tübingen, Germany

[6]Interfaculty Research Institute for Sports and Physical Activity, University of Tübingen, Tübingen, Germany

[7]Department of Psychology, University of Zurich, Zürich, Switzerland

Corresponding Author:
Kai Sassenberg, Dr rer nat
Social Processes Lab
Leibniz-Institut für Wissensmedien
Schleichstrasse 6
Tübingen, 72076
Germany
Phone: 49 7071 979 220
Fax: 49 7071 979 200
Email: k.sassenberg@iwm-tuebingen.de

Abstract

Background: Physical exercise has been found to assert a positive impact on many muscular conditions. Exercise under face-to-face supervision is the gold standard, but access to it is limited, for instance, for economic reasons. App-guided therapy is an intervention that is more affordable and easily accessible. However, attitude toward technology is a key predictor for media adoption and is therefore expected to shape user experience during app-guided therapy. This might be of particular importance for mastery experience, which is crucial for promoting exercise-related self-efficacy and perceived usefulness of the interaction. Both should empower patients to continuously exercise.

Objective: This study sought to test whether attitudes toward technology predict mastery experience and perceived usefulness of the interaction after an app- versus a physiotherapist-guided treatment. We expect that attitudes toward technology positively predict both outcomes in case of the app-guided but not in case of the physiotherapist-guided treatment.

Methods: Patients (n=54) with clinically diagnosed hip osteoarthritis participated in 2 training sessions with the same exercise intervention, once guided by an app on a tablet computer and once guided by a physiotherapist in a German university hospital. The order of the sessions was randomized. Attitude toward technology was assessed as predictor before the first session, while mastery experience and the global perceived usefulness of interaction as self-reported outcomes after each session.

Results: In line with our hypotheses, attitude toward technology predicted mastery experience (b=0.16, standard error=0.07, *P*=.02) and usefulness of interaction (b=0.17, standard error=0.06, *P*=.01) after the app-based training but not after the training delivered by a physiotherapist (*P*>.3 in all cases). Mastery experience was lower for the app-based training but reached a very similar level as the physiotherapist-guided training for those holding a very positive attitude toward technology.

Conclusions: The attitude toward technology predicts the extent of mastery experience after app-guided exercise therapy. As mastery experience is highly important for self-efficacy and future exercise behavior, attitudes toward technology should be considered when delivering app-guided exercise treatments.

Trial Registration: German Clinical Trials Register DRKS00015759; https://www.drks.de/drks_web/navigate.do?navigationId=trial.HTML&TRIAL_ID=DRKS00015759

(JMIR Form Res 2022;6(2):e28913) doi:10.2196/28913

KEYWORDS

mobile app; exercise; mastery experience; self-efficacy; attitudes toward technology; osteoarthritis

Introduction

Background

According to the World Health Organization (WHO), musculoskeletal conditions are the leading contributor to disability worldwide [1]. Their prevalence increases across the lifetime. For many of these conditions such as osteoarthritis (OA), physical activity and exercise contribute to the reduction of symptoms [2]. Therefore, physical exercise is recommended in many treatment guidelines [3-6]. Unfortunately, many patients do not follow these guidelines [7]. A reason among others is the fear of deterioration of their symptoms due to an incorrect execution of the exercises [8,9]. In other words, low self-efficacy is a major barrier for physical exercise.

This barrier can be overcome by supervision [10,11]. Exercises should initially be instructed by a health or exercise professional [5]. An alternative cost-efficient means to provide guidance regarding physical exercise is via digital apps on tablet computers. As tablets are highly mobile, they can be conveniently used in locations that allow to exercise [12-16]. They have a screen of sufficient size for video-based instructions and most importantly, older people are more likely to use tablets than smartphones [17]. However, the attitude toward digital technologies and with it the willingness of older adults to adopt digital technology for health purposes vary [18,19].

Therefore, this study sought to compare an app-based intervention with the gold standard of an intervention supervised by a physiotherapist. We tested the effect of both treatments on *mastery experience*, which is known to facilitate exercise-related self-efficacy and, in turn, continuous exercising [20]. In addition, we studied the perceived *usefulness of the interaction*, a key predictor of the attitudes toward the intervention which should also be related to continuous exercise [21]. To conduct a fair test, both interventions rely on the same evidence-based exercise intervention for patients with hip OA [22,23]. To do justice to the older age of the target group of this intervention and the varying acceptance of technology-based health interventions in that group, we considered attitudes toward technology as an additional predictor.

Theoretical Underpinning of (Digital) Exercise Interventions

The model of physical activity–related health competence (PAHCO) [24,25] guided the development of the examined app as well as this study. The core idea of this health education model is that exercising in a health-effective and low-risk manner requires a set of competences, namely, movement competence, control competence, and self-regulation competence. *Movement competence* includes motor abilities and skills as well as movement and body awareness. To train this competence, we ensured that our app provided detailed instructions regarding movement as well as body signals and allowed for the repeated viewing of videos until instructions were well understood. *Control competence* requires activity-related knowledge and the ability to perceive and interpret body signals (eg, to sense muscle soreness and adjust exercise intensity based on it). To direct users' attention to this aspect, the app contained questions about pain and intensity after each exercise and provided feedback on how to adapt the exercise to ensure optimal dose parameters.

Finally, *self-regulation competence* summarizes motivational and volitional determinants of regular exercise including self-efficacy, which refers to the feeling that exercise can be executed independently and, thus, key for its uptake [26]. Self-efficacy is developed through the experience that the exercise session empowers the user to execute the exercises effectively, called *mastery experience* [20,27]. Mastery experience can refer to the movement-related demands (which relates to movement competence), the self-directed control of physical loads (which relates to control competence), or—most relevant in the current context—the app- or physiotherapist-guided exercise instructions. Given that mastery experience is decisive for the adoption of regular exercise [26,28] and the use of digital devices more generally [29], we focused particularly on this indicator in this study.

Attitudes Toward Technology Among Older Adults

Attitudes toward technology and acceptance of eHealth vary substantially among older adults [30,31]. Research has demonstrated that the general attitude toward technology relate positively to the judgment of health-related technologies including (1) the perceived usefulness and (2) the self-efficacy regarding the use of the specific technology [32]. This suggests that attitudes toward technology might relate to perceived usefulness and self-efficacy, because the attitudes color the experience during technology use, including perceived usefulness of the interaction and mastery experience. Given that perceived usefulness and mastery experience both contribute substantially to the adoption of the technology [21,33] and, thus, in the current context to health behavior, knowledge about the relation between these variables is highly relevant. At the same time, there is no reason to assume that attitudes toward technology predict the usefulness of the interaction and the mastery experience in the context of interventions delivered by a human instructor. Accordingly, we hypothesized:

> *Attitudes toward technology is positively related to (1) the mastery experience and (2) the perceived usefulness of the interaction regarding an app-guided treatment but not regarding a treatment delivered by a human instructor.*

The Current Research

These hypotheses were tested using the data collected in a larger training study, parts of which have been reported by Durst et al [34]. In this experimental study, patients with hip OA received the same evidence-based exercise intervention [22,23] once

delivered by an app on a tablet computer and once by a physiotherapist with the order of sessions being randomized between participants. Attitudes toward technology, mastery experience, and usefulness of the interaction were assessed after both sessions.

Methods

Design and Participants

Parts of this section correspond to those of a previous publication on this study [34] given that both articles describe the same study. We conducted a randomized crossover trial (see Multimedia Appendix 1 for the CONSORT checklist) with a 2 (treatment: app vs physiotherapist—within participants) × 2 (sequence—between participants) design. The attitude toward technology was assessed as additional continuous predictor. Participants were randomly assigned in a 1:1 allocation ratio to the 2 exercise treatment sequences. Randomization was based on a list generation with an online tool [35].

The AP (app–physiotherapist) group first had a training session using a tablet computer–based app and later a second session supervised by a physiotherapist, whereas the PA (physiotherapist–app) group was supervised by the physiotherapist in the first session and had the app-based training in the second session. For each participant the 2 intervention sessions were scheduled 4-6 weeks apart to allow for a sufficiently strong washout of treatment effects. The analyses reported by Durst et al [34] show that washout was only partly successful regarding movement competence. Therefore, we include sequence as a factor in the analyses reported below. Ethical approval for this study was obtained from the Ethical Committee of Tuebingen University Hospital. The study was registered in the German Clinical Trials Register (DRKS00015759). This preregistration did not include the hypothesis tested here.

Participants with diagnosed hip OA were recruited via advertisements in regional newspapers, by an email sent out via the employee list-serve of the University of Tuebingen and the Tuebingen University Hospital, and via flyers distributed by orthopedic surgeons and physiotherapists. In a telephone call interested individuals were screened for eligibility (for exclusion criteria, see Textbox 1). Eligible people were then randomly allocated to 1 of the 2 treatment sequences (determined by the next free slot in the randomization list) and informed about (1) the positive effects of exercise therapy for hip OA, (2) the details of the treatment, and (3) the research questions. Finally, the 2 treatment sessions at the Tuebingen University Hospital were scheduled.

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

1. 50 years and older
2. Self-reported lifetime prevalence of hip osteoarthritis diagnosed by a medical practitioner
3. Informed consent to study participation

Exclusion criteria

1. Comorbidities leading to major impairments in everyday life and representing contraindications for physical activities
2. Self-reported acute illness
3. Significantly established osteoporosis requiring treatment, previous spontaneous or low-impact fracture
4. Musculoskeletal surgery at the lower extremity within the last 3 months
5. Regular use of gait aids (eg, walker, crutch)
6. Insufficient German language skills for self-administered questionnaires
7. Previous experience from hip exercise groups

In case of an artificial joint replacement at the other hip or the knee joints:

1. Artificial joint replacement at the knee or hip joint or both within the last 6 months, with unstable anchoring or with known radiological signs of implant loosening
2. Current pain at rest or with activity due to artificial joint replacement
3. Luxation as an adverse event of artificial hip replacement
4. Acute joint inflammation at the knee or hip joint or both

Trial Interventions

Overview

The interventions (physiotherapist and app) used in this study were extracted from an evidence-based 12-week exercise program that was specifically designed for patients with hip OA [23,36,37]. Four exemplary exercises and their instructions were selected from this program. Both types of training sessions lasted 45-60 minutes. Participants were asked to report perceived exertion and OA-related pain after each set using a 10-point Likert scale.

Physiotherapist-Guided Exercises

The physiotherapist had 5 years of work experience. She introduced the exercises, corrected deficient or improper execution, and asked to adjust the exercise to the planned level of intensity, and in case of increasing pain according to the used target values of physical exhaustion and pain that had been implemented in the algorithm of the app to modify exercise intensity instructions. The physiotherapist also adapted the intensity level for the participant on an individual basis, as applicable.

App-Guided Exercises

The app was designed for and presented on a 9.7-in. (24.64 cm) tablet computer, which was mounted on a holder in a convenient position. All instructions were given on the tablet, after the app had been started by the experimenter. In line with the PAHCO model, the app supports practical exercises, cognitive and motor learning, and the processing of personal experience with movement [38]. The app consists of 5 components: (1) technical introduction, (2) creation of an individual user profile, (3) pedagogical agent, (4) exercise introductions, and (5) feedback-based dose adjustments and further instructions. Videos and acoustic signals are implemented in the software to guide the different exercises and to support the participant during the exercises. The videos combine long shots and close-ups based on interviews in a pretest. In addition, the camera's perspective and the choice of actors were optimized based on the results of the pretest to render the starting position and the movements easily visible. Movement speeds for exercise repetition are set using an auditory signal and visually supported by the actor in the video. For details about the elements and the algorithms of the app, see Multimedia Appendix of Durst et al [34].

Measures

Sample Characteristics

Sociodemographic, anthropometric, personal, OA-related variables, and additional measures unrelated to the current research question were assessed before the first training session.

Attitudes Toward Technology

A validated 19-item scale for attitudes toward technology (German: "Technikaffinität" TA-EG [39]) was presented before the first training session. Participants had to indicate their agreement to each item (eg, "I enjoy trying out electronic devices"; α=.83) on a 5-point scale (1=*does not apply at all*, 5=*exactly applies*). Ratings were averaged and summarized in one index by averaging the values after recoding negatively worded items. Higher values indicate a more positive attitude toward technology (see Table 1 for descriptive statistics).

Table 1. Baseline data for the complete sample differentiated according to treatment sequence.

Characteristics	Total (n=54)	PA[a] (n=26)	AP[b] (n=28)	*P* value
Age (years), mean (SD)	62.4 (8.2)	62.5 (8.0)	62.3 (8.5)	.91
Gender				.74
Female, n (%)	32 (59)	16 (62)	16 (57)	
Male, n (%)	22 (41)	10 (39)	12 (43)	
Education				.19
Academic education, n (%)	22 (41)	8 (31)	14 (50)	
Vocational education, n (%)	31 (57)	18 (69)	13 (46)	
No vocational education, n (%)	1 (2)	0 (0)	1 (4)	
Work situation				.44
Employed, n (%)	32 (59)	14 (54)	18 (64)	
Retired, n (%)	22 (41)	12 (46)	10 (36)	
Experience with exercise groups (1-5), median (IQR)	3.00 (1.0)	3.00 (1.0)	3.00 (2.0)	.31
Daily everyday activity (minutes of cycling and walking/week), median (IQR)	215 (360)	215 (330)	225 (458)	.49
Sports activity (minutes/week), median (IQR)	209 (273)	229 (309)	184 (308)	.26
Attitudes toward technology, median (IQR)	3.16 (0.5)	3.13 (0.5)	3.20 (0.5)	.63

[a]PA: physiotherapist–app.

[b]AP: app–physiotherapist.

Mastery Experience and Perceived Usefulness of the Interaction

Mastery experience and perceived usefulness of the interaction were each measured once for the physiotherapist and once for the app. Four items were used to assess the *mastery experience* regarding the exercise after each session, of which 2 were adopted from the subscale *Competence* of the Need Satisfaction in Exercise Scale [40] (eg, "I had the impression that I was executing the exercise effectively"; internal consistency: α_{APP}=.88; α_{PHYSIO}=.67). Four additional items assessing the *usefulness of the interaction* were self-developed and 1 was adopted from the usability measure by Harder et al [41] (eg, "The instructions were helpful"; α_{APP}=.85; α_{PHYSIO}=.54). The

internal consistency for the usefulness of the interaction with the physiotherapist was not satisfying. Given that it could not be improved by dropping an item and that we aimed at parallel measures for both interventions, we did use the scale nonetheless. Both scales used a 4-point scale (1=*does not apply at all*, 4=*exactly applies*; see Table 2 for descriptive statistics). For all items, see Multimedia Appendix 2.

Table 2. Mean (SD) of mastery experience and usefulness of interaction by sequence and treatment.

Measure	Total (n=54)	PA[a] (n=26)	AP[b] (n=28)
Usefulness of interaction			
Physio (n=49)	3.84 (0.24)	3.84 (0.27)	3.84 (0.22)
App (n=51)	3.32 (0.68)	3.43 (0.57)	3.22 (0.75)
Mastery experience			
Physio (n=49)	3.51 (0.32)	3.55 (0.27)	3.47 (0.36)
App (n=51)	3.16 (0.56)	3.34 (0.40)	3.00 (0.64)

[a]PA: physiotherapist–app.

[b]AP: app–physiotherapist.

Sample Size

We planned to collect data from at least 40 participants.

Statistical Analysis

Participant characteristics are summarized for the whole sample and for the 2 sequence conditions (Table 1). We tested for differences between sequence conditions using Pearson chi-square test for categorical data, independent Student *t* test for indices from rating scales, or Mann–Whitney *U* test. The latter was used if the assumption of normally distributed data was violated.

The main hypothesis was tested using a linear mixed design analysis of variance (mixed analysis of variance) with participant as random factor (nested within sequence of treatment order) and treatment (P and A), sequence (PA and AP), period (T1 and T2), and attitude toward technology (mean centered) as well as their (2- and 3-way) interactions as fixed factors separately for mastery experience and usefulness of interaction. Effects of interactions were resolved using simple slope analyses. We report the results based on analyses assuming normal distribution of the variables. In cases where this assumption was violated, we repeated the analyses after normalization of scores and these scores were used for further data analysis. Results of both analyses were virtually identical.

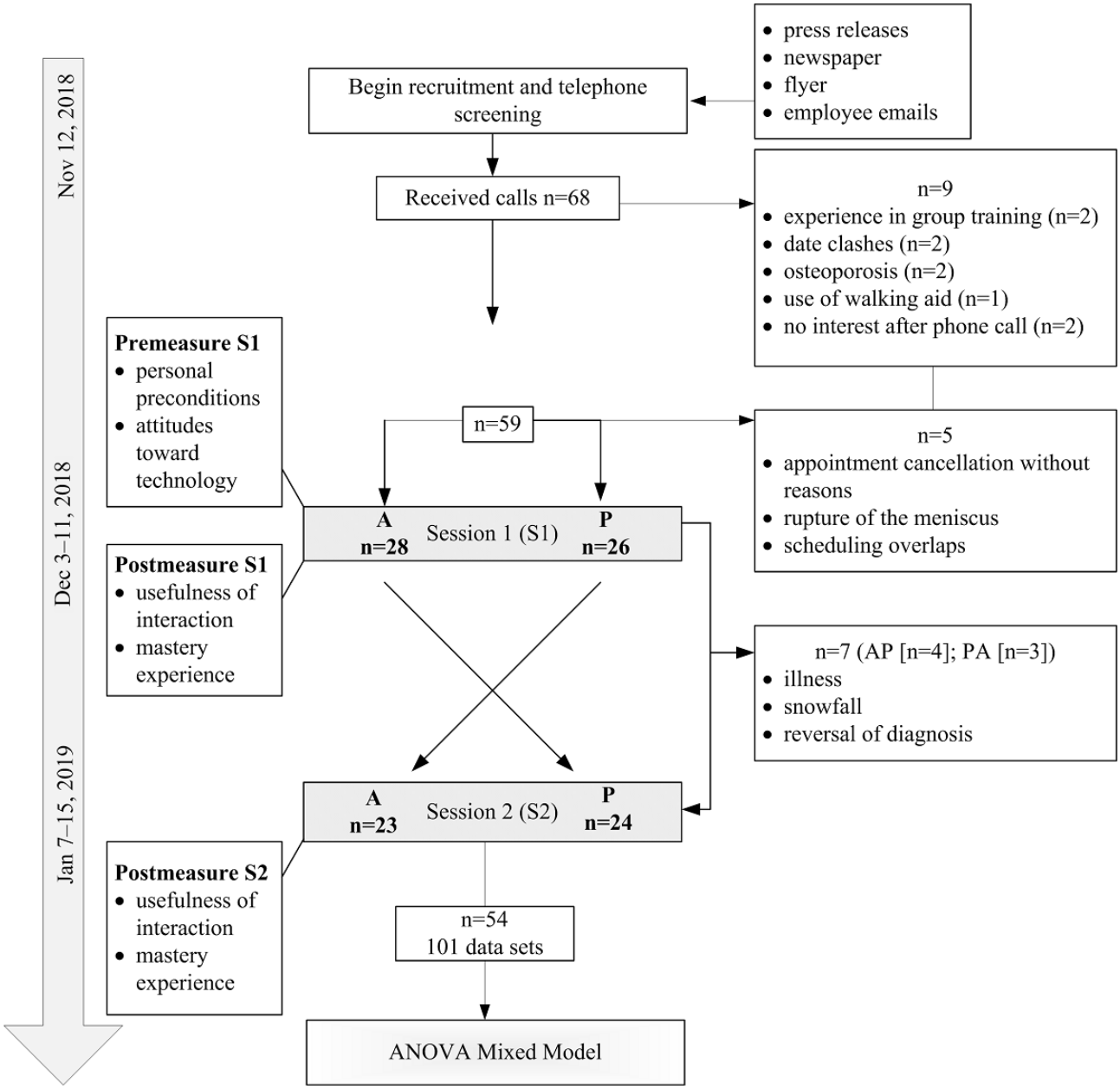
The level of statistical significance was set at the conventional level of α=.05. All data were analyzed using SPSS version 25 (IBM) and R version 3.6.1 (The R Foundation).

Results

Participants

Among 68 people, 59 fulfilled our inclusion criteria, contacted the study staff, and made an appointment. Five individuals canceled the first training appointment. Of the remaining 54 participants who completed the first training session, 7 could not attend the second session. One participant did not provide the ratings of the physiotherapist in the first session. Therefore, this case drops out of all analyses including this measure. Further details on flow of participants are depicted in Figure 1. The individual period between T1 and T2 ranged from 27 to 42 days, with an average interval of 34.7 days.

Figure 1. Study flowchart. A: app; AP: app-guided followed by a physiotherapist-guided sequence; P: physiotherapist; PA: physiotherapist-guided followed by an app-guided sequence.



Baseline Data

The key baseline characteristics including physical activity and exercise-related experiences did not differ between participants allocated to the 2 treatment sequences (Table 1). For additional information, see Durst et al [34].

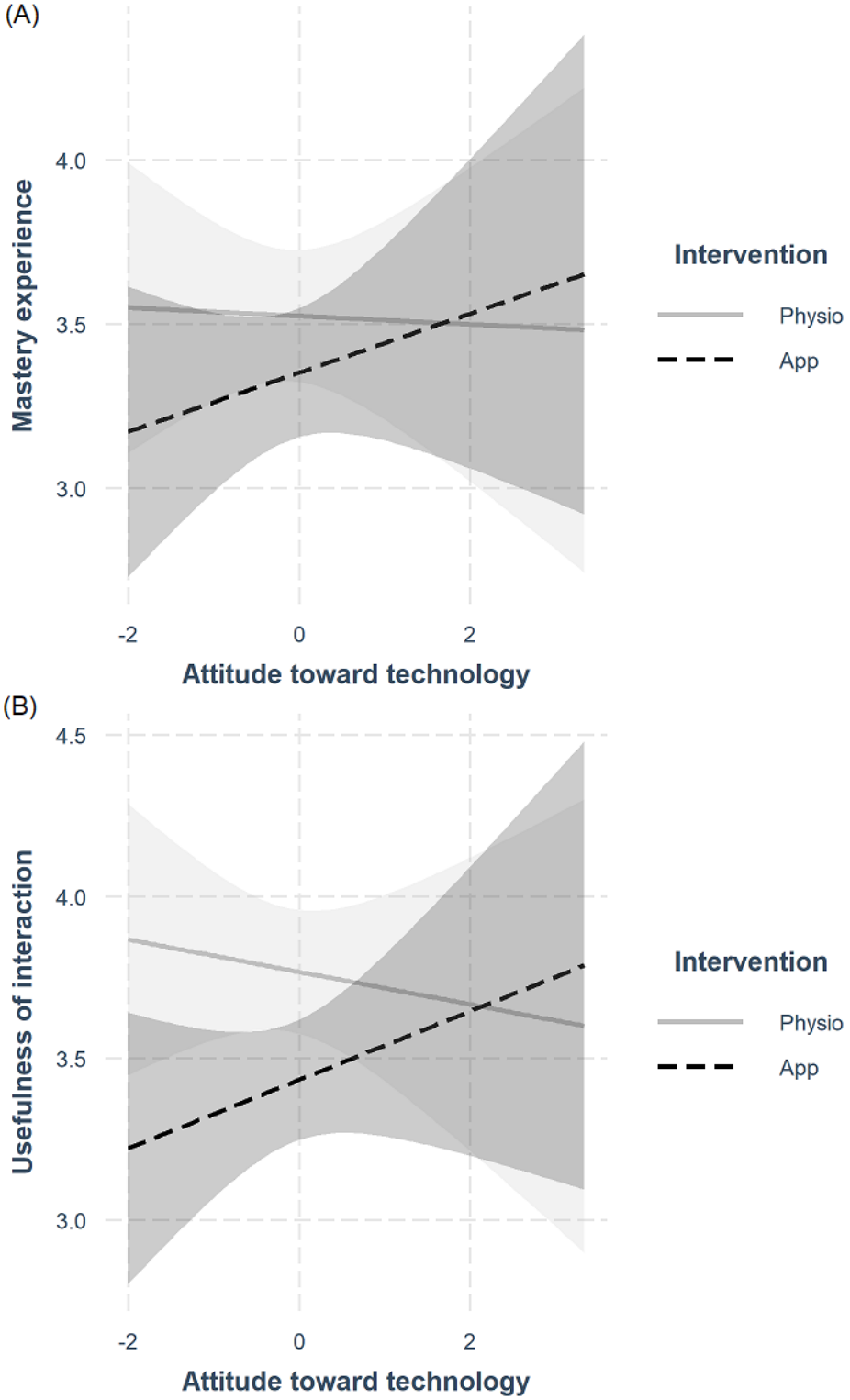
Hypothesis Testing

The analysis for *mastery experience* revealed a main effect of treatment, $F_{1,41.9}$=14.89, P<.001, η^2_{part}=0.26, $CI_{b\text{-}90\%}$ of 0.09-0.43, which was again qualified by the expected treatment × attitudes toward technology interaction, $F_{1,41.6}$=5.95, P=.02, η^2_{part}=0.12, $CI_{b\text{-}90\%}$ of 0.01-0.29. In addition, there was a main effect of sequence factor, $F_{1,42.8}$=4.16, P=.05, η^2_{part}=0.09, $CI_{b\text{-}90\%}$ of 0.00-0.24. In the PA condition the mastery experience was perceived more positive across both treatments than in the AP condition, which is mostly driven by the judgment of the app (Table 2). The other main and interaction effects were not significant (in all cases, F<3.6 [41<*df*<43], P>.05).

Similarly, the analysis for *usefulness of interaction* revealed a main effect of treatment, $F_{1,42}$=26.98, P<.001, η^2_{part}=.38, $CI_{b\text{-}90\%}$ of 0.20-0.54, which was qualified by the predicted treatment × attitudes toward technology interaction, $F_{1,41.7}$=4.88, P=.03, η^2_{part}=0.10, $CI_{b\text{-}90\%}$ of 0.00-0.26. All other main effects or interactions were not significant (in all cases, F<2.2 [41<*df*<43], P>.10).

Simple slope analyses revealed that a more positive attitude toward technology correlated with a more positive mastery experience, b=0.16, standard error=0.07, t_{43}=2.46, P=.02, $CI_{b\text{-}95\%}$ of 0.03-0.29, and a higher usefulness of the interaction, b=0.17, standard error=0.06, t_{43}=2.76, P=.01, $CI_{b\text{-}95\%}$ of 0.05-0.30, regarding the app-based intervention, but not regarding the physio, both $|t_{43}|$ <1 and P>.3 in all cases (Figure 2).

Figure 2. (A) Usefulness of interaction and (B) mastery experience by attitudes toward technology and intervention. Shaded areas represent 95% CIs.



Discussion

Principal Findings

This study aimed at investigating the role of attitudes toward technology for the development of PAHCO (ie, mastery experience and the usefulness of the interaction) comparing app- and physiotherapist-guided exercise. We hypothesized that attitudes toward technology would predict the mastery experience and the perceived usefulness of the interaction regarding app-guided exercise but not regarding physiotherapist-guided exercise. The results supported this prediction.

Overall mastery experience and usefulness of the interaction were lower as an outcome of app-guided exercise than as an outcome of physiotherapist-guided exercise. However, this main effect of intervention type was qualified by the predicted attitudes toward technology × intervention type interaction. For

people with a less positive attitude toward technology both outcomes were lower after the app-guided intervention than after the physiotherapist-guided intervention. This difference was substantially reduced for people with more positive attitudes toward technology and descriptively disappeared 2 SD above the mean, suggesting that only a few people with a very positive attitude toward technology might benefit to a similar extent from an app-based intervention as from a physiotherapist-guided intervention (but see the "Limitations" section). It should be noted, however, that as reported in Durst et al [34] the movement performance (at least for more complex exercises) is higher after physiotherapist-guided exercise compared with app-guided exercise.

Consistent with a recent review mainly referring to qualitative studies [42], our quantitative study approach provides additional evidence for the importance of attitudes toward technology in the process of implementing app-guided exercise interventions (and potentially also health apps including other interventions). People holding a less positive attitude toward technology benefit less in their health competence from the use of an app-guided intervention. This will not only work against the persistent use of such apps but also undermine the long-term health benefits that using such an app could have. In an environment where policy makers stress the self-reliance of patients and a rapidly growing amount of health technologies become available (and partly also replace other interventions), this is an important finding to be considered. Those holding a less positive attitude toward technology might face disadvantages. One intervention that might help to increase positive responses to app-based interventions among those holding a less positive attitude toward technology is a session in which the app is introduced face-to-face. This might result in increased self-efficacy, and therefore most likely also app use. Thus, combining the app with a face-to-face intervention might prevent disadvantages of those with negative attitudes toward technology that might otherwise occur.

What might drive the effects of attitudes toward technology? We assume that people with a negative attitude focus on different (ie, more negative) experiences while using a new technology than people holding a positive attitude. This attention-based explanation effect is speculative and should, thus, be tested in future research.

Limitations

This study has some limitations that should be noted. The means for the outcome measures after both interventions, but in particular after the physiotherapist-guided intervention, are very high. We are, thus, potentially dealing with a ceiling effect for both outcomes. This will most likely lead to an underestimation of effect sizes. Moreover, the intersections between regression lines should be interpreted with caution. Further research with larger samples and measures capturing the variance in the upper range of the scale in a more differentiated manner should be conducted before drawing conclusions about the level of attitudes toward technology from which an equality of both interventions could be assumed.

The usefulness of interaction scale for the physiotherapist had a low internal consistency, but analysis based on single items do not result in a different pattern compared with the reported analysis. This indicates that the current results are stable even though the internal consistency of 1 indicator was low. The low internal consistency most likely results from the richer impression formation process for humans than for technology. The more differentiated impression people have about the physiotherapist might have contributed to a lower correlation between the aspects summarized in the usefulness of the interaction with the physiotherapist scale (compared with the app). At the same time, similar scales are required to compare the outcomes of both types of intervention. Future research might opt for a more differentiated measurement approach making up for this issue.

Strengths

Confronting each participant with both the app- and physiotherapist-guided intervention in randomized order is a strength of this study. It should, however, be noted that this might lead to carryover effects, that is, the outcome of the second intervention might be affected by the first intervention. In the reported analysis, the relevant sequence × treatment interactions were not significant. However, for transparency reasons we would like to note that the attitudes toward technology × treatment interactions are descriptively stronger when the app is presented first. If this difference is replicated in future research, it would indicate that the attitude toward technologies is less relevant and the use of exercise apps is particularly beneficial after exercise sessions guided by a physiotherapist—for instance, as a refresher or as an extension. The results of the movement performance data point in the same direction [34].

A further strength of this study is that we compared the outcomes of using an exercise-related app with the gold standard of a physiotherapist-guided exercise, whereas many studies only focus on the evaluation of apps (often comparing it with a no intervention control condition or paper instructions only). Our comparison sets a very high standard and in this light the difference between both interventions is not surprisingly high. This might in part result from another strength of this study, namely the fact that the exercise program implemented in the app is an evidence-based exercise intervention [22,23]. Finally, it should be noted that the study was conducted among diagnosed patients of an age group that is usually considered as being less technology savvy.

Conclusion

This study provided evidence for the impact of attitudes toward technology for the outcomes of app-guided but not of physiotherapist-guided physical exercise interventions regarding PAHCO. A positive attitude toward technology predicted higher mastery experience after an app-guided but not after a physiotherapist-guided intervention, which is most likely beneficial for task self-efficacy. Therefore, attitudes toward technology should be considered when prescribing and implementing app-based interventions to ensure task self-efficacy and beneficial effects on competencies for a healthy, physically active lifestyle.

Acknowledgments

The authors wish to thank the study participants. They also express their thanks to Simone Schweda, Hendrik Maier, Phillip Weber, and Georg Haupt for their assistance in the preparation and conduction of the study. The project is funded by the Leibniz-Wissenschafts-Campus Tuebingen “Cognitive Interface” with funds from the Ministry of Science, Research and the Arts Baden-Württemberg. The app was designed for a 9.7-in. (24.64 cm) Apple iPad and was developed by Ambigate GmbH (Tuebingen, Germany) according to the specifications of the authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 9770 KB - formative_v6i2e28913_app1.pdf]

Multimedia Appendix 2
Items assessing mastery experience and usefulness of interaction.
[DOCX File , 13 KB - formative_v6i2e28913_app2.docx]

References

1. World Health Organization. Musculoskeletal conditions. World Health Organization. 2021 Feb 08. URL: https://www.who.int/news-room/fact-sheets/detail/musculoskeletal-conditions [accessed 2022-02-02]
2. Briggs AM, Woolf AD, Dreinhöfer K, Homb N, Hoy DG, Kopansky-Giles D, et al. Reducing the global burden of musculoskeletal conditions. Bull World Health Organ 2018 May 01;96(5):366-368 [FREE Full text] [doi: 10.2471/BLT.17.204891] [Medline: 29875522]
3. Deutsche Gesellschaft für Orthopädie und Orthopädische Chirurgie e.V. (DGOOC), Stöve J. S2k-Leitlinie Gonarthrose (Registernummer 033 - 004). AWMF. Berlin, Germany: AWMF Online; 2018 Jan 18. URL: https://www.awmf.org/leitlinien/detail/ll/033-004.html [accessed 2021-03-17]
4. Deutsche Gesellschaft für Orthopädie und Orthopädische Chirurgie e.V. (DGOOC), Matziolis G. S2k-Leitlinie Koxarthrose (Registernummer 033-001). Berlin, Germany: AWMF Online; 2019 Jul 08. URL: https://www.awmf.org/leitlinien/detail/ll/033-001.html [accessed 2021-03-17]
5. Fernandes L, Hagen KB, Bijlsma JWJ, Andreassen O, Christensen P, Conaghan PG, European League Against Rheumatism (EULAR). EULAR recommendations for the non-pharmacological core management of hip and knee osteoarthritis. Ann Rheum Dis 2013 Jul;72(7):1125-1135 [FREE Full text] [doi: 10.1136/annrheumdis-2012-202745] [Medline: 23595142]
6. McAlindon TE, Bannuru RR, Sullivan MC, Arden NK, Berenbaum F, Bierma-Zeinstra SM, et al. OARSI guidelines for the non-surgical management of knee osteoarthritis. Osteoarthritis Cartilage 2014 Mar;22(3):363-388 [FREE Full text] [doi: 10.1016/j.joca.2014.01.003] [Medline: 24462672]
7. Wallis JA, Webster KE, Levinger P, Taylor NF. What proportion of people with hip and knee osteoarthritis meet physical activity guidelines? A systematic review and meta-analysis. Osteoarthritis Cartilage 2013 Nov;21(11):1648-1659 [FREE Full text] [doi: 10.1016/j.joca.2013.08.003] [Medline: 23948979]
8. Gunn AH, Schwartz TA, Arbeeva LS, Callahan LF, Golightly Y, Goode A, et al. Fear of Movement and Associated Factors Among Adults With Symptomatic Knee Osteoarthritis. Arthritis Care Res (Hoboken) 2017 Dec;69(12):1826-1833 [FREE Full text] [doi: 10.1002/acr.23226] [Medline: 28371481]
9. Danbjørg DB, Villadsen A, Gill E, Rothmann MJ, Clemensen J. Usage of an Exercise App in the Care for People With Osteoarthritis: User-Driven Exploratory Study. JMIR Mhealth Uhealth 2018 Jan 11;6(1):e11 [FREE Full text] [doi: 10.2196/mhealth.7734] [Medline: 29326092]
10. Bennell KL, Dobson F, Hinman RS. Exercise in osteoarthritis: moving from prescription to adherence. Best Pract Res Clin Rheumatol 2014 Feb;28(1):93-117. [doi: 10.1016/j.berh.2014.01.009] [Medline: 24792947]
11. McCarthy CJ, Mills PM, Pullen R, Roberts C, Silman A, Oldham JA. Supplementing a home exercise programme with a class-based exercise programme is more effective than home exercise alone in the treatment of knee osteoarthritis. Rheumatology (Oxford) 2004 Jul;43(7):880-886. [doi: 10.1093/rheumatology/keh188] [Medline: 15113993]
12. Daly RM, Gianoudis J, Hall T, Mundell NL, Maddison R. Feasibility, Usability, and Enjoyment of a Home-Based Exercise Program Delivered via an Exercise App for Musculoskeletal Health in Community-Dwelling Older Adults: Short-term Prospective Pilot Study. JMIR Mhealth Uhealth 2021 Jan 13;9(1):e21094 [FREE Full text] [doi: 10.2196/21094] [Medline: 33439147]

13. Kiss N, Baguley BJ, Ball K, Daly RM, Fraser SF, Granger CL, et al. Technology-Supported Self-Guided Nutrition and Physical Activity Interventions for Adults With Cancer: Systematic Review. JMIR Mhealth Uhealth 2019 Feb 12;7(2):e12281 [FREE Full text] [doi: 10.2196/12281] [Medline: 30747720]
14. Mehra S, van den Helder J, Visser B, Engelbert RHH, Weijs PJM, Kröse BJA. Evaluation of a Blended Physical Activity Intervention for Older Adults: Mixed Methods Study. J Med Internet Res 2020 Jul 23;22(7):e16380 [FREE Full text] [doi: 10.2196/16380] [Medline: 32459652]
15. Robertson MC, Tsai E, Lyons EJ, Srinivasan S, Swartz MC, Baum ML, et al. Mobile Health Physical Activity Intervention Preferences in Cancer Survivors: A Qualitative Study. JMIR Mhealth Uhealth 2017 Jan 24;5(1):e3 [FREE Full text] [doi: 10.2196/mhealth.6970] [Medline: 28119278]
16. Rossen S, Kayser L, Vibe-Petersen J, Christensen JF, Ried-Larsen M. Cancer Survivors' Receptiveness to Digital Technology-Supported Physical Rehabilitation and the Implications for Design: Qualitative Study. J Med Internet Res 2020 Aug 05;22(8):e15335 [FREE Full text] [doi: 10.2196/15335] [Medline: 32755892]
17. Ofcom. Rise of the social seniors revealed. Ofcom. 2017 Jun 20. URL: https://www.ofcom.org.uk/about-ofcom/latest/features-and-news/rise-social-seniors [accessed 2021-03-17]
18. Cajita MI, Hodgson NA, Lam KW, Yoo S, Han H. Facilitators of and Barriers to mHealth Adoption in Older Adults With Heart Failure. Comput Inform Nurs 2018 Aug;36(8):376-382 [FREE Full text] [doi: 10.1097/CIN.0000000000000442] [Medline: 29742549]
19. Chen TL, Bhattacharjee T, Beer JM, Ting LH, Hackney ME, Rogers WA, et al. Older adults' acceptance of a robot for partner dance-based exercise. PLoS One 2017;12(10):e0182736 [FREE Full text] [doi: 10.1371/journal.pone.0182736] [Medline: 29045408]
20. Warner LM, Schüz B, Knittle K, Ziegelmann J, Wurm S. Sources of perceived self-efficacy as predictors of physical activity in older adults. Applied Psychology: Health and Well-Being 2011;3(2):172-192. [doi: 10.1111/j.1758-0854.2011.01050.x]
21. Rahimi B, Nadri H, Lotfnezhad Afshar H, Timpka T. A Systematic Review of the Technology Acceptance Model in Health Informatics. Appl Clin Inform 2018 Jul;9(3):604-634 [FREE Full text] [doi: 10.1055/s-0038-1668091] [Medline: 30112741]
22. Haupt G, Janßen P, Krauß I, Steinhilber B. Das Tübinger Hüftkonzept. Essen, Germany: Hellblau-Verlag; Feb 10, 2014.
23. Krauß I, Steinhilber B, Haupt G, Miller R, Martus P, Janßen P. Exercise therapy in hip osteoarthritis--a randomized controlled trial. Dtsch Arztebl Int 2014 Sep 01;111(35-36):592-599 [FREE Full text] [doi: 10.3238/arztebl.2014.0592] [Medline: 25249361]
24. Sudeck G, Pfeifer K. Physical activity-related health competence as an integrative objective in exercise therapy and health sports – conception and validation of a short questionnaire. Sportwiss 2016 May 17;46(2):74-87. [doi: 10.1007/s12662-016-0405-4]
25. Carl J, Sudeck G, Pfeifer K. Competencies for a Healthy Physically Active Lifestyle-Reflections on the Model of Physical Activity-Related Health Competence. J Phys Act Health 2020 May 29;17(7):688-697. [doi: 10.1123/jpah.2019-0442] [Medline: 32473589]
26. Higgins TJ, Middleton KR, Winner L, Janelle CM. Physical activity interventions differentially affect exercise task and barrier self-efficacy: a meta-analysis. Health Psychol 2014 Aug;33(8):891-903 [FREE Full text] [doi: 10.1037/a0033864] [Medline: 23957904]
27. 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]
28. Marcus BH, Eaton CA, Rossi JS, Harlow LL. Self-Efficacy, Decision-Making, and Stages of Change: An Integrative Model of Physical Exercise1. J Appl Social Pyschol 1994 Mar;24(6):489-508. [doi: 10.1111/j.1559-1816.1994.tb00595.x]
29. Behringer N, Sassenberg K. Introducing social media for knowledge management: Determinants of employees' intentions to adopt new tools. Computers in Human Behavior 2015 Jul;48:290-296. [doi: 10.1016/j.chb.2015.01.069]
30. Anderberg P, Eivazzadeh S, Berglund JS. A Novel Instrument for Measuring Older People's Attitudes Toward Technology (TechPH): Development and Validation. J Med Internet Res 2019 May 23;21(5):e13951. [doi: 10.2196/13951] [Medline: 31124467]
31. Currie M, Philip LJ, Roberts A. Attitudes towards the use and acceptance of eHealth technologies: a case study of older adults living with chronic pain and implications for rural healthcare. BMC Health Serv Res 2015 Apr 16;15:162 [FREE Full text] [doi: 10.1186/s12913-015-0825-0] [Medline: 25888988]
32. Kim J, Park H. Development of a health information technology acceptance model using consumers' health behavior intention. J Med Internet Res 2012 Oct 01;14(5):e133 [FREE Full text] [doi: 10.2196/jmir.2143] [Medline: 23026508]
33. Legris P, Ingham J, Collerette P. Why do people use information technology? A critical review of the technology acceptance model. Information & Management 2003 Jan;40(3):191-204. [doi: 10.1016/s0378-7206(01)00143-4]
34. Durst J, Roesel I, Sudeck G, Sassenberg K, Krauss I. Effectiveness of Human Versus Computer-Based Instructions for Exercise on Physical Activity-Related Health Competence in Patients with Hip Osteoarthritis: Randomized Noninferiority Crossover Trial. J Med Internet Res 2020 Sep 28;22(9):e18233 [FREE Full text] [doi: 10.2196/18233] [Medline: 32985991]
35. True random number service. RANDOM.ORG. URL: https://www.random.org [accessed 2022-02-02]

36. Krauss I, Steinhilber B, Haupt G, Miller R, Grau S, Janssen P. Efficacy of conservative treatment regimes for hip osteoarthritis--evaluation of the therapeutic exercise regime "Hip School": a protocol for a randomised, controlled trial. BMC Musculoskelet Disord 2011 Nov 24;12:270 [FREE Full text] [doi: 10.1186/1471-2474-12-270] [Medline: 22114973]
37. Steinhilber B, Haupt G, Miller R, Janssen P, Krauss I. Exercise therapy in patients with hip osteoarthritis: Effect on hip muscle strength and safety aspects of exercise-results of a randomized controlled trial. Mod Rheumatol 2017 May;27(3):493-502 [FREE Full text] [doi: 10.1080/14397595.2016.1213940] [Medline: 27486681]
38. Pfeifer K, Sudeck G, Geidl W, Tallner A. Bewegungsförderung und Sport in der Neurologie? Kompetenzorientierung und Nachhaltigkeit. Neurol Rehab 2013 Jan 01;19(1):7-19.
39. Karrer K, Glaser C, Clemens C, Bruder C. Technikaffinität erfassen-der Fragebogen TA-EG. Der Mensch im Mittelpunkt Technischer Systeme 2009 Oct 07;8:196-201.
40. Rackow P, Scholz U, Hornung R. The German Psychological Need Satisfaction in Exercise Scale. Swiss Journal of Psychology 2013 Jan;72(3):137-148. [doi: 10.1024/1421-0185/a000107]
41. Harder H, Holroyd P, Burkinshaw L, Watten P, Zammit C, Harris PR, et al. A user-centred approach to developing bWell, a mobile app for arm and shoulder exercises after breast cancer treatment. J Cancer Surviv 2017 Dec;11(6):732-742 [FREE Full text] [doi: 10.1007/s11764-017-0630-3] [Medline: 28741202]
42. Hosseinpour M, Terlutter R. Your Personal Motivator is with You: A Systematic Review of Mobile Phone Applications Aiming at Increasing Physical Activity. Sports Med 2019 Sep;49(9):1425-1447 [FREE Full text] [doi: 10.1007/s40279-019-01128-3] [Medline: 31144235]

Abbreviations

AP: app–physiotherapist
OA: osteoarthritis
PA: physiotherapist–app
PAHCO: physical activity–related health competence
WHO: World Health Organization

Edited by A Mavragani; submitted 18.03.21; peer-reviewed by K Halttu, S Rostam Niakan Kalhori; comments to author 30.07.21; revised version received 26.08.21; accepted 22.12.21; published 18.02.22.

Please cite as:
Sassenberg K, Roesel I, Sudeck G, Bernecker K, Durst J, Krauss I
The Relation of Attitude Toward Technology and Mastery Experience After an App-Guided Physical Exercise Intervention: Randomized Crossover Trial
JMIR Form Res 2022;6(2):e28913
URL: https://formative.jmir.org/2022/2/e28913
doi:10.2196/28913
PMID:35179500

Original Paper

Consistency With and Disengagement From Self-monitoring of Weight, Dietary Intake, and Physical Activity in a Technology-Based Weight Loss Program: Exploratory Study

Chelsea A Carpenter[1], MPH, PhD; Abraham Eastman[1], MS; Kathryn M Ross[1], MPH, PhD

Department of Clinical and Health Psychology, College of Public Health and Health Professions, University of Florida, Gainesville, FL, United States

Corresponding Author:
Kathryn M Ross, MPH, PhD
Department of Clinical and Health Psychology
College of Public Health and Health Professions
University of Florida
PO Box 100165
Gainesville, FL, 32610
United States
Phone: 1 (352) 294 8433
Email: kmross@phhp.ufl.edu

Abstract

Background: Digital self-monitoring tools offer promise to improve adherence to self-monitoring of weight and weight-related behaviors; however, less is known regarding the patterns of participant consistency and disengagement with these tools.

Objective: This study characterizes the consistency of use and time to disengagement with digital self-monitoring tools during a 6-month weight loss intervention and investigates whether the provision of phone-based intervention improved self-monitoring adherence.

Methods: Participants were 54 adults with overweight or obesity (mean age 49.6 years, SD 12.4 years; mean BMI 32.6 kg/m^2, SD 3.2 kg/m^2) enrolled in a pilot trial assessing the impact of self-monitoring technology (Fitbit Zip, Aria scale, and smartphone app), with and without additional interventionist contact, on weight loss. All participants received weight loss education and were asked to self-monitor weight, dietary intake, and physical activity daily throughout the 6-month program. Consistency was defined as the number of weeks that participants adhered to self-monitoring recommendations (7 out of 7 days). Disengagement was defined as the first of 2 consecutive weeks that the 7-day self-monitoring adherence goal was not met. Wilcoxon signed-rank tests were used to examine differences in consistency and disengagement by behavioral targets. *t* tests (2-tailed) and Cox proportional hazards models were used to examine whether providing additional interventionist contact would lead to significant improvements in consistency and time to disengagement from self-monitoring tools, respectively. Linear regressions were used to examine associations between consistency, time to disengagement, and weight loss.

Results: Participants consistently self-monitored physical activity for more weeks (mean 17.4 weeks, SD 8.5 weeks) than weight (mean 11.1 weeks, SD 8.5 weeks) or dietary intake (mean 10.8 weeks, SD 8.7 weeks; $P<.05$). Similarly, participants had a significantly longer time to disengagement from self-monitoring of physical activity (median 19.5 weeks) than weight (4 weeks) or dietary intake (10 weeks; $P<.001$). Participants randomized to receive additional interventionist contact had significantly greater consistency and longer time to disengagement for self-monitoring of dietary intake compared with participants who did not ($P=.006$); however, there were no statistically significant differences between groups for self-monitoring of weight or physical activity ($P=.24$ and $P=.25$, respectively). Greater consistency and longer time to disengagement were associated with greater weight loss for self-monitoring of weight and dietary intake ($P<.001$ and $P=.004$, respectively) but not for physical activity ($P=.57$).

Conclusions: Results demonstrated that self-monitoring adherence differed by behavioral target, with greater consistency and longer time to disengagement associated with lower-burden tools (ie, self-monitoring of physical activity). Consistent with supportive accountability theory, additional interventionist contact improved consistency and lengthened time to disengagement from self-monitoring of dietary intake. Given the observed associations between consistency, disengagement, and weight loss outcomes, it is important to identify additional methods of increasing consistency and engagement with digital self-monitoring tools.

(JMIR Form Res 2022;6(2):e33603) doi:10.2196/33603

KEYWORDS

self-monitoring; adherence; weight loss; digital tools; mobile phone

Introduction

Background

Self-monitoring of weight and weight-related behaviors (eg, dietary intake and physical activity) is considered a cornerstone of evidence-based weight loss programs [1,2]. Research has consistently demonstrated that greater adherence to self-monitoring is associated with better weight loss outcomes [3,4]; however, rates of adherence to self-monitoring are often suboptimal and tend to decrease over time [5-8].

Digital self-monitoring tools, including smartphone apps, wearables, and smart scales, offer promise to improve adherence to self-monitoring. These tools reduce the time needed to complete self-monitoring records compared with traditional paper-and-pencil tools, capitalize on tools that many individuals already carry or wear throughout the day, and may be more acceptable to use in social situations [9]. Promisingly, evidence to date has demonstrated that these technologies can produce greater adherence to self-monitoring for participants in behavioral weight management programs when compared with traditional self-monitoring tools [7,10,11].

Digital self-monitoring tools also allow researchers to examine patterns in self-monitoring behavior on levels not previously possible. Before the introduction of these tools, self-monitoring adherence was typically assessed either via a self-report questionnaire (eg, with items asking participants if they self-monitored daily, weekly, monthly, or at other frequencies) or by counting days of self-monitoring reported on logs returned to study staff. Data from self-report questionnaires have known biases related to retrospective recall [12] and tend to have only moderate to poor association with measured self-monitoring, such that individuals tend to overreport self-monitoring behavior on these questionnaires [13]. Counts of self-monitoring logs may also lead to inaccurate estimates of self-monitoring; one study that used an unobtrusive monitoring device to assess when paper self-monitoring logs were opened found that participants self-reported monitoring their dietary intake more often than indicated by the monitoring device [14]. Moreover, missing data from these counts could indicate either a day that self-monitoring did not occur or a day that the self-monitoring was not reported (eg, when records or self-monitoring summary sheets were not returned to the study team). In contrast, digital tools automatically send data directly to cloud-based storage, reducing missingness from failure to self-report self-monitoring. Although this difference may seem trivial, a recent study demonstrated only moderate agreement between frequency of self-weighing as assessed by self-report logs and that by a digital smart scale that sent weights directly back to research servers [15].

Beyond the simple use of frequency of self-monitoring (ie, the number of times a person self-monitors during a specific period) as a marker of adherence, recent research has suggested that certain patterns of self-monitoring may be important. A total of 2 studies have demonstrated that, after controlling for frequency of self-monitoring, the consistency of self-monitoring (ie, the amount of time that a person self-monitors at a certain frequency) may matter [16,17]. Peterson et al [16] found that, after the end of an initial weight loss program, a higher frequency of self-monitoring dietary intake during a 12-month extended care program was associated with less weight regain only when coupled with higher consistency of self-monitoring (with consistency defined as the number of weeks wherein the participant self-monitored dietary intake at least 3 out of 7 days in a week). Brockmann et al [17] found no association between the total frequency of self-monitoring weight and weight regain during a 9-month observation period following a 3-month weight loss program but found that greater consistency of self-monitoring weight (defined as the number of weeks that participants self-weighed at least 6 or 7 days in a week) was associated with less weight regain. Importantly, both studies were conducted during *maintenance* periods following the end of initial weight loss programs; less is known regarding the associations between consistency of self-monitoring and weight loss during the initial weight loss period.

Furthermore, few studies have examined the concept of *disengagement* (eg, the point at which a person stops using self-monitoring tools as recommended), which may be particularly relevant when assessing patterns of self-monitoring with digital tools. Disengagement appears to occur after the initial novelty (eg, excitement related to the use of new technology) of digital self-monitoring tools wears off, especially when these tools are implemented without the support of a structured weight management program [18-21]. Studies of physical activity wearables have shown that most participants disengage with these devices after as little as 2 weeks [19,21] to 3 months [18]. Similarly, studies of commercial mobile health apps that allow for self-monitoring of dietary intake found that participants typically disengage with these apps after approximately a month [22,23]. Promisingly, research has demonstrated that provision of additional intervention support can improve the frequency of self-monitoring in technology-based interventions [24-26]; however, no studies have investigated the impact of this support on consistency with or disengagement from self-monitoring.

Objectives

To address these gaps, this study aims to characterize patterns of consistency with and disengagement from self-monitoring technology for weight, dietary intake, and physical activity during a 6-month weight loss intervention. Data for this study were gathered from a randomized pilot trial (NCT01999244) that assessed the impact of digital self-monitoring tools, provided with and without additional phone-based interventionist contact, on weight loss in adults with overweight and obesity [26]. The primary aim of this study is to characterize the patterns of consistency and disengagement in the use of

digital self-monitoring tools and to test the hypothesis that participants randomized to receive additional phone-based interventionist contact would demonstrate greater consistency and a longer time until disengagement compared with participants who did not receive this additional contact. As a secondary aim, we examine whether greater consistency and longer time to disengagement were associated with greater weight loss during the 6-month intervention period. Finally, as an exploratory aim, we explore whether participants re-engaged with digital tools after the first period of disengagement and descriptively explored whether patterns of re-engagement were different between groups.

Methods

Parent Study Design

This study was a secondary analysis of data collected from a randomized pilot study investigating the impact of digital self-monitoring tools, provided with and without additional phone-based interventionist contact, on weight loss over 6 months in adults with overweight and obesity [26]. Participants in the parent study were randomized to one of three treatment conditions: (1) standard, in which participants were asked to self-monitor dietary intake, physical activity, and weight via traditional self-monitoring tools (with the study providing a printed calorie reference book, a standard pedometer, paper records, and a standard bathroom scale if one was not already owned by the participant); (2) technology-based self-monitoring tools (TECH), in which participants were asked to self-monitor via digital self-monitoring tools (with the study providing a Fitbit Zip activity monitor, a Fitbit Aria smart scale, and access to the Fitbit website and smartphone app, which allowed participants to self-monitor dietary intake and observe the synced data from their activity monitor and smart scale); or (3) TECH plus additional phone-based intervention (TECH+PHONE), in which participants were asked to self-monitor using the same tools as participants randomized to the TECH condition but were also provided with 14 phone-based intervention sessions over the 6-month intervention period.

Participants

Participants in the parent study were adults (aged 18-70 years) with overweight or obesity (BMIs between 27 and 40 kg/m^2) who had access to a computer and Wi-Fi internet at home [26]. Potential participants were recruited through local advertisements and flyers and asked to complete a web-based prescreen assessment before attending an in-person orientation visit, during which detailed information about the study was provided and written informed consent was obtained. Potential participants were excluded from the parent study if they reported any physical limitations that prevented walking 402.3 meters without stopping, they were participating in another weight loss program, they were taking weight loss medication, they were pregnant or planned to become pregnant during the study period, or they had any medical conditions that would contraindicate participation in a weight loss program (eg, uncontrolled type 2 diabetes, uncontrolled hypertension, or history of coronary heart disease).

Given that only participants randomized to the TECH and TECH+PHONE conditions received digital self-monitoring technology, participants randomized to the standard group were excluded from this study. Full demographic details for the 54 participants randomized to TECH (27/54, 50%) or TECH+PHONE (27/54, 50%) have been published previously [26]. In brief, participants had an average age of 49.6 (SD 12.4) years and an average BMI of 32.6 (SD 3.2) kg/m^2 at baseline. In addition, 87% (47/54) of the participants were identified as women, and in terms of race and ethnicity, 85% (46/54) identified as non-Hispanic White, 7% (4/54) as Hispanic, 2% (1/54) as non-Hispanic Black, and 6% (3/54) identified as another category or indicated multiple race and ethnicity categories.

Intervention Components

All participants included in this study (ie, those randomized to the TECH and TECH+PHONE conditions) received access to the Fitbit website and smartphone app, a Fitbit Aria smart scale (which synced participants' weight directly to the Fitbit website and smartphone app), and a Fitbit Zip (a small digital pedometer attached to participants' waistbands and synced data directly to the Fitbit website and smartphone app). The Fitbit website and smartphone apps were used to self-monitor dietary intake and allowed participants to track the foods and beverages they consumed by searching for them in a comprehensive web-based database. This database listed dietary intake information for all items and allowed individuals to easily adjust portion sizes for logged items. Unless participants chose to opt out of emails from Fitbit, participants also received an automated summary email from Fitbit each Sunday that provided information regarding average steps per day, average number of calories consumed each day, weight change over the previous week, and updates and marketing content from Fitbit (this was an automated email sent by Fitbit and sent regardless of adherence to self-monitoring; it was not intended to serve as an intervention component but is discussed as it may have impacted adherence). No additional reminders to self-monitor were sent.

All participants received the same weight loss education content during the initial Weight Loss 101 session. During this session, participants were provided with a daily calorie goal (ranging from 1200 to 1500 kcal per day, based on their baseline weight) and were encouraged to consume <30% of calories from fat each day. Participants were also given two physical activity goals: (1) to gradually increase daily steps to achieve 10,000 steps per day and (2) to gradually increase engagement in moderate-intensity exercise (eg, walking at a brisk pace) to 250 min/week. Participants were then taught how to use the study-provided digital self-monitoring tools and practiced wearing the Fitbit pedometer and using the Fitbit smartphone app and website to self-monitor an example meal. During this time, support was provided for any participant who had difficulty using the tools. Starting the day after the Weight Loss 101 visit, participants were asked to self-monitor their weight, dietary intake, and physical activity each day. Participants who experienced technical challenges after the end of this Weight Loss 101 session were encouraged to call a study number to

speak with a noninterventionist staff member or to contact Fitbit customer service directly if a higher level of support was needed.

Participants randomized to the TECH+PHONE condition also received 14 structured phone calls (8 weekly, 4 biweekly, and 2 monthly calls) with a trained behavioral weight loss interventionist. Each call started with a check-in assessing the frequency of self-monitoring weight, dietary intake, and physical activity, and progress toward goals since the previous call. Reinforcement was provided for goals that were successfully met, and structured problem-solving strategies [27] were used to address any participant-reported barriers to goal attainment. A brief discussion would then focus on a specific behavioral weight management topic (eg, goal-setting [28], stimulus control [29], seeking social support [30], and relapse prevention [31]). Each call ended with structured goal-setting for the following week, including specific goals for dietary intake, physical activity, and self-monitoring of weight, dietary intake, and physical activity. Overall, calls were anticipated to last for 10-15 minutes (with 3-5 minutes devoted to the check-in, 4-5 minutes devoted to the session topic, and 3-5 minutes focused on goal-setting). Participants randomized to the TECH condition received no additional intervention contact during the 6-month study period. Additional details regarding the Weight Loss 101 sessions and TECH+PHONE call content have been published previously [26].

Intervention Outcomes

Intervention outcomes have been published previously [26]. Of the 80 participants randomized in the parent study, 27 (34%) were randomized to TECH and 27 (34%) to TECH+PHONE. Retention at the 6-month visit was 93% (25/27) for TECH and 89% (24/27) for TECH+PHONE, with no differences between groups. From baseline to the 6-month assessment visit, participants in the TECH group lost an average (mean) of −4.04 (SD 1.37) kg (a loss of 4.35%, SD 1.29% of baseline weight), whereas those in the TECH+PHONE group lost an average of −6.40 (SD 1.17) kg (7.37%, SD 1.29% of baseline weight), with no significant differences between groups.

Measures

Sociodemographic characteristics were assessed at baseline via REDCap (Research Electronic Data Capture; Vanderbilt University) surveys [32]. Height was assessed at baseline by a trained research assistant, to the nearest 0.1 cm, using a wall-mounted stadiometer. Weight was assessed at baseline and 6 months by trained research assistants, to the nearest 0.1 kg, using a calibrated digital scale. For measurements of height and weight, participants were asked to remove shoes and wear only light indoor clothing, with pockets emptied.

All self-monitoring data were collected via the study-provided digital self-monitoring tools, with data pulled from the Fitbit servers via Fitabase [33], a third-party research data management platform. Even though participants could manually input physical activity and weight data, only data passively collected by the Fitbit Zip and Aria scale were used in this study's analyses.

Statistical Analyses

Given that all participants were instructed to self-monitor weight, dietary intake, and physical activity daily throughout the 6-month program, consistency of self-monitoring was defined as the number of weeks that participants met this goal (self-monitoring 7 out of 7 days), calculated separately for each type of self-monitoring. A valid day of self-monitoring weight was defined as a day on which at least one weight value was captured via the Fitbit Aria scale (the scale automatically filtered out weights that were likely to be from other users or pets). A valid day of self-monitoring dietary intake was defined as a day that any foods or beverages were logged via the Fitbit website or smartphone app, which is consistent with the approach used in previous studies [34,35]. Similarly, even though participants were encouraged to meet both an exercise minute and step goal, a valid day of self-monitoring physical activity was defined as a day that any steps were recorded via the Fitbit Zip (as this Fitbit version only tracked steps and did not allow for tracking of exercise or minutes of moderate- and vigorous-intensity physical activity). Disengagement was defined as the point at which a participant failed to self-monitor consistently (7 out of 7 days) for a period of 2 consecutive weeks (with the first week of the 2 weeks counted as the week in which disengagement began, calculated separately for each method of self-monitoring). A 2-week period was used as the threshold for disengagement, as it provided some allowance for brief disturbances (eg, illness or travel) that may have interrupted participant engagement without resulting in prolonged disengagement. For the exploratory aim, re-engagement was defined as any point at which a participant met the 7 out of 7-day self-monitoring goal after the initial 2-week period of disengagement.

Descriptive statistics were used to characterize consistency and time to disengagement (eg, the number of weeks before a participant disengaged with each self-monitoring tool) overall and by treatment condition. For the primary aim, Wilcoxon signed-rank tests were used to examine whether there were significant differences in consistency and time to disengagement among the different modalities of self-monitoring. In addition, independent 2-tailed *t* tests (*df*=53) were used to assess differences in consistency by treatment group, and Cox proportional hazards models were used to examine whether the time to disengagement from each self-monitoring method differed significantly between the 2 treatment groups. For the secondary aim, linear regressions were used to assess associations between consistency and disengagement and weight loss during the 6-month intervention. For the exploratory aim, descriptive statistics were used to characterize patterns of re-engagement after initial disengagement (eg, proportion of participants who re-engaged after disengaging, number of weeks until re-engagement, and number of weeks of re-engagement), and chi-square analyses and independent 2-tailed *t* tests were used to examine if there were differences in re-engagement among groups.

Analyses were conducted using an intent-to-treat approach, with multiple imputation used to manage missing 6-month weight data [26]. Because of the small sample size and absence of significant differences in demographic characteristics among groups [26], no additional variables were controlled for in the

analyses. SAS (version 9.4; SAS Institute Inc) for Windows was used for descriptive statistics, linear regressions, and independent 2-tailed *t* tests. SPSS (version 26; IBM Corp) was used for the Cox proportional hazards models.

Ethics Approval

The study was approved by The Miriam Hospital IRB (206813), and all participants gave their written informed consent before participation in the study.

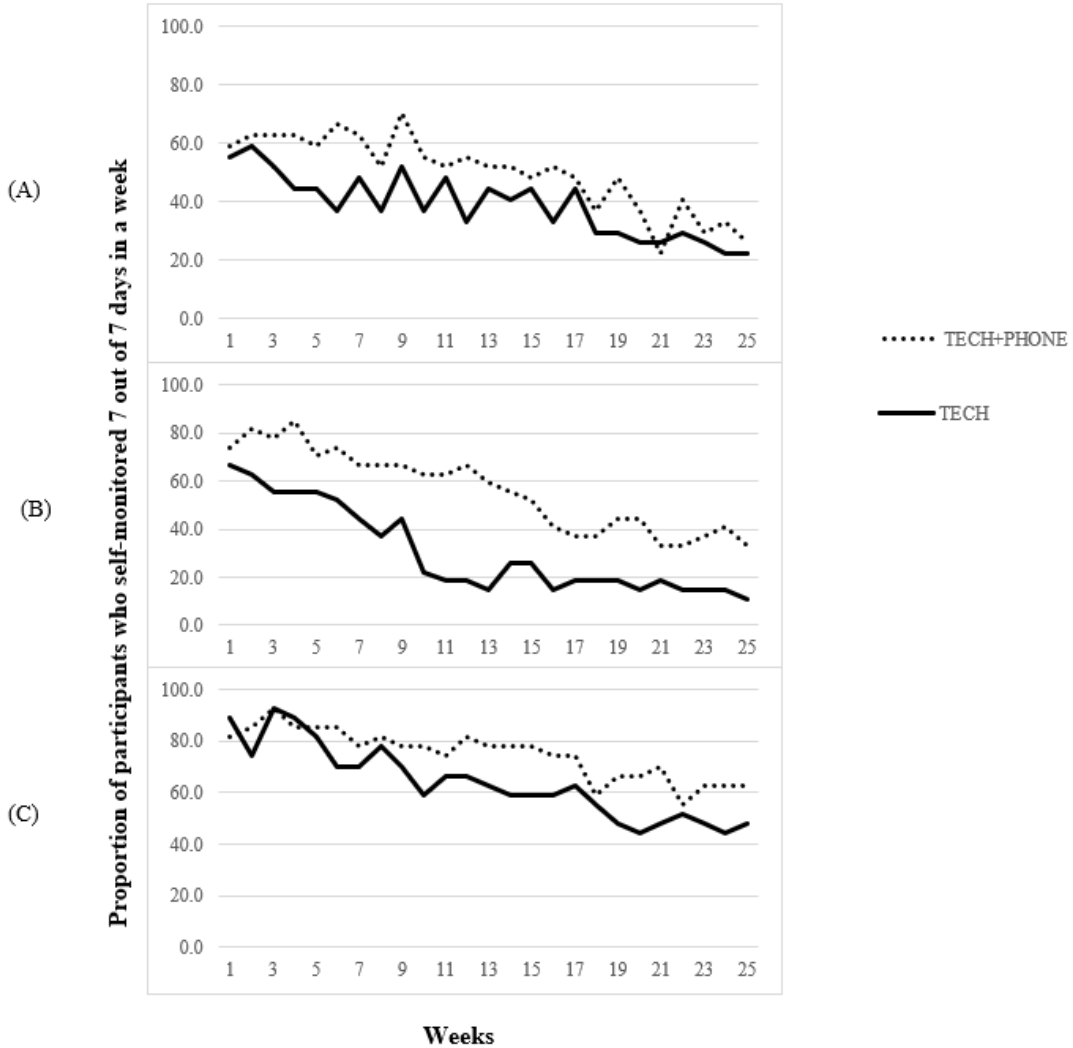
Results

Consistency

Across the total sample of participants (N=54) randomized to either the TECH (27/54, 50%) or TECH+PHONE group (27/54, 50%), participants consistently self-monitored physical activity on more weeks (mean 17.4 weeks, SD 8.5 weeks out of 25 possible weeks) than weight (mean 11.1 weeks, SD 8.5 weeks; *P*=.02) or dietary intake (mean 10.8 weeks, SD 8.7 weeks; *P*<.001); however, there was no significant difference in consistency of self-monitoring for dietary intake versus weight (*P*=.76). Figure 1 shows the trends in the proportion of participants who consistently self-monitored weight, dietary intake, and physical activity over time, by intervention group. Approximately 57% (31/54) of the participants consistently self-monitored weight in the first week of the program, a rate that fell below 24% (13/54) by the end of the 6-month program. Similarly, approximately 70% (38/54) of the participants consistently self-monitored dietary intake during the first week of the program, whereas 22% (12/54) did so by the end of the program. Finally, 85% (46/54) of the participants consistently self-monitored physical activity during the first week of the program, and over half still did so by the end of the program.

Investigating differences between treatment groups, TECH+PHONE participants consistently self-monitored dietary intake on a greater number of weeks than TECH participants (mean 14.0 weeks, SD 8.7 weeks vs mean 7.6 weeks, SD 7.5 weeks; t_{52}=−2.88; *P*=.006; Cohen *d*=0.785). There were no significant differences between TECH+PHONE and TECH in the mean number of weeks that participants consistently self-monitored weight (mean 12.5 weeks, SD 8.1 weeks vs mean 9.7 weeks, SD 10.0 weeks; *P*=.24; Cohen *d*=0.327) or physical activity (mean 18.7 weeks, SD 8.5 weeks vs mean 16.0 weeks, SD 8.4 weeks; *P*=.25; Cohen *d*=0.320).

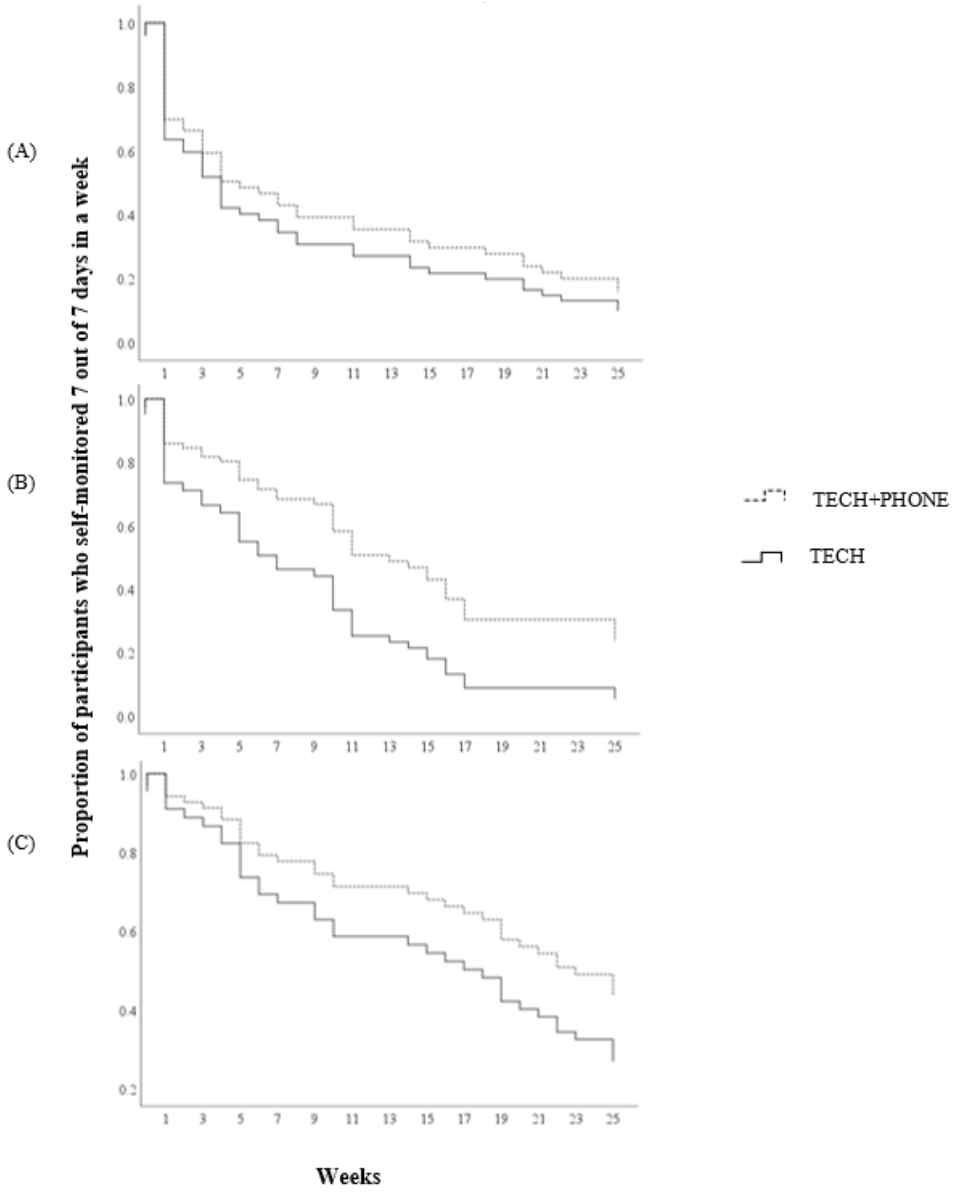
Figure 1. Consistency of (A) weight, (B) dietary intake, and (C) physical activity self-monitoring over time by intervention group. TECH: technology-based self-monitoring tools; TECH+PHONE: technology-based self-monitoring tools plus additional phone-based intervention.



Disengagement

Similar to consistency, participants had a significantly longer time to disengagement from self-monitoring of physical activity (median 19.5 weeks, IQR 6-26 weeks) than weight (median 4 weeks, IQR 1-18 weeks; $P<.001$) or dietary intake (median 10 weeks, IQR 3-17 weeks; $P<.001$); however, there was no significant difference in time to disengagement between dietary intake and weight ($P=.06$). Figure 2 shows survival curves for time to disengagement for self-monitoring of dietary intake, weight, and physical activity by intervention group. TECH+PHONE participants had a longer time to disengagement from self-monitoring of dietary intake compared with TECH participants (at a median of 15.5 vs 7.5 weeks, respectively; $\chi^2_1=5.5$; $P=.02$); however, there were no significant differences in time to disengagement between TECH+PHONE and TECH participants for self-monitoring of weight (4.3 vs 11.0 weeks, respectively; $P=.43$) or physical activity (19.0 vs 25.0 weeks, respectively; $P=.18$).

Figure 2. Time to disengagement from self-monitoring for (A) weight, (B) dietary intake, and (C) physical activity by intervention group. TECH: technology-based self-monitoring tools; TECH+PHONE: technology-based self-monitoring tools plus additional phone-based intervention.



Associations Between Consistency and Disengagement and 6-Month Weight Change

There was a significant association between 6-month weight change and greater consistency of self-monitoring weight ($R^2=0.19$; $F_{1,52}=12.34$; $P<.001$), such that each additional week of consistent self-monitoring of weight was associated with an average (mean −0.33 kg, SE 0.10 kg) greater weight loss. There was also a significant association between weight change and greater consistency of self-monitoring dietary intake ($R^2=0.15$; $F_{1,52}=8.92$; $P=.004$), such that each additional week of consistent self-monitoring of dietary intake was associated with an average −0.29 (SD 0.10) kg greater weight loss. However, there was no significant association between consistency of physical activity self-monitoring and weight ($P=.57$).

Similarly, there was a significant association between time to disengagement of self-monitoring and weight loss for weight

(R^2=0.18; $F_{1,52}$=11.12; P=.002) and dietary intake (R^2=0.16; $F_{1,52}$=9.63; P=.003), such that each additional week that participants remained engaged with self-monitoring of weight and dietary intake resulted in an average (mean −0.29 kg, SE 0.09 kg and mean −0.29 kg, SE 0.09 kg greater weight loss, respectively). There was no significant association between weight change during the intervention and time to disengagement for self-monitoring of physical activity (P=.58).

Re-engagement

Of the 46 participants who disengaged from self-monitoring of weight, 30 (65%) re-engaged for at least 1 week. Of the 46 participants who disengaged from self-monitoring of dietary intake, 18 (39%) re-engaged for at least 1 week. Of the 35 participants who disengaged from self-monitoring of physical activity, 21 (60%) re-engaged for at least 1 week. Differences in re-engagement between the TECH and TECH+PHONE participants are presented in Table 1; the only significant difference among the groups was for the number of weeks until re-engagement for self-monitoring of physical activity, which was significantly shorter for individuals in the TECH+PHONE compared with individuals in the TECH group (P=.009).

Table 1. Re-engagement in self-monitoring for weight, dietary intake, and physical activity by intervention group (N=54)[a].

Self-monitoring target	TECH[b] (n=27)	TECH+PHONE[c] (n=27)	*P* value
Weight			
Participants who disengaged, n (%)	23 (85)	23 (85)	.99
Participants who re-engaged, n (% of disengaged)	12 (44)	18 (67)	.10
Weeks until re-engagement, mean (SD)	2.8 (0.9)	4 (3.6)	.18
Weeks of re-engagement, mean (SD)	8 (6.2)	5.2 (5.0)	.18
Dietary intake			
Participants who disengaged, n (%)	25 (93)	21 (78)	.25
Participants who re-engaged, n (% of disengaged)	8 (30)	10 (37)	.77
Weeks until re-engagement, mean (SD)	3.1 (1.1)	3.3 (1.8)	.82
Weeks of re-engagement, mean (SD)	3.5 (5.0)	4.2 (3.5)	.71
Physical activity			
Participants who disengaged, n (%)	20 (74)	15 (56)	.15
Participants who re-engaged, n (% of disengaged)	13 (48)	8 (30)	.16
Weeks until re-engagement, mean (SD)	4.6 (2.4)	2.4 (1.1)	.009
Weeks of re-engagement, mean (SD)	4.6 (4.3)	6.5 (7.1)	.41

[a]Most periods of re-engagement were not consecutive.

[b]TECH: technology-based self-monitoring tools.

[c]TECH+PHONE: technology-based self-monitoring tools plus additional phone-based intervention.

Discussion

Principal Findings

This study characterized the patterns of adherence to digital tools for self-monitoring weight and weight-related behaviors within 2 technology-based weight loss programs. As expected, consistency of self-monitoring (defined as the number of weeks that participants met intervention self-monitoring goals) decreased over time. Interestingly, overall levels of consistency and rates of decline appeared to differ by target outcome, such that adherence to self-monitoring of weight and dietary intake declined more substantially than physical activity over the 6-month period. Consistency and time to disengagement (defined as a 2-week period of nonadherence to self-monitoring goals) were also significantly different between treatment groups for dietary intake, such that the provision of additional phone-based interventionist contact led to significantly greater consistency and a longer time to disengagement from self-monitoring of dietary intake; however, the provision of this additional support did not significantly impact consistency or time to disengagement for self-monitoring of weight or physical activity. The results also demonstrated that, for both weight and dietary intake, greater consistency of self-monitoring and longer time to disengagement were significantly associated with greater weight loss; however, consistency and time to disengagement for self-monitoring of physical activity were not associated with weight change. Finally, although most participants disengaged with at least one of the digital tools for self-monitoring at some point during the 6-month intervention, over half of those who disengaged re-engaged in self-monitoring of weight (30/45, 65%) and physical activity (21/35, 60%) for at least 1 week; however, <40% (18/46) re-engaged in self-monitoring of dietary intake.

The results of this study have several important implications. First, patterns of consistency and disengagement from self-monitoring with digital technologies were different by

behavior, with participants demonstrating significantly greater consistency and longer time to disengagement for self-monitoring of physical activity compared with dietary intake and weight. Similar patterns were also seen in a recent study by Butryn et al [35] and may be related to differences in participant burden when using digital self-monitoring tools. Passive self-monitoring tools (such as the wearable device used to measure physical activity in this study) can require substantially less effort compared with more active methods (eg, the use of a website or smartphone app to record dietary intake, which can take up to 15-20 minutes each day [36]). The increased effort and time burden associated with active self-monitoring may thus adversely affect participant consistency and engagement. Indeed, other research has also demonstrated higher rates of engagement for behaviors that are passively versus actively self-monitored [37,38]. In addition, although self-monitoring weight via smart scales does not represent a similar time burden as self-monitoring of dietary intake, engagement may be impacted by other barriers such as negative emotions. Research has shown that daily self-weighing does not cause adverse psychological outcomes [39-41] but that some individuals may avoid the scale because of shame or denial [42]. For example, 1 study found that individuals tended to avoid self-weighing after a day with above their average dietary intake [43]. Thus, self-monitoring of both dietary intake and weight via digital tools may represent higher-burden behaviors compared with self-monitoring of physical activity via digital tools, which may have adversely influenced participant engagement.

Conversely, a certain level of burden may be necessary for behavior change. A study by Turner-McGrievy et al [44] found that participants who used more passive methods for dietary self-monitoring (eg, photo-based apps and wearable bite counters) reported lower rates of habit formation for self-monitoring dietary intake than participants who entered foods or drinks consumed via smartphone apps. Another study by Silberman et al [45] found that participants in a digital weight management program who weighed themselves using traditional, in-home scales and then manually entered those weights into the study website lost more weight over a 12-month period compared with participants who used smart scales to automatically sync weights with the study website. It may be that manual tracking of some self-monitoring outcomes (eg, dietary intake and weight) may be necessary to improve awareness of the target outcome and promote behavior change. Therefore, determining other ways to increase adherence to digital self-monitoring tools, beyond alleviation of burden, may be important for future weight management research.

Accordingly, the results of this study demonstrated that provision of additional interventionist support may also be key for promoting consistency and preventing disengagement with digital self-monitoring tools, especially for higher-burden behaviors such as self-monitoring of dietary intake. These results were consistent with supportive accountability theory, which posits that additional interventionist support can increase adherence to and engagement with mobile health interventions [46]. The intervention phone calls offered to the TECH+PHONE group provided participants with accountability and feedback on progress (coupling encouragement with opportunities to engage in problem solving to overcome barriers experienced in relation to behavior change), which likely supported the higher levels of engagement with dietary self-monitoring tools observed in this study. Although there were no statistically significant differences between TECH and TECH+PHONE participants for consistency and time to disengagement for self-monitoring of physical activity or weight, there was a small or moderate effect size favoring the TECH+PHONE group. It is possible that, as a pilot study, this study was underpowered to detect effects of this size. Thus, future studies with larger samples should investigate whether additional interventionist support can improve engagement with digital tools for self-monitoring weight and physical activity.

On a final note, the results demonstrated suboptimal adherence to self-monitoring across all behavioral targets. Adherence was defined in relation to intervention goals, which included daily self-monitoring of dietary intake, weight, and physical activity (ie, 7 out of 7 days each week). Alternatively, there may be important thresholds below this level that represent clinically meaningful adherence to self-monitoring. Little research, however, has been conducted to establish these thresholds, and evidence from the weight maintenance literature suggests that clinically meaningful thresholds may differ by behavior. Peterson et al [16] found associations between less weight regain and greater consistency of self-monitoring dietary intake when consistency was defined as self-monitoring intake on ≥3 days per week. Brockmann et al [17] found that greater consistency of self-monitoring weight was only significantly associated with less weight regain when defined as ≥6 days of weight self-monitoring. No studies have investigated clinically relevant thresholds for adherence to self-monitoring of physical activity during the maintenance period after initial weight loss programs, and no studies have examined thresholds for adherence to any of these behaviors during initial weight loss programs. Moving forward, it is important for future studies to identify these minimum thresholds to support the development of clinically meaningful definitions of adherence.

Strengths and Limitations

This study was the first to examine the effect of providing additional phone-based interventionist contact on engagement with digital tools for the self-monitoring of dietary intake, physical activity, and weight, independently. A major strength of this study was the use of objective data from the Fitbit Zip digital pedometers and Fitbit Aria smart scales, which reduced the risk of retrospective recall bias. Moreover, issues related to missing data from nonreturn of paper self-monitoring records were minimized because of the collection of self-monitoring data in real time, directly from these devices and the Fitbit website or smartphone app.

This study also had a number of important limitations. First, disengagement was defined as a 2-week period in which intervention self-monitoring goals were not met; however, a prolonged period of adherence to self-monitoring before disengagement does not ensure that participants were adequately *engaged* with the digital self-monitoring technology or even fully adherent to the full process of self-monitoring (ie,

observing change in a target behavior to assess progress toward a goal). For example, a participant could have stepped on the scale for the sake of the study without looking at their weight or have worn the activity monitor without paying attention to their step count. Unfortunately, more precise data on engagement were not collected during this study; future studies should attempt to better disentangle the constructs of adherence versus engagement. Second, as the parent study was conceptualized as a pilot trial, the sample size of both the TECH and TECH+PHONE groups were small and predominately consisted of women who identified as White and non-Hispanic. Thus, the results may not generalize to broader populations, including groups disproportionally burdened by obesity (eg, adults who identify as Black and Hispanic) [47]. Additional research on engagement with digital tools in larger and more diverse samples, especially those including larger numbers of men and individuals from backgrounds historically underrepresented in clinical research, is critical. Third, the parent trial did not include measures assessing participants' history of or comfort with using digital self-monitoring tools, which could have influenced adherence to self-monitoring in this study. However, even though the data on the exact number of technical errors are not available, few technical issues were experienced and were easily addressed. Finally, this study did not investigate week-to-week associations between changes in weight, dietary intake, and physical activity and later changes in self-monitoring behaviors; however, some emerging research suggests that there may be important bidirectional associations between self-monitoring and changes in weight and weight-related behaviors [48]. Future studies with larger samples should characterize the nature of these associations across time.

Conclusions

This study examined consistency with and disengagement from digital tools for self-monitoring of weight and weight-related behaviors in a technology-based weight management intervention. The results demonstrated that patterns of consistency and disengagement varied by behavior, such that participants demonstrated greater consistency and longer time to disengagement from self-monitoring of physical activity compared with weight and dietary intake. Furthermore, the results demonstrated that the provision of additional phone-based interventionist contact led to greater consistency in self-monitoring of dietary intake. Importantly, greater consistency and longer time to disengagement from self-monitoring of dietary intake and weight were associated with greater weight loss during the 6-month intervention. Given the key role of self-monitoring for weight loss and weight loss maintenance, future studies should replicate these results in larger, more generalizable samples and focus on identifying novel methods to improve individuals' engagement with digital self-monitoring tools.

Conflicts of Interest

None declared.

References

1. Butryn ML, Webb V, Wadden TA. Behavioral treatment of obesity. Psychiatr Clin North Am 2011;34(4):841-859 [FREE Full text] [doi: 10.1016/j.psc.2011.08.006] [Medline: 22098808]
2. Wing RR. Behavioral approaches to the treatment of obesity. In: Bray GA, Bouchard C, editors. Handbook of obesity: clinical applications. 2nd edition. New York, NY: Marcel Dekker; 2004:147-167.
3. Burke LE, Wang J, Sevick MA. Self-monitoring in weight loss: a systematic review of the literature. J Am Diet Assoc 2011;111(1):92-102 [FREE Full text] [doi: 10.1016/j.jada.2010.10.008] [Medline: 21185970]
4. Patel ML, Wakayama LN, Bennett GG. Self-monitoring via digital health in weight loss interventions: a systematic review among adults with overweight or obesity. Obesity (Silver Spring) 2021;29(3):478-499. [doi: 10.1002/oby.23088] [Medline: 33624440]
5. Tate DF, Jackvony EH, Wing RR. A randomized trial comparing human e-mail counseling, computer-automated tailored counseling, and no counseling in an Internet weight loss program. Arch Intern Med 2006;166(15):1620-1625. [doi: 10.1001/archinte.166.15.1620] [Medline: 16908795]
6. Krukowski RA, Harvey-Berino J, Bursac Z, Ashikaga T, West DS. Patterns of success: online self-monitoring in a web-based behavioral weight control program. Health Psychol 2013;32(2):164-170 [FREE Full text] [doi: 10.1037/a0028135] [Medline: 22545978]
7. Burke LE, Conroy MB, Sereika SM, Elci OU, Styn MA, Acharya SD, et al. The effect of electronic self-monitoring on weight loss and dietary intake: a randomized behavioral weight loss trial. Obesity (Silver Spring) 2011;19(2):338-344 [FREE Full text] [doi: 10.1038/oby.2010.208] [Medline: 20847736]
8. Greaney ML, Sprunck-Harrild K, Bennett GG, Puleo E, Haines J, Viswanath KV, et al. Use of email and telephone prompts to increase self-monitoring in a Web-based intervention: randomized controlled trial. J Med Internet Res 2012;14(4):e96 [FREE Full text] [doi: 10.2196/jmir.1981] [Medline: 22842775]
9. Burke LE, Swigart V, Warziski Turk M, Derro N, Ewing LJ. Experiences of self-monitoring: successes and struggles during treatment for weight loss. Qual Health Res 2009;19(6):815-828 [FREE Full text] [doi: 10.1177/1049732309335395] [Medline: 19365099]

10. Burke LE, Styn MA, Sereika SM, Conroy MB, Ye L, Glanz K, et al. Using mHealth technology to enhance self-monitoring for weight loss: a randomized trial. Am J Prev Med 2012;43(1):20-26 [FREE Full text] [doi: 10.1016/j.amepre.2012.03.016] [Medline: 22704741]
11. Wang J, Sereika SM, Chasens ER, Ewing LJ, Matthews JT, Burke LE. Effect of adherence to self-monitoring of diet and physical activity on weight loss in a technology-supported behavioral intervention. Patient Prefer Adherence 2012;6:221-226 [FREE Full text] [doi: 10.2147/PPA.S28889] [Medline: 22536058]
12. Ross KM, Wing RR. "Memory bias" for recall of experiences during initial weight loss is affected by subsequent weight loss outcome. J Behav Med 2018;41:130-137 [FREE Full text] [doi: 10.1007/s10865-017-9896-1] [Medline: 29080115]
13. Crane MM, Gavin K, Wolfson J, Linde JA. How accurate are recalls of self-weighing frequency? Data from a 24-month randomized trial. Obesity (Silver Spring) 2018;26(8):1296-1302 [FREE Full text] [doi: 10.1002/oby.22239] [Medline: 30070045]
14. Burke LE, Sereika SM, Music E, Warziski M, Styn MA, Stone A. Using instrumented paper diaries to document self-monitoring patterns in weight loss. Contemp Clin Trials 2008;29(2):182-193 [FREE Full text] [doi: 10.1016/j.cct.2007.07.004] [Medline: 17702667]
15. Ross KM, Eastman A, Wing RR. Accuracy of self-report versus objective smart-scale weights during a 12-week weight management intervention. Obesity (Silver Spring) 2019;27(3):385-390 [FREE Full text] [doi: 10.1002/oby.22400] [Medline: 30703282]
16. Peterson ND, Middleton KR, Nackers LM, Medina KE, Milsom VA, Perri MG. Dietary self-monitoring and long-term success with weight management. Obesity (Silver Spring) 2014;22(9):1962-1967 [FREE Full text] [doi: 10.1002/oby.20807] [Medline: 24931055]
17. Brockmann AN, Eastman A, Ross KM. Frequency and consistency of self-weighing to promote weight-loss maintenance. Obesity (Silver Spring) 2020;28(7):1215-1218 [FREE Full text] [doi: 10.1002/oby.22828] [Medline: 32437055]
18. Shin G, Feng Y, Jarrahi MH, Gafinowitz N. Beyond novelty effect: a mixed-methods exploration into the motivation for long-term activity tracker use. JAMIA Open 2019;2(1):62-72 [FREE Full text] [doi: 10.1093/jamiaopen/ooy048] [Medline: 31984346]
19. Lee H, Lee Y. A look at wearable abandonment. In: Proceedings of 18th IEEE International Conference on Mobile Data Management. 2017 Presented at: MDM '17; May 29-June 1, 2017; Daejeon, South Korea. [doi: 10.1109/mdm.2017.70]
20. Shih PC, Han K, Poole ES, Rosson MB, Carroll JM. Use and adoption challenges of wearable activity trackers. In: Proceedings of iConference. 2015 Presented at: iConference '15; March 24-27, 2015; Newport Beach, CA.
21. Fausset CB, Mitzner TL, Price CE, Jones BD, Fain WB, Rogers WA. Older adults' use of and attitudes toward activity monitoring technologies. Proc Hum Factors Ergon Soc Annu Meet 2013;57(1):1683-1687 [FREE Full text] [doi: 10.1177/1541931213571374] [Medline: 31263349]
22. Laing BY, Mangione CM, Tseng CH, Leng M, Vaisberg E, Mahida M, et al. Effectiveness of a smartphone application for weight loss compared with usual care in overweight primary care patients: a randomized, controlled trial. Ann Intern Med 2014;161(10 Suppl):S5-12 [FREE Full text] [doi: 10.7326/M13-3005] [Medline: 25402403]
23. Serrano KJ, Coa KI, Yu M, Wolff-Hughes DL, Atienza AA. Characterizing user engagement with health app data: a data mining approach. Transl Behav Med 2017;7(2):277-285 [FREE Full text] [doi: 10.1007/s13142-017-0508-y] [Medline: 28616846]
24. Lemstra M, Bird Y, Nwankwo C, Rogers M, Moraros J. Weight loss intervention adherence and factors promoting adherence: a meta-analysis. Patient Prefer Adherence 2016;10:1547-1559 [FREE Full text] [doi: 10.2147/PPA.S103649] [Medline: 27574404]
25. Tate DF, Jackvony EH, Wing RR. Effects of Internet behavioral counseling on weight loss in adults at risk for type 2 diabetes: a randomized trial. JAMA 2003;289(14):1833-1836. [doi: 10.1001/jama.289.14.1833] [Medline: 12684363]
26. Ross KM, Wing RR. Impact of newer self-monitoring technology and brief phone-based intervention on weight loss: a randomized pilot study. Obesity (Silver Spring) 2016;24(8):1653-1659 [FREE Full text] [doi: 10.1002/oby.21536] [Medline: 27367614]
27. D'Zurilla TJ, Goldfried MR. Problem solving and behavior modification. J Abnorm Psychol 1971;78(1):107-126. [doi: 10.1037/h0031360] [Medline: 4938262]
28. Cullen KW, Baranowski T, Smith SP. Using goal setting as a strategy for dietary behavior change. J Am Diet Assoc 2001;101(5):562-566. [doi: 10.1016/S0002-8223(01)00140-7] [Medline: 11374350]
29. Foreyt JP, Goodrick GK. Evidence for success of behavior modification in weight loss and control. Ann Intern Med 1993;119(7 Pt 2):698-701. [doi: 10.7326/0003-4819-119-7_part_2-199310011-00014] [Medline: 8363200]
30. Wing RR, Jeffery RW. Benefits of recruiting participants with friends and increasing social support for weight loss and maintenance. J Consult Clin Psychol 1999;67(1):132-138. [doi: 10.1037//0022-006x.67.1.132] [Medline: 10028217]
31. Marlatt GA, Witkiewitz K. Relapse prevention for alcohol and drug problems. In: Marlatt GA, Donovan DM, editors. Relapse prevention: maintenance strategies in the treatment of addictive behaviors. 2nd edition. New York, NY: The Guilford Press; 2005:1-44.

32. 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;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
33. Fitabase research device data and analytics. Fitabase. 2021. URL: https://www.fitabase.com/ [accessed 2020-12-11]
34. Turner-McGrievy GM, Wilcox S, Boutté A, Hutto BE, Singletary C, Muth ER, et al. The dietary intervention to enhance tracking with mobile devices (DIET mobile) study: a 6-month randomized weight loss trial. Obesity (Silver Spring) 2017;25(8):1336-1342 [FREE Full text] [doi: 10.1002/oby.21889] [Medline: 28600833]
35. Butryn ML, Godfrey KM, Martinelli MK, Roberts SR, Forman EM, Zhang F. Digital self-monitoring: does adherence or association with outcomes differ by self-monitoring target? Obes Sci Pract 2019;6(2):126-133 [FREE Full text] [doi: 10.1002/osp4.391] [Medline: 32313670]
36. Harvey J, Krukowski R, Priest J, West D. Log often, lose more: electronic dietary self-monitoring for weight loss. Obesity (Silver Spring) 2019;27(3):380-384 [FREE Full text] [doi: 10.1002/oby.22382] [Medline: 30801989]
37. Patel ML, Brooks TL, Bennett GG. Consistent self-monitoring in a commercial app-based intervention for weight loss: results from a randomized trial. J Behav Med 2020;43:391-401. [doi: 10.1007/s10865-019-00091-8] [Medline: 31396820]
38. Cordeiro F, Epstein DA, Thomaz E, Bales E, Jagannathan A, Abowd GD, et al. Barriers and negative nudges: exploring challenges in food journaling. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. 2015 Presented at: Association for Computing Machinery Conference on Human Factors in Computing Systems; April 18 - 23, 2015; Seoul Republic of Korea p. 1159-1162 URL: http://europepmc.org/abstract/MED/26894233 [doi: 10.1145/2702123.2702155]
39. Steinberg DM, Tate DF, Bennett GG, Ennett S, Samuel-Hodge C, Ward DS. Daily self-weighing and adverse psychological outcomes: a randomized controlled trial. Am J Prev Med 2014;46(1):24-29 [FREE Full text] [doi: 10.1016/j.amepre.2013.08.006] [Medline: 24355668]
40. Zheng Y, Klem ML, Sereika SM, Danford CA, Ewing LJ, Burke LE. Self-weighing in weight management: a systematic literature review. Obesity (Silver Spring) 2015;23(2):256-265 [FREE Full text] [doi: 10.1002/oby.20946] [Medline: 25521523]
41. Wing RR, Tate DF, Gorin AA, Raynor HA, Fava JL, Machan J. STOP regain: are there negative effects of daily weighing? J Consult Clin Psychol 2007;75(4):652-656. [doi: 10.1037/0022-006X.75.4.652] [Medline: 17663619]
42. Hartmann-Boyce J, Boylan AM, Jebb SA, Aveyard P. Experiences of self-monitoring in self-directed weight loss and weight loss maintenance: systematic review of qualitative studies. Qual Health Res 2019;29(1):124-134. [doi: 10.1177/1049732318784815] [Medline: 29984630]
43. Tanenbaum ML, Ross KM, Wing RR. Overeat today, skip the scale tomorrow: an examination of caloric intake predicting nonadherence to daily self-weighing. Obesity (Silver Spring) 2016;24(11):2341-2343 [FREE Full text] [doi: 10.1002/oby.21650] [Medline: 27619935]
44. Turner-McGrievy GM, Yang CH, Monroe C, Pellegrini C, West DS. Is burden always bad? Emerging low-burden approaches to mobile dietary self-monitoring and the role burden plays with engagement. J Technol Behav Sci 2021;6(3):447-455. [doi: 10.1007/s41347-021-00203-9]
45. Silberman JM, Kaur M, Sletteland J, Venkatesan A. Outcomes in a digital weight management intervention with one-on-one health coaching. PLoS One 2020;15(4):e0232221 [FREE Full text] [doi: 10.1371/journal.pone.0232221] [Medline: 32353035]
46. Mohr DC, Cuijpers P, Lehman K. Supportive accountability: a model for providing human support to enhance adherence to eHealth interventions. J Med Internet Res 2011;13(1):e30 [FREE Full text] [doi: 10.2196/jmir.1602] [Medline: 21393123]
47. Hales CM, Carroll MD, Fryar CD, Ogden CL. Prevalence of obesity and severe obesity among adults: United States, 2017-2018. NCHS Data Brief 2020(360):1-8 [FREE Full text] [Medline: 32487284]
48. Goldstein SP, Goldstein CM, Bond DS, Raynor HA, Wing RR, Thomas JG. Associations between self-monitoring and weight change in behavioral weight loss interventions. Health Psychol 2019;38(12):1128-1136 [FREE Full text] [doi: 10.1037/hea0000800] [Medline: 31556659]

Abbreviations

REDCap: Research Electronic Data Capture
TECH: technology-based self-monitoring tools
TECH+PHONE: technology-based self-monitoring tools plus additional phone-based intervention

Edited by T Leung; submitted 15.09.21; peer-reviewed by M Stansbury, R Krukowski; comments to author 11.11.21; revised version received 23.12.21; accepted 09.01.22; published 18.02.22.

Please cite as:
Carpenter CA, Eastman A, Ross KM
Consistency With and Disengagement From Self-monitoring of Weight, Dietary Intake, and Physical Activity in a Technology-Based Weight Loss Program: Exploratory Study
JMIR Form Res 2022;6(2):e33603
URL: https://formative.jmir.org/2022/2/e33603
doi:10.2196/33603
PMID:35179513

Original Paper

Asynchronous Remote Assessment for Cognitive Impairment: Reliability Verification of the Neurotrack Cognitive Battery

Jennifer Rae Myers[1], DPhil; Jordan M Glenn[1,2], DPhil; Erica N Madero[1], MSci, MPH; John Anderson[1], BSc; Rachel Mak-McCully[1], DPhil; Michelle Gray[2], DPhil; Joshua L Gills[2], MSc; John E Harrison[3,4,5], DPhil

[1]Neurotrack Technologies Inc, Redwood City, CA, United States

[2]University of Arkansas, Fayetteville, AR, United States

[3]Metis Cognition Ltd, Kilmington Common, United Kingdom

[4]Alzheimer Center, VU Medical Center, Vrije Universiteit Amsterdam, Amsterdam, Netherlands

[5]Institute of Psychiatry, Psychology & Neuroscience, King's College, London, United Kingdom

Corresponding Author:
Jennifer Rae Myers, DPhil
Neurotrack Technologies Inc
399 Bradford St. #101
Redwood City, CA, 21216
United States
Phone: 1 650 549 8566
Email: jennifer@neurotrack.com

Abstract

Background: As evidenced by the further reduction in access to testing during the COVID-19 pandemic, there is an urgent, growing need for remote cognitive assessment for individuals with cognitive impairment. The Neurotrack Cognitive Battery (NCB), our response to this need, was evaluated for its temporal reliability and stability as part of ongoing validation testing.

Objective: The aim of this study is to assess the temporal reliability of the NCB tests (5 total) across a 1-week period and to determine the temporal stability of these measures across 3 consecutive administrations in a single day.

Methods: For test-retest reliability, a range of 29-66 cognitively healthy participants (ages 18-68 years) completed each cognitive assessment twice, 1 week apart. In a separate study, temporal stability was assessed using data collected from 31 different cognitively healthy participants at 3 consecutive timepoints in a single day.

Results: Correlations for the assessments were between 0.72 and 0.83, exceeding the standard acceptable threshold of 0.70 for temporal reliability. Intraclass correlations ranged from 0.60 to 0.84, indicating moderate to good temporal stability.

Conclusions: These results highlight the NCB as a brief, easy-to-administer, and reliable assessment for remote cognitive testing. Additional validation research is underway to determine the full magnitude of the clinical utility of the NCB.

(JMIR Form Res 2022;6(2):e34237) doi:10.2196/34237

KEYWORDS

cognition; screening; remote testing; psychometric; challenge; validation; assessment; impairment; access; reliability; stability; testing; utility

Introduction

Background

Remote cognitive assessment, through the use of digital tools, represents an efficient means for individuals to assess their levels of function, without needing to visit an in-person clinic. These tools allow for the “benchmarking” of current levels of function, as well as the detection of change over time. Monitoring changes in performance over time enables the detection of progressive decline that might be indicative of neurological disease. It also enables the detection of improvement from interventions targeting lifestyle changes and modifiable risk factors aimed at improving cognitive health [1-3]. Remote assessment also plays a role in areas such as screening for clinical trial participation, postmarketing surveillance of licensed therapies, and the collection of data in large prospectively ascertained cohorts. Clinical research has yielded a number of candidate measures for indexing key cognitive skills, with experts largely agreeing that, in studies

of individuals living with Alzheimer disease, assessments should measure attention, memory, and executive function [4,5].

Development of the Neurotrack Cognitive Battery

Although there are other brief and computerized cognitive assessments available, such as the Cambridge Neuropsychological Test Automated Battery (CANTAB), Savonix, and BrainCheck, the Neurotrack Cognitive Battery (NCB) offers several distinct platform features to address common concerns associated with cognitive and remote assessment. These include web camera capability, objective scoring via algorithms, and the ability to be administered without the need for a trained health care professional. Regarding assessment design, there are several recommended properties of an ideal cognitive assessment tool, which includes the need to assess the full range of relevant cognitive processes, be sensitive to aging and cognitive deficits, contain equivalent versions for repeat administration, have a reasonable testing duration, and have good reliability and validity metrics [6]. The application of recommended test selection properties led us to develop digital versions of Part B of the Trail Making Test [7] as a measure of executive function, as well as a computerized novel variant of the Digit Symbol Substitution (DSST) paradigm. The DSST enjoys the virtues of brevity and reliability, as well as being a well-known general measure of cognitive function that is sensitive to subtle changes in cognition [8]. To further index these functions, we selected the Erikson flanker task and the go/no-go test [9]. To extend the coverage of the assessment to include episodic memory, we also included a paired associate learning task in which individuals are required to pair shopping list items with the associated prices. These assessments were combined with a previously validated associative and recognition memory visual paired comparison paradigm [10]. This task is based on research conducted by Zola et al [11] and utilizes eye tracking to determine novelty preference as an index of memory.

Measures of cognitive function must be reliable, sensitive, and valid to show meaningful change over time [12]. In the early stages of test development, we have focused on ensuring the scientific validity of the Neurotrack assessments through standard psychometric testing. Thus, the aims of our research were to (1) assess the temporal reliability of the NCB tests across a 1-week period and (2) determine the temporal stability of these measures across 3 consecutive administrations in a single day.

Methods

Participants

Temporal Reliability

Participants were workers recruited through Amazon Mechanical Turk (MTurk) and Prolific, crowdsourcing websites used for research recruitment and testing. The use of MTurk and Prolific has shown to be comparable to traditional research methods and allow for greater access to hard-to-reach and diverse populations [13,14]. Up to 150 subjectively cognitively healthy participants in the United States were recruited for each assessment separately. Individuals who successfully completed an assessment at time point 1 were granted access to retake the assessment 1 week later at time point 2. Participants were compensated up to US $2.85 for each assessment completed. Participant characteristics for each assessment are outlined in Table 1.

Table 1. Participant characteristics for temporal reliability.

Characteristics	NCB[a] assessment and participants				
	Path points, n=66	Symbol match, n=29	Arrow match, n=46	Light reaction, n=46	Item price, n=51
Age (years), mean (SD)	32.82 (7.56)	28.97 (9.08)	33.39 (7.73)	39.50 (11.34)	36.53 (10.18)
Sex, n (%)					
Female	15 (23)	12 (41)	11 (24)	19 (41)	14 (27)
Male	51 (77)	17 (59)	35 (76)	27 (59)	37 (73)
Race, n (%)					
Nonwhite	12 (18)	8 (28)	11 (24)	8 (17)	11 (22)
White	54 (82)	21 (72)	35 (76)	38 (83)	40 (78)

[a]NCB: Neurotrack Cognitive Battery.

Temporal Stability

Potential participants were recruited through a research interest listserv and by word of mouth; these individuals were selected from a different pool than the original group, with no overlap present between the two. A total of 55 individuals who were subjectively cognitively healthy expressed interest in participating in the study. Of the 55 individuals, 31 completed the study in its entirety. The mean age of the participants was 51 (SD 17.61) years. Regarding other participant characteristics, 41 of the 55 participants (75%) had a college degree, 41 (75%) were female, and 21 (38%) identified as a person of color. Participants were asked to complete the entire NCB 3 consecutive times in a single day, which took approximately 60 minutes. Participants who completed the study received a US $50 electronic gift card.

Measures and Procedure

As previously mentioned, the following measures were selected for use based on their relative ease of administration, brevity, and tendency to represent reliable measures of cognition:

- Path points is a 2-minute assessment of executive function. This task requires participants to connect dots alternating between a number and a letter in ascending order. The primary assessment score is based on completion time.
- Symbol match is a 2-minute assessment of processing speed. Participants are instructed to determine whether 2 symbols are equal or unequal based on a legend with 9 number/symbol pairs. Participants must complete as many trials as they can in 2 minutes. Primary scores are based on accuracy and speed.
- Arrow match is a 3-minute assessment of attention. This task requires participants to indicate which direction the center arrow is pointing (left or right) among 4 distractor arrows. Primary scores are based on accuracy and speed.
- Light reaction is a 3-minute assessment of inhibition. This task requires participants to respond when they see a green light and refrain from responding when they see a red light. The primary assessment score is based on accuracy and speed.
- Item price is a 3-minute assessment of associative learning. This task requires participants to learn the prices of various produce items (eg, bananas, carrots) and identify the correct price during the recognition trials. Primary scores are based on accuracy.

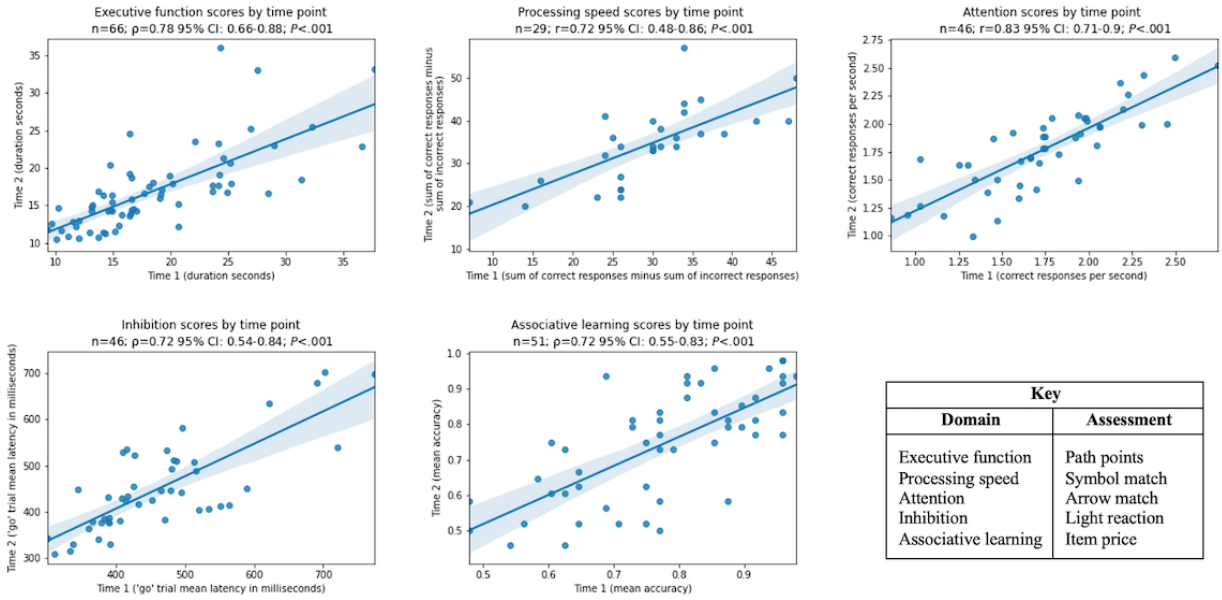
For temporal reliability and stability, data were collected separately, using new participants for each, in an unsupervised remote setting. Participants were instructed to complete the assessments when feeling rested and in a private room to optimize focus. Participants were also instructed to complete the assessments using a laptop or desktop computer with a physical keyboard, mouse, and camera, with their device charged to at least 20%, and on a stable internet connection. Study protocols were approved by the University of Arkansas Institutional Review Board and participants provided informed consent prior to study enrollment.

Results

Temporal Reliability

For each assessment, the test-retest reliability of the scores from time point 1 and 2 were assessed using both Pearson and Spearman correlations (Figure 1). In instances where a visual inspection of the data suggested a general monotonic relationship, the Spearman correlation coefficient was selected. Outliers, defined as scores more or less than 5 standard deviations from the mean, were removed from the final analyses. Correlations for the assessments ranged from 0.72 to 0.83, indicating acceptable to good test-retest reliability of the NCB.

Figure 1. Scatterplots of test-retest scores for the Neurotrack Cognitive Battery (NCB) assessments. The Pearson correlation coefficient is represented by r. The Spearman correlation coefficient represented by the Greek letter ρ.



Key	
Domain	**Assessment**
Executive function	Path points
Processing speed	Symbol match
Attention	Arrow match
Inhibition	Light reaction
Associative learning	Item price

Temporal Stability

Temporal stability was examined by calculating estimates of within-subject standard deviation (s_w) and intraclass correlation coefficients (ICCs) for each assessment. The s_w is used to quantify measurement error in repeated measurements as a single overall measure. Results are outlined in Table 2.

Table 2. Within-subject mean and standard deviation (s_w) and Kendall rank correlation coefficient (Kendall τ) values for Neurotrack Cognitive Battery (NCB) assessments (n=31 participants).

NCB assessment	Within-subject mean	s_w	Kendall's τ	95% CI
Path points[a]	13.47	0.27	–0.42	2.48-3.57
Symbol match[b]	29.80	4.32	–0.02	3.54-5.10
Arrow match[c]	1.51	0.15	0.15	0.12-0.17
Light reaction[d]	476.69	52.25	–0.11	42.53-61.96
Item price[e]	0.74	0.08	–0.20	0.07-1.00

[a]Score is measured in seconds.

[b]Score is the sum of correct response minus the sum of incorrect responses.

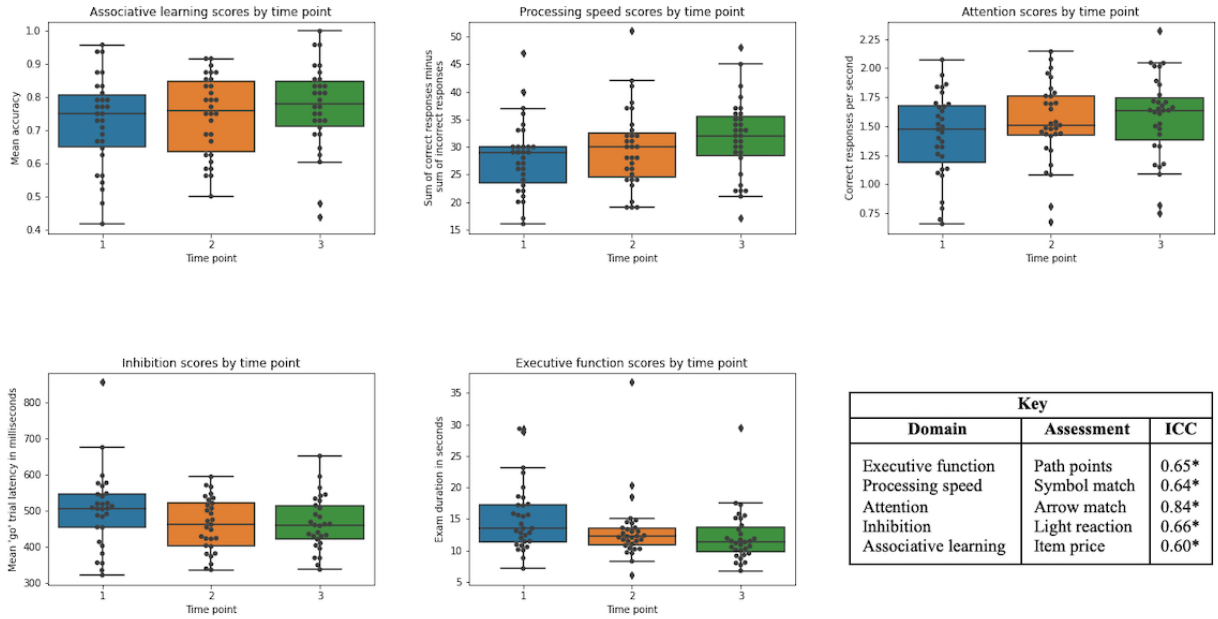
[c]Score is the number of correct responses per second.

[d]Score is measured in milliseconds.

[e]Score is the mean accuracy of responses.

It should be noted that s_w is based on the assumption that the s_w is independent from the within-subject mean (assessed using the Kendall rank correlation coefficient [Kendall τ]). Thus, for the path points (executive function) and item price (associative learning) assessments, there is the possibility of overestimation or underestimation in how close scores are to the mean. ICCs were also calculated to assess variation due to measurement error. The ICC (2,1) was selected as it is the recommended ICC form to use for test-retest metrics [15,16]. Correlations for the assessments ranged from 0.60 to 0.84, indicating moderate to good reliability. Boxplots for each assessment are depicted in Figure 2.

Figure 2. Boxplots representing scores from each assessment separated by time point and intraclass correlation coefficients (ICCs). The asterisk indicates *P*<.001. n=31 participants.



Discussion

Principal Findings

The results of this initial validation study indicate that the NCB is a reliable set of assessments, measuring key cognitive domains as largely accepted by the neuropsychological field. The examination of temporal reliability yielded test-retest reliability correlations which exceed the standard psychometry threshold of 0.7 for acceptable temporal reliability [17]. The NCB also demonstrated favorable temporal stability by traditional standards given the ICC values were greater than 0.5 [15].

Limitations

The primary aim of developing the NCB has been to provide an instrument capable of reliably and remotely assessing cognition. We have sought to imbue the NCB with characteristics that facilitate the evaluation of cognitive change in group studies, as well as at the individual level. Such an approach has been advocated for some time [18] and emphasis has been placed on the need for the use of reliable measures

[12]. Although the sample was diverse, the limitations of this study are the small group sample sizes, the mean age of the participants, and the use of convenience sampling to obtain participants, which impacts the generalizability of the results. All but one assessment sample met the recommended sample size of at least 30 participants [15]; nonetheless, the use of larger sample sizes and older sample populations, as well as probability sampling methods in future validation studies is warranted.

Conclusion

This study has demonstrated that the NCB can be successfully delivered reliably and remotely, with promising clinical implications. As evidenced by the COVID-19 pandemic, there is a critical need for feasible and valid assessments for remote testing. As it has been well established that the lack of access to cognitive assessments can result in delayed diagnoses, less effective treatment, and missed lifestyle and other health-related interventions [19], the NCB addresses this critical gap as a brief, reliable, and easy-to-administer assessment battery. Additional research regarding psychometric properties, usability, and feasibility is currently underway to determine the magnitude of the clinical validity of the NCB as part of a clinician's diagnosis process.

Acknowledgments

The authors would like to thank the participants for their time. This work was supported by funding from Neurotrack Technologies Inc.

Conflicts of Interest

JRM, JMG, ENM, JA, and RM report income and equity received from Neurotrack Technologies Inc. JEH reports personal fees from AstraZeneca PLC, AXON Neuroscience SE, Sio Gene Therapies Inc, Biogen Idec Ltd, Boehringer Ingelheim International GmbH, Signant Health, CRF Health, Eisai Co Ltd, Eli Lilly and Company, Games for Health Europe, Heptares Therapeutics Ltd, Kaasa Health GmbH, MyCognition, Neurocog, Neurotrack Technologies Inc, Novartis International AG, Nutricia, Vivoryon Therapeutics AG, Regeneron Pharmaceuticals Inc, Sanofi SA, Servier Laboratories, Takeda Pharmaceutical Company Ltd, vTv Therapeutics Inc, H. Lundbeck A/S, Compass Pathways PLC, C4X Discovery, Cognition Therapeutics Inc, AlzeCure Pharma AB, Recognify Life Sciences Inc, BlackThorn Therapeutics, Winterlight Labs, Rodin Therapeutics Inc, Lysosome Therapeutics Inc, Syndesi Therapeutics SA, Vivoryon Therapeutics NV, Neurodyn Inc, Aptinyx Inc, Athira Pharma Inc, EIP Pharma Inc, Cerecin Inc, Neurocentria Inc, CuraSen Therapeutics Inc, Biosplice Therapeutics Inc, Cognition Therapeutics Inc, ReMynd NV, Ki-Elements, and the National Health Service, outside the submitted work. MG and JLG have no conflicts of interest to disclose.

References

1. Bott N, Kumar S, Krebs C, Glenn JM, Madero EN, Juusola JL. A remote intervention to prevent or delay cognitive impairment in older adults: design, recruitment, and baseline characteristics of the Virtual Cognitive Health (VC Health) study. JMIR Res Protoc 2018 Aug 13;7(8):e11368 [FREE Full text] [doi: 10.2196/11368] [Medline: 30104186]
2. Glenn J, Madero EN, Gray M, Fuseya N, Ikeda M, Kawamura T, et al. Engagement with a digital platform for multimodal cognitive assessment and multidomain intervention in a Japanese population: pilot, quasi-experimental, longitudinal study. JMIR Mhealth Uhealth 2019 Oct 25;7(10):e15733 [FREE Full text] [doi: 10.2196/15733] [Medline: 31654567]
3. Ngandu T, Lehtisalo J, Solomon A, Levälahti E, Ahtiluoto S, Antikainen R, et al. A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): a randomised controlled trial. Lancet 2015 Jun 06;385(9984):2255-2263. [doi: 10.1016/S0140-6736(15)60461-5] [Medline: 25771249]
4. Ritchie K, Ropacki M, Albala B, Harrison J, Kaye J, Kramer J, et al. Recommended cognitive outcomes in preclinical Alzheimer's disease: consensus statement from the European Prevention of Alzheimer's Dementia project. Alzheimers Dement 2017 Feb;13(2):186-195. [doi: 10.1016/j.jalz.2016.07.154] [Medline: 27702619]
5. Vellas B, Andrieu S, Sampaio C, Coley N, Wilcock G, European Task Force Group. Endpoints for trials in Alzheimer's disease: a European task force consensus. Lancet Neurol 2008 May;7(5):436-450. [doi: 10.1016/S1474-4422(08)70087-5] [Medline: 18420157]
6. Whitehouse PJ. Harmonization of Dementia Drug Guidelines (United States and Europe): a report of the International Working Group for the Harmonization for Dementia Drug Guidelines. Alzheimer Dis Assoc Disord 2000;14 Suppl 1:S119-S122. [doi: 10.1097/00002093-200000001-00018] [Medline: 10850740]
7. Lezak M. Neuropsychological Assessment, 3rd Edition. Oxford, United Kingdom: Oxford University Press; 1995.
8. Jaeger J. Digit Symbol Substitution Test: the case for sensitivity over specificity in neuropsychological testing. J Clin Psychopharmacol 2018 Oct;38(5):513-519 [FREE Full text] [doi: 10.1097/JCP.0000000000000941] [Medline: 30124583]
9. Donders F. On the speed of mental processes. Acta Psychol (Amst) 1969;30:412-431. [doi: 10.1016/0001-6918(69)90065-1] [Medline: 5811531]
10. Gills J, Glenn JM, Madero EN, Bott NT, Gray M. Validation of a digitally delivered visual paired comparison task: reliability and convergent validity with established cognitive tests. Geroscience 2019 Aug;41(4):441-454 [FREE Full text] [doi: 10.1007/s11357-019-00092-0] [Medline: 31463649]

11. Crutcher MD, Calhoun-Haney R, Manzanares CM, Lah JJ, Levey AI, Zola SM. Eye tracking during a visual paired comparison task as a predictor of early dementia. Am J Alzheimers Dis Other Demen 2009 Feb;24(3):258-266 [FREE Full text] [doi: 10.1177/1533317509332093] [Medline: 19246573]
12. Harrison JE. Measuring the mind: detecting cognitive deficits and measuring cognitive change in patients with depression. In: McIntyre RS, editor. Cognitive Impairment in Major Depressive Disorder. Cambridge, United Kingdom: Cambridge University Press; 2016.
13. Casler K, Bickel L, Hackett E. Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. Comput Human Behav 2013;29(6):2156-2160 [FREE Full text] [doi: 10.1016/j.chb.2013.05.009]
14. 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]
15. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016 Jun;15(2):155-163 [FREE Full text] [doi: 10.1016/j.jcm.2016.02.012] [Medline: 27330520]
16. Liljequist D, Elfving B, Skavberg Roaldsen K. Intraclass correlation - A discussion and demonstration of basic features. PLoS One 2019;14(7):e0219854 [FREE Full text] [doi: 10.1371/journal.pone.0219854] [Medline: 31329615]
17. Kline P. The Handbook of Psychological Testing, 2nd Edition. Hove, United Kingdom: Psychology Press; 2000.
18. Harrison J, Maruff P. Measuring the mind: assessing cognitive change in clinical drug trials. Expert Rev Clin Pharmacol 2008 Jul;1(4):471-473. [doi: 10.1586/17512433.1.4.471] [Medline: 24410548]
19. Geddes MR, O'Connell ME, Fisk JD, Gauthier S, Camicioli R, Ismail Z, Alzheimer Society of Canada Task Force on Dementia Care Best Practices for COVID‐19. Remote cognitive and behavioral assessment: report of the Alzheimer Society of Canada Task Force on dementia care best practices for COVID-19. Alzheimers Dement (Amst) 2020;12(1):e12111 [FREE Full text] [doi: 10.1002/dad2.12111] [Medline: 32999916]

Abbreviations

CANTAB: Cambridge Neuropsychological Test Automated Battery
DSST: Digit Symbol Substitution
ICC: intraclass correlation coefficient
Kendall τ: Kendall rank correlation coefficient
MTurk: Amazon Mechanical Turk
NCB: Neurotrack Cognitive Battery
s_w: within-subject standard deviation

Edited by A Mavragani; submitted 13.10.21; peer-reviewed by M O'Connell, Y Yu; comments to author 03.11.21; revised version received 07.12.21; accepted 30.12.21; published 18.02.22.

Please cite as:
Myers JR, Glenn JM, Madero EN, Anderson J, Mak-McCully R, Gray M, Gills JL, Harrison JE
Asynchronous Remote Assessment for Cognitive Impairment: Reliability Verification of the Neurotrack Cognitive Battery
JMIR Form Res 2022;6(2):e34237
URL: https://formative.jmir.org/2022/2/e34237
doi:10.2196/34237
PMID:35179511

Original Paper

Telehealth Adoption and Discontinuation by US Hospitals: Results From 2 Quasi-Natural Experiments

Aaron Baird[1,2], MBA, PhD; Yichen Cheng[3], PhD; Yusen Xia[3], MS, PhD

[1]Institute of Health Administration, Georgia State University, Atlanta, GA, United States

[2]Department of Computer Information Systems, Robinson College of Business, Georgia State University, Atlanta, GA, United States

[3]Institute for Insight, Robinson College of Business, Georgia State University, Atlanta, GA, United States

Corresponding Author:
Aaron Baird, MBA, PhD
Institute of Health Administration
Georgia State University
35 Broad Street
Suite 805
Atlanta, GA, 30303-3083
United States
Phone: 1 4044137639
Fax: 1 4044137639
Email: abaird@gsu.edu

Abstract

Background: Prior US hospital telehealth (video visit) studies have focused on describing factors that influence telehealth adoption or performance effects for specific patient segments, hospital systems, or geographic regions. To our knowledge, a larger-scale, national-level (US) study has yet to be conducted on the causal impacts of hospital telehealth adoption as well as discontinuation.

Objective: The aim of this study is to understand the causal impact of US hospital telehealth adoption or discontinuation on hospital performance from 2016 to 2018.

Methods: We analyzed impacts of telehealth adoption or discontinuation by US hospitals on emergency department visits, total ambulatory visits (minus emergency department visits), outpatient services revenue, total facility expenses, and total hospital revenue for the 2016-2018 period. We specifically focused on performance effects for hospitals that switched from not having telehealth to adopting telehealth, or vice versa, during the 2016-2018 period, thus exploiting 2 quasi-natural experiments. We applied a difference-in-differences research design to each of the 2 main analyses. We compared hospitals that have made a telehealth change to groups of hospitals with similar characteristics that did not make a telehealth change, which established a counterfactual. To appropriately match hospitals between treatment and control groups, we applied propensity score matching. Our primary data were from the American Hospital Association Annual Survey and the Healthcare Cost Report Information System data. Several control variables were obtained from additional sources, including the Area Health Resource File and the Federal Communications Commission.

Results: We found that telehealth adoption by US hospitals during the 2016-2018 period resulted in, on average, an increased number of total ambulatory visits (P=.008), increased total facility expenses (P<.001), and increased hospital revenue (P=.004) compared with the control group. We found that telehealth discontinuation during the same period resulted in, on average, decreased outpatient services revenue (P=.02) compared with the control group.

Conclusions: Our findings suggest that telehealth adoption increases use but has mixed impacts on performance, given that cost and revenue increase. However, once telehealth is offered, removing it can have a negative impact on performance, implying that returning to prior performance levels, if telehealth is removed, may be challenging.

(JMIR Form Res 2022;6(2):e28979) doi:10.2196/28979

KEYWORDS

telehealth; hospitals; adoption; discontinuation; health information system

Introduction

Background

Telehealth, in the form of video visits between health care providers and patients (*telehealth*, henceforth), is used by hospitals and their affiliated clinics to maintain or improve access to postdischarge follow-up, continuity of care, and care for nonurgent issues [1-5]. Although a number of studies have evaluated the impacts of telehealth on outcomes [2,6-10], such studies have primarily focused on either the determinants of telehealth adoption [11] or effects of telehealth primarily for patient populations limited to specific hospital systems or regions [4,12-14]. Larger-scale studies exploiting national-level natural variation in telehealth adoption as well as discontinuation over multiple years have yet to be conducted.

Overall, although many view telehealth with optimism, we do not yet fully understand the impact on hospital-level outcomes when telehealth is adopted or, in the case of challenges, discontinued. Thus, this study seeks to understand such impacts, including impacts resulting from telehealth discontinuation, which is not an aspect of telehealth that has been considered yet in the literature. For instance, in regard to challenges that may lead to discontinuation of telehealth, it is well known that telehealth can be especially difficult to sustain and integrate with workflows designed for in-person interactions [7] and can result in variable outcomes [15,16]. Particular challenges for hospitals offering telehealth include prioritization of the success of telehealth; engagement by providers, patients, and leaders; and continuous improvement [17]. Many times, telehealth is initially viewed with optimism, but the reality is that many clinicians have stopped using it in the past after a few visits [17]. Especially important to mitigate such issues are deliberate efforts to create protocols, develop appropriate scheduling techniques, and formalize an understanding for when telehealth is and is not appropriate [18], which, if not addressed, can lead to significant challenges, resistance, or program failure. Furthermore, the effects of telehealth have been found to have mixed or even positive effects on costs [3,19]. In the case of telehealth substituting for expensive in-person visits such as visits to the emergency department (ED) or in-patient admissions, telehealth can be cost-effective [20,21]. However, when offering video-based consultations to patients, it is also possible that increased access to health care increases provider costs and the number of visits requested by patients, which can result in less revenue, especially if telehealth is reimbursed at a lower rate than in-person visits [3].

Finally, telehealth is a particularly interesting case because it can be technically relatively easy to adopt or discontinue, especially if using a vendor-supported or cloud-based system, but, as discussed previously, can simultaneously result in significant and costly workflow challenges [8,22]. It is well known that telehealth use is an excellent opportunity to enhance access to care, but it is also well known that inadequate barrier identification and management can doom telehealth pilots [17,23]. Furthermore, given the variety of factors that may influence telehealth adoption, use, and potential discontinuation, several factors, including hospital and regional characteristics, must be controlled for. Thus, this study comprehensively examines both telehealth adoption and discontinuation in the United States from 2016 to 2018 through analysis of 2 quasi-natural experiments (ie, one for adoption and one for discontinuation), while controlling for several potential confounding variables. We also conduct robustness checks to validate our findings.

Implications

Our primary findings are as follows: (1) telehealth adoption by US hospitals during the period studied resulted in increased ambulatory visits, increased facility expenses, and increased hospital revenue in comparison with the control group, and (2) telehealth discontinuation resulted in decreased outpatient services revenue in comparison with the control group. The implications are that adopting telehealth increases use of ambulatory services, which implies greater access, but these findings also suggest that profit performance will likely be mixed. Furthermore, removing telehealth once offered can negatively affect future performance, implying that performance levels likely will not simply return to what they were before telehealth was adopted and then subsequently discontinued. Further implications are discussed later.

Methods

Overview

To address our research objectives, we analyzed the impact of telehealth adoption or discontinuation by US hospitals from 2016 to 2018 using difference-in-differences estimation of 2 quasi-natural experiments: (1) US hospital telehealth *adoption* during the period considered and (2) US hospital telehealth *discontinuation* during the same period. We specifically considered impacts of telehealth adoption or discontinuation during this period on ED visits, total ambulatory visits (minus ED visits), outpatient services revenue, total facility expenses, and hospital revenue (a more detailed description of these dependent variables is available in Multimedia Appendix 1).

Data

Data on which US hospitals continued to offer, or discontinued, telehealth were obtained from the American Hospital Association (AHA) Annual Survey for 2016-2018 (although data quality may be a concern, prior studies such as the one by Adler-Milstein et al [11] have found the AHA data to be highly consistent with the data from the Healthcare Information and Management Systems Society data set, suggesting high data quality). Outcome data for ED visits, total ambulatory visits, outpatient services revenue, total facility expenses, and hospital revenue per US hospital were obtained from the 2016-2018 AHA Annual Survey and the AHA's Centers for Medicare & Medicaid Services Healthcare Cost Report Information System (HCRIS) data (ie, AHA's version of the Centers for Medicare & Medicaid Services HCRIS data). Covariates used for propensity score matching and controls were obtained from the AHA data sets and from the US county-level data available from the Area Health Resource File, as well as the Area Deprivation Index (ADI) sourced from BroadStreet, health rankings data from the University of Wisconsin Population

Health Institute, and supplementary data from the Federal Communications Commission for broadband speeds per county. We included several controls from these data sources to account for rival explanations. Controls and covariates were derived from a literature review [24-30]. Tables 1 and 2 describe the relevant variables.

Table 1. Telehealth adoption sample descriptive statistics averaged for 2016-2018 (for 135 US hospitals that did not have telehealth for all 3 years or started to adopt telehealth in 2017 or 2018).

Group and variable	Description	Value, N	Values, mean (SD; range)
Telehealth adoption and outcomes			
Telehealth video-based consultation	Whether a hospital adopted telehealth in a given year	405	0.39 (0.49; 0-1)
EDVisits[a] (in thousands)	Number of emergency department visits	405	47.02 (32.98; 0-174.96)
TotAmbVisits[a] (in thousands)	Total number of ambulatory visits (minus emergency department visits)	405	196 (209.36; 2.09-1488.13)
OutpatSerRev[a] (in millions, US $)	Outpatient services revenue	393	179 (312.35; 0-2831.15)
TotFacExp[a] (in millions, US $)	Total facility expenses	405	321.61 (342.3; 18.97-2687.47)
HospRev[a] (in millions, US $)	Total hospital revenue	393	152.34 (283.84; 0-2399.62)
Hospital-level variables			
SystemOwned	System ownership	405	0.71 (0.45; 0-1)
WageIndx	Index of hospital labor market wages	393	1.01 (0.15; 0.72-1.35)
HITAssetCost (in millions, US $)	Health information technology asset acquisition costs	393	0.94 (3.81; 0-29.8)
TotAdmAndVsts (in thousands)	Sum of inpatient admissions and outpatient visits	405	303.62 (276.89; 5.19-1853.46)
Herfindahl-Hirschman Index	Competition index (1=monopoly) per hospital referral region	405	0.13 (0.09; 0.03-0.56)
PercMdcdElig	Percentage Medicaid eligibility	405	0.22 (0.08; 0.07-0.5)
COTH	Teaching hospital	405	0.09 (0.29; 0-1)
Own_FP	For-profit ownership	405	0.12 (0.32; 0-1)
Own_NP	Not-for-profit ownership	405	0.77 (0.42; 0-1)
Own_Gov	Government ownership	405	0.11 (0.31; 0-1)
PercCapit	Percentage of net patient revenue capitated	375	0.79 (4.17; 0-53)
PercRsk	Percentage of net patient revenue shared risk	347	1.77 (5.41; 0-42)
Region_MW	Midwestern region	405	0.31 (0.46; 0-1)
Region_S	Southern region	405	0.37 (0.48; 0-1)
Region_W	Western region	405	0.02 (0.14; 0-1)
Region_NE	Northeast region	405	0.3 (0.46; 0-1)
CMI	Case Mix Index	393	1.74 (0.33; 0.93-2.93)
Urban	1 if urban location	405	1 (0)
County-level variables			
CntyHlthRank	Normalized within state county health rankings for health outcomes, 1 being best	405	0.4 (0.32; 0-1)
CntyPercPop65	Percentage of population aged >65 years	405	0.15 (0.03; 0.09-0.26)
CntyPercBlack	Percentage of population Black	405	15.04 (11.97; 0.6-63.7)
CntyPercNative	Percentage of population Native	405	0.6 (1.53; 0.1-17.5)
CntyPercLatino	Percentage of population Latino	405	12.89 (11.89; 0.8-60.6)

Group and variable	Description	Value, N	Values, mean (SD; range)
CntyPercDpPov	Percentage of population in deep poverty	405	6.64 (2.59; 2.2-16)
CntyPercDisabled	Percentage of population disabled	405	10.01 (3.05; 5.1-17.2)
CntyBBMaxUP	Maximum advertised broadband upload speed	405	26.68 (32.56; 1.56-160.98)
CntyHsholdIntUse	Percentage of households who report using the internet	405	0.87 (0.04; 0.72-0.97)
CntyADI	Area Deprivation Index (10=most deprived)	405	5.04 (1.51; 1.53-8.4)

[a]More details about the outcome variables are available in Multimedia Appendix 1.

Table 2. Telehealth discontinuation sample descriptive statistics averaged for 2016-2018 (for 524 US hospitals that had telehealth for all 3 years or started to remove telehealth in 2017 or 2018).

Group and variable	Description	Value, N	Values, mean (SD; range)
Telehealth adoption and outcomes			
Telehealth video-based consultation	Whether a hospital adopted telehealth in a given year	1572	0.93 (0.26; 0 to 1)
EDVisits[a] (in thousands)	Number of emergency department visits	1572	61.16 (49.82; 0 to 617.78)
TotAmbVisits[a] (in thousands)	Total number of ambulatory visits (minus emergency department visits)	1572	263.74 (413.69; 1.02 to 6497.28)
OutpatSerRev[a] (in millions, US $)	Outpatient services revenue	1550	247.07 (476.36; –1443.98 to 6717.17)
TotFacExp[a] (in millions, US $)	Total facility expenses	1572	440.55 (565.66; 13.91 to 6004.75)
HospRev[a] (in millions, US $)	Total hospital revenue	1550	188.32 (377.16; 0 to 7055.45)
Hospital-level variables			
SystemOwned	System ownership	1572	0.86 (0.35; 0 to 1)
WageIndx	Index of hospital labor market wages	1550	0.98 (0.14; 0.71 to 1.44)
HITAssetCost (in millions, US $)	Health information technology asset acquisition costs	1550	2.97 (12.94; 0 to 175.42)
TotAdmAndVsts (in thousands)	Sum of inpatient admissions and outpatient visits	1572	407.36 (493.07; 3.82 to 6989.63)
Herfindahl-Hirschman Index	Competition index (1=monopoly) per hospital referral region	1572	0.15 (0.12; 0.03 to 0.96)
PercMdcdElig	Percentage Medicaid eligibility	1572	0.21 (0.07; 0.05 to 0.5)
COTH	Teaching hospital	1572	0.16 (0.37; 0 to 1)
Own_FP	For-profit ownership	1572	0.11 (0.31; 0 to 1)
Own_NP	Not-for-profit ownership	1572	0.8 (0.40; 0 to 1)
Own_Gov	Government ownership	1572	0.09 (0.29; 0 to 1)
PercCapit	Percentage of net patient revenue capitated	1476	0.53 (2.62; 0 to 40)
PercRsk	Percentage of net patient revenue shared risk	1381	2.28 (7.33; 0 to 81)
Region_MW	Midwestern region	1572	0.33 (0.47; 0 to 1)
Region_S	Southern region	1572	0.42 (0.49; 0 to 1)
Region_W	Western region	1572	0.03 (0.17; 0 to 1)
Region_NE	Northeast region	1572	0.21 (0.41; 0 to 1)
CMI	Case Mix Index	1550	1.68 (0.26; 0.99 to 2.68)
Urban	1 if urban location	1572	1 (0; 1 to 1)
County-level variables			
CntyHlthRank	Normalized within state county health rankings for health outcomes, 1 being best	1572	0.39 (0.3; 0 to 1)
CntyPercPop65	Percentage of population aged >65 years	1572	0.15 (0.04; 0.09 to 0.35)
CntyPercBlack	Percentage of population Black	1572	14.19 (12.71; 0.4 to 69.1)
CntyPercNative	Percentage of population Native	1572	0.64 (2.38; 0.1 to 38.4)

Group and variable	Description	Value, N	Values, mean (SD; range)
CntyPercLatino	Percentage of population Latino	1572	12.2 (13.79; 0.5 to 90.6)
CntyPercDpPov	Percentage of population in deep poverty	1572	6.6 (2.59; 1.8 to 19.9)
CntyPercDisabled	Percentage of population disabled	1572	10.2 (2.92; 4.2 to 20.7)
CntyBBMaxUP	Maximum advertised broadband upload speed	1572	22.02 (26.32; 1.26 to 160.98)
CntyHsholdIntUse	Percentage of households who report using the internet	1572	0.87 (0.05; 0.6 to 0.97)
CntyADI	Area Deprivation Index (10=most deprived)	1572	5.31 (1.35; 2.01 to 8.93)

[a]More details about the outcome variables are available in Multimedia Appendix 1.

Statistical Analyses

We applied difference-in-differences (DID) estimation with propensity score matching at the firm (hospital) unit of analysis to understand the effect of telehealth adoption and discontinuation by US hospitals during the 2016-2018 period. We conducted 2 primary analyses that exploited 2 quasi-natural experiments. The first DID analysis focused on telehealth adoption and evaluated impacts on performance for hospitals that went from no telehealth to offering telehealth during this period. The second DID analysis focused on telehealth discontinuation and evaluated impacts for hospitals that went from offering telehealth to discontinuing telehealth during this period. Control group selection and formation is discussed later in this section. This design followed other notable studies that assessed the impact of health information technology adoption and use on outcomes [28-31] as well as recommendations on effectively estimating causal effects by means of observational data [32,33]. This design is appropriate for estimating causal effects when pre- and posttreatment observational data are available, treatment and control groups with sufficiently balanced covariates and common trends before treatment can be established, and exogenous shocks can be assumed to be consistent between groups [34].

For the telehealth adoption analysis, treatment hospitals are those that first did not offer telehealth but then offered telehealth in a subsequent year. As we have 3 years of data that include the telehealth video visit (yes or no) question, we restricted our focus to video visits for chronic conditions or postsurgical follow-up as opposed to also including consideration of telehealth related to remote patient monitoring and mental health and addiction as separately measured in the AHA Annual Survey. For all US hospitals surveyed by the AHA for this quasi-natural experiment, treatment hospitals are those that (1) did not offer telehealth in 2016 but started in 2017 or 2018 (group 1, n=71) or (2) did not offer telehealth in 2016 or 2017 but then started offering it in 2018 (group 2, n=14). Control hospitals are those that did not offer telehealth in all 3 years (n=50).

For the telehealth discontinuation analysis, treatment hospitals are those that offered telehealth but then discontinued it in a subsequent year. For this quasi-natural experiment, the treatment hospitals are those that (1) offered telehealth in 2016 but discontinued in 2017 or 2018 (group 1, n=12) or (2) offered telehealth in both 2016 and 2017 but discontinued in 2018 (group 2, n=80). Control hospitals are those that offered telehealth in all 3 years (n=432).

To balance the covariates between the treatment and control groups in each of these analyses, we applied propensity scoring and, subsequently, matching. Propensity scoring is applied by first determining the propensity of a hospital being in the treatment group, given observable covariates [35,36]. Then, to reduce selection bias, a matching technique is used to find control group participants (hospitals, in this case) that ultimately result in no observable significant covariate differences between treatment and control groups [35]. Similar to Oh et al [30] and Bao et al [29], we calculated propensity scores by means of logistic regression for each of the analyses (ie, for the adoption analysis and then again for the discontinuation analysis), as explained further in this section. Our covariates consisted of both hospital-level variables and county-level variables, with SEs clustered at the hospital level to account for repeated county-level observations for hospitals within the same county. The logistic regression analysis results for propensity scores are reported in Multimedia Appendix 1.

Using the scores that resulted from obtaining predicted values per hospital, we applied one-to-many matching using both the propensity score and covariates (a one-to-one matching procedure was also tested, as reported in Multimedia Appendix 1, and the results were similar). Matched hospitals belonged to the same teaching, urban, and system status. In addition, we matched hospitals with similar sizes by restricting hospital size (total admissions plus visits) to a difference of no more than a factor of 1.5 and a difference in propensity scores of no more than 0.1. Therefore, for each treatment hospital, we had a cluster of hospitals as the control. For telehealth adoption, the result was a treatment group consisting of 85 hospitals and a matched sample control group consisting of 85 hospital clusters, with an average size of 2 controls and a median size of 1 control in each hospital cluster. For telehealth discontinuation, the result was a treatment group consisting of 92 hospitals and a matched sample control group consisting of 92 hospital clusters, with an average size of 28 controls and a median size of 17 controls in each hospital cluster. We used averaged outcomes (ED visits, total ambulatory visits, outpatient services revenue, total facility expenses, and hospital revenue) for each observed control

cluster. The matching for hospitals in group 1 was conducted based on the propensity score and covariates observed at year 2016, and the matching for group 2 was based on observations at year 2017. Comparison of covariates between the 2 groups resulted in no significant differences.

To obtain the propensity score, we conducted a logistic regression analysis using treatment group membership (1 for yes and 0 for no) as the dependent variable for the adoption analysis and then again for the discontinuation analysis. We applied a collection of hospital and county level characteristics as the independent variables for each analysis, with the same control variables being used in each propensity score model. Let p_{it}=*P*(*hospital i in the treatment group*) with the following formula:

$$\ln(p_{it}/1-p_{it}) = \beta_0 + \beta'_1 X_{it}$$

β_0 is the constant and X_{it} represents factors that affect a hospital's decision of whether telehealth existed for the adoption analysis (1=hospital is in the treatment group and therefore adopted telehealth in 2017 or 2018) or was discontinued for the discontinuation analysis (1=hospital is in the treatment group and therefore discontinued telehealth in 2017 or 2018). β'_1 is the coefficient vector.

Next, identification of the change in ED visits, total ambulatory visits, outpatient services revenue, total facility expenses, and hospital revenue after telehealth adoption and discontinuation was derived through the following DID model, applied once to the adoption analysis and once to the discontinuation analysis. Note that when conducting the analyses, we combined hospitals from group 1 and group 2 as the treatment group. β_0 is the constant, β_1 is the effect from the treatment group, β_2 represents posttreatment periods, and β_3 is the treatment effect (ie, the DID effect), which is the expected value difference in the time trend as well as the difference between treatment and control groups after treatment. We included hospital fixed effects (μ_i) to address any time-invariant hospital heterogeneity and time fixed effects to address time trends (ϑ_t). We performed an estimation using ordinary least squares [31]. The DID equation representing our model is as follows:

$$Y_{it} = \beta_0 + \beta_1 T_{it} + \beta_2 p_{it} + \beta_3 (T_{it} \times p_{it}) + \mu_i + \vartheta_t + \varepsilon_{it}$$

Robustness

Threats to validity could include endogeneity of telehealth adoption and decision-making around discontinuation, especially if our sample was subject to selection bias. We addressed this concern by also conducting Heckman analyses. Furthermore, nonrandom market changes, after treatment, may differentially affect outcomes [37]. For instance, perhaps broadband infrastructure or household use of the Internet expanded or contracted at different rates between the control and treatment groups in or after 2017 or 2018. These threats were addressed with our propensity scoring and matching approach that included county-level maximum broadband speeds and household internet use as covariates in the logistic regression analysis, in addition to several other covariates considered when scoring and matching. For instance, we also included the ADI in our propensity score matching procedure to address regional economic states and, potentially, changes over time such as changes after treatment that are not fully addressed in a DID model. Overall, we included several hospital-level (eg, Case Mix Index, hospital size, and market competition) and county-level covariates (eg, maximum broadband speeds, household internet use, county health ranking, and ADI) to address a variety of potential threats to validity (eg, differences in broadband penetration affecting telehealth adoption or outcomes). Finally, we also tested whether outcomes change in the years *after* treatment to provide additional explanatory information.

Results

Common Trends

For testing common trends, we plotted the averages of ED visits, total ambulatory visits, outpatient services revenue, total facility expenses, and hospital revenue for each of the groups at points in time (years) *relative* to when telehealth was adopted (Figure 1) or discontinued (Figure 2). Note that throughout the paper, the numbers of visits are shown in thousands, whereas expenses and revenue are shown in millions (US $).

To test the common trends assumption statistically, we also interacted pretreatment values with corresponding time dummies within the DID model (Figure 1). None of the coefficients were significant, suggesting that the trends are sufficiently common.

Again, to test the common trends assumption statistically, we also interacted pretreatment values with corresponding time dummies within the DID model (Figure 2). None of the coefficients were significant, suggesting that the trends are sufficiently common.

Figure 1. Common trends per outcome per year relative to when telehealth was adopted. EDVisits: emergency department visits; HospRev: hospital revenue; OutpatSerRev: outpatient services revenue; TotAmbVisits: total ambulatory visits; TotFacExp: total facility expenses.

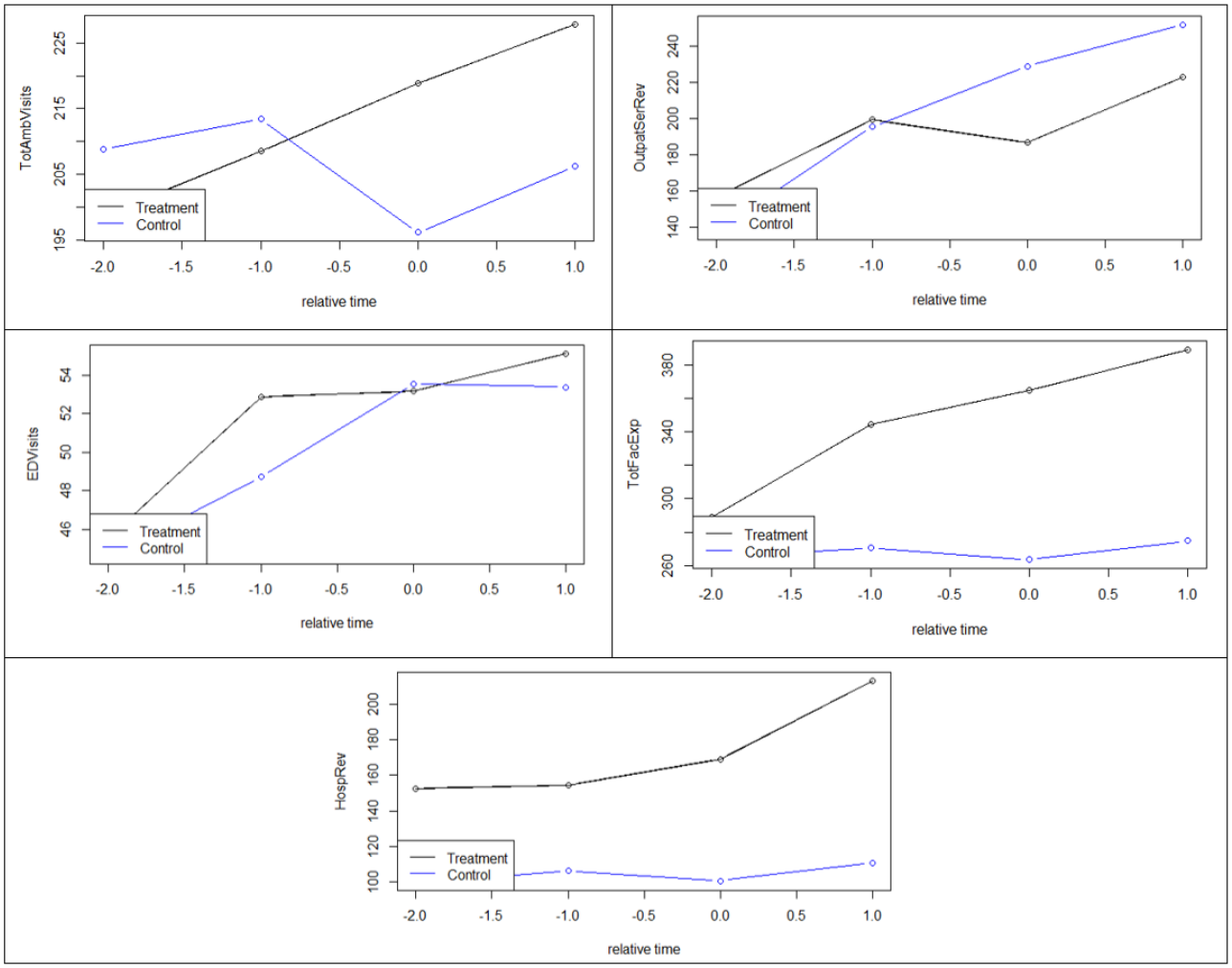
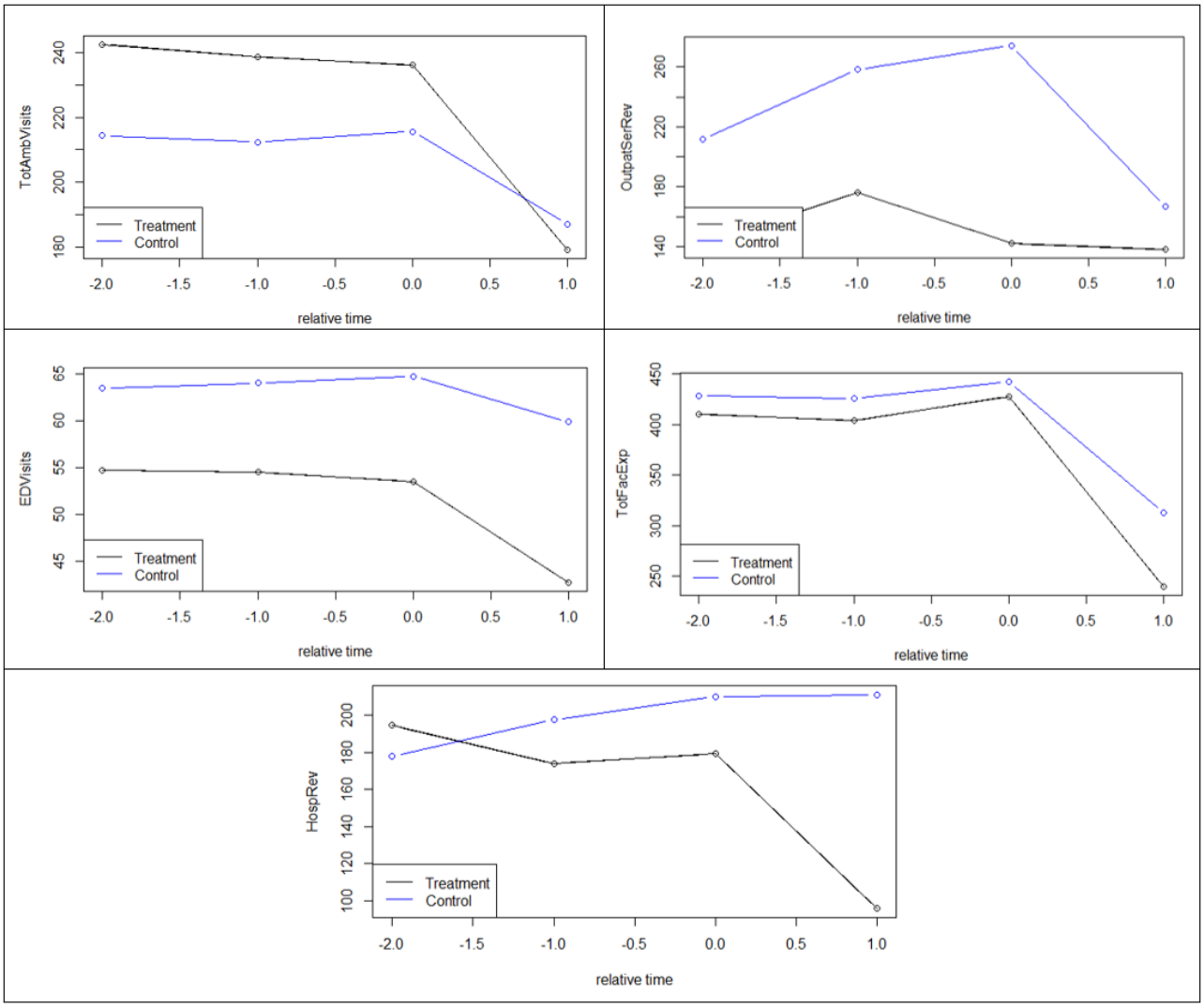


Figure 2. Common trends per outcome per year relative to when telehealth was discontinued. EDVisits: emergency department visits; HospRev: hospital revenue; OutpatSerRev: outpatient services revenue; TotAmbVisits: total ambulatory visits; TotFacExp: total facility expenses.



Estimations

The estimation results are reported below in Table 3 (for adoption) and Table 4 (for discontinuation; model analyses were conducted with R [The R Foundation for Statistical Computing]). For brevity, control variables are not included in the tables, but they were included in all regressions along with hospital and time fixed effects. The interaction terms represent the DID effect, which represents the expected value of the additional difference between the treatment and control groups after treatment (ie, the end of the time trend), when first accounting for the differences in time trends and difference in treatment and control groups.

For *telehealth adoption*, we found the DID interaction term for total ambulatory visits to be positive and significant (P=.008). This means that the expected value of total ambulatory visits was higher in the treatment group than in the control group, even after accounting for the time and group differences, as well as several covariates discussed earlier and also in Multimedia Appendix 1. The average total ambulatory visits, as reported earlier, was 196 (thousand; SD 209 [thousand]). Given that the DID coefficient is 24.53 (thousand), this effect represents a significant increase in total ambulatory visits. Thus, we conclude that telehealth adoption resulted in more ambulatory visits for the adopting US hospitals during the period studied.

We further found the DID interaction term to be positive (P<.001) for the effect on total facility expenses. Thus, the expected value of total facility expenses was higher in the treatment group (ie, those that adopted telehealth) than in the control group (ie, similar hospitals that did not have, and did not adopt, telehealth during the same period). The average total facility expenses in our sample, as reported earlier, was (in millions) US $321.61 (SD US $342.3). The coefficient (in millions) is US $33.39 ($P$<.001), which represents a substantial average increase in the expenses when telehealth was adopted.

We also found the DID interaction term to be positive (P=.004) for the effect on hospital revenue, which suggests higher total revenue on average for those in the treatment group. The average total hospital revenue in our sample, as reported earlier, was (in millions) US $152.34. The coefficient (in millions) is $32.60 ($P$=.004), which represents a substantial average increase in the revenue when telehealth was adopted. However, we also note that this coefficient is slightly lower than that of the average

increase in total facility expenses, suggesting that profits are likely to be negative or minimal when telehealth is first adopted.

The impact on ED visits (*P*=.36) was nonsignificant. The impact on outpatient services revenue was marginally significant (*P*=.01) and negative, suggesting that adoption led to at least a temporary drop in revenue, on average, in comparison with the control group.

For *telehealth discontinuation*, we found the DID interaction term (trt×post) to be significant and negative (*P*=.02) for the effect on outpatient services revenue. This means that the expected value for outpatient services revenue, ceteris paribus, was lower in the treatment group (ie, the group that discontinued telehealth) than in the control group, after accounting for the time trend and the assumed trend for the counterfactual. We also note that many control variables and fixed effects, to account for an unobserved time invariant heterogeneity, were accounted for. The average outpatient service revenue in our sample, as reported earlier, was (in millions) US $247.07. The coefficient (millions) is –US $65.37 (*P*=.02), which represents a substantial average drop in revenue compared with the control group when telehealth was discontinued.

We also found the DID interaction term to be negative and marginally significant (*P*=.09) for the effect on hospital revenue. This again means that the expected value, given all the aforementioned trends and variables, was lower in the treatment group than in the control group after treatment. The average hospital revenue for our sample, as reported earlier, was (in millions) US $188.32. Given that the coefficient (in millions) is –US $13.22, this represents a substantial potential average drop in total hospital revenue when telehealth was discontinued.

The impacts on ED visits (*P*=.10), total ambulatory visits (*P*=.28), and total facility expenses (*P*=.35) were nonsignificant.

Table 3. Difference-in-differences results for telehealth adoption.

	Emergency department visits (in thousands)[a]	*P* value	Total ambulatory visits (in thousands)[a]	*P* value	Outpatient services revenue (in millions, US $)[a]	*P* value	Total facility expenses (in millions, US $)[a]	*P* value	Hospital revenue (in millions, US $)[a]	*P* value
trt[b]	88.19 (6.19[c])	<.001	64.10 (10.93)	<.001	–159.01 (19.86)	<.001	360.58 (5.49)	<.001	59.45 (24.76)	.02
post[d]	1.00 (1.52)	.51	–15.18 (6.96)	.03	29.85 (19.39)	.13	–16.29 (7.78)	.04	6.68 (21.85)	.77
trt×post	–1.49 (1.63)	.36	24.53 (9.15)	.008	–34.64 (20.71)	.10	33.39 (6.36)	<.001	32.60 (11.24)	.004
Hospital fixed effects	✓	N/A[e]	✓	N/A	✓	N/A	✓	N/A	✓	N/A
Time fixed effects	✓	N/A	✓	N/A	✓	N/A	✓	N/A	✓	N/A
(Intercept)	29.94 (2.63)	<.001	74.04 (5.60)	<.001	153.46 (16.46)	<.001	108.13 (3.80)	<.001	62.73 (23.61)	.008
n	499[f]	N/A	510	N/A	502[g]	N/A	510	N/A	502[h]	N/A
R^2	0.96	N/A	0.97	N/A	0.90	N/A	0.99	N/A	0.95	N/A
F-statistic	44.46	<.001	53.56	<.001	16.32	<.001	313	<.001	36.6	<.001

[a]All the dependent variables are Winsorized at 0.01 and 0.99 level.

[b]trt: treatment group.

[c]Robust SEs clustered at the hospital level (in parentheses).

[d]post: posttreatment time periods.

[e]N/A: not applicable.

[f]A total of 11 observations that did not have a mention of an emergency department visit were omitted from the model for emergency department visits, which is why the n is 499 instead of 510.

[g]A total of 8 observations that did not have a mention of outpatient services revenue were omitted from the model for outpatient services revenue, which is why the n is 502 instead of 510.

[h]A total of 8 observations that did not have a mention of hospital revenue were omitted from the model for hospital revenue, which is why the n is 502 instead of 510.

Table 4. Difference-in-differences results for telehealth discontinuation.

	Emergency department visits (in thousands)[a]	*P* value	Total ambulatory visits (in thousands)[a]	*P* value	Outpatient services revenue (in millions, US $)[a]	*P* value	Total facility expenses (in millions, US $)[a]	*P* value	Hospital revenue (in millions, US $)[a]	*P* value
trt[b]	26.86 (2.06[c])	<.001	26.72 (4.98)	<.001	170.08 (37.33)	<.001	58.12 (12.30)	<.001	–104.41 (18.51)	<.001
post[d]	–0.25 (1.11)	.83	2.68 (6.23)	.67	–10.91 (29.27)	.71	0.07 (5.06)	.99	9.85 (10.40)	.34
trt×post	–1.65 (1.01)	.11	–7.84 (7.30)	.28	–65.37 (26.76)	.02	7.15 (7.65)	.35	–13.22 (7.70)	.09
Hospital fixed effects	✓	N/A[e]	✓	N/A	✓	N/A	✓	N/A	✓	N/A
Time fixed effects	✓	N/A	✓	N/A	✓	N/A	✓	N/A	✓	N/A
(Intercept)	32.49 (0.42)	<.001	29.91 (2.74)	<.001	22.70 (22.99)	.32	123.88 (8.24)	<.001	252.04 (11.35)	<.001
n	551[f]	N/A	552	N/A	552	N/A	552	N/A	549[g]	N/A
R^2	0.98	N/A	0.99	N/A	0.76	N/A	0.99	N/A	0.98	N/A
F-statistic	94.6	<.001	154.3	<.001	6.10	<.001	517.5	<.001	110.8	<.001

[a]All the dependent variables are Winsorized at 0.01 and 0.99 level.

[b]trt: treatment group.

[c]Robust SEs clustered at the hospital level (in parentheses).

[d]post: posttreatment time periods.

[e]N/A: not applicable.

[f]An observation that did not have a mention of an emergency department visit was omitted from the model for emergency department visits, which is why the n is 551 instead of 552.

[g]A total of 3 observations that did not have a mention of hospital revenue were omitted from the model for hospital revenue, which is why the n is 549 instead of 552.

Robustness Checks

We conducted additional tests to address potential endogeneity issues and threats to validity. First, hospital management, not some central regulatory authority, makes telehealth adoption and discontinuation decisions. Thus, our sample has a potential self-selection endogeneity issue. To address this statistically, beyond the use of propensity score matching, we used a Heckman model [38,39]. The Heckman model consists of 2 stages and is designed to control for those omitted from the sample. The first stage models the self-selection decision, that is, whether a hospital adopts or discontinues telehealth. The second stage models the treatment effect while taking into consideration the selection decision by including the inverse mills ratio calculated from the first stage. The results of this robustness check are available in Multimedia Appendix 1 and are consistent with our primary results.

To test whether the outcomes were different for different years *after* treatment, we conducted 2-sample *t* tests using the 2018 data for group 1 (2 years after the treatment) versus the 2018 data of group 2 (1 year after the treatment). Recall that both group 1 and group 2 consist of treatment hospitals. Hospitals in group 1 are those that did not receive treatment in 2016, then received treatment in 2017 and 2018. Hospitals in group 2 are those that did not receive treatment in 2016 and 2017, then received treatment in 2018. The results are reported in Table 5 (for adoption) and Table 6 (for discontinuation).

We observed that after telehealth was adopted, there was an upward trend for the number of visits, expenses, and revenue when comparing year 2 to year 1 after the treatment (Table 5), although none are significant.

We observed that after telehealth was discontinued, there was no significant difference for most of the outcome variables, except for ED visits and total facility expenses (Table 6). For ED visits, we observed that the number of ED visits decreased further 2 years after the treatment compared with the previous year. The same trend of a further decrease 2 years after the treatment was found for total facility expenses.

Table 5. Results of comparison of outcome variables after treatment (year 1 vs year 2) after telehealth was adopted.

	Emergency department visits (in thousands)	Total ambulatory visits (in thousands)	Total facility expenses (in millions, US $)	Outpatient services revenue (in millions, US $)	Hospital revenue (in millions, US $)
Year 1, n, mean (SD); median	14, 46.03 (24.08); 44.61	14, 184.86 (163.96); 161.47	14, 329.75 (436.80); 193.85	14, 196.18 (88.89); 91.62	14, 137.33 (209.57); 42.19
Year 2, n, mean (SD); median	71, 55.15 (29.52); 47.47	71, 227.88 (256.83); 138.87	71, 389.40 (389.39); 323.32	71, 222.58 (407.19); 119.02	71, 213.31 (366.04); 96.29
t test, difference (SE)	9.12 (7.34)	43.02 (53.37)	59.66 (125.55)	26.40 (102.30)	75.99 (71.67)

Table 6. Results of comparison of outcome variables after treatment (year 1 vs year 2) after telehealth was discontinued.

	Emergency department visits (in thousands)	Total ambulatory visits (in thousands)	Total facility expenses (in millions, US $)	Outpatient services revenue (in millions, US $)	Hospital revenue (in millions, US $)
Year 1, n, mean (SD); median	80, 55.01 (35.37); 48.72	80, 243.16 (341.72); 111.67	80, 456.60 (589.83); 247.90	80, 172.74 (224.79); 91.62	80, 193.38 (333.23); 77.20
Year 2, n, mean (SD); median	12, 42.72 (18.65); 39.10	12, 179.07 (134.64); 134.65	12, 239.61 (155.71); 199.68	12, 150.55 (103.77); 162.36	12, 95.81 (78.31); 75.31
t test, difference (SE), *P* value	–12.29 (6.68), .03	–64.09 (54.61), .12	–216.98 (79.80), .003	–22.19 (38.47), .28	–97.57 (43.58), .01

Discussion

Overview

This study assessed the impact of telehealth video visit consultation adoption or discontinuation by US hospitals from 2016 to 2018 through analysis of 2 quasi-natural experiments (ie, one for adoption and one for discontinuation). After conducting a number of robustness checks to validate our findings, we can conclude that, for this period, telehealth adoption resulted in an average increase in total ambulatory visits, total facility expenses, and hospital revenue in comparison with the control group of similar hospitals that neither offered, nor had adopted, telehealth services during this same period. Telehealth discontinuation resulted in an average reduction in outpatient services revenue compared with the control group of similar hospitals that did not discontinue telehealth during this period. Furthermore, in our robustness check, we found telehealth discontinuation to reduce total facility expenses over time, suggesting that telehealth investments are costly and cannot simply rely on existing communications infrastructure (ie, it is not the case that little to no additional costs are involved).

Principal Findings

First, we found that telehealth adoption for US hospitals from 2016 to 2018 resulted in increased visits, expenses, and revenue in comparison with the control group. These findings are similar to those of another study that found telehealth not only increased use (ie, resulted in more visits), but also increased costs [3]. However, this previous study focused on direct-to-consumer telehealth for a payer-based patient population in California as opposed to telehealth offered by hospitals throughout the United States. Thus, we contribute by demonstrating a similar trend at the national level and for hospital-based (provider-based) telehealth as opposed to payer-supported direct-to-consumer telehealth. The implications of our findings are that providers switching from not offering telehealth to offering telehealth can expect higher visit volumes but not necessarily significant increases in profits, especially given that the coefficient for increased expenses (US $33.39 million) is slightly higher than the coefficient for increased revenue (US $32.60 million) in our telehealth adoption results. The results make sense because it has been found that offering telehealth can increase provider workload [40], reduce workflow efficiency (at first) [23,41], and result in billing and payment issues [42]. Furthermore, given that payment parity laws are only now becoming more commonplace for telehealth and are still subject to significant variability [43], revenue from additional telehealth visits may be less than expected, especially if visits that were typically in person are now being replaced with video-based visits. Thus, telehealth adoption may provide more convenience for patients but may have mixed impacts on provider performance, likely requiring a significant investment by providers in overcoming barriers at least in the short term, as was also found in other telehealth studies such as those in the area of telestroke [44].

Second, we found that telehealth discontinuation had a negative impact on outpatient services revenue. The implication is that once telehealth is offered, performance may subsequently suffer if it is discontinued. Thus, careful thought must be given to what might happen with patient expectations once telehealth is offered, even if only for a short time. However, we also note that, although the observed decline in visit volume might be

expected to be responsible for loss in revenue, we did not find a significant impact on total ambulatory visits in comparison with the control group when telehealth was discontinued. This means that the revenue loss may be attributable to a decline in other outpatient services such as wellness and prevention programs, observation programs, supplies, laboratory tests, or other services, which suggests a spillover effect. Future research could examine this effect in more detail to gain a deeper understanding of potential spillover effects between discontinuing a digital service and other outpatient services offered. Most importantly, spillover effects aside, our results demonstrate that offering a digital service may change expectations, which cannot simply be reverted if telehealth is then no longer offered in a future period.

Limitations

We note that this study is limited by the binary nature of the response variable in that telehealth is a yes or no variable rather than an extent of use or assimilation variable. We also note that our data dates to before the COVID-19 pandemic period in which telehealth adoption and use significantly increased at first but subsequently significantly declined [8,45]. Future research could consider whether the effects found in this study are consistent with the postpandemic period, once more data are available. This study is also limited by a lack of detail in regard to the mechanisms that cause the effects we observed. This is also a significant opportunity for future research. Finally, our data are limited to the United States.

Additional Thoughts on Future Research

In addition to studying the spillover effects of telehealth adoption and discontinuation decisions, as well as determination of whether the effects found here remain consistent after the pandemic once more data are available, future research could consider price optimization for service channel differences such as in-person versus video visits and establish recommendations for optimal mixes of visit types, conditional on patient conditions and provider expertise. Given that the relationship among telehealth use, costs, and revenue is complex, uncertain, and mixed, more research is needed on service mix optimization.

We further note that our results are specific to US hospitals. Future research could consider whether these results are consistent with telehealth being adopted and discontinued in other countries and regions, as well as any unique conditions that may affect telehealth differently in other areas.

Finally, telehealth impacts, especially from adoption of telehealth, are likely to change over time. For instance, costs associated with telehealth may decrease in some ways as efficiencies are gained over time but increase in other ways such as potentially more technical and scheduling staff being required to support a mix of in-person and telehealth visits. Therefore, an excellent future area for future research will be a more fine-grained analysis of telehealth-specific costs over a longer period of time.

Conclusions

In conclusion, this study offers insights into the effects of telehealth adoption and discontinuation by US hospitals from 2016 to 2018. It is our hope that these results will inform health care providers, administrators, and policy makers regarding expected performance outcomes when telehealth adoption and discontinuation decisions are made.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional descriptions, first stage results, and robustness checks.
[DOCX File , 49 KB - formative_v6i2e28979_app1.docx]

References

1. Tuckson RV, Edmunds M, Hodgkins ML. Telehealth. N Engl J Med 2017 Oct 19;377(16):1585-1592. [doi: 10.1056/nejmsr1503323]
2. Shigekawa E, Fix M, Corbett G, Roby DH, Coffman J. The current state of telehealth evidence: a rapid review. Health Aff (Millwood) 2018 Dec;37(12):1975-1982. [doi: 10.1377/hlthaff.2018.05132] [Medline: 30633674]
3. Ashwood JS, Mehrotra A, Cowling D, Uscher-Pines L. Direct-to-consumer telehealth may increase access to care but does not decrease spending. Health Aff (Millwood) 2017 Mar 01;36(3):485-491. [doi: 10.1377/hlthaff.2016.1130] [Medline: 28264950]
4. Slightam C, Gregory AJ, Hu J, Jacobs J, Gurmessa T, Kimerling R, et al. Patient perceptions of video visits using veterans affairs telehealth tablets: survey study. J Med Internet Res 2020 Apr 15;22(4):e15682 [FREE Full text] [doi: 10.2196/15682] [Medline: 32293573]
5. Liu N, Huang R, Baldacchino T, Sud A, Sud K, Khadra M, et al. Telehealth for noncritical patients with chronic diseases during the COVID-19 pandemic. J Med Internet Res 2020 Aug 07;22(8):e19493 [FREE Full text] [doi: 10.2196/19493] [Medline: 32721925]
6. Kalankesh L, Pourasghar F, Nicholson L, Ahmadi S, Hosseini M. Effect of telehealth interventions on hospitalization indicators: a systematic review. Perspect Health Inf Manag 2016;13(Fall):1h [FREE Full text] [Medline: 27843425]

7. Kruse CS, Krowski N, Rodriguez B, Tran L, Vela J, Brooks M. Telehealth and patient satisfaction: a systematic review and narrative analysis. BMJ Open 2017 Aug 03;7(8):e016242 [FREE Full text] [doi: 10.1136/bmjopen-2017-016242] [Medline: 28775188]
8. Monaghesh E, Hajizadeh A. The role of telehealth during COVID-19 outbreak: a systematic review based on current evidence. BMC Public Health 2020 Aug 01;20(1):1193 [FREE Full text] [doi: 10.1186/s12889-020-09301-4] [Medline: 32738884]
9. Totten AM, Womack DM, Eden KB, McDonagh MS, Griffin JC, Grusing S, et al. Telehealth: Mapping the Evidence for Patient Outcomes From Systematic Reviews. Rockville, MD: Agency for Healthcare Research and Quality (US); 2016.
10. Wade VA, Karnon J, Elshaug AG, Hiller JE. A systematic review of economic analyses of telehealth services using real time video communication. BMC Health Serv Res 2010 Aug 10;10(1):233 [FREE Full text] [doi: 10.1186/1472-6963-10-233] [Medline: 20696073]
11. Adler-Milstein J, Kvedar J, Bates DW. Telehealth among US hospitals: several factors, including state reimbursement and licensure policies, influence adoption. Health Aff (Millwood) 2014 Mar;33(2):207-215. [doi: 10.1377/hlthaff.2013.1054] [Medline: 24493762]
12. Hwa K, Wren SM. Telehealth follow-up in lieu of postoperative clinic visit for ambulatory surgery: results of a pilot program. JAMA Surg 2013 Sep 01;148(9):823-827. [doi: 10.1001/jamasurg.2013.2672] [Medline: 23842982]
13. Lonergan PE, Washington Iii SL, Branagan L, Gleason N, Pruthi RS, Carroll PR, et al. Rapid utilization of telehealth in a comprehensive cancer center as a response to COVID-19: cross-sectional analysis. J Med Internet Res 2020 Jul 06;22(7):e19322 [FREE Full text] [doi: 10.2196/19322] [Medline: 32568721]
14. Woods LW, Snow SW. The impact of telehealth monitoring on acute care hospitalization rates and emergency department visit rates for patients using home health skilled nursing care. Home Healthc Nurse 2013 Jan;31(1):39-45. [doi: 10.1097/NHH.0b013e3182778dd3] [Medline: 23238623]
15. Schoenfeld AJ, Davies JM, Marafino BJ, Dean M, DeJong C, Bardach NS, et al. Variation in quality of urgent health care provided during commercial virtual visits. JAMA Intern Med 2016 May 01;176(5):635-642 [FREE Full text] [doi: 10.1001/jamainternmed.2015.8248] [Medline: 27042813]
16. Steventon A, Bardsley M, Billings J, Dixon J, Doll H, Hirani S, Whole System Demonstrator Evaluation Team. Effect of telehealth on use of secondary care and mortality: findings from the Whole System Demonstrator cluster randomised trial. BMJ 2012 Jun 21;344(jun21 3):e3874 [FREE Full text] [doi: 10.1136/bmj.e3874] [Medline: 22723612]
17. Ellimoottil C, An L, Moyer M, Sossong S, Hollander JE. Challenges and opportunities faced by large health systems implementing telehealth. Health Aff (Millwood) 2018 Dec;37(12):1955-1959. [doi: 10.1377/hlthaff.2018.05099] [Medline: 30633667]
18. Croymans D, Hurst I, Han M. Telehealth: the right care, at the right time, via the right medium. NEJM Catalyst Innov Care Delivery 2020;1(6):1-12 [FREE Full text]
19. Cho S, Mathiassen L, Gallivan M. Crossing the diffusion chasm: from invention to penetration of a telehealth innovation. Info Technol People 2009 Nov 13;22(4):351-366. [doi: 10.1108/09593840911002450]
20. Langabeer JR, Champagne-Langabeer T, Alqusairi D, Kim J, Jackson A, Persse D, et al. Cost–benefit analysis of telehealth in pre-hospital care. J Telemed Telecare 2016 Dec 05;23(8):747-751. [doi: 10.1177/1357633x16680541]
21. Hill RD, Luptak MK, Rupper RW, Bair B, Peterson C, Dailey N, et al. Review of veterans health administration telemedicine interventions. Am J Manag Care 2010 Dec;16(12 Suppl HIT):e302-e310. [Medline: 21322300]
22. Wosik J, Fudim M, Cameron B, Gellad Z, Cho A, Phinney D, et al. Telehealth transformation: COVID-19 and the rise of virtual care. J Am Med Inform Assoc 2020 Jun 01;27(6):957-962 [FREE Full text] [doi: 10.1093/jamia/ocaa067] [Medline: 32311034]
23. Lin CC, Dievler A, Robbins C, Sripipatana A, Quinn M, Nair S. Telehealth in health centers: key adoption factors, barriers, and opportunities. Health Aff (Millwood) 2018 Dec;37(12):1967-1974. [doi: 10.1377/hlthaff.2018.05125] [Medline: 30633683]
24. Baird A, Furukawa MF, Raghu TS. Understanding contingencies associated with the early adoption of customer-facing web portals. J Manag Inf Syst 2014 Dec 08;29(2):293-324. [doi: 10.2753/mis0742-1222290210]
25. Atasoy H, Chen P, Ganju K. The spillover effects of health IT investments on regional healthcare costs. Manag Sci 2018 Jun;64(6):2515-2534. [doi: 10.1287/mnsc.2017.2750]
26. Adler-Milstein J, Everson J, Lee SD. EHR adoption and hospital performance: time-related effects. Health Serv Res 2015 Dec;50(6):1751-1771 [FREE Full text] [doi: 10.1111/1475-6773.12406] [Medline: 26473506]
27. McCullough J, Parente S, Town R. Health information technology and patient outcomes: the role of organizational and informational complementarities. National Bureau of Economic Research. URL: https://www.nber.org/system/files/working_papers/w18684/w18684.pdf [accessed 2022-01-31]
28. Atasoy H, Greenwood BN, McCullough JS. The digitization of patient care: a review of the effects of electronic health records on health care quality and utilization. Annu Rev Public Health 2019 Apr 01;40:487-500. [doi: 10.1146/annurev-publhealth-040218-044206] [Medline: 30566385]
29. Bao C, Bardhan IR, Singh H, Meyer BA, Kirksey K. Patient–provider engagement and its impact on health outcomes: a longitudinal study of patient portal use. MISQ 2020 Jun 1;44(2):699-723. [doi: 10.25300/misq/2020/14180]

30. Oh JC, Zheng ZE, Bardhan IR. Sooner or later? Health information technology, length of stay, and readmission risk. Prod Oper Manag 2017 Aug 03;27(11):2038-2053. [doi: 10.1111/poms.12748]
31. Ganju KK, Atasoy H, McCullough J, Greenwood B. The role of decision support systems in attenuating racial biases in healthcare delivery. Manag Sci 2020 Nov;66(11):5171-5181. [doi: 10.1287/mnsc.2020.3698]
32. Basu S, Meghani A, Siddiqi A. Evaluating the health impact of large-scale public policy changes: classical and novel approaches. Annu Rev Public Health 2017 Mar 20;38:351-370 [FREE Full text] [doi: 10.1146/annurev-publhealth-031816-044208] [Medline: 28384086]
33. Imbens GW, Wooldridge JM. Recent developments in the econometrics of program evaluation. J Economic Literature 2009 Mar 01;47(1):5-86. [doi: 10.1257/jel.47.1.5]
34. Wing C, Simon K, Bello-Gomez RA. Designing difference in difference studies: best practices for public health policy research. Annu Rev Public Health 2018 Apr 01;39:453-469. [doi: 10.1146/annurev-publhealth-040617-013507] [Medline: 29328877]
35. Guo S, Fraser MW. Propensity Score Analysis: Statistical Methods and Applications. Thousand Oaks, CA: SAGE Publications; 2014.
36. Caliendo M, Kopeinig S. Some practical guidance for the implementation of propensity score matching. J Economic Surveys 2008 Feb;22(1):31-72. [doi: 10.1111/j.1467-6419.2007.00527.x]
37. Kahn-Lang A, Lang K. The promise and pitfalls of differences-in-differences: reflections on 16 and pregnant and other applications. J Business Econ Stat 2019 Apr 02;38(3):613-620. [doi: 10.1080/07350015.2018.1546591]
38. Heckman JJ. Sample selection bias as a specification error. Econometrica 1979 Jan;47(1):153. [doi: 10.2307/1912352]
39. Certo ST, Busenbark JR, Woo H, Semadeni M. Sample selection bias and Heckman models in strategic management research. Strat Manag J 2016 Feb 09;37(13):2639-2657. [doi: 10.1002/smj.2475]
40. Sharma U, Clarke M. Nurses' and community support workers' experience of telehealth: a longitudinal case study. BMC Health Serv Res 2014 Apr 10;14(1):164 [FREE Full text] [doi: 10.1186/1472-6963-14-164] [Medline: 24721599]
41. Zulman DM, Wong EP, Slightam C, Gregory A, Jacobs JC, Kimerling R, et al. Making connections: nationwide implementation of video telehealth tablets to address access barriers in veterans. JAMIA Open 2019 Oct;2(3):323-329 [FREE Full text] [doi: 10.1093/jamiaopen/ooz024] [Medline: 32766533]
42. Sisk B, Alexander J, Bodnar C, Curfman A, Garber K, McSwain SD, et al. Pediatrician attitudes toward and experiences with telehealth use: results from a national survey. Acad Pediatr 2020 Jul;20(5):628-635 [FREE Full text] [doi: 10.1016/j.acap.2020.05.004] [Medline: 32437881]
43. Shachar C, Engel J, Elwyn G. Implications for telehealth in a postpandemic future: regulatory and privacy issues. JAMA 2020 Jun 16;323(23):2375-2376. [doi: 10.1001/jama.2020.7943] [Medline: 32421170]
44. Singh R, Baird A, Mathiassen L. Collaboration risk management in IT-enabled asymmetric partnerships: evidence from telestroke networks. Inf Organization 2018 Dec;28(4):170-191. [doi: 10.1016/j.infoandorg.2018.10.001]
45. Early impact of CMS expansion of medicare telehealth during COVID-19. Health Affairs Forefront. URL: https://www.healthaffairs.org/do/10.1377/forefront.20200715.454789/full/ [accessed 2022-01-31]

Abbreviations

ADI: Area Deprivation Index
AHA: American Hospital Association
DID: difference-in-differences
ED: emergency department
HCRIS: Healthcare Cost Report Information System

Edited by A Mavragani; submitted 20.03.21; peer-reviewed by E Jussupow, I Mircheva, M Rusll, M Jones; comments to author 24.06.21; revised version received 10.08.21; accepted 04.01.22; published 18.02.22.

Please cite as:
Baird A, Cheng Y, Xia Y
Telehealth Adoption and Discontinuation by US Hospitals: Results From 2 Quasi-Natural Experiments
JMIR Form Res 2022;6(2):e28979
URL: https://formative.jmir.org/2022/2/e28979
doi:10.2196/28979
PMID:35179503

Original Paper

The Attitudes of Egyptian Web-Based Health Information Seekers Toward Health Information Provided Through the Internet: Qualitative Study

Mayada Ghweeba[1,2], MD; Antje Lindenmeyer[1], PhD; Sobhi Shishi[2], MD; Amani Waheed[2], MD; Mostafa Kofi[2], MD; Shaymaa Amer[2], MD

[1]Institute of Clinical Sciences, Birmingham, United Kingdom

[2]Community Medicine Department, Faculty of Medicine, Suez Canal University, Ismailia, Egypt

Corresponding Author:
Antje Lindenmeyer, PhD
Institute of Clinical Sciences
University of Birmingham Edgbaston
Birmingham, B15 2TT
United Kingdom
Phone: 44 (0)121 414 5390
Email: a.lindenmeyer@bham.ac.uk

Abstract

Background: The internet has become an established source of health information for many Egyptian internet users. Understanding users' attitudes toward the benefits and limitations of web-based health information will explain the influence of this information on users' health-related behavior and decisions.

Objective: This qualitative study aims to understand the attitude of Egyptian internet users toward internet health information and to explore the impact of obtained health information on users' behavior and on their physician-patient relationship.

Methods: For this qualitative study, semistructured interviews were conducted with a total of 49 participants (41/49, 84% Egyptian internet users and 8/49, 16% physicians) who participated in focus groups or individual interviews. We used a thematic analysis approach to explain and demonstrate participants' views, thoughts, and experiences in using web-based health information.

Results: The internet has become an important source of health information in comparison with other health information sources and is the central theme that has emerged across the thematic analysis. The attitude toward the use of internet health was classified into three main themes: feeling toward web-based health information (with subthemes: favoring, disliking, neutral, or having ambivalence feelings), motivators to seek internet health information, and behavioral changes using internet health information (subthemes: confidence, satisfaction, and improved knowledge). Themes that emerged from physicians' interviews included the accessibility of the internet health information, good communication, and coordination of care between patients and their physicians, and the active engagement of patients with their management plan.

Conclusions: The internet has become an essential source of health information for Egyptian adults. Internet health information can improve the patient-physician relationship, especially when users discuss the obtained health information with their physician. Internet health information provided seekers with social support and self-confidence when making health decisions.

(JMIR Form Res 2022;6(2):e30108) doi:10.2196/30108

KEYWORDS

Egyptian internet users; online health information; doctor–patient relationship

Introduction

Searching for online health information (OHI) has become one of the most popular web-based activities among Egyptians [1,2]. However, the use of OHI in developed countries is greater than that in low or middle socioeconomic populations [3]. Traditional modes of health information such as asking local experts or seeking mass media still play pivotal roles in disseminating health information for many health information seekers in developing countries [4]. In Egypt, previous studies have found that women with higher educational achievement, with the presence of children in the household and having chronic health

conditions were more likely to search for health information on the internet [5].

The internet provides a wide range of health topics from healthy eating and fitness to information about self-management of chronic health problems [6,7]. The internet presents medical information through a variety of mediums including websites, listservs, online support groups, chat rooms, instant messaging, and emails [8]. In addition, the internet serves as a medium for users to interact frequently with health care professionals, access health information on the web, and participate in online support groups [3,7].

The accessibility and convenience of the OHI motivate users to search for health information through the internet [9,10]. Moreover, OHI has enabled users to find support from people with similar health problems, which had a substantial impact on their physical and psychological well-being. The availability of OHI is known to empower patients toward healthier behavior changes, become active collaborators in their health decisions, and become more adherent to their treatment [11]. Nevertheless, many concerns have been raised about using the OHI, such as the quality, completeness, and accuracy of OHI, especially in the Arabic language, which is difficult to ascertain by OHI users [2,12]. Previous studies have found a direct relationship between OHI seeking and health literacy [13]. They found that OHI seekers should have a minimal level of health literacy and guidance to be able to review and evaluate the quality of the available health information on the internet, particularly in the absence of proper communication with their physician [14,15].

In Egypt, the need for affordable, accessible, and quality health information source is growing in accordance with the growth of the Egyptian population. Although the internet has become an important source of health information for many internet users, little information is available about Egyptians' OHI-seeking behavior. In this study, we used a qualitative method to understand how and why people seek OHI and explore the influence of this information on users' behavior and their relationship with their physicians. The qualitative approach allowed for in-depth insights into the participants' experiences, behavior, and feelings associated with health-seeking behaviors to better recognize their needs and identify the required improvements in the area of internet health information quality and availability [16].

Methods

Ethical Approval, Study Design, and Participants

Participants of this qualitative study were drawn from Egyptian internet users who participated in our previous cross-sectional study to identify the personal characteristics of Egyptian OHI seekers [1]. The inclusion criteria for the study were being Egyptian, aged ≥18 years, living in Egypt at the time of the study, and previously used the internet to search for health information for themselves or others. A total of 49 participants (41/49, 84% Egyptian internet users and 8/49, 16% physicians) participated in 9 focus group discussions (FGDs) and 11 individual interviews. The 8 physicians who participated in this study were purposively selected from physicians who had at least 1 year of cumulative experience of providing OHI through different health information websites. They all received a summary of the study objectives and an explanation of the qualitative study procedures. The FGDs were held at different community locations convenient to the participants; participants' travel expenses were reimbursed. The 11 interviews, including physicians' interviews, were conducted through telecommunications apps such as Skype or phone calls.

Consent was obtained from all participants before the FGD and the interviews commenced; participants were reassured that their data would be anonymized. They had the option to choose a nickname to be used during the interviews and data analysis. At the beginning of the FGD and the interviews, participants were asked to complete a brief questionnaire that included demographic information (age, gender, marital status, level of education, current occupation, and self-assessed general health status) and questions about health information-seeking methods such as their preferred source of health information and the frequency of seeking OHI.

Semistructured interviews with open-ended questions were conducted to allow the participants to explain and demonstrate their views, thoughts, and experiences in their own words. We used 2 interview topic guides, one for internet users and another for physicians, which were revised iteratively, allowing for further exploration of new issues raised [10]. The interview guides started with a warm-up question about participants' perception of the common health topics people are looking for on the internet. The interview guide included questions about the frequency of OHI seeking, influence of web-based interaction with others who have similar health concerns, personal reasons that encourage or prevent them from seeking OHI, perceived impact of OHI-seeking activities on the relationship with their physicians, and how they personally assessed the quality and reliability of health information on the internet. Similarly, the physicians' interview started with a demographic questionnaire about their gender, years in practice, their medical specialty, and any health information website they used, followed by open-ended questions to obtain insights from their experience of the benefits and challenges of OHI, OHI seekers' motives and challenges, and how OHI might affect the physician-patient relationship. Probes were used according to the interview topic guides when the participants' narratives ended or deviated significantly from the topic of interest.

The duration of the FGDs ranged between 56 and 82 (mean 68.7, SD 7.5) minutes, and the range for the interviews was 22 and 46 (mean 34.3, SD 7.1) minutes.

The main data sources were the transcribed audio-recorded interviews, the notes summary collected in each interview, and the questionnaire form used at the beginning of the interviews.

Ethical approval for the study was obtained from the Medical Research Ethics Committee of the Faculty of Medicine at Suez Canal University and granted approval for conducting this study on January 8, 2017 (approval #2298).

Analysis

Questionnaire results were analyzed to provide descriptive statistics for the demographic characteristics and general use of

OHI. Data were analyzed using SPSS Statistics (version 25, IBM Corp) predictive analytics software. Qualitative data were managed and analyzed manually using a thematic analysis approach [16]. The first stage involved translation into English, which included cultural and interpretive insights, and then transcription of all the recorded interviews. Transcripts and notes were read repeatedly to focus on preserving as much detail as possible. The transcripts were then systematically coded. The initial codes were sorted into themes and subthemes, and then these themes were re-evaluated (merged, deleted, or refined) to ensure that each theme had sufficient supporting data and the data cohered meaningfully. The data were then displayed in a graphic format to organize the data and show connections between different themes [17].

Results

Overview

As previously outlined, a total of 49 participants (41/49, 84% Egyptian OHI seekers and 8/49, 16% physicians) took part in the study. The demographic characteristics of the OHI seekers are shown in Table 1. The age of the OHI seekers ranged between 19 and 57 years, with a mean age of 36.3 years. Table 1 shows that 63% (26/41) of the OHI seekers were women and 73% (30/41) participants had a university degree. Most of the OHI seekers had health insurance coverage (28/41, 68%), and 17% (7/41) participants had chronic health conditions. Nearly half of the OHI seekers (21/41, 51%) rated their general health status as good.

Regarding the physicians who participated in this study (n=8), 5 (63%) participants were men and 4 (50%) participants had more than 10 years of medical practice, whereas most of the participants had 1 to 2 years of experience in providing OHI through health websites (Table 2).

Analysis of the qualitative data showed that *the use of the internet as a source of health information has grown in comparison to other health information sources* has emerged as a central or main theme across all the interviews. We used a graphic framework to summarize our findings into three major themes: participants' feelings, OHI-seeking behavior, and the influence of OHI on users' health and their relationship with their physicians (Figure 1). Although each theme was presented and expanded separately, they were often linked to each other owing to the complexity of the research topic. The extracts were cited to illustrate the range of participants' responses for each theme.

Table 1. Demographic and internet use information of the online health information seekers (n=41 internet users).

Sociodemographic characteristics	Value, n (%)
Age (years)	
15-34	17 (41)
35-49	18 (44)
≥50	6 (15)
Gender	
Male	15 (37)
Female	26 (63)
Level of education	
University student	7 (17)
University graduate	30 (73)
Postgraduate degree	4 (10)
Social status	
Never married before	9 (22)
Married	28 (68)
Divorced, widow, or widower	4 (10)
Employment	
Employed	20 (49)
Unemployed[a]	14 (34)
Student	7 (17)
Self-reported general health status	
Excellent or very good	21 (51)
Good or fair	18 (44)
Poor	2 (5)
Having a chronic health condition	
Yes[b]	7 (17)
No	34 (83)
Having health insurance coverage	
Yes	28 (68)
No	13 (32)

[a]Unemployed also included housewife or retired.

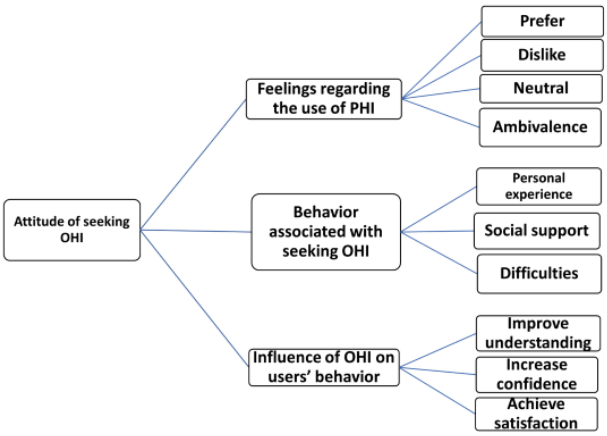
[b]Having one or more chronic health problem.

Table 2. Demographic characteristics of the physicians (n=8 physicians).

Characteristics of the physicians	Value, n (%)
Gender	
Male	5 (63)
Female	3 (37)
Years in practice	
0-5	2 (25)
6-10	2 (25)
>10	4 (50)
Medical specialty	
Internal medicine	4 (50)
Family medicine	1 (12.5)
Surgery	2 (25)
Pediatrics	1 (12.5)
Years in working for OHI[a] websites	
1-2 years	4 (50)
3-4 years	2 (25)
>4 years	2 (25)

[a]OHI: online health information.

Figure 1. The attitude of Egyptian online health information seekers toward using internet health information. OHI: online health information; PHI: personal health information.



Feelings Toward the Use of OHI

Participants' feelings toward the use of OHI were classified into four themes: favoring, disliking, neutral, or having ambivalent feelings toward the use of the OHI.

Participants mentioned various reasons to favor the use of OHI. For instance, a female participant explained that OHI made her feel more comfortable when discussing sensitive topics about her health. She mentioned that she wished to avoid the predictable discomfort of face-to-face interactions with her physicians with potentially embarrassing health complaints. Another 38-year-old male participant added that the ability to hide his physical disability made him feel more self-assured and able to talk with others about his health concern confidently. He noted as follows:

> *I can use the internet, and no one can see what is behind the screen. The good thing about it is that no one will know what you want to hide about yourself.*

A 42-year-old female participant noted as follows:

> *I prefer the internet websites to discuss sensitive or personal things, and no one will know it is you. You can say what you want without embarrassing yourself in front of your doctor.*

In addition, participants mentioned that finding other people with similar health problems gave them a sense of reassurance as well as alleviating their stress and guiding them to make better health decisions.

A 27-year-old female participant noted as follows:

> *After we found out that my husband had cancer. I was shocked. I felt that my life was ended. I can't explain how the support I found on the internet health groups helped me to regain my strength again and be able to give him the help and care he needs.*

Another group of participants expressed their dislikes regarding the use of OHI. They explained that they felt uncomfortable using the internet, especially in front of younger people owing to their limited computer and internet skills. Many participants who were aged ≥55 years mentioned that they preferred to call their physician when they had any health concerns. When they could not reach a physician, they asked for assistance from a friend or a family member to find the OHI they needed. A 58-year-old female participant noted as follows:

> *I feel that using the Internet is not for me. I am not very good at technology. That's why I save my time and ask one of my sons to find it for me...I found this more relaxing.*

Other participants expressed some concerns when posting their personal information on private health information websites, which may have commercial agendas. They believe that these websites would use their personal information to promote their products. A 36-year-old female participant noted as follows:

> *Some health websites may have hidden aims to push you to a particular clinic or specific product they promote. You may find them claim to cure incurable diseases like Hepatitis C or AIDS.*

Another group of participants expressed a neutral feeling regarding the use of OHI. *A 47-year-old woman* explained that she sought health information from any available health information resource without preference:

> *I understand that the health information on the internet is a good thing, but for me, I feel all other health information is similar.*

In this theme, many participants expressed positive feelings toward seeking OHI. They mentioned that OHI helped them feel confident and empowered to make decisions regarding their health. However, other participants linked their dislike of OHI to their limited health-related knowledge, limited internet skills, and a lack of trust.

OHI Seeking Behavior

In this theme, participants explained how their personal experience in accessing and using OHI influenced their behavior toward further use of the internet as a source of health information.

As previously outlined, many participants expressed their preference for using OHI because of its convenience and accessibility. They also mentioned that financial constraints or having a busy schedule encouraged them to prefer OHI in comparison to other, more traditional sources of health information.

A 34-year-old female participant noted as follows:

> *For me, the most important thing is to avoid transport. I have two baby daughters, so I like to use the internet to find what I want easily.*

A 42-year-old male participant noted as follows:

> *I prefer to find health information from the internet than from any other source of health information. For me, it's always the first place to go.*

Participants mentioned that the limited consultation time in hospitals or clinics encouraged them to seek OHI to find answers to their health enquires. Some participants preferred to seek OHI before their consultation to be able to discuss the information they got with their physicians. Others preferred seeking OHI after their consultation to find supplementary health information or a second opinion regarding their health problems.

A 42-year-old male participant noted as follows:

> *Sometimes the doctor does not have enough time to discuss or explain everything to me, especially when I see them in the hospital, so I usually look for more information on the Internet after I see him.*

Finding trustworthy Arabic health information websites on the internet was a major challenge faced by most of the study participants. Many participants looked for health information on websites run by governmental organizations, as they felt more confident about the reliability of the provided health information. A 47-year-old male participant noted as follows:

> *We need to understand that not all health information pages are trustable. In my opinion, health websites sponsored by the Government are trustable sources of health information.*

In contrast, a 27-year-old male participant mentioned that the accessibility of the internet service and the connection speed represent another constraint to limit an entire subgroup of the population living in rural areas to access web-based health information. He noted as follows:

> *In my village, the internet is slower and more expensive than it is at the university. If you need to access the internet to find important health advice, you can't.*

In this theme, the study participants mentioned many other factors that encouraged them to seek OHI, such as the convenience of OHI, limited consultation time when seeing a physician, and the variability of health topics on the internet. Another group of participants demonstrated many constraints, which prevented them from using the OHI such as their limited access to internet services and the difficulty in finding trustworthy health information websites.

Influence of OHI on Participants' Health

The influence of OHI on participants' health was categorized into three subthemes; OHI helps users to (1) understand their health concerns, (2) improve their confidence when communicating with their physicians, and (3) achieve higher levels of satisfaction with health services. Many participants mentioned that the use of OHI enhanced their knowledge and understanding of their health concerns more than traditional

health information sources, such as leaflets or information provided by physicians. A male participant mentioned that being able to understand his health problem better helped him enjoy a healthier life. Another 27-year-old male participant mentioned that he sought OHI to improve his understanding of his newly diagnosed health condition. He noted as follows:

> *At least after I read it [refers to the online health information] I know how to ask the doctor about my next steps. If I do not understand my health problem, I will not be able to discuss it with my doctor.*

The participants agreed that OHI helped them become more confident when making a health decision, discussing their management options, or sharing their feelings with their physicians or family members. Another participant explained that he kept a list of questions regarding the health information he found on the internet to discuss it with his physician in his next visit. A 29-year-old female participant noted as follows:

> *After searching the internet for health information, I can say that it has helped me to understand my health problem more and gave me the courage to make better decisions for my treatment.*

Most of the participants reported achieving satisfaction with the health information obtained from the internet, especially when they were encouraged by their physicians to seek OHI. They also preferred to use health information from health websites that are recommended by their physicians. A 35-year-old female participant noted as follows:

> *When I asked my doctor about managing my condition (the participant is diabetic), he encourages me to search the internet. If I find health information, I print it and discuss it with him on my next visit.*

A 45-year-old male participant noted as follows:

> *I feel lucky that my doctor always encourages me to search health information on the internet and I believe this helped me a lot to be more confident about my treatment plan.*

In this domain, participants explained how the use of OHI improved their confidence and understanding of their health problems, particularly when they discussed their health management with their physicians. It also has a positive influence on their relationship with their physicians.

Physicians Interview Findings

Physicians in this study expressed different views based on their experience in providing OHI to Egyptian OHI seekers. We presented physicians' quotes in Table 3 owing to the small number of participating physicians in this study in comparison to the number of OHI seekers. Their personal experiences and opinions were essential to understand the impact of OHI on the physician-patient relationship. Themes emerging from physicians' interviews include the accessibility of OHI, good communication, and coordination of care between patients and their physicians, and the active engagement of patients with their management plan has a great impact on the effectiveness of the OHI and a better physician–patient relationship. Many physicians agreed that the benefits of using OHI outweighed its drawbacks when working with Egyptian internet users. They explained that some OHI users simply wanted to be informed about their condition, while others wanted to have full control over all medical decisions. However, this may result in incorrect health decisions or stress to the OHI seekers. A 48-year-old surgeon noted as follows:

> *To be honest, I love patients' sense of curiosity about their health and how they want to know more, but not every weight loss is cancer. Their online searches can sometimes go very far resulting in unnecessary stress.*

Physicians described the impact of OHI on the physician-patient relationship as a variable. One explained that some physicians advise their patients to trust their expertise and not to use OHI, as they feel that patients are challenging their knowledge when they seek OHI. In contrast, others encourage their patients to seek OHI to become more aware of their health problems and more involved in their health management. Another participant explained that guiding patients toward reliable health information is a part of their responsibilities.

A 48-year-old male participant noted as follows:

> *Delivery of validated health information through the internet should be our {physicians who provide health information through the internet} priority. We should provide our patients with a trustable source of health information.*

A 38-year-old pediatrician noted as follows:

> *We know that the Internet is accessible, and patients will use it. So, we should probably reduce the harm and refer them to reliable and high-quality health information websites.*

In this section, we explain how the physician-patient relationship has changed because of patients' use of OHI. Therefore, physicians and patients should collaborate in finding reliable health information websites.

Table 3. Major themes for physicians' attitude toward the OHI[a].

Themes or subtheme and attitude toward OHI	Example quotes
Accessibility	
Positive	"Waiting times in the hospital or clinics in general can be very long. So, when patients search the internet for health advice, they can avoid unnecessary travel time and get the information they need."
Coordination of care	
Positive	"Instead of the doctor being the only manager of patient care, internet health has emerged in which patients and their doctors are partners in managing their care.
Positive	"the Internet can actually empower patients and enrich the patient–doctor relationship."
Technical competence	
Positive	"I believe doctors should know the available sources of OHI and be aware of the updates."
Negative	"We should not assume that all patients know how to use the internet and access reliable health information sources."
Communication	
Negative	"Some physicians have concerns regarding the over informed patient."
Negative	"Some patients may give negative impression on their physicians when they discuss the obtained OHI with them. They feel challenged by their patients."
Negative	"I believe the intention to provide online health information was to grant doctors better conditions, to be faster with patients. but with over-informed patients, we spent more time with them to convince them with the treatment options and answer their endless questions."
Engagement	
Positive	"patients who read more about their health problem, I found them more engaged in their management plan."
Positive	"I think doctors has an important role which is encouraging patients to read more about their health problem and to take part in their health care plan."

[a]OHI: online health information.

Discussion

Principal Findings

In this study, we used a qualitative approach to provide an insight into the attitudes of Egyptian internet users and their OHI-seeking behaviors and the perspectives of physicians engaged with OHI. We used a graphic representation to show the key elements of our findings (Figure 1). Overall, the participants expressed a preference for using the internet as a source of health information in comparison with other health information sources, which was consistent with previous studies with both English-and Arabic-speaking internet users [3,12]. Participants provided convenience, anonymity, and accessibility of OHI as reasons for using OHI. Convenience and accessibility were also mentioned as important to young people who search the internet for mental health resources [18]. Notably, we found that traditional sources of health information (eg, physicians and television) are still the primary sources for other health information seekers, especially older adults who prefer personal interaction rather than using the internet, which is consistent with previous studies [2].

In this study, participants explained that OHI helped them understand their health problems more clearly, especially when they had a newly diagnosed health condition. The findings also suggest that OHI provides users with social support that encourages them to further use of health information on the internet. They explained that finding social support from persons with similar health problems facilitated their emotional adjustment and practical abilities to care for and access services for their health or other family members. In similar studies, health advice provided on the internet gave positive feedback on coping with emotional distress especially for OHI seekers with cancer and encouraged them to join the internet health communities [19,20].

Finding social support from patients with similar health problems has been reported in this study to reduce anxiety, especially for patients who are immobile and homebound because of their debilitating illness [21]. In contrast, Fergus et al [22] found that repeated searches for health information on the internet could exacerbate health anxiety for some users, especially in the absence of physicians' guidance. As reported in the physicians' interviews, the use of OHI may lead to confusion and cause unrealistic expectations from the obtained OHI, consequently increasing medical litigation against health care providers. Another study found that seeking OHI may increase the cost of health care by adopting inappropriate health advice, making the patients overconcerned about their health, and it could disturb the time efficiency of the physician consultation [23,24].

In this study, participants mentioned that limited internet service and connection speed in their residence were key barriers to prevent them from seeking OHI. Similar to our findings, a study

by McCloud et al [25] found that poor internet connections and access restrictions are significant barriers to seek OHI by adults, especially in rural areas. In this study, participants mentioned difficulties in finding trustable Arabic health websites owing to a shortage of this type of information, as they believe that English language is the dominant internet-based language. They explained that there are few certified Arabic medical websites, except websites of international organizations that provide translated content into different languages. Another study explained that the availability of OHI in the native language may enhance the utility of the internet for health information [26]. In addition, the quality of Arabic language health information websites was questioned by Alhajj et al [27], who concluded that the quality of web-based Arabic health information was poor regardless of being readable by internet users. The development of more credible Arabic websites offering more evidence-based and regularly updated information is recommended [28].

In this study, physicians expressed a general positive attitude toward the effects of OHI on users' health outcomes. They explained that sharing and discussing the obtained OHI with their patients will help internet users be more confident about their health decisions. Laugesen et al [29] explained that physicians have a greater impact on patient's health outcomes and compliance than internet health information. Concerns expressed by physicians in this study about the quality of health information on the internet and access to reliable websites followed similar studies [12,30]. Broom et al [24] recommended that to promote the impact of OHI on health, physicians should consider it as a complementary tool, rather than reflecting a breakdown in their authority or status. Hence, Egyptian health care providers need to take actions to ensure the dissemination of correct and reliable health information on the internet that matches the needs of Egyptian OHI seekers.

Limitations

Participants of this study were Egyptian adults who had already used the internet as a source of health information. We did not explore the perspectives of Egyptians who did not use the internet for health information. In addition, the personal characteristics of the participants were that we did not medically verify their general health status as it was self-assessed. Although participant numbers are usually small for this type of qualitative research, a larger number of participants could have uncovered a wider range of perspectives on the topic of study. Therefore, further research with a larger, purposive sample of Egyptian internet users would be valuable. In addition, we could not medically verify participants' general health status as it was self-assessed, and we did not have access to their medical records.

Conclusions

Our study is the first qualitative study to evaluate the attitude of Egyptians toward using OHI and how it influences their health behavior and their relationship with their physicians. The study participants demonstrated numerous benefits and challenges of using Arabic health information websites. Participants explained how OHI improved their knowledge and confidence, particularly when making health decisions. In addition, the findings from this study point to a range of concerns regarding the quality of the provided OHI and the limited availability of reliable Arabic health information websites. We recommend that health care providers provide more guidance and support to Egyptian OHI seekers to find good-quality health information websites. The responsible authorities should also ensure the availability of reliable Arabic language health websites.

Conflicts of Interest

None declared.

References

1. Ghweeba M, Lindenmeyer A, Shishi S, Abbas M, Waheed A, Amer S. What predicts online health information-seeking behavior among egyptian adults? A cross-sectional study. J Med Internet Res 2017 Jun 22;19(6):e216 [FREE Full text] [doi: 10.2196/jmir.6855] [Medline: 28642216]
2. Hassan S, Masoud O. Online health information seeking and health literacy among non-medical college students: gender differences. J Public Health (Berl.) 2020 Mar 09;29(6):1267-1273. [doi: 10.1007/s10389-020-01243-w]
3. Choudhury SM, Arora T, Alebbi S, Ahmed L, Aden A, Omar O, et al. How do Qataris source health information? PLoS One 2016;11(11):e0166250 [FREE Full text] [doi: 10.1371/journal.pone.0166250] [Medline: 27832112]
4. AlQarni ZA, Yunus F, Househ MS. Health information sharing on Facebook: an exploratory study on diabetes mellitus. J Infect Public Health 2016;9(6):708-712 [FREE Full text] [doi: 10.1016/j.jiph.2016.08.015] [Medline: 27618634]
5. Jamal A, Khan SA, AlHumud A, Al-Duhyyim A, Alrashed M, Bin Shabr F, et al. Association of online health information-seeking behavior and self-care activities among type 2 diabetic patients in Saudi Arabia. J Med Internet Res 2015 Aug 12;17(8):e196 [FREE Full text] [doi: 10.2196/jmir.4312] [Medline: 26268425]
6. Fiksdal AS, Kumbamu A, Jadhav AS, Cocos C, Nelsen LA, Pathak J, et al. Evaluating the process of online health information searching: a qualitative approach to exploring consumer perspectives. J Med Internet Res 2014 Oct 07;16(10):e224 [FREE Full text] [doi: 10.2196/jmir.3341] [Medline: 25348028]
7. Asiri E, Khalifa M, Shabir S, Hossain MN, Iqbal U, Househ M. Sharing sensitive health information through social media in the Arab world. Int J Qual Health Care 2017 Feb 01;29(1):68-74. [doi: 10.1093/intqhc/mzw137] [Medline: 28003369]

8. Chu JT, Wang MP, Shen C, Viswanath K, Lam TH, Chan SS. How, when and why people seek health information online: qualitative study in Hong Kong. Interact J Med Res 2017 Dec 12;6(2):e24 [FREE Full text] [doi: 10.2196/ijmr.7000] [Medline: 29233802]
9. Maslen S, Lupton D. "You can explore it more online": a qualitative study on Australian women's use of online health and medical information. BMC Health Serv Res 2018 Dec 03;18(1):916 [FREE Full text] [doi: 10.1186/s12913-018-3749-7] [Medline: 30509261]
10. Powell J, Inglis N, Ronnie J, Large S. The characteristics and motivations of online health information seekers: cross-sectional survey and qualitative interview study. J Med Internet Res 2011 Feb 23;13(1):e20 [FREE Full text] [doi: 10.2196/jmir.1600] [Medline: 21345783]
11. Nangsangna RD, da-Costa Vroom F. Factors influencing online health information seeking behaviour among patients in Kwahu West Municipal, Nkawkaw, Ghana. Online J Public Health Inform 2019;11(2):e13 [FREE Full text] [doi: 10.5210/ojphi.v11i2.10141] [Medline: 31632607]
12. Al-Ak'hali MS, Fageeh HN, Halboub E, Alhajj MN, Ariffin Z. Quality and readability of web-based Arabic health information on periodontal disease. BMC Med Inform Decis Mak 2021 Feb 04;21(1):41 [FREE Full text] [doi: 10.1186/s12911-021-01413-0] [Medline: 33541345]
13. Rosenbaum JE, Johnson BK, Deane AE. Health literacy and digital media use: assessing the health literacy skills instrument - short form and its correlates among African American college students. Digit Health 2018;4:2055207618770765 [FREE Full text] [doi: 10.1177/2055207618770765] [Medline: 29942630]
14. 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]
15. Sun Y, Zhang Y, Gwizdka J, Trace CB. Consumer evaluation of the quality of online health information: systematic literature review of relevant criteria and indicators. J Med Internet Res 2019 May 02;21(5):e12522 [FREE Full text] [doi: 10.2196/12522] [Medline: 31045507]
16. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
17. Verdinelli S, Scagnoli NI. Data display in qualitative research. Int J Qual Methods 2013 Feb 01;12(1):359-381. [doi: 10.1177/160940691301200117]
18. Kayrouz R, Dear BF, Karin E, Fogliati VJ, Gandy M, Keyrouz L, et al. Acceptability of mental health services for anxiety and depression in an Arab sample. Community Ment Health J 2018 Aug;54(6):875-883. [doi: 10.1007/s10597-018-0235-y] [Medline: 29368132]
19. Solberg LB. The benefits of online health communities. Virtual Mentor 2014 Apr 01;16(4):270-274 [FREE Full text] [doi: 10.1001/virtualmentor.2014.16.04.stas1-1404] [Medline: 24735575]
20. Kemp E, Koczwara B, Butow P, Turner J, Girgis A, Schofield P, et al. Online information and support needs of women with advanced breast cancer: a qualitative analysis. Support Care Cancer 2018 Oct;26(10):3489-3496. [doi: 10.1007/s00520-018-4206-1] [Medline: 29693203]
21. Singh K, Fox JR, Brown RJ. Health anxiety and internet use: a thematic analysis. Cyberpsychology 2016 Jul 01;10(2). [doi: 10.5817/cp2016-2-4]
22. Fergus TA, Dolan SL. Problematic internet use and internet searches for medical information: the role of health anxiety. Cyberpsychol Behav Soc Netw 2014 Dec;17(12):761-765. [doi: 10.1089/cyber.2014.0169] [Medline: 25412398]
23. Kim J, Kim S. Physicians' perception of the effects of internet health information on the doctor-patient relationship. Inform Health Soc Care 2009 Sep;34(3):136-148. [doi: 10.1080/17538150903102422] [Medline: 19670004]
24. Broom A. Medical specialists' accounts of the impact of the internet on the doctor/patient relationship. Health (London) 2005 Jul;9(3):319-338. [doi: 10.1177/1363459305052903] [Medline: 15937035]
25. McCloud RF, Okechukwu CA, Sorensen G, Viswanath K. Beyond access: barriers to internet health information seeking among the urban poor. J Am Med Inform Assoc 2016 Nov;23(6):1053-1059 [FREE Full text] [doi: 10.1093/jamia/ocv204] [Medline: 27206459]
26. Kyriacou A, Sherratt C. Online health information-seeking behavior by endocrinology patients. Hormones (Athens) 2019 Dec;18(4):495-505 [FREE Full text] [doi: 10.1007/s42000-019-00159-9] [Medline: 31749117]
27. Alhajj M, Mashyakhy M, Ariffin Z, Ab-Ghani Z, Johari Y, Salim N. Quality and readability of web-based arabic health information on denture hygiene: an infodemiology study. J Contemp Dent Pract 2020;21(9). [doi: 10.5005/jp-journals-10024-2918]
28. Alkhateeb JM, Alhadidi MS. Information about epilepsy on the internet: an exploratory study of Arabic websites. Epilepsy Behav 2018 Jan;78:288-290. [doi: 10.1016/j.yebeh.2017.07.012] [Medline: 29153965]
29. Laugesen J, Hassanein K, Yuan Y. The impact of internet health information on patient compliance: a research model and an empirical study. J Med Internet Res 2015 Jun 11;17(6):e143 [FREE Full text] [doi: 10.2196/jmir.4333] [Medline: 26068214]
30. 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]

Abbreviations

FGD: focus group discussion
OHI: online health information
PHI: personal health information

Edited by A Mavragani; submitted 02.05.21; peer-reviewed by M Alhajj, M Tizzani; comments to author 28.06.21; revised version received 05.07.21; accepted 12.12.21; published 18.02.22.

Please cite as:
Ghweeba M, Lindenmeyer A, Shishi S, Waheed A, Kofi M, Amer S
The Attitudes of Egyptian Web-Based Health Information Seekers Toward Health Information Provided Through the Internet: Qualitative Study
JMIR Form Res 2022;6(2):e30108
URL: https://formative.jmir.org/2022/2/e30108
doi:10.2196/30108
PMID:35179505

Original Paper

Use of Live Community Events on Facebook to Share Health and Clinical Research Information With a Minnesota Statewide Community: Exploratory Study

Jinhee Cha[1], MPH; Ian W West[2], MEd, MS; Tabetha A Brockman[1], MA; Miguel Valdez Soto[2], BA; Joyce E Balls-Berry[3], MPE, PhD; Milton Eder[1,4], PhD; Christi A Patten[2], PhD; Elisia L Cohen[1], PhD

[1]Clinical and Translational Science Institute, University of Minnesota, Minneapolis, MN, United States

[2]Center for Clinical and Translational Sciences, Mayo Clinic, Rochester, MN, United States

[3]Knight Alzheimer's Disease Research Center, Washington University in St. Louis, St Louis, MO, United States

[4]Department of Family Medicine and Community Health, University of Minnesota, Minneapolis, MN, United States

Corresponding Author:
Jinhee Cha, MPH
Clinical and Translational Science Institute
University of Minnesota
717 Delaware Street SE
Minneapolis, MN, 55414
United States
Phone: 1 612 625 2874
Email: cha00003@umn.edu

Abstract

Background: Community engagement can make a substantial difference in health outcomes and strengthen the capacity to deal with disruptive public health events such as the COVID-19 pandemic. Social media platforms such as Facebook are a promising avenue to reach the broader public and enhance access to clinical and translational science, and require further evaluation from the scientific community.

Objective: This study aims to describe the use of live community events to enhance communication about clinical and health research through a Facebook platform case study (*Minnesota [MN] Research Link*) with a Minnesota statewide community. We examined variables associated with video engagement including video length and type of posting.

Methods: From June 2019 to February 2021, *MN Research Link* streamed 38 live community events on its public Facebook page, *MN Research Link*. Live community events highlighted different investigators' clinical and health research in the areas of mental health, health and wellness, chronic diseases, and immunology/infectious diseases. Facebook analytics were used to determine the number of views, total minutes viewed, engagement metrics, and audience retention. An engagement rate was calculated by the total number of interactions (likes, shares, and comments) divided by the total length of the live event by the type of live community event.

Results: The 38 live community events averaged 23 minutes and 1 second in duration. The total time viewed for all 38 videos was 10 hours, 44 minutes, and 40 seconds. Viewers' watch time averaged 23 seconds of content per video. After adjusting for video length, promotional videos and research presentations had the highest engagement and retention rates. Events that included audience participation did not have higher retention rates compared to events without audience participation.

Conclusions: The use of live community events showed adequate levels of engagement from participants. A view time of 23 seconds on average per video suggests that short informational videos engage viewers of clinical and translational science content. Live community events on Facebook can be an effective method of advancing health promotion and clinical and translational science content; however, certain types of events have more impact on engagement than others.

(JMIR Form Res 2022;6(2):e30973) doi:10.2196/30973

KEYWORDS

social media; virtual; digital; community engagement; engagement; retention; Facebook; health information; information sharing; communication; participation; retention; eHealth

Introduction

Community engagement is essential for advancing clinical and translational science (CTS) to improve public health [1]. Community engagement can make a substantial difference in health outcomes, strengthen the capacity to deal with disruptive public health events such as the COVID-19 pandemic, and empower communities [2]. In addition, community engagement that advances and disseminates CTS is critical to alleviating health disparities within marginalized communities and improving health outcomes [3,4]. Social media platforms such as Facebook provide a promising avenue to reach the broader public. The use of social media to access CTS content requires further evaluation by the scientific community [5].

Though Facebook is used as a platform for community engagement, the literature regarding its use to disseminate information about both research and health topics is still in its infancy [5,6]. Developing a Facebook community and using social networks to build engagement for health promotion and CTS content has been recognized as a valuable endeavor [7]. Evaluations of pre-existing Facebook communities involved with increasing community engagement and retention have been useful for future researchers interested in disseminating CTS content online [8-10]. In a previous program evaluation, our collaborative partnership between the Mayo Clinic (Clinic Center for Clinical and Translational Science [CCaTS]) and University of Minnesota's Clinical and Translational Science Institute (CTSI) evaluated the feasibility of a virtual statewide Facebook community platform to enhance public engagement with health research in Minnesota [1]. Our partnership found that Facebook can be used as a community platform and is feasible in engaging Minnesotan residents in conversations around health and research topics. Nevertheless, additional research is needed to evaluate different types of Facebook posts and their potential to reach, engage, and retain community members with content posted on Facebook [1].

Facebook is one of the most widely used social media sites in the United States; 69% of adults report using Facebook, and its use is spread similarly across sociodemographic groups and geographic locations (urban/rural) [11]. For this reason, Facebook's breadth and reach has made it particularly useful in reaching underserved or disadvantaged communities [12-14]. The use of Facebook in health interventions has demonstrated an improvement in participant retention and engagement with research studies [8,15]. For researchers, Facebook offers tools for measuring engagement from their audience through its built-in engagement metrics, including reactions (ie, *likes*), shares, or comments [16]. Many built-in Facebook capabilities such as livestreaming are used for public health promotion. For instance, Facebook Live allows users to schedule and then livestream broadcasts to a wide-ranging audience. Live streaming of health and education content can reach specific audiences through users' social media routines and daily habits [8,17]. Therefore, Facebook, as a broad and established social media platform remains a viable tool for public health engagement that can be disseminated seamlessly into users' social media routines and daily habits [8].

Learning about Facebook algorithms could assist researchers in building engagement. Facebook "pages" host information about an organization and broadcast updates to Facebook members based on unknown algorithms. On these pages, Facebook users have the option to create different types of posts and content (eg, status updates, videos, images, and live community events) [6]. Nevertheless, the algorithm dictates the amount of exposure an individual post may receive and, thus, controls the visibility of the Facebook page [6]. This algorithm is informed by an individual user's unique interactions and behaviors with Facebook's content [17]. Facebook incentivizes users to build virtual communities, maintain interests, and populate their "front page" on Facebook [18,19]. Though researchers operate without a broad understanding of Facebook's proprietary algorithms (which can change without announcement), the literature suggests that certain types of content seem to gain preferential treatment. Posts and videos seem to gain more "likes" and exposure in comparison to status updates [6,20]. Additionally, posts that garner more attention and interactions from users also seem to gain preferential treatment from Facebook's algorithms. For instance, encouraging users to post feedback or asking the audience to participate with the page may encourage users to do so, leading to further dissemination on the platform [5].

Research examining types of social media posts and community engagement is in its infancy. Testimonials with positive emotional and informative posts are associated with higher engagement compared to negative emotional posts [6]. In an analysis among federal US health agencies' Facebook posts, posts with more visual cues such as photos or videos, or those that express positive sentiment generated more engagement [21]. Evaluative studies of clinical research and health posts are not readily evident, although work on business posting assessing positive and negative posts found negatively valanced posts garnering more comments [22]. Recent discussions of *leaked* internal Facebook documents further suggest increased user exposure to *polarizing* content increases engagement with the platform [23]. Nevertheless, what is clear to social media strategists from these reports is that Facebook continues to be a widely used platform for content developers to engage with audiences as its news feed, advertisements, post likes, comments, and livestreaming opportunities continue to drive engagement. Since Facebook's deployment of livestreaming, little research has investigated the role of live community events in fostering engagement or evaluating the use of livestreaming to promote health. The growth in its use has been in the field of strategic communication, and its impact to promote content has not yet been thoroughly investigated [24]. Thus, the question remains: how can researchers use the tools Facebook has to offer to create engaging content, use Facebook's algorithms, and promote CTS content?

To address this question, a partnership between the Mayo Clinic CCaTS and University of Minnesota CTSI used Facebook's tools to create the page *Minnesota (MN) Research Link*. *MN Research Link*'s mission was to provide credible information for Minnesota residents (1400 followers; n=434, 31% rural residents) in conversations about health and research topics, including health communication, health education, psychology,

behavioral science, epidemiology, public health, and COVID-19 [1]. Content and posts on *MN Research Link* included status updates, videos, images (eg, research articles related to health research), and interactive postings/events (eg, live interviews with researchers and polls) [1].

To address the central research question, this study uses *MN Research Link* as a case to examine whether Minnesotans engage with live community events (ie, videos) on a Facebook page (*MN Research Link*) and whether live events can be used to build followers (audience) to support CTS. This case study serves to generate hypotheses for future researchers interested in using live community events on Facebook to generate community engagement and promote CTS content.

Methods

Study Platform

To examine how live community events can promote CTS content, a Facebook page, *MN Research Link*, was developed by the Mayo Clinic (CCaTS) and the University of Minnesota (CTSI) National Institutes of Health–funded CTS Community Engagement Programs. *MN Research Link* was created to serve as a community platform for public engagement on health research statewide in Minnesota [1]. The project was deemed exempt by both the Mayo Clinic and University of Minnesota institutional review boards.

Designing Live Community Events

Based on expert recommendations for disseminating CTS content, members of the research team constructed an internal calendar of live community events to broadcast on *MN Research Link* [3]. Experts were recruited from the University of Minnesota, the Mayo Clinic, and from local organizations who had experience in developing CTS content for social media. Based on their recommendations, specific health topics were featured on *MN Research Link* monthly. Members of the research team also met weekly to plan daily posts that were relevant to *MN Research Link*'s health topics and in accordance with the Centers for Disease Control and Prevention's (CDC) social media guidelines [21]. Monthly health topics featured on *MN Research Link* included topics such as mental health, health and wellness (eg, nutrition and physical activity), chronic diseases (eg, cardiovascular disease), and immunology/infectious diseases. Over time, the research team adapted its monthly health topics to address important and relevant topics for followers. The final schedule included content on COVID-19, exercise, mental health, and well-being during the pandemic.

From June 2019 to February 2021 (20 months), four social media managers (authors JC, IWW, and MVS) facilitated 38 live community events on *MN Research Link*. Live community events of the following types were broadcasted on *MN Research Link* through the Facebook Live Event feature: brief interviews with experts, public health communications, research presentations, promotional videos for *MN Research Link*, and panels with expert public health professionals (Textbox 1). Through the Facebook Live Event feature and depending on the content and type of live community event, participants had the option of attending the live event through Zoom or watching on *MN Research Link*. During the live community events, social media managers facilitated discussions, moderated, or responded to participant questions on the *MN Research Link* page and through the Zoom chat. Guest speakers were recruited from the University of Minnesota, Mayo Clinic, and local organizations to participate in our live community events. When communicating with guest speakers, the social media managers emphasized that live community events were targeted for Minnesotans who may not have a background in science. Therefore, social media managers encouraged guest speakers to speak at a fifth-grade reading level.

Textbox 1. Key terms and their descriptions.

Interview

A live event where a social media manager of *Minnesota (MN) Research Link* speaks with a presenter about their research and expertise over Facebook.

Public health communication

A live event where a social media manager of *MN Research Link* speaks about a public health subject.

Research presentation

A live event where a social media manager of *MN Research Link* attends a research conference and interviews poster presenters on their research.

Promotional video

A live event where a social media manager of *MN Research Link* promotes *MN Research Link* events and posts or encourages engagement.

Panel

A live event where a social media manager of *MN Research Link* facilitates a conversation with more than one presenter on a public health subject.

Audience participation

A feature of live community events where a social media manager of *MN Research Link* encourages viewers to engage with the event by asking questions in the chat, asking questions live, or asking for feedback on the event.

Reactions (likes)

An engagement metric from Facebook Insights. User-generated indicators of post engagement that includes "like," "love," "haha," "wow," "sad," and "angry."

Shares

An engagement metric from Facebook Insights. Additionally a feature of Facebook posts where users may share the post/content on their own social media page and groups or with other users.

Comments

An engagement metric from Facebook Insights. User-generated comments, including replies to an individual post. Does not include comments from shares.

Clicks

An engagement metric from Facebook Insights. The number of unique clicks a user has made to play a live event.

Audience retention

An engagement metric from Facebook Insights. The average length of time viewers watch a video.

Promotion of Live Community Events

For interviews and panels, promotional materials were posted on *MN Research Link* at least 1 week before the date of the live event. Promotion materials included a post featuring graphics with details of the events, details on how to join the live event, and a promotion video from one of our social media managers briefly talking about the event. To create promotional videos, a research assistant uploaded a brief video (around 2 minutes) explaining the event, introducing the guest speaker, and sharing details of the live community events at least 2 days prior to the live event. Before a live event, approximately 1.5 hours were spent between the meetings communicating with the presenter or presenters, creating relevant posts, and creating a promotional video.

Advertisements or targeted "boosts" (Facebook's term for creating ads for posts across Facebook pages) [16] were used to expand the reach of the page to new audiences and to promote attendance for the live community events among Minnesotan Facebook users 18 years or older. Posts on the platform that receive paid boosts demonstrate greater reach and engagement with new audiences than posts that do not [22] and are an appropriate use of research funds to enhance audience reach. Approximately one-third of events were boosted 1 week prior to their start date. We used boosts to attract new followers and did not conceptualize boosts as a direct way to increase live event–specific engagement or retention rates. Boosts for *MN Research Link* posts occurred approximately 5 to 7 days prior to the event and were targeted for Minnesotan Facebook users 18 years and older. University of Minnesota and Mayo Clinic staff members, researchers, stakeholders, and community partners were asked to share events hosted by the *MN Research Link* page on their organization's Facebook page. Additionally, these members also shared events with individuals within their social network.

Procedure

Ideas for live community events were generated by weekly meetings between the Mayo Clinic CCaTS and University of Minnesota CTSI based on monthly health topics. After generating ideas for live community events, professors and researchers at the University of Minnesota and Mayo Clinic were invited as subject matter experts to participate in the live community events. For instance, when the health topic was "exercise during COVID-19," we invited a university presenter to talk about her research on exercise and how to exercise during the pandemic. Invitations were sent by email at least 2 weeks

prior to the suggested live event date with information about presenters and the project. Additionally, we encouraged presenters to share their research, in any form, ahead of time on *MN Research Link*. For instance, for the live event planned on "exercise during COVID-19," the presenter provided exercise science research to share on *MN Research Link*. Before the live event, the social media manager facilitated the live event to make sure the technology was working properly and that the presenter was ready. The length of the live events differed by type; interviews aimed to last less than 15 minutes in duration, whereas public health communications, research presentations, and promotional videos for *MN Research Link* were shorter (some less than 5 minutes in duration). Panels featuring expert public health professionals were less than an hour long.

Engagement Metrics and Analysis

To examine how live community events on *MN Research Link* can promote CTS content, a variety of metrics were used to measure audience engagement. Broadly considered, engagement metrics are metrics such as the number of followers, "likes," shares, clicks, and comments on individual live community events (Textbox 1). Although Facebook includes other reactions aside from "likes" (eg, "love," "care," or "laugh") as part of its engagement reaction metrics, for the purposes of our study, we operationalized all of these types of reactions to posts as a part of a singular Facebook engagement metric (Textbox 1).

The study used engagement metrics produced using Facebook Insights. Facebook Insights provides quantitative information of engagement metrics for all posts and content on a Facebook page [16]. Facebook Insights were used to generate descriptive statistics per live community events such as the total minutes viewed and audience retention. The audience retention rate was defined by Facebook as the average length of time (in seconds) viewers watch a video (Textbox 1). We also formulated our own "engagement score" by summarizing the total number of reactions (likes, shares, and comments) with an individual post (Textbox 1). The engagement score was used to summarize the reach and participation with *MN Research Link* through June 2019 to February 2021 (20 months) when live community events were implemented. Certain types of live community events may attract more attention than others because of Facebook's algorithms [18]. Moreover, longer live events offer more opportunities to engage with the content during the live broadcast. Therefore, we controlled for video length to determine the engagement and retention by the type of live community event.

To examine how audience participation on *MN Research Link* can promote CTS content, we compared engagement rate and audience retention rate across live community events with and without audience participation. We defined events with audience participation as events that featured the audience participating with the chat, asking questions live, or asking for feedback on the event. For example, for a panel on "Mental Health during COVID-19," the social media managers encouraged the audience to ask questions or make comments in the chat.

Results

Key terms and their descriptions are included in Textbox 1. A total of 38 live community events were hosted by *MN Research Link* using the Facebook platform (Table 1). Events averaged 23 minutes and 1 second. The total time viewed for all 38 videos was 10 hours, 44 minutes, and 40 seconds. Viewers spent an average of 23 seconds watching each piece of video content (Table 1).

Table 1. Descriptive statistics for all live video events, June 2019 to November 2020 (n=38).

Descriptive statistic	Value
Video length (hh:mm:ss), mean	00:23:01
Total length of video viewed (hh:mm:ss)	10:44:40
Audience retention (hh:mm:ss), mean	00:00:23

Interviews were the most frequent type of live community event hosted on *MN Research Link* (n=19; total video length 5:51:39), whereas promotional videos and panels were the least frequent type of live community events (n=3 and n=2; total video length 00:05:43 and 00:05:54, respectively; Table 2). After controlling for video length, promotional videos had the highest average engagement rate (0.404), followed by research presentation (0.224). After adjusting for video length, promotional videos and research presentations had the highest retention rate (0.096 and 0.090, respectively). Panels had the lowest engagement and retention rate, after adjusting for video length (0.001 and 0.020, respectively; Table 2).

Table 2. Types of live community events and engagement metrics adjusted by video lengths.

Type of event	Events, n (%)	Video length (hh:mm:ss)	Engagement rate[a]	Retention rate[b]
Interview	19 (50)	05:51:39	0.045	0.021
Public health communication	10 (26)	01:30:06	0.017	0.040
Research presentation	4 (11)	00:05:43	0.224	0.090
Promotional video	3 (8)	00:05:54	0.404	0.096
Panel	2 (5)	01:53:40	0.001	0.020

[a]Engagement rate is calculated by the sum engagement divided by the sum video length in seconds.

[b]Retention rate is calculated by the sum retention divided by the sum video length in seconds.

Of 38 live community events, 29 of them did not ask for or include audience participation in terms of questions and answers (Table 3). However, event engagement rates between events with audience participation and events without audience participation looked similar (0.042 and 0.049, respectively; Table 3). In contrast, events that did not include audience participation had higher audience retention rate scores compared to events with no audience participation (0.026 and 0.016, respectively; Table 3).

Table 3. Audience participation and engagement metrics adjusted by video length in seconds.

Audience participation type	Events, n (%)	Video length (hh:mm:ss)	Engagement rate[a]	Retention rate[b]
No audience participation included	29 (76)	07:27:28	0.042	0.026
Audience participation included	9 (24)	03:17:12	0.049	0.016

[a]Engagement rate is calculated by the sum engagement divided by the sum video length in seconds.

[b]Retention rate is calculated by the sum retention divided by the sum video length in seconds.

Discussion

Principal Results

To our knowledge, our study is the first to examine the use of live community events over Facebook to promote CTS content through community engagement. This case study examined whether Minnesotans would engage with live community events (ie, live videos) on a Facebook page (*MN Research Link*) and whether live community events could be used to build followers (audience) to support CTS. Using a novel engagement metric and tools by Facebook Insights, we summarized both the reach and engagement with followers on *MN Research Link*. Understanding how live community events can engage the public and how audience participation may enhance the promotion of CTS content (both in terms of their reactions to it and its stick-with-it-ness) will better inform researchers on how to reach communities via social media [20].

The findings indicate that live community events engaged users for CTS content. Facebook reports that the average retention for videos is 16.7 seconds [25], whereas the average retention for live community events on *MN Research Link* was 23 seconds (Table 1). Our retention time suggests that live community events are a better medium for retention than regular videos.

"Likes" or "reactions" to posts on Facebook suggests that users appreciated, liked, or were impacted with the post more so than usual [13,14]. The engagement rate score summed all reactions on each post to identify which types of CTS content garnered the most community engagement. The findings suggest that certain types of live community events have more of an impact on engagement than others; promotional videos (which generally contain minimal CTS research content) had the highest engagement, followed by research presentations, whereas public health communications and panels had the lowest amount of engagement (Table 3). Similarly, promotional videos had the highest retention rate, followed by research presentations. This suggests that social media users from Minnesota, on average, engaged with promotional videos the most and stayed engaged for longer periods of time compared to other types of live community events.

Generally, events with no audience participation garnered more retention for live community events but had minimal impact on engagement. More specifically, live community events with no audience participation element had 1.63 times the retention metric compared to live community events with audience participation (Table 3). This suggests that including audience participation does not engage social media users through "likes," "reactions," or "shares" and may even decrease how long audience members may remain attentive toward live community events. We offer two possible explanations as to why retention was lower for live community events with audience participation. First, audience participation requires more effort; viewers may not want to remain attentive to an event if they are asked to participate. Additionally, live community events with audience participation included panels that were significantly longer than other types of content. Since average retention was calculated by sum retention divided by sum video length in seconds, having longer videos would increase the denominator, resulting in a smaller quotient.

Implications

Our results add to the literature by suggesting that live community events on Facebook can not only be an effective method of advancing CTS content but also be better than

conventional posts with just pictures or videos. Our research builds on prior research suggesting the use of photos and videos garner more community engagement compared to posts that do not include photos and videos [6]. It suggests that though video engagement is important, the nature of the content matters for engagement, and researchers should consider the length of videos posted for users for CTS content promotion.

Developing and promoting live community events was a time intensive process; planning for a live event alone took approximately 1.5 hours. However, our case study demonstrates these events were only watched for an average of 23 seconds. This presents researchers with an interesting challenge; condensing CTS information into 30 seconds or less may be difficult for research dissemination. Ultimately, short form live community events may be more cost- and time-effective compared to longer videos for audience retention and engagement [25]. Ensuring viewers are greeted with a story and getting straight to the point, even if it is short, may also result in greater engagement and retention for the video [25]. Additionally, when hosting a longer live community event, asking viewers to participate may not result in increased retention. Thus, these findings raise appreciable evidence of interest for medical researchers and public health promotion experts seeking to leverage Facebook and other social media to improve audience engagement.

Findings from this case study confirm related findings on how brief informational video interventions, particularly those that are interactive and engaging, play important roles in health promotion. Data from this case study also suggests that short video content (eg, TikTok length) may be repurposed for effective use on Facebook to engage with the audience. This confirms the CDC's 2018 Health Communication Playbook, which identifies this strategy for consumer communication: keep information short and concise, avoid jargon, and use a content calendar to strategically engage your audience [26].

Conclusions

Although this case study demonstrates the ways that the *MN Research Link* page engages with its audience, the metrics provided are limited with respect to the constraints of the platform [16]. Future research could also evaluate how different health topics engage or retain Facebook users for CTS content. For example, because likes and shares are public [14], posts with stigmatizing topics such as mental health may or may not gain engagement compared to less stigmatizing topics. Such research could also examine how short form live community events or videos can be time-effective and offer more condensed information all while retaining audience retention. More work is needed, however, to determine how CTS content can be promoted through live community events though other social media networks (eg, Instagram Live or Twitch). These findings emphasize the need for ongoing CTS promotional research to ensure that accessible relevant health information is readily available for all social media users [2,12,27].

Acknowledgments

This project was supported by grants UL1 TR002494 and UL1 TR002377 from the National Center for Advancing Translational Sciences. Its contents are solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest

None declared.

References

1. Patten CA, Balls-Berry J, Cohen EL, Brockman TA, Valdez Soto M, West IW, et al. Feasibility of a virtual Facebook community platform for engagement on health research. J Clin Transl Sci 2021 Feb 18;5(1):e85 [FREE Full text] [doi: 10.1017/cts.2021.12] [Medline: 34007468]
2. Van den Broucke S. Why health promotion matters to the COVID-19 pandemic, and vice versa. Health Promot Int 2020 Apr 01;35(2):181-186 [FREE Full text] [doi: 10.1093/heapro/daaa042] [Medline: 32297931]
3. Michener L, Cook J, Ahmed SM, Yonas MA, Coyne-Beasley T, Aguilar-Gaxiola S. Aligning the goals of community-engaged research: why and how academic health centers can successfully engage with communities to improve health. Acad Med 2012 Mar;87(3):285-291 [FREE Full text] [doi: 10.1097/ACM.0b013e3182441680] [Medline: 22373619]
4. Woolf SH, Zimmerman E, Haley A, Krist AH. Authentic engagement of patients and communities can transform research, practice, and policy. Health Aff (Millwood) 2016 Apr;35(4):590-594 [FREE Full text] [doi: 10.1377/hlthaff.2015.1512] [Medline: 27044956]
5. Park H, Rodgers S, Stemmle J. Health organizations' use of Facebook for health advertising and promotion. J Interactive Advertising 2011 Sep;12(1):62-77. [doi: 10.1080/15252019.2011.10722191]
6. Kite J, Foley BC, Grunseit AC, Freeman B. Please like me: Facebook and public health communication. PLoS One 2016;11(9):e0162765 [FREE Full text] [doi: 10.1371/journal.pone.0162765] [Medline: 27632172]
7. Bensley R, Brookins-Fisher J. Community and public health education methods: a practical guide. University of Minnesota. Burlington, VT: Jones & Bartlett Learning, LLC; 2018. URL: http://ebookcentral.proquest.com/lib/umn/detail.action?docID=5303625ISBN:978-1-284-15686-7 [accessed 2021-10-01]

8. Laranjo L, Arguel A, Neves AL, Gallagher AM, Kaplan R, Mortimer N, et al. The influence of social networking sites on health behavior change: a systematic review and meta-analysis. J Am Med Inform Assoc 2015 Jan;22(1):243-256 [FREE Full text] [doi: 10.1136/amiajnl-2014-002841] [Medline: 25005606]
9. Woolley P, Peterson M. Efficacy of a health-related Facebook social network site on health-seeking behaviors. Soc Marketing Q 2012 May 02;18(1):29-39. [doi: 10.1177/1524500411435481]
10. Platt T, Platt J, Thiel DB, Kardia SLR. Facebook advertising across an engagement spectrum: a case example for public health communication. JMIR Public Health Surveill 2016 May 30;2(1):e27 [FREE Full text] [doi: 10.2196/publichealth.5623] [Medline: 27244774]
11. Perrin R, Anderson M. Share of U.S. adults using social media, including Facebook, is mostly unchanged since 2018. Pew Research Center. 2018. URL: https://www.pewresearch.org/fact-tank/2019/04/10/share-of-u-s-adults-using-social-media-including-facebook-is-mostly-unchanged-since-2018/ [accessed 2021-10-01]
12. Sendall MC, McCosker LK, Crane P, Rowland B, Fleming M, Biggs HC. Using Facebook for health promotion in "Hard-to-Reach" truck drivers: qualitative analysis. J Med Internet Res 2018 Nov 01;20(11):e286 [FREE Full text] [doi: 10.2196/jmir.9689] [Medline: 30389653]
13. Noonan D, Simmons LA. Navigating nonessential research trials during COVID-19: the push we needed for using digital technology to increase access for rural participants? J Rural Health 2021 Jan;37(1):185-187 [FREE Full text] [doi: 10.1111/jrh.12446] [Medline: 32282959]
14. Onitsuka K. How social media can foster social innovation in disadvantaged rural communities. Sustainability 2019 May 12;11(9):2697. [doi: 10.3390/su11092697]
15. Mychasiuk R, Benzies K. Facebook: an effective tool for participant retention in longitudinal research. Child Care Health Dev 2012 Sep;38(5):753-756. [doi: 10.1111/j.1365-2214.2011.01326.x] [Medline: 21985571]
16. Insights. Facebook. URL: https://www.facebook.com/help/794890670645072 [accessed 2021-02-07]
17. McGee M. EdgeRank is dead: Facebook's news feed algorithm now has close to 100K weight factors. MarTech. 2013. URL: https://marketingland.com/edgerank-is-dead-facebooks-news-feed-algorithm-now-has-close-to-100k-weight-factors-55908 [accessed 2021-02-06]
18. Bucher T. The friendship assemblage. Television N Media 2012 Aug 24;14(6):479-493. [doi: 10.1177/1527476412452800]
19. Bucher T. The algorithmic imaginary: exploring the ordinary affects of Facebook algorithms. Inf Commun Soc 2016 Feb 25;20(1):30-44. [doi: 10.1080/1369118x.2016.1154086]
20. Kite J, Grunseit A, Li V, Vineburg J, Berton N, Bauman A, et al. Generating engagement on the make healthy normal campaign Facebook page: analysis of Facebook analytics. JMIR Public Health Surveill 2019 Jan 14;5(1):e11132 [FREE Full text] [doi: 10.2196/11132] [Medline: 31344679]
21. Bhattacharya S, Srinivasan P, Polgreen P. Social media engagement analysis of U.S. Federal health agencies on Facebook. BMC Med Inform Decis Mak 2017 Apr 21;17(1):49 [FREE Full text] [doi: 10.1186/s12911-017-0447-z] [Medline: 28431582]
22. Smith A. New research shows negative posts on Facebook business pages outweigh positive posts 2 to 1: the impact it's having on brands and profitability. INFORMS. URL: https://tinyurl.com/25tx6vpx [accessed 2021-11-16]
23. Hwang K. Here are 5 major allegations whistle-blower Frances Haugen has leveled against Facebook. MSN. URL: https://www.msn.com/en-us/news/us/here-are-5-major-allegations-whistle-blower-frances-haugen-has-leveled-against-facebook/ar-AAP8eC2 [accessed 2021-11-16]
24. Woodcock J, Johnson MR. Live streamers on Twitch.tv as social media influencers: chances and challenges for strategic communication. Int J Strategic Commun 2019 Sep 04;13(4):321-335. [doi: 10.1080/1553118x.2019.1630412]
25. Rabkin M. New medium, new rules: video advertising in the mobile age. Facebook. URL: https://www.facebook.com/business/news/new-medium-new-rules-video-advertising-in-the-mobile-age [accessed 2021-11-11]
26. Health communication playbook. Centers for Disease Control and Prevention. 2018. URL: https://stacks.cdc.gov/view/cdc/103379 [accessed 2022-02-07]
27. Saud M, Mashud M, Ida R. Usage of social media during the pandemic: seeking support and awareness about COVID-19 through social media platforms. J Public Affairs 2020 Sep 15:e02417. [doi: 10.1002/pa.2417]

Abbreviations

CCaTS: Clinic Center for Clinical and Translational Science
CDC: Centers for Disease Control and Prevention
CTS: clinical and translational science
CTSI: Clinical and Translational Science Institute
MN: Minnesota

Edited by A Mavragani; submitted 07.06.21; peer-reviewed by N Finn, W Buente, M Salimi, Y Zhang; comments to author 24.09.21; revised version received 18.11.21; accepted 18.12.21; published 18.02.22.

Please cite as:
Cha J, West IW, Brockman TA, Soto MV, Balls-Berry JE, Eder M, Patten CA, Cohen EL
Use of Live Community Events on Facebook to Share Health and Clinical Research Information With a Minnesota Statewide Community: Exploratory Study
JMIR Form Res 2022;6(2):e30973
URL: https://formative.jmir.org/2022/2/e30973
doi:10.2196/30973
PMID:35179514

Original Paper

A Digital Mental Health Intervention in an Orthopedic Setting for Patients With Symptoms of Depression and/or Anxiety: Feasibility Prospective Cohort Study

Ashwin J Leo[1], BA; Matthew J Schuelke[2], DPhil; Devyani M Hunt[3], MD; John P Metzler[3], MD; J Philip Miller[2], BA; Patricia A Areán[4], DPhil; Melissa A Armbrecht[3]; Abby L Cheng[3], MD, MPHS

[1]Washington University in St Louis School of Medicine, St Louis, MO, United States

[2]Division of Biostatistics, Washington University in St Louis School of Medicine, St Louis, MO, United States

[3]Division of Physical Medicine and Rehabilitation, Department of Orthopedic Surgery, Washington University in St Louis School of Medicine, St Louis, MO, United States

[4]Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, United States

Corresponding Author:
Abby L Cheng, MD, MPHS
Division of Physical Medicine and Rehabilitation
Department of Orthopedic Surgery
Washington University in St Louis School of Medicine
660 South Euclid Avenue
St Louis, MO, 63110
United States
Phone: 1 3147472823
Email: chengal@wustl.edu

Abstract

Background: Symptoms of depression and anxiety commonly coexist with chronic musculoskeletal pain, and when this occurs, standard orthopedic treatment is less effective. However, mental health intervention is not yet a routine part of standard orthopedic treatment, in part because of access-related barriers. Digital mental health intervention is a potential scalable resource that could be feasibly incorporated into orthopedic care.

Objective: This study's primary purpose was to assess the feasibility of introducing a digital mental health intervention (Wysa) in an outpatient orthopedic setting to patients with coexisting symptoms of depression and/or anxiety. The secondary purpose was to perform a preliminary effectiveness analysis of the intervention.

Methods: In this single-arm, prospective cohort study, participants included adult patients (18 years and older) who presented to a nonsurgical orthopedic specialist at a single tertiary care academic center for evaluation of a musculoskeletal condition and who self-reported symptoms of depression and/or anxiety (Patient-Reported Outcomes Measurement Information System [PROMIS] Depression and/or Anxiety score ≥55). Face-to-face enrollment was performed by a research coordinator immediately after the participant's encounter with an orthopedic clinician. Participants were provided 2 months of access to a mobile app called Wysa, which is an established, multicomponent digital mental health intervention that uses chatbot technology and text-based access to human counselors to deliver cognitive behavioral therapy, mindfulness training, and sleep tools, among other features. For this study, Wysa access also included novel, behavioral activation–based features specifically developed for users with chronic pain. Primary feasibility outcomes included the study recruitment rate, retention rate, and engagement rate with Wysa (defined as engagement with a therapeutic Wysa tool at least once during the study period). Secondary effectiveness outcomes were between-group differences in mean longitudinal PROMIS mental and physical health score changes at 2-month follow-up between high and low Wysa users, defined by a median split.

Results: The recruitment rate was 29.3% (61/208), retention rate was 84% (51/61), and engagement rate was 72% (44/61). Compared to low users, high users reported greater improvement in PROMIS Anxiety scores (between-group difference −4.2 points, 95% CI −8.1 to −0.2; *P*=.04) at the 2-month follow-up. Between-group differences in PROMIS Depression (−3.2 points, 95% CI −7.5 to 1.2; *P*=.15) and Pain Interference scores (−2.3 points, 95% CI −6.3 to 1.7; *P*=.26) favored high users but did not meet statistical significance. Improvements in PROMIS Physical Function scores were comparable between groups.

Conclusions: Delivery of a digital mental health intervention within the context of orthopedic care is feasible and has the potential to improve mental health and pain-related impairment to a clinically meaningful degree. Participants' engagement rates exceeded industry standards, and additional opportunities to improve recruitment and retention were identified. Further pilot study followed by a definitive, randomized controlled trial is warranted.

Trial Registration: ClinicalTrials.gov NCT04640090; https://clinicaltrials.gov/ct2/show/NCT04640090

(JMIR Form Res 2022;6(2):e34889) doi:10.2196/34889

KEYWORDS

digital health; mental health; depression; anxiety; chronic pain; musculoskeletal; orthopedic; pain management; health intervention; mobile phone

Introduction

Background

Of the over 70 million Americans who have chronic musculoskeletal pain, up to half of them have a comorbid diagnosis of depression and/or anxiety [1-4]. Chronic pain and psychological impairment such as depression and anxiety are driven by shared biological pathways and neurotransmitters, and when a person has both conditions simultaneously, the effectiveness of standard treatment for either condition in isolation is reduced [5-7]. For example, among people who seek care for chronic musculoskeletal pain, those with coexisting depression and/or anxiety report worse pain, greater physical impairment, increased postoperative opioid use, and a lower rate of return to work than those without depression and anxiety [8-11].

Despite strong evidence regarding the interconnection between physical pain and mental health, psychological assessment and intervention is not yet considered a routine part of orthopedic treatment. This is, in part, because of both provider- and patient-related barriers. Orthopedic clinicians do not feel equipped with the necessary time and referral resources to address mental health [12], and patients report barriers to seeking mental health treatment even outside an orthopedic setting. Among people who recognize their need for mental health treatment, 64% are resistant to reaching out to another person for assistance and 23% have financial, transportation, or time-related barriers to seeking care [13,14]. Even if all patients had the resources to seek appropriate care, unmet need would persist because of a national shortage of qualified mental health clinicians [15].

Digital mental health interventions address the accessibility-, cost-, convenience-, and stigma-related barriers to traditional in-person mental health treatment [16]. Furthermore, a growing body of evidence supports the effectiveness of digital mental health interventions in improving depression and anxiety symptoms, sometimes to a degree equivalent to that obtained from in-person treatment [17-19]. With the widespread, everyday use of smartphones, these emerging digital tools could function as efficient, widely available mental health resources for orthopedic clinicians to offer their patients [16,20-22].

Objectives

The primary purpose of this study was to assess the feasibility of introducing a digital mental health intervention (Wysa) in an outpatient orthopedic setting to patients with musculoskeletal pain and coexisting symptoms of depression and/or anxiety. The secondary purpose was to perform a preliminary effectiveness analysis of the intervention. We hypothesized that delivery of a digital mental health intervention would be feasible in an orthopedic clinic setting. Furthermore, compared to "low users" of the digital mental health intervention, "high users" would report greater improvements in mental and physical health symptoms at the 2-month follow-up.

Methods

Design

This was a single-arm, prospective cohort, pilot feasibility study performed at a tertiary care academic medical center in the United States. Institutional review board approval was obtained prior to participant recruitment (IRB #202005219), and the study was registered through ClinicalTrials.gov (NCT04640090). Participants were enrolled from December 8, 2020, through July 14, 2021, with some interruptions related to the COVID-19 pandemic.

Participants

Eligibility Criteria

Participants were recruited from among patients who presented to an orthopedic department for evaluation and treatment of a musculoskeletal condition. To be eligible, patients had to be adults (18 years or older) presenting to a nonsurgical orthopedic provider. Additionally, patients had to have symptoms of depression and/or anxiety as indicated by a score of 55 or higher on the Patient-Reported Outcomes Measurement Information System (PROMIS) Depression and Anxiety measures; this is collected as standard clinical care on check-in for each outpatient appointment at this orthopedic department. Patients who were actively planning to start in-person mental health treatment were excluded from participation. Patients without access to a mobile device were also not eligible.

Recruitment Process

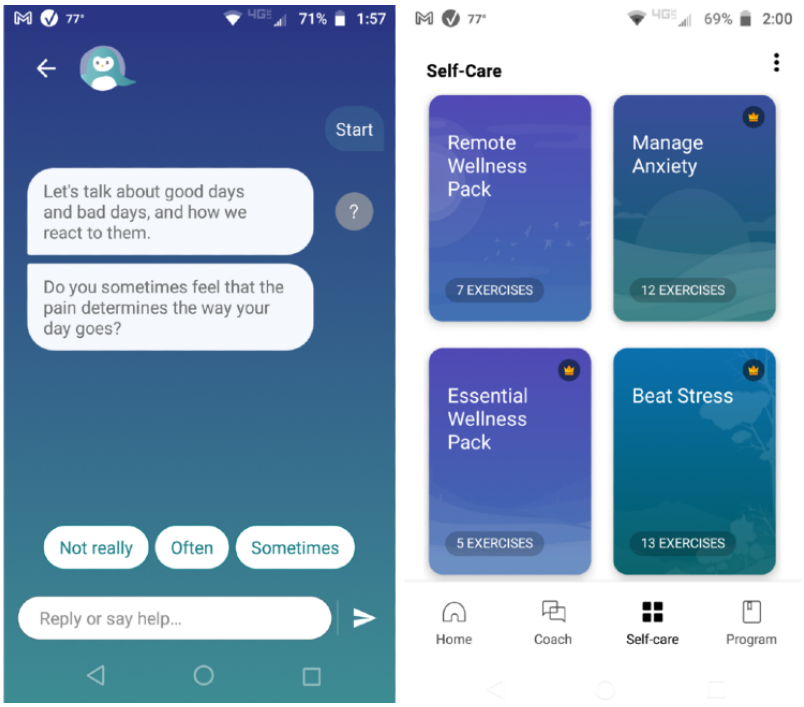
We recruited participants from among patients who presented to 6 orthopedic clinicians, 5 of whom were board-certified physical medicine and rehabilitation physicians (physiatrists) with subspecialty training in spine and sports medicine and 1 who was a nurse practitioner with 11 years of nonsurgical orthopedics experience. Prior to the clinician's encounter with each patient, the study coordinator notified the clinician of

patients who met the age and PROMIS score criteria for the study. Then, at the end of a potentially eligible patient's standard clinic encounter, the clinician briefly introduced the study. For those who expressed initial interest, the research coordinator approached the patient, provided further study details, and completed the final screening, consent, and enrollment process for those who were interested. When possible, participant enrollment was completed in person before the participant left the orthopedic clinic office. Patients who were interested in the study but did not have time to enroll immediately were contacted over the phone by the research coordinator to complete the enrollment process. No compensation besides access to the study intervention was provided to participants for their participation in the study.

Intervention

All study participants received a complimentary, 2-month subscription to Wysa, which is a commercially available digital mental health intervention. Wysa is a multicomponent intervention that uses artificial intelligence–based chatbot technology and human "coaches" (counselors) with master's degrees in psychology to deliver therapeutic content such as cognitive behavioral therapy, dialectical behavioral therapy, motivational interviewing, mindfulness training, deep breathing techniques, and sleep meditations [23,24]. Commercially, basic chatbot features are freely available, and premium digital features and text-based sessions with a human coach are available for a monthly fee. In this study, participants received access to a novel version of Wysa that was specifically developed for users interested in tools to manage chronic pain (Figure 1). Using behavioral activation and pain acceptance principles, this pain-specific version of Wysa encourages users to "engage with things that bring joy, despite having pain" [25-28]. Added features include daily check-ins, weekly reports, the ability to unlock premium "reward" tool packs by engaging with the weekly reports, and a progress roadmap. The chatbot and human coach features are also still embedded within the pain version of Wysa. This version is not yet part of the Wysa commercial product.

Figure 1. Screenshots of Wysa's chronic pain version.



Variables

Data were obtained from 3 sources: electronic medical records, participants' self-report, and usage reports from Wysa. Data collected as standard clinical care were obtained from the electronic medical record; this included basic sociodemographic variables, medical history, and pain medication use. Additionally, on the day of the orthopedic clinic visit and study enrollment, patients completed Adult PROMIS CAT Depression v1.0, Anxiety v1.0, Pain Interference v1.1, and Physical Function v2.0 measures as standard care prior to meeting with the clinician. These measures were completed on a tablet computer (iPad Mini; Apple Inc). Each PROMIS domain is normalized to the general United States population with a mean (SD) score of 50 (SD 10) [29]. Higher scores represent more of the domain; for instance, a score of 60 on PROMIS Anxiety indicates anxiety symptoms that are worse than those in the general population, but a score of 60 on PROMIS Physical Function indicates physical function that is better than that in the general population.

On study enrollment, participants completed additional self-reported measures on the tablet computers. These included participants' pain duration, primary pain location, height and weight (to calculate BMI), history of tobacco use, exercise habits, Brief Resilience Scale [30], and Importance, Readiness, and Confidence to Change rulers [31].

At the 1- and 2-month follow-up, participants were emailed a questionnaire that included PROMIS Depression, Anxiety, Pain Interference, and Physical Function domains; participants

received automated email reminders as needed 2 and 4 days later. For the 2-month follow-up, they also received up to 3 phone call reminders if they had not completed the questionnaire by the sixth day after the follow-up period. All electronic medical record and self-reported data were stored in a secure electronic Research Electronic Data Capture (REDCap; Vanderbilt University) database [32,33].

Wysa provided user-specific, timestamped usage data for each app interaction by participants. Individual users were identified with randomly generated codes, and only study personnel had access to the participant key. No personally identifiable data were shared with Wysa by the study team or vice versa.

Outcomes

Feasibility

The primary study outcomes were related to the feasibility of introducing a digital mental health intervention (Wysa) in an outpatient orthopedic setting and conducting a fully powered, randomized controlled effectiveness trial in this setting. Outcomes included the study recruitment rate (defined as the number of patients who signed the study informed consent document divided by the number of patients approached for participation), study retention rate (defined as the proportion of recruited participants who completed the 2-month follow-up PROMIS measures), and the rate of engagement with Wysa (defined as the proportion of recruited participants who interacted with Wysa at least once after onboarding to the app). Other engagement metrics were also evaluated, including the weekly engagement rate (defined as the proportion of recruited participants who completed at least 8 interactions with Wysa over the 8-week period, because this would approximately mirror the intensity of weekly sessions with an in-person therapist), the coach engagement rate (defined as the proportion of recruited participants who completed at least one session with a human coach), and the most frequently used features within Wysa identified in a descriptive analysis. An exploratory analysis was also conducted to assess for differences in patient demographics (ie, age, sex) between those who enrolled in this study and those who did not.

Effectiveness

A secondary purpose of the study was to perform a preliminary effectiveness analysis of the intervention. To account for clinical improvements that could be related to usual orthopedic care or natural variations in symptom severity over time, the effectiveness outcomes considered were the between-group differences (between high and low Wysa users) in 2-month changes on PROMIS Depression, Anxiety, Pain Interference, and Physical Function measures. High versus low Wysa usage was defined by a participant's number of total Wysa interactions lying above versus below the median engagement level among participants. Minimum clinically meaningful effect sizes were a priori chosen in accordance with previously published literature involving patients with chronic musculoskeletal pain who were managed nonoperatively. Effect sizes also had to exceed the SE of measurement for each PROMIS computer adaptive test at the study institution to be considered meaningful. Therefore, clinically meaningful effect sizes were defined as at least 3.2 points for PROMIS Depression, 3.0 points for Anxiety, 2.0 points for Pain Interference, and 2.2 points for Physical Function [34-36].

Statistical Analysis

Univariate descriptive statistics were calculated for all study variables. Age and sex were compared between those who did and did not enroll using an unequal-variances *t* test and a two-sample chi-square test of proportions, respectively. A median split on the total number of Wysa interactions was chosen to categorize participants as high and low Wysa users because density plots of the distributions of Wysa total interactions, chatbot interactions, and coach sessions failed to identify any empirical cut points. For the effectiveness analyses, linear mixed-effects models were fit predicting each of the 4 PROMIS measures from Wysa usage status (high versus low) and month. Adjusting for age as a covariate in the models did not meaningfully influence the model results, so unadjusted models are reported. Missing data were omitted. Significance was set a priori at $P<.05$, and sample size for this feasibility study was determined by the availability of resources. Statistical analyses were performed using R (v4.0.2; R Core Team).

Results

Feasibility

The study recruitment rate was 29.3% (61/208; Figure 2). Among the 61 patients who enrolled in the study, the median age was 55 years (range 18-83 years) and 87% (53/61) were female (Tables 1 and 2). The most common reasons patients did not enroll were that they did not want to use the app (104/147, 70.7%) or did not want to complete the research surveys (26/147, 17.7%). No statistically significant differences were observed in the age or sex distribution of patients who enrolled versus those who did not enroll in the study ($P=.21$; Table 3).

The retention rate of participants who consented to participate in the study was 84% (51/61). More participants completed the 2-month follow-up measures (51/61, 84% using email and phone reminders) than the 1-month follow-up measures (36/61, 59% using email reminders only).

The overall rate of engagement with Wysa among participants who consented to participate in the study was 72% (44/61). The weekly engagement rate was 57% (35/61), and the coach engagement rate was 33% (20/61). The therapeutic mechanisms underlying the tools most commonly used by participants were related to mindfulness (122/351, 34.8%), cognitive behavioral therapy (76/351, 21.7%), and sleep (53/351, 15.1%; Multimedia Appendix 1). Of the rewards unlocked by participants, sleep-related rewards were the most common (16/48, 33%) Table 4).

Figure 2. Participant flow, from initial study contact through final follow-up.

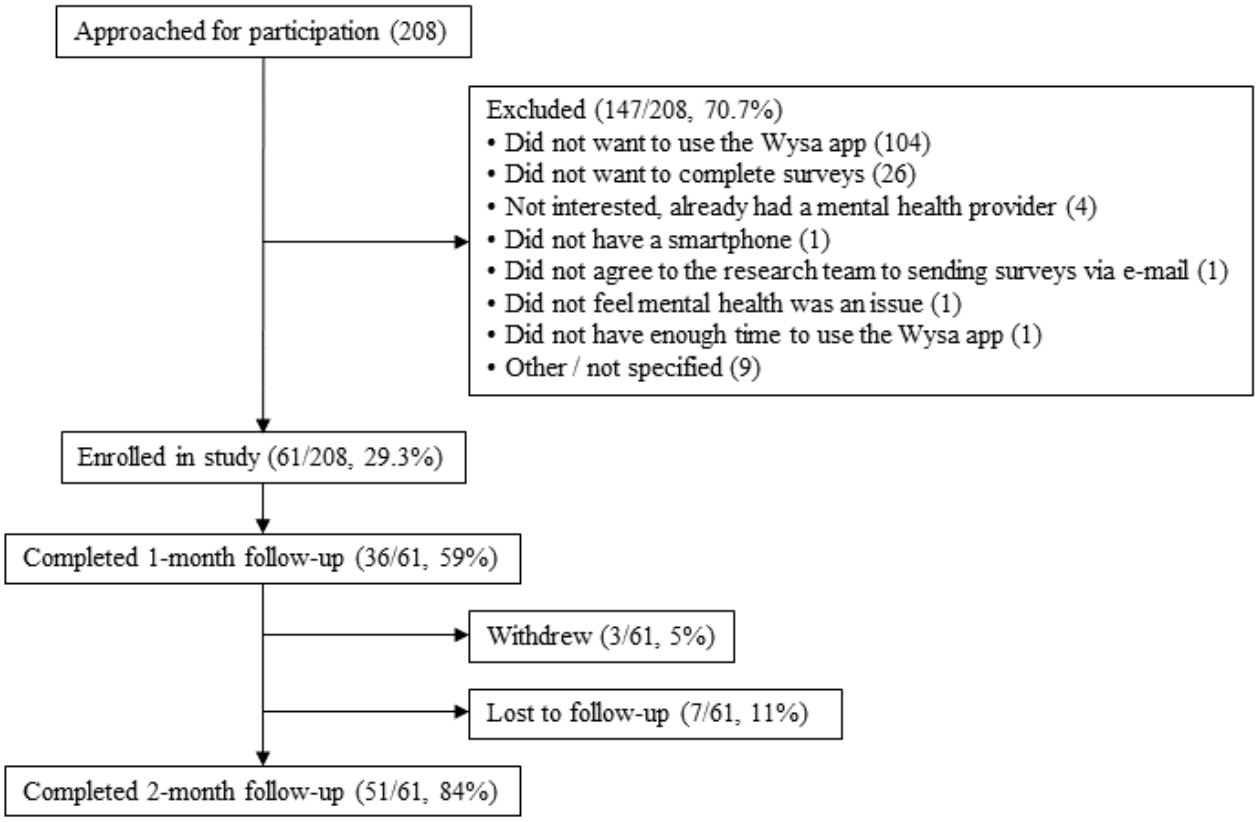


Table 1. Baseline characteristics of patients who presented to an orthopedic clinic for a musculoskeletal condition, reported coexisting symptoms of depression and/or anxiety, and enrolled in a research study that provided access to a digital mental health intervention (Wysa) (N=61).

Characteristic	Value
Age (years), median (IQR)	55 (42-64)
Sex, n (%)	
Female	53 (87)
Male	7 (12)
Transgender	1 (2)
Race, n (%)	
White/Caucasian	55 (90)
Black/African American	6 (10)
Ethnicity, n (%)	
Hispanic	2 (3)
Not Hispanic	59 (97)
Area Deprivation Index[a], n (%)	
Quartile 1 (least deprived)	19 (31)
Quartile 2	15 (25)
Quartile 3	14 (23)
Quartile 4 (most deprived)	13 (21)
Pain duration (years), median (IQR)	2 (0.9-5.9)
Pain location[b]	
Low back	31 (51)
Leg	19 (31)
Neck	18 (30)
Arm	8 (13)
Generalized pain	3 (5)
BMI (kg/m^2), mean (SD)	29 (7)
Tobacco use, n (%)	4 (7)
Moderate/strenuous exercise (minutes/week), median (IQR)	40 (0-120)
Exercise limited by pain, n (%)	49 (85)
Exercise limited by enjoyment, n (%)	6 (11)
Brief Resilience Scale score[c], median (IQR)	3.2 (2.5-3.7)
Potential for behavior change[d], median (IQR)	
Importance of change	76 (68-90)
Readiness for change	69 (50-82)
Confidence in the ability to change	68 (50-74)
Any pain medication use, n (%)	46 (75)
Specific pain medication use[e], n (%)	
Opioid	7 (15)
Nonsteroidal anti-inflammatory drug	15 (33)
Neuropathic	22 (48)
Other	20 (44)
Medical history, n (%)	

Characteristic	Value
Hypertension	18 (30)
Hyperlipidemia	17 (28)
Cardiovascular disease	7 (11)
Diabetes	8 (13)
Sleep apnea	9 (15)
Depression	24 (39)
Anxiety	23 (38)

[a]The national Area Deprivation Index is a community-level measure of social disadvantage based on a person's 9-digit zip code [37,38].

[b]Some patients reported multiple pain locations.

[c]The Brief Resilience Scale is scored from 1 to 5, with higher scores representing greater resilience [30].

[d]Potential for change measures are scored from 0 to 100, and higher scores are favorable [39].

[e]Number of patients among the 46 patients who used any pain medication.

Table 2. Summary of mental and physical health scores measured by the Patient-Reported Outcomes Measurement Information System over the 2-month follow-up in patients who presented to an orthopedic clinic for a musculoskeletal condition, reported coexisting symptoms of depression and/or anxiety, and enrolled in a research study that provided access to a digital mental health intervention (Wysa).

PROMIS[a] scores	Baseline (n=61), mean (SD)	2-month follow-up (n=51), mean (SD)
Depression	58.1 (7.3)	54.7 (8.7)
Anxiety	62.2 (5.9)	58.0 (7.8)
Pain interference	65.2 (6.5)	62.1 (7)
Physical function	35.9 (6.6)	39.5 (6.7)

[a]PROMIS: Patient-Reported Outcomes Measurement Information System.

Table 3. Demographic comparison of patients who presented to an orthopedic clinic for a musculoskeletal condition, reported coexisting symptoms of depression and/or anxiety, and enrolled versus those who did not enroll in a research study that provided access to a digital mental health intervention (Wysa) (N=208).

Demographic	Enrolled (n=61)	Did not enroll (n=147)	Between-group difference	
			Mean difference or % difference (95% CI)	*P* value
Age (years), mean (SD)	53 (14)	56 (18)	3 (−2 to 8)	.21
Female sex, n (%)	53 (87)	115 (78.2)	9 (−3 to 21)	.21

Table 4. Summary of engagement with a digital mental health intervention (Wysa) by patients who presented to an orthopedic clinic for a musculoskeletal condition and reported coexisting symptoms of depression and/or anxiety (N=61).

Engagement metric	Median (IQR)	Range
Total interactions	18 (0-64)	0-544
Interactions with chatbot	9 (0-29)	0-181
Daily check-ins completed	2 (0-16)	0-81
Weekly reports viewed	0 (0-2)	0-14
Rewards unlocked	0 (0)	0-8
Text-based sessions with a human coach	0 (0-1)	0-12
Total messages to a human coach	0 (0-27)	0-382

Effectiveness

Compared to low Wysa users, high Wysa users reported greater improvement to a clinically meaningful degree in PROMIS Anxiety scores at the 2-month follow-up (between-group difference −4.2 points, 95% CI −8.1 to −0.2; *P*=.04; Table 5). Between-group differences in 2-month improvements in PROMIS Depression (−3.2 points, 95% CI −7.5 to 1.2; *P*=.15) and Pain Interference scores (−2.3 points, 95% CI −6.3 to 1.7; *P*=.26) favored high Wysa users and met clinically meaningful

thresholds but did not meet statistical significance. In contrast, both groups reported clinically meaningful and statistically significant improvements in PROMIS Physical Function scores from baseline to 2 months (*P*=.99), but there was no between-group difference in this domain (Figures 3-6).

Table 5. Summary of mental and physical health changes (measured by the Patient-Reported Outcomes Measurement Information System) across the 2-month follow-up in patients who presented to an orthopedic clinic for a musculoskeletal condition and reported coexisting symptoms of depression and anxiety subgrouped by frequency of use of a digital mental health intervention (Wysa).

PROMIS[a] domain	Baseline (n=61), mean (95% CI)	2-month follow-up (n=51)[b], mean (95% CI)	Within-group longitudinal change, mean (95% CI)	Between-group change[c]	
				Mean (95% CI)	*P* value
Depression				−3.2 (−7.5 to 1.2)	.15
High users[d]	57.5 (54.5 to 60.4)	52.8 (49.9 to 55.8)	−4.7 (−7.5 to −1.8)		
Low users[e]	58.7 (55.8 to 61.7)	57.3 (53.9 to 60.7)	−1.5 (−4.7 to 1.8)		
Anxiety				−4.2 (−8.1 to −0.2)	.04
High users	62.1 (59.6 to 64.7)	56.6 (54.0 to 59.1)	−5.6 (−8.2 to −3.0)		
Low users	62.4 (59.8 to 64.9)	61.0 (58.0 to 63.9)	−1.4 (−4.4 to 1.6)		
Pain interference				−2.3 (−6.3 to 1.7)	.26
High users	64.5 (62.1 to 66.8)	60.6 (58.2 to 63.0)	−3.9 (−6.5 to −1.2)		
Low users	65.9 (63.5 to 68.3)	64.3 (61.6 to 67.1)	−1.5 (−4.6 to 1.5)		
Physical function				0.0 (−3.4 to 3.5)	.99
High users	37.1 (34.6 to 39.5)	40.5 (38.0 to 42.9)	3.4 (1.2 to 5.7)		
Low users	34.7 (32.3 to 37.1)	38.0 (35.3 to 40.8)	3.4 (0.8 to 6.0)		

[a]PROMIS: Patient-Reported Outcomes Measurement Information System.

[b]The 2-month follow-up PROMIS measures were available for 30 high and 21 low Wysa users.

[c]Clinically meaningful effect sizes are defined as at least 3.2 points for PROMIS Depression, 3.0 points for Anxiety, 2.0 points for Pain Interference, and 2.2 points for Physical Function [34-36].

[d]There were 30 patients in the high Wysa use subgroup.

[e]There were 31 patients in the low Wysa use subgroup.

Figure 3. Mean longitudinal change in Patient-Reported Outcomes Measurement Information System (PROMIS) Depression scores over 2-month follow-up in patients who presented to an orthopedic clinic for a musculoskeletal condition, reported coexisting symptoms of depression and/or anxiety, and were high users (n=30, green circles) versus low users (n=31, blue triangles) of a digital mental health intervention (Wysa). Error bars represent standard errors.

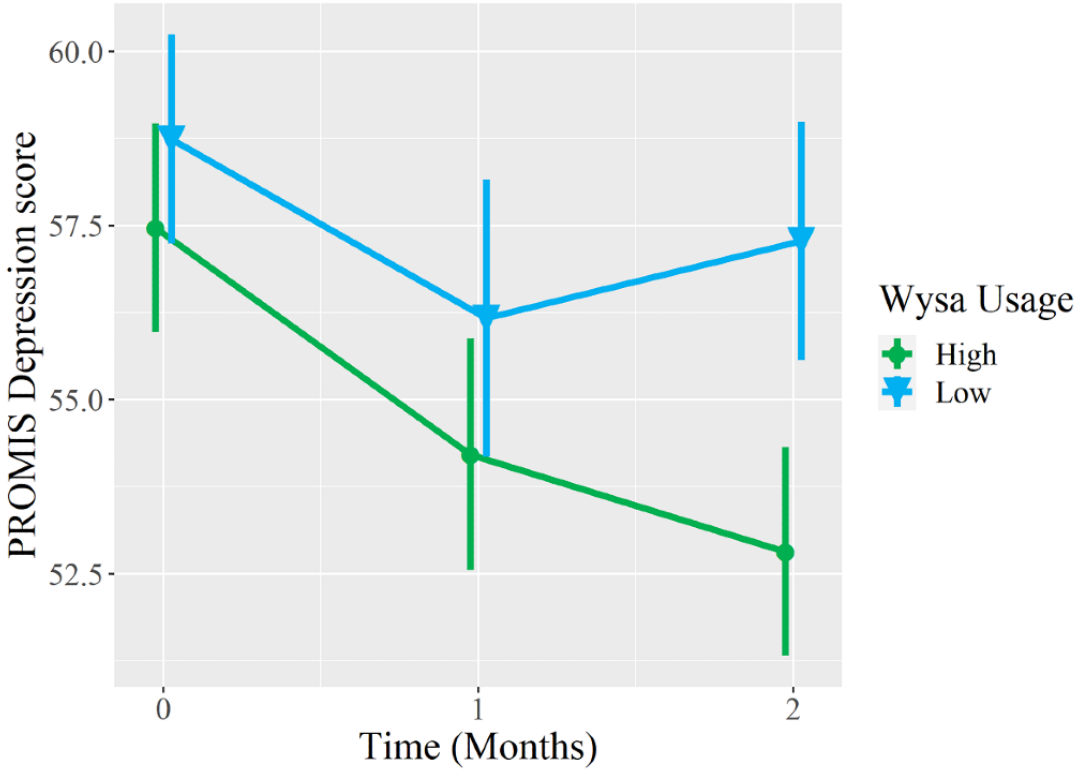


Figure 4. Mean longitudinal change in Patient-Reported Outcomes Measurement Information System (PROMIS) Anxiety scores over 2-month follow-up in patients who presented to an orthopedic clinic for a musculoskeletal condition, reported coexisting symptoms of depression and/or anxiety, and were high users (n=30, green circles) versus low users (n=31, blue triangles) of a digital mental health intervention (Wysa). Error bars represent standard errors. Asterisk represents a statistically significant between-group difference in longitudinal score changes from baseline to 2-month follow-up (*P*<.05).

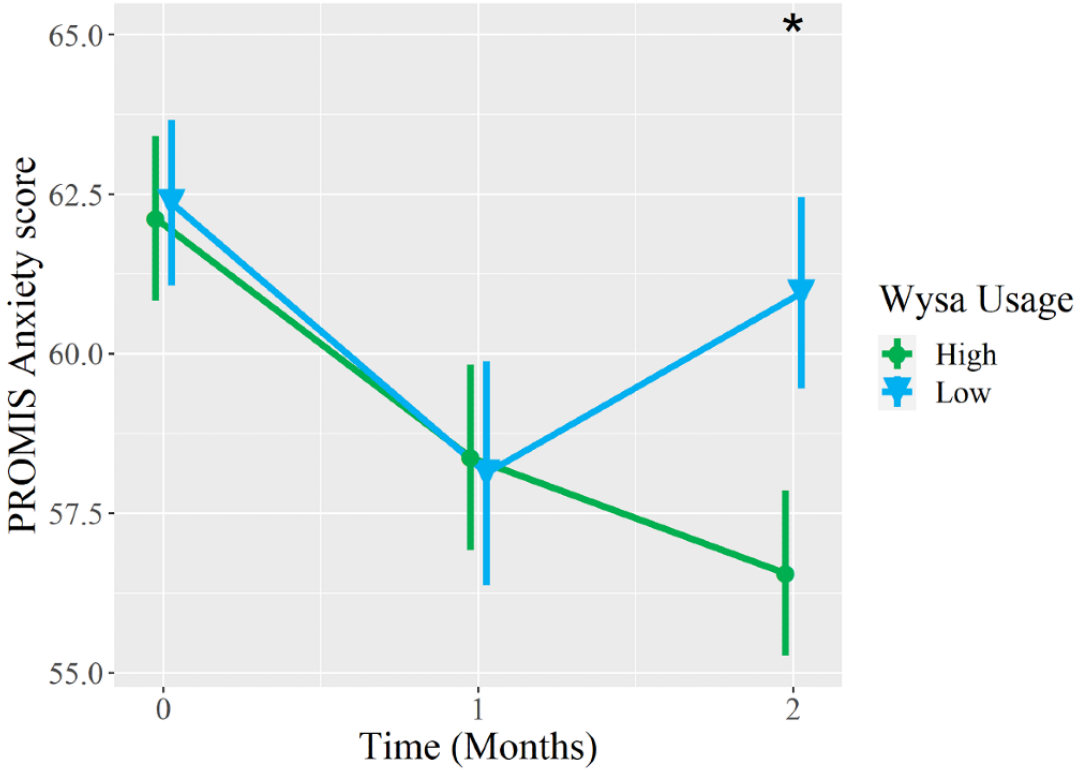


Figure 5. Mean longitudinal change in Patient-Reported Outcomes Measurement Information System (PROMIS) Pain Interference scores over 2-month follow-up in patients who presented to an orthopedic clinic for a musculoskeletal condition, reported coexisting symptoms of depression and/or anxiety, and were high users (n=30, green circles) versus low users (n=31, blue triangles) of a digital mental health intervention (Wysa). Error bars represent standard errors.

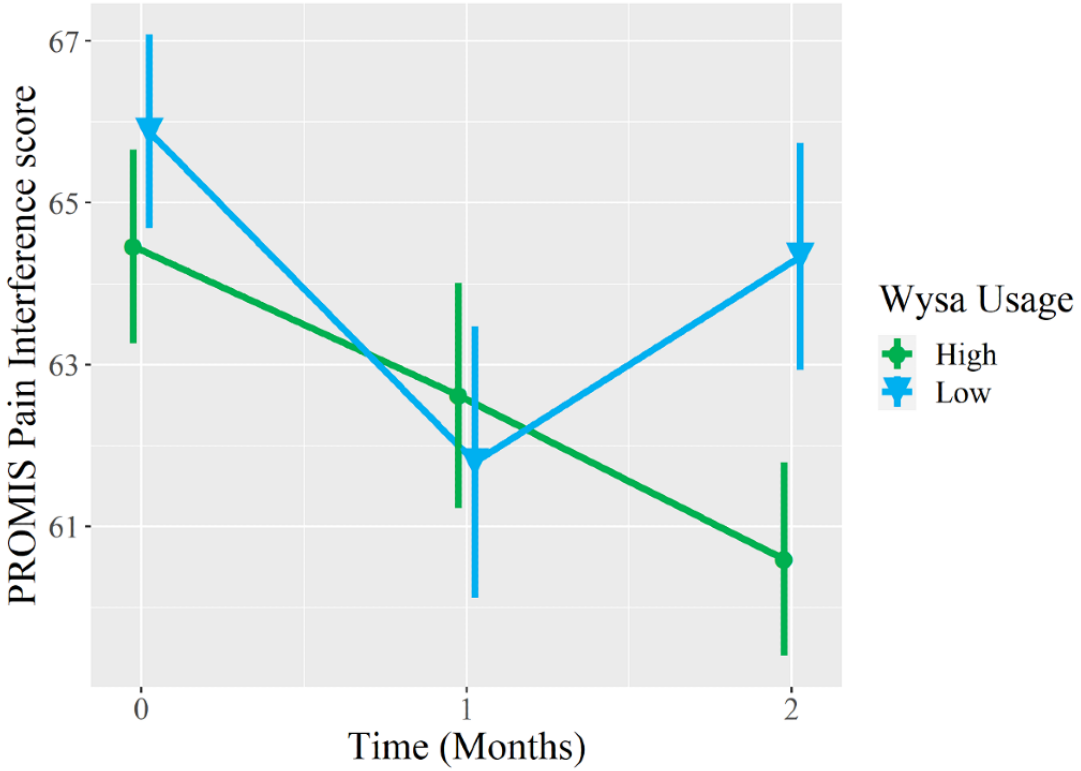
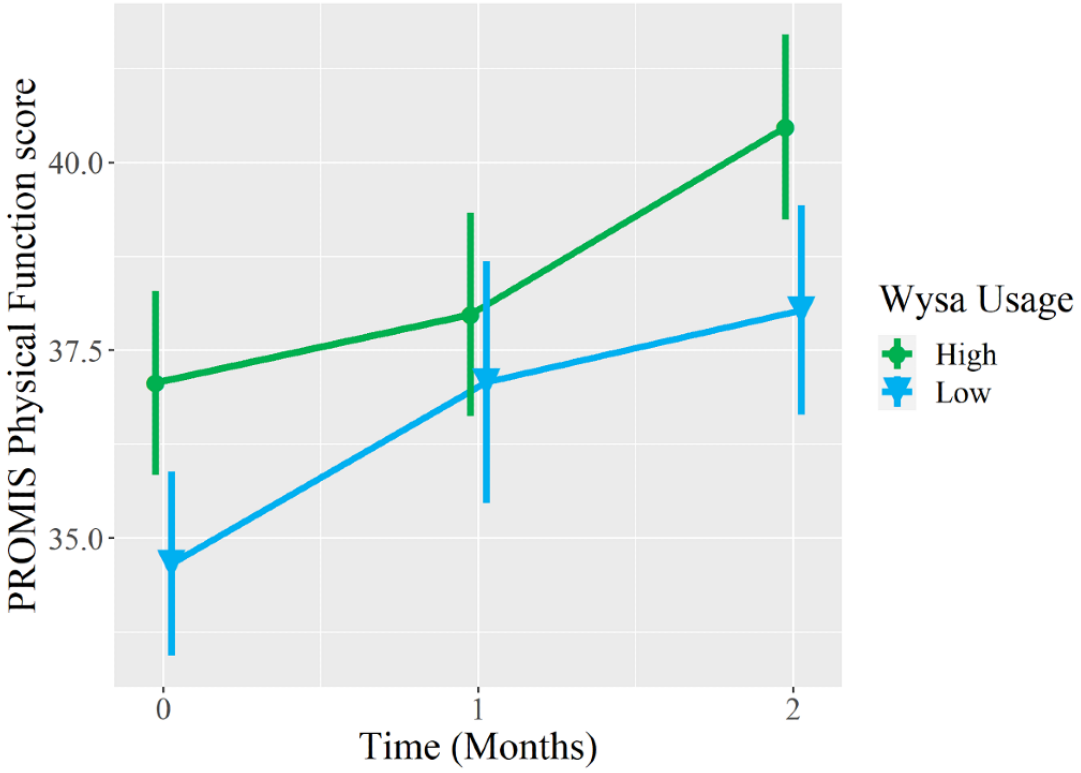


Figure 6. Mean longitudinal change in Patient-Reported Outcomes Measurement Information System (PROMIS) Physical Function scores over 2-month follow-up in patients who presented to an orthopedic clinic for a musculoskeletal condition, reported coexisting symptoms of depression and/or anxiety, and were high users (n=30, green circles) versus low users (n=31, blue triangles) of a digital mental health intervention (Wysa). Error bars represent standard errors..



Discussion

Principal Results

Intervention for mental health, which is a comorbidity that so often accompanies musculoskeletal pain, can improve both mental health- and pain-related impairment in patients who present to an orthopedic clinic for a musculoskeletal condition. To address mental health in the orthopedic setting, an acceptable, scalable, effective intervention must be available. In this pilot study, we investigated the feasibility and potential for effectiveness of delivering the digital mental health intervention Wysa in an orthopedic setting to patients with coexisting symptoms of depression and/or anxiety. The recruitment rate was 29.3% (61/208) across 6 orthopedic providers and the retention rate was 84% (51/61) with respect to completion of the 2-month outcome measures. Of the enrolled participants, 72% (44/61) engaged with Wysa at least once during the study period. Compared to low Wysa users, high Wysa users reported greater improvements by 4.2 points in PROMIS Anxiety scores at the 2-month follow-up. This difference is statistically significant (P=.044) and meets the predetermined threshold for clinical meaningfulness. Between-group differences in 2-month improvement also favored high Wysa users for PROMIS Depression (by 3.2 points) and PROMIS Pain Interference (by 2.3 points); these differences met clinically meaningful thresholds, but they did not meet statistical significance in this pilot feasibility study.

Implications

These findings suggest that a fully powered, randomized controlled trial is feasible and warranted to assess the clinical benefit of offering a digital mental health intervention in the context of orthopedic care for patients with coexisting symptoms of depression and/or anxiety. Engagement in our study was notably higher than digital health industry standards, which measure engagement on the order of days instead of weeks or months [40]. The high engagement rates and positive clinical findings in this study are consistent with those in a study of 10,000 participants by Bailey et al [41], in which people with chronic low back and/or knee pain were given a 12-week digital health intervention sponsored by their employers [41]. In that study, 73% of participants continued engagement with the program into the third month. We hypothesize that introducing a digital health intervention through a trusted resource such as a physician or employer may contribute to these higher engagement rates. Furthermore, if an orthopedic provider referred a patient to a digital mental health intervention as part of standard care (as opposed to a research study), we hypothesize that rates of onboarding to the intervention would be even greater than the enrollment rates we observed in this study. Patients who present to an orthopedic clinic for a musculoskeletal condition have demonstrated a willingness to engage with digital (physical) health interventions [42], and incorporation of digital mental health support is a logical next step to improving clinical outcomes in orthopedic care.

This line of study may also have implications for ambulatory care settings other than orthopedic clinics. We believe that clinical encounters related to chronic pain are a particularly opportune setting to address mental health because when patients understand that their physical pain experience is directly linked to their mental health, their desire to reduce pain becomes a tangible (and potentially nonstigmatizing) motivator for addressing mental health. This approach of inserting a mental health access point and resources could be applied to other frequently used ambulatory care settings that address chronic pain (eg, primary care, rheumatology, gastroenterology, headache clinics) [43], and it would create opportunities to provide mental health screening and intervention for patients with symptoms of depression and/or anxiety who might not have otherwise sought mental health treatment. In our experience, the orthopedic setting poses some especially challenging barriers to the discussion of mental health with patients, which include patients' expectations regarding appropriate content to discuss during the clinic encounter and providers' lack of training and possible lack of inherent interest or budgeted time to discuss mental health with patients. Therefore, evidence of the feasibility of delivering mental health intervention within the orthopedic setting is certainly encouraging for potential success in other ambulatory care settings that address chronic pain. Furthermore, although no coordination was made with patients' established mental health care providers in this study (for those who were already receiving some form of mental health treatment), better care coordination and communication between nontraditional mental health "access points" (eg, within orthopedics) and patients' established mental health providers should be focused on in the future.

Lessons Learned

During this feasibility study, we gained insight into participant recruitment and retention that we anticipate can be used to further improve these rates in a subsequent fully powered randomized controlled trial. Regarding participant recruitment, although every orthopedic provider that we approached was willing to assist in recruitment, it was clear that their comfort in initiating a conversation about mental health with patients varied widely. Providers who were initially less comfortable were receptive to brief training with short scripts that could be used during patient-provider discussions, and their comfort level subjectively improved through the recruitment period. The research coordinator's conversation with patients also evolved during the study as she learned the language and content that were most acceptable and compelling to patients (ie, "this app is designed to help with stress and well-being" as opposed to "this is a mental health app"). We are pursuing further pilot work to obtain a more comprehensive understanding of (surgical and nonsurgical) orthopedic provider and patients' needs regarding discussion of mental health in an orthopedic setting, both in a clinical and research context. This additional work will be essential to facilitate widespread patient engagement with a mental health intervention that is introduced in an orthopedic context.

Regarding participant retention, integration of phone call reminders improved the follow-up survey response rate compared to the use of email reminders alone. However, many participants were still hesitant to answer phone calls and voicemails were left when possible. For future trials, we plan

to integrate SMS text messaging–based reminders as well. We will also compensate participants not for use of the intervention but for completion of each follow-up survey. We anticipate these protocol changes to further improve our recruitment and retention rates in subsequent trials.

Strengths and Limitations

The primary strength of this feasibility study is that to our knowledge, it is the first investigation to incorporate a digital mental health intervention into an orthopedic care setting. The primary limitation is that there was no true control arm because of a lack of available resources. To account for longitudinal improvements that may have resulted from the usual orthopedic care that patients received during the study period (eg, physical therapy, injections, pain medications), we compared high versus low users of Wysa in our preliminary effectiveness analysis. However, healthy participant bias could still have been present; that is, high Wysa users may have been more ready than low Wysa users to address their mental health, and the same phenomenon could have occurred between patients who chose to participate and those who did not participate in the study. Nevertheless, successful completion of this feasibility study suggests that at least some orthopedic patients are interested in receiving mental health intervention as a component of their orthopedic care.

Generalizability

While the study findings offer insight into important next steps for this research, the generalizability of the results of this current single-center study is somewhat limited. Collection of mental health screening measures is becoming more commonplace in orthopedic clinics, but it is still not performed universally [44], and collection of these measures may have facilitated the recruitment process for our study. Comfort level and success in discussing mental health in an orthopedic setting are also likely provider-dependent [12]. The recruitment of participants from among patients who presented to 6 providers at our institution improves the generalizability of our findings, but all providers were nonsurgical specialists. Furthermore, we observed variations among providers with respect to the introduction of the study to the patients and the longitudinal refinement of each provider's approach during the study. From an implementation standpoint, further work is also needed to understand how a digital mental health intervention could most effectively and efficiently be delivered to patients who present to an orthopedic clinic for a musculoskeletal condition when a research coordinator is not available to assist the patient with onboarding to the app. We are actively investigating these provider-level and implementation-related limitations to generalizability in further pilot work.

Conclusions

In summary, this pilot study explored the feasibility and potential for effectiveness of delivering the digital mental health intervention Wysa in an orthopedic setting to patients with coexisting symptoms of depression and/or anxiety. We demonstrated that it is possible for orthopedic providers to introduce a mental health intervention to these patients. Furthermore, high users of the intervention reported greater improvement in anxiety symptoms than low users at the 2-month follow-up. High users of the intervention may also achieve meaningfully greater 2-month improvements in depression and pain interference, but this study was not powered to detect these differences. It is valuable to continue this line of research because it offers potential to improve both mental health and pain-related impairment in people with chronic musculoskeletal pain and coexisting depression and/or anxiety in a manner that is feasible and scalable to deliver in an orthopedic setting. More pilot work is needed (and is ongoing) to optimize the ability of orthopedic providers to efficiently discuss mental health with these patients in a manner that is mutually acceptable. Ultimately, a fully powered randomized controlled trial is warranted, and this approach of inserting mental health access points and resources into non–mental health ambulatory care settings could also be considered in primary care and other specialty clinics that often address chronic pain.

Acknowledgments

This study was funded by a Washington University/BJC HealthCare Big Ideas innovation grant. Work on this study was also supported by the National Institutes of Health (K23AR074520) and the Doris Duke Charitable Foundation (to ALC). The study team purchased participant access codes for Wysa's novel version tailored for people with chronic pain (which is not commercially available). Wysa provided participants' deidentified app usage data to the study team free of charge.

Authors' Contributions

ALC conceived the study idea, supervised the entire project, and contributed to writing the final version of the manuscript. DMH and JP Metzler contributed to participant recruitment, data collection, and editing of the manuscript. MAA performed all participant enrollment, contributed to follow-up data collection, and edited the final manuscript. AJL contributed to follow-up data collection and writing of the manuscript. PAA and JP Miller contributed to the statistical methods, data presentation, and editing of the final manuscript. MJS performed the statistical analysis and contributed to writing of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Distribution of therapeutic mechanisms underlying the Wysa tools used by 61 patients who presented to an orthopedic clinic for a musculoskeletal condition and coexisting symptoms of depression and anxiety. CBT: cognitive behavioral therapy.
[PNG File , 76 KB - formative_v6i2e34889_app1.png]

References

1. Zelaya C, Dahlhamer J, Lucas J, Connor E. Chronic pain and high-impact chronic pain among U.S. adults, 2019. NCHS Data Brief 2020 Nov(390):1-8. [Medline: 33151145]
2. Bair MJ, Robinson RL, Katon W, Kroenke K. Depression and pain comorbidity: a literature review. Arch Intern Med 2003 Nov 10;163(20):2433-2445. [doi: 10.1001/archinte.163.20.2433] [Medline: 14609780]
3. McWilliams LA, Cox BJ, Enns MW. Mood and anxiety disorders associated with chronic pain: an examination in a nationally representative sample. Pain 2003 Nov;106(1-2):127-133. [doi: 10.1016/s0304-3959(03)00301-4] [Medline: 14581119]
4. Yelin E, Weinstein S, King T. An update on the burden of musculoskeletal diseases in the U.S. Semin Arthritis Rheum 2019 Aug;49(1):1-2. [doi: 10.1016/j.semarthrit.2019.04.010] [Medline: 31133279]
5. Gambassi G. Pain and depression: the egg and the chicken story revisited. Arch Gerontol Geriatr 2009;49 Suppl 1:103-112. [doi: 10.1016/j.archger.2009.09.018] [Medline: 19836622]
6. DeVeaugh-Geiss AM, West SL, Miller WC, Sleath B, Gaynes BN, Kroenke K. The adverse effects of comorbid pain on depression outcomes in primary care patients: results from the ARTIST trial. Pain Med 2010 May;11(5):732-741. [doi: 10.1111/j.1526-4637.2010.00830.x] [Medline: 20353408]
7. Ang DC, Bair MJ, Damush TM, Wu J, Tu W, Kroenke K. Predictors of pain outcomes in patients with chronic musculoskeletal pain co-morbid with depression: results from a randomized controlled trial. Pain Med 2010 Apr;11(4):482-491. [doi: 10.1111/j.1526-4637.2009.00759.x] [Medline: 20002592]
8. Cheng AL, Schwabe M, Doering MM, Colditz GA, Prather H. The effect of psychological impairment on outcomes in patients with prearthritic hip disorders: a systematic review and meta-analysis. Am J Sports Med 2020 Aug;48(10):2563-2571. [doi: 10.1177/0363546519883246] [Medline: 31829034]
9. Edwards RR, Klick B, Buenaver L, Max MB, Haythornthwaite JA, Keller RB, et al. Symptoms of distress as prospective predictors of pain-related sciatica treatment outcomes. Pain 2007 Jul;130(1-2):47-55. [doi: 10.1016/j.pain.2006.10.026] [Medline: 17156925]
10. Lee H, Hübscher M, Moseley GL, Kamper SJ, Traeger AC, Mansell G, et al. How does pain lead to disability? A systematic review and meta-analysis of mediation studies in people with back and neck pain. Pain 2015 Jun;156(6):988-997. [doi: 10.1097/j.pain.0000000000000146] [Medline: 25760473]
11. Alattas SA, Smith T, Bhatti M, Wilson-Nunn D, Donell S. Greater pre-operative anxiety, pain and poorer function predict a worse outcome of a total knee arthroplasty. Knee Surg Sports Traumatol Arthrosc 2017 Nov;25(11):3403-3410 [FREE Full text] [doi: 10.1007/s00167-016-4314-8] [Medline: 27734110]
12. Vranceanu AM, Beks RB, Guitton TG, Janssen SJ, Ring D. How do orthopaedic surgeons address psychological aspects of illness? Arch Bone Jt Surg 2017 Jan;5(1):2-9 [FREE Full text] [Medline: 28271080]
13. Andrade LH, Alonso J, Mneimneh Z, Wells JE, Al-Hamzawi A, Borges G, et al. Barriers to mental health treatment: results from the WHO World Mental Health surveys. Psychol Med 2014 Apr;44(6):1303-1317 [FREE Full text] [doi: 10.1017/S0033291713001943] [Medline: 23931656]
14. Kozloff N, Sommers BD. Insurance coverage and health outcomes in young adults with mental illness following the Affordable Care Act dependent coverage expansion. J Clin Psychiatry 2017 Jul;78(7):e821-e827. [doi: 10.4088/JCP.16m11357] [Medline: 28703947]
15. Weil TP. Insufficient dollars and qualified personnel to meet United States mental health needs. J Nerv Ment Dis 2015 Apr;203(4):233-240. [doi: 10.1097/NMD.0000000000000271] [Medline: 25816044]
16. 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]
17. 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]
18. Firth J, Torous J, Nicholas J, Carney R, Rosenbaum S, Sarris J. Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. J Affect Disord 2017 Aug 15;218:15-22 [FREE Full text] [doi: 10.1016/j.jad.2017.04.046] [Medline: 28456072]
19. Linardon J, Cuijpers P, Carlbring P, Messer M, Fuller-Tyszkiewicz M. The efficacy of app-supported smartphone interventions for mental health problems: a meta-analysis of randomized controlled trials. World Psychiatry 2019 Oct;18(3):325-336 [FREE Full text] [doi: 10.1002/wps.20673] [Medline: 31496095]
20. Mohr DC, Burns MN, Schueller SM, Clarke G, Klinkman M. Behavioral intervention technologies: evidence review and recommendations for future research in mental health. Gen Hosp Psychiatry 2013;35(4):332-338 [FREE Full text] [doi: 10.1016/j.genhosppsych.2013.03.008] [Medline: 23664503]

21. Abd-Alrazaq AA, Alajlani M, Ali N, Denecke K, Bewick BM, Househ M. Perceptions and opinions of patients about mental health chatbots: scoping review. J Med Internet Res 2021 Jan 13;23(1):e17828 [FREE Full text] [doi: 10.2196/17828] [Medline: 33439133]
22. Torous J, Bucci S, Bell IH, Kessing LV, Faurholt-Jepsen M, Whelan P, et al. The growing field of digital psychiatry: current evidence and the future of apps, social media, chatbots, and virtual reality. World Psychiatry 2021 Oct;20(3):318-335 [FREE Full text] [doi: 10.1002/wps.20883] [Medline: 34505369]
23. Wysa. Mental Health Support, For Everyone. Wysa. URL: http://wysa.io [accessed 2021-03-10]
24. Inkster B, Sarda S, Subramanian V. An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: real-world data evaluation mixed-methods study. JMIR Mhealth Uhealth 2018 Nov 23;6(11):e12106 [FREE Full text] [doi: 10.2196/12106] [Medline: 30470676]
25. Burkhardt HA, Alexopoulos GS, Pullmann MD, Hull TD, Areán PA, Cohen T. Behavioral activation and depression symptomatology: longitudinal assessment of linguistic indicators in text-based therapy sessions. J Med Internet Res 2021 Jul 14;23(7):e28244 [FREE Full text] [doi: 10.2196/28244] [Medline: 34259637]
26. Jensen MP, Ehde DM, Day MA. The behavioral activation and inhibition systems: implications for understanding and treating chronic pain. J Pain 2016 May;17(5):529.e1-529.e18. [doi: 10.1016/j.jpain.2016.02.001] [Medline: 27052644]
27. Kratz AL, Davis MC, Zautra AJ. Pain acceptance moderates the relation between pain and negative affect in female osteoarthritis and fibromyalgia patients. Ann Behav Med 2007 Jun;33(3):291-301 [FREE Full text] [doi: 10.1007/BF02879911] [Medline: 17600456]
28. McCracken LM. Learning to live with the pain: acceptance of pain predicts adjustment in persons with chronic pain. Pain 1998 Jan;74(1):21-27. [doi: 10.1016/S0304-3959(97)00146-2] [Medline: 9514556]
29. Cella D, Yount S, Rothrock N, Gershon R, Cook K, Reeve B, PROMIS Cooperative Group. The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. Med Care 2007 May;45(5 Suppl 1):S3-S11 [FREE Full text] [doi: 10.1097/01.mlr.0000258615.42478.55] [Medline: 17443116]
30. Smith BW, Dalen J, Wiggins K, Tooley E, Christopher P, Bernard J. The brief resilience scale: assessing the ability to bounce back. Int J Behav Med 2008;15(3):194-200. [doi: 10.1080/10705500802222972] [Medline: 18696313]
31. Zimmerman GL, Olsen CG, Bosworth MF. A 'stages of change' approach to helping patients change behavior. Am Fam Physician 2000 Mar 01;61(5):1409-1416 [FREE Full text] [Medline: 10735346]
32. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: Building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
33. 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]
34. Lee AC, Driban JB, Price LL, Harvey WF, Rodday AM, Wang C. Responsiveness and Minimally Important Differences for 4 Patient-Reported Outcomes Measurement Information System Short Forms: Physical Function, Pain Interference, Depression, and Anxiety in Knee Osteoarthritis. J Pain 2017 Sep;18(9):1096-1110 [FREE Full text] [doi: 10.1016/j.jpain.2017.05.001] [Medline: 28501708]
35. Kroenke K, Yu Z, Wu J, Kean J, Monahan PO. Operating characteristics of PROMIS four-item depression and anxiety scales in primary care patients with chronic pain. Pain Med 2014 Nov;15(11):1892-1901 [FREE Full text] [doi: 10.1111/pme.12537] [Medline: 25138978]
36. Chen CX, Kroenke K, Stump TE, Kean J, Carpenter JS, Krebs EE, et al. Estimating minimally important differences for the PROMIS pain interference scales: results from 3 randomized clinical trials. Pain 2018 Apr;159(4):775-782 [FREE Full text] [doi: 10.1097/j.pain.0000000000001121] [Medline: 29200181]
37. Kind AJH, Jencks S, Brock J, Yu M, Bartels C, Ehlenbach W, et al. Neighborhood socioeconomic disadvantage and 30-day rehospitalization: a retrospective cohort study. Ann Intern Med 2014 Dec 02;161(11):765-774 [FREE Full text] [doi: 10.7326/M13-2946] [Medline: 25437404]
38. Singh GK. Area deprivation and widening inequalities in US mortality, 1969-1998. Am J Public Health 2003 Jul;93(7):1137-1143. [doi: 10.2105/ajph.93.7.1137] [Medline: 12835199]
39. Bertholet N, Gaume J, Faouzi M, Gmel G, Daeppen J. Predictive value of readiness, importance, and confidence in ability to change drinking and smoking. BMC Public Health 2012 Aug 29;12:708 [FREE Full text] [doi: 10.1186/1471-2458-12-708] [Medline: 22931392]
40. Baumel A, Muench F, Edan S, Kane JM. Objective user engagement with mental health apps: systematic search and panel-based usage analysis. J Med Internet Res 2019 Sep 25;21(9):e14567 [FREE Full text] [doi: 10.2196/14567] [Medline: 31573916]
41. Bailey JF, Agarwal V, Zheng P, Smuck M, Fredericson M, Kennedy DJ, et al. Digital care for chronic musculoskeletal pain: 10,000 participant longitudinal cohort study. J Med Internet Res 2020 May 11;22(5):e18250 [FREE Full text] [doi: 10.2196/18250] [Medline: 32208358]

42. 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]
43. Bartels SJ, Coakley EH, Zubritsky C, Ware JH, Miles KM, Areán PA, PRISM-E Investigators. Improving access to geriatric mental health services: a randomized trial comparing treatment engagement with integrated versus enhanced referral care for depression, anxiety, and at-risk alcohol use. Am J Psychiatry 2004 Aug;161(8):1455-1462. [doi: 10.1176/appi.ajp.161.8.1455] [Medline: 15285973]
44. Ring D. Mental and social health are inseparable from physical health. J Bone Joint Surg Am 2021 Jun 02;103(11):951-952. [doi: 10.2106/JBJS.21.00121] [Medline: 33877056]

Abbreviations

PROMIS: Patient-Reported Outcomes Measurement Information System
REDCap: Research Electronic Data Capture

Edited by A Mavragani; submitted 15.12.21; peer-reviewed by D Ring, H Burkhardt; comments to author 07.01.22; revised version received 09.01.22; accepted 17.01.22; published 21.02.22.

Please cite as:
Leo AJ, Schuelke MJ, Hunt DM, Metzler JP, Miller JP, Areán PA, Armbrecht MA, Cheng AL
A Digital Mental Health Intervention in an Orthopedic Setting for Patients With Symptoms of Depression and/or Anxiety: Feasibility Prospective Cohort Study
JMIR Form Res 2022;6(2):e34889
URL: https://formative.jmir.org/2022/2/e34889
doi:10.2196/34889
PMID:35039278

Original Paper

Web-Based Lifestyle Interventions for Survivors of Cancer: Usability Study

Victoria Williams[1], PhD; Nashira Brown[1], MS; Justin Xavier Moore[2], PhD; David Farrell[3], MPH; Suzanne Perumean-Chaney[4], PhD; Erica Schleicher[1], MS; Kevin Fontaine[1], PhD; Wendy Demark-Wahnefried[5], PhD; Dori Pekmezi[1], PhD

[1]Department of Health Behavior, University of Alabama at Birmingham, Birmingham, AL, United States

[2]Department of Population Health Sciences, Augusta University, Augusta, GA, United States

[3]People's Design Inc., Durham, NC, United States

[4]Department of Biostatistics, University of Alabama at Birmingham, Birmingham, AL, United States

[5]Department of Nutritional Sciences, University of Alabama at Birmingham, Birmingham, AL, United States

Corresponding Author:
Victoria Williams, PhD
Department of Health Behavior
University of Alabama at Birmingham
1665 University Boulevard
Birmingham, AL, 35233
United States
Phone: 1 205 934 4993
Email: contact@drvaw.com

Abstract

Background: Internet-based lifestyle programs are increasingly being used to deliver health behavior change interventions to survivors of cancer. However, little is known about website use in this population or its association with healthy lifestyle changes.

Objective: The aim of this study is to describe lifestyle intervention website use (log-ins, time on website, and page views) among survivors of cancer and patterns of use by participant characteristics. In addition, associations were explored between website use and changes in healthy lifestyle knowledge and practice.

Methods: A total of 35 survivors of cancer were recruited between August 2017 and 2018 to participate in a 2-week, single-arm pilot test of the SurvivorSHINE lifestyle intervention website. Knowledge and practices related to healthy diet and physical activity behaviors were measured at baseline and follow-up. Website use (eg, time spent on the website, frequency of log-ins, and page views) were collected from the SurvivorSHINE administrative site during the intervention period. Patterns of use were examined by participants' gender and race. Correlations between website use and changes in healthy lifestyle knowledge, physical activity, diet, and weight were explored. Mann–Whitney *U* tests were used to compare demographic factors on website use.

Results: Participants logged into the SurvivorSHINE intervention website an average of 3.2 (SD 2) times over the 2-week period and spent a total average of 94 (SD 56) minutes viewing the website during the intervention. Examining website activity, 1905 page views were logged. The *User Profile* (344 page views) and *Home* sections (301 page views) were the most frequently visited components. No associations were observed between the frequency of log-ins or the total time on the website, improvements in knowledge related to healthy lifestyles, or changes in body weight or dietary intake. However, the total time on the website was positively correlated with improvements in accelerometer-measured physical activity (r=0.74; P=.02) and self-reported physical activity (r=0.35; P=.04).

Conclusions: Survivors of cancer demonstrated clear interest in a diet and exercise intervention website, as evidenced by their frequency of log-ins, page views on numerous features, and total viewing time. Moreover, increased website use was correlated with improvements in physical activity.

(JMIR Form Res 2022;6(2):e30974) doi:10.2196/30974

KEYWORDS

cancer survivors; diet; exercise; lifestyle; internet; physical activity; web-based; website; weight management; digital health; cancer; online health

Introduction

Background

Emerging evidence shows that web-delivered interventions can improve diet [1-3], physical activity [1-16], and weight management [10,17,18] in survivors of cancer [19-21]. Previous studies have reported positive associations between the lifestyle intervention website use by survivors of cancer and improvements in diet (ie, vegetable consumption) [2], physical activity (ie, aerobic and resistance-based) [13,22], and behavioral outcomes [3,23,24]. However, to date, most of these studies have been conducted predominantly on survivors of breast or prostate cancer who are White [25]. As survivors of cancer are growing in number and becoming more diverse by cancer type, age, and race [26], this study fills a critical research gap by investigating lifestyle intervention website use in a more heterogeneous sample in which racial minorities and various cancer types beyond the breast and prostate are represented. Moreover, cancer is a disease associated with aging; however, extant studies have primarily enrolled younger survivors of cancer [24,27]. Thus, conducting more web-delivered lifestyle interventions within an older and more diverse cancer survivor population provides a more thorough understanding of how to better design these interventions to engage survivors of a variety of cancers to make a larger impact.

Most survivors of cancer in the United States are older adults (>65 years), as the median age of cancer diagnosis in the United States is 66 years [28]. Although older adults have lagged behind younger adults in internet use in the past [29], more recent studies have found that older adults are the fastest-growing segment of internet users and the digital divide is closing [30]. However, disparities still exist among older adults with regard to internet use, with affluent, well-educated adults using the internet more than other subgroups [31]. However, initial concerns that internet-delivered interventions might not be appropriate for survivors of cancer, as most are older adults appear unfounded [29]. Questions remain about which segments of survivors are most suited for web-based healthy lifestyle interventions, how such programs work, and whether website use is associated with health benefits among survivors of cancer. Given the recent increases in survivors of cancer as well as technology adoption, a better understanding of healthy lifestyle website use patterns and outcomes among older survivors of cancer would help inform future optimization of the experience and related outcomes.

Objectives

The purpose of this research is to describe the use (log-ins, time on the website, and page views) of a web-based healthy lifestyle program, SurvivorSHINE [32], in a pilot study for adult survivors of cancer. Furthermore, we examined patterns of use by participant characteristics (gender and race) and associations between website use and changes in knowledge, body weight, diet, and physical activity. It was hypothesized that (1) upon log-in to SurvivorSHINE, survivors of cancer will visit key features, such as body weight, diet, and physical activity sections of the website, similar to use rates found in past studies with survivors of breast or prostate cancers who are predominantly White [19-21]; (2) survivors of cancer with higher website use will have greater improvements in healthy lifestyle knowledge, and exhibit changes in physical activity, diet quality, and weight, given positive associations between website use and behavior change (increased physical activity) in previous research [33]; and (3) survivors who are non-Hispanic White and men will have higher website use compared with survivors who are non-Hispanic Black and women, given Pew Research Center data indicating particularly high computer and internet use rates in these groups [34,35].

Methods

Research Design

Data for this study were collected from survivors of cancer participating in a single-arm pilot study of the SurvivorSHINE lifestyle intervention website. Initial findings have already supported feasibility, acceptability, and behavior change (improvements in physical activity, healthy lifestyle knowledge, etc) in response to SurvivorSHINE [32]. The current analyses focus on website metrics (time spent on websites, log-ins, and page visits) over the 2 weeks of use and their relationship with demographics at baseline and lifestyle factors (healthy lifestyle knowledge, body weight, diet, and physical activity) at baseline and after the intervention.

Participants

Participants were survivors of cancers with >80% 5-year survival (eg, cancers of the breast, prostate, and thyroid). The eligibility criteria were (1) non-Hispanic White and non-Hispanic Black adults, (2) residence within a 20-mile radius of the University of Alabama at Birmingham, (3) self-reported internet access and regular use, and (4) ability to speak and read English. Recruitment was completed using the following methods: (1) ascertaining patients with cancer from the University of Alabama at Birmingham Cancer Registry and sending a letter of invitation, (2) contacting local cancer survivor support groups and related community organizations, (3) local news advertisements, and (4) word of mouth.

Protocol Overview

Participants completed the assessments (anthropometrics, healthy lifestyle knowledge, and current physical activity and dietary practices) in person with study staff, whereas screening was conducted by telephone. Participants were then given access to the SurvivorSHINE website and encouraged to review and use all of its components (ie, update user profile and visit different sections of the website) as frequently as possible (ideally, daily) over the 2-week website use period. Participants were instructed to create a profile using a username or personal email along with a password. Moreover, they were directed to

access the website via their personal profiles each time. At the completion of the intervention period, participants completed follow-up assessments and exit interviews in person [32]. Full details of the trial design, participants, protocol, measures, and intervention are described in a previous report [32] but are briefly outlined below.

Intervention Website

SurvivorSHINE is a theoretically-driven (social cognitive theory) [32], web-delivered intervention promoting physical activity, a healthy diet, and weight management among survivors of cancer. SurvivorSHINE was largely adapted from the written materials that were used in the Reach-out to Enhance Wellness intervention, which demonstrated proven efficacy in improving diet quality and physical activity, as well as body weight and physical functioning among 641 older survivors of cancer [36]. The website was designed using input obtained from 4 sets of focus groups stratified by gender (male or female) and race (White or Black) [32]. It includes six main sections: *My Profile*, *Home Page*, *Healthy Weight*, *Healthy Eating*, *Exercise*, and *News You Can Use*. For this study, users created an account and entered data on their demographics and cancer type in the *My Profile* section to guide the personal tailoring of the diet, weight management, and exercise information. Direct links for the *Healthy Eating*, *Healthy Weight*, *Exercise*, and *News You Can Use* sections of the website were provided on the *Home Page*. The *Healthy Eating* section generated personalized feedback on the users' diet and information on a healthy diet, fast food, serving sizes, and tips to promote healthy eating behavior. Moreover, specific guidance was given to diet-related domains that corresponded with the American Cancer Society guidelines for survivors of cancer [37] and encouraged to achieve the following dietary goals for vegetables and fruits (5 daily servings), whole grains (≥50% of grain consumption), added sugar (≤6 teaspoon per day for women and ≤9 teaspoon per day for men), red and processed meat (<0.51 kg per week), saturated fat (<10% of total calories), and alcohol consumption (to achieve an eventual goal of ≤1 drink per day for women and ≤2 drinks per day for men).

Individualized feedback was also provided in the *Exercise* section based on the physical activity information users entered in their profile. More specifically, participants were provided with tailored recommendations based on physical activity information input in the user profile. Incremental increases in the frequency and duration of endurance exercise resulted in a net increase of 30 minutes per week (ie, 150 minutes per week) and frequency of strength training (ie, 2-3 times per week) to meet the American Cancer Society guidelines for survivors of cancer [33]. Moreover, this section highlights leg strengthening exercises, the benefits of exercise, the use of pedometers and accelerometers, the importance of setting SMART (specific, measurable, attainable, relevant, and timely) goals [38], and a calorie burning guide. Feedback on users' weight was provided in the *Healthy Weight* section, which featured a BMI calculator, calorie calculator, and sample meal plans. Finally, easy-to-read summaries of recent cancer-related healthy lifestyle research were provided in the *News You Can Use* section; this section was highly endorsed by focus group participants [32].

Measures

Demographic information assessed at baseline included age, gender, race or ethnicity, the highest level of educational attainment, marital status, current employment, and income. Healthy lifestyle knowledge, physical activity, diet, and weight were measured at baseline and follow-up. Survivors' healthy lifestyle knowledge was evaluated using a 10-item questionnaire about the American Cancer Society recommendations for diet, physical activity, and weight management [32]. Physical activity was assessed subjectively using the Godin Leisure-Time Exercise Questionnaire and assessed objectively using an accelerometer [39]. The Godin Leisure-Time Exercise Questionnaire has acceptable reliability, internal consistency, and similar validity with more objective measures of physical activity levels [40-42]. ActiGraph accelerometers (ActiGraph GT3X+) were used to collect objective measures of physical activity and have been validated using heart rate telemetry [43] and total energy expenditure [44]. The ActiGraph protocol used has been previously reported [32]; in brief, the minimum threshold for moderate intensity was set at 1952 counts per minute [45]. In addition, the minimum valid wear time was set at 4 days for at least 600 minutes of wear. The Automated Self-Administered 24-Hour Dietary Assessment, a validated web-based tool [46], was used to capture dietary recalls for 2 days, including 1 weekday and 1 weekend day. The Automated Self-Administered 24-Hour Dietary Assessment tool generates a nutrient analysis of beverages and food consumed with variables of interest, including total kilocalories, total fat, saturated fat, added sugar, alcohol, and servings of vegetables and fruits, whole grains, and red meat. Weight was measured without shoes and in light clothing using the Health O Meter Floor Scale (894KLTEA).

Website use was measured by use statistics of time spent on the website, frequency of log-ins, and page views. Log-ins, or the number of times a user signs in, were collected each time a user signed in using a username or email address and password. Page views, or the number of times a user visits a page of the website, were determined based on website logs that indicated pages users visited (eg, navigating to the *Healthy Eating* section of the website required users to visit a specific link). Website use data and logs were obtained from the administrative site of SurvivorSHINE, which contained time-stamped activity records for participant accounts (eg, participant 27 clicked the *Healthy Eating* section at 2:34 PM on December 15, 2019) [32]. The website use data were connected to trial participants through website usernames and passwords associated with each study identifier. Although the website was publicly available, only website use data connected to actual trial participants were included in the analyses.

Data Collection and Analyses

Although this study was exploratory, power calculations suggest that given a sample size of 35, this study has 99% power to detect strong correlations of 0.7 and 68.1% power to detect moderate correlations of 0.4. Power analyses were computed using SAS (version 9.4; SAS Institute Inc). The analyses included descriptive statistics and frequencies of website use. Owing to the nonnormality of data and small sample size,

Spearman rank correlations examined the relationships between website use and healthy behavior knowledge, as well as behavior change (physical activity, dietary intake, and weight). Mann–Whitney *U* tests were used to compare demographic factors on website use. All analyses were performed using SPSS (version 25; IBM Corp). For this study, data from participants who completed both baseline and follow-up measures were included (35/41,85%).

Ethical Considerations

The protocol was approved by the University of Alabama at Birmingham Institutional Review Board (IRB-140428003), and signed informed consent was obtained from all participants.

Results

Participants

The participants were survivors of many types of cancers, but breast and prostate cancers were the most common diagnoses. The average age of the sample was 62.1 (SD 11.9) years, and most were women (19/35, 54%), non-Hispanic White (22/35, 63%), married (24/35, 69%), and retired (19/35, 54%). Income and educational levels were higher than state averages, with more than half of the sample reporting annual household incomes surpassing US $50,536 and baccalaureate or higher degrees [47] (Table 1).

Table 1. Sample characteristics (N=35).

Variable	Values
Age (years), mean (SD)	62.1 (11.9)
Gender, n (%)	
Female	19 (54)
Race, n (%)	
Black	13 (37)
White	22 (63)
Cancer type[a], n (%)	
Breast	16 (46)
Prostate	12 (34)
Myeloma	2 (6)
Skin	2 (6)
Thyroid	2 (6)
Head and neck	2 (6)
Marital status, n (%)	
Married or civil union	24 (69)
Widowed	5 (14)
Single, never married	3 (9)
Divorced	3 (9)
Education, n (%)	
High-school degree of equivalent (eg, General Educational Development)	4 (11)
Some college but no degree	4 (11)
Associate degree	8 (23)
Bachelor's degree	11 (31)
Graduate degree	8 (23)
Income (US $), n (%)	
≤49,999	13 (37)
50,000-124,999	10 (29)
≥125,000	9 (26)
Refused	3 (9)
Employment, n (%)	
Employed, working full-time	9 (26)
Employed, working part-time	3 (9)
Disabled, not able to work	4 (11)
Retired	19 (54)

[a]Participants may select multiple cancers.

Intervention Website Use

On average, participants logged 3.2 (SD 2) times and spent an average of 31 (SD 16) minutes per log-in. Thus, the total time spent on the website during the 2-week period was 94 (SD 56) minutes. On average, users visited 1905 different areas of the website, as evidenced by the 1905 page views (Table 2). Over the 2-week period, page views ranged from 2 to 153 page views per user. Participants averaged a total of 57.7 (SD 30.7) page views per person and 20 (SD 10.2) page views per log-in. Overall, the initial *User Profile* (344 page views) and the *Home* sections (301 page views) were the most visited components of SurvivorSHINE (Table 2, column 2). The *News You Can Use* section was the least visited section with only 91 overall total page views. When these data were examined at the individual level (page views per user), a more thorough website exploration was found. For example, participants also frequently viewed

the *Healthy Eating*, *Healthy Weight*, *Exercise*, and *News You Can Use* sections (32 user page views; Table 2, column 3). Overall, participants spent the most time on the *User Profile* section updating personal exercise, weight, and eating information (62,912 seconds). On average, the most time per participant was spent on the *User Profile* (mean 183 seconds, SD 173 seconds), in the *Healthy Weight* section viewing sample meal plans (mean 110 seconds, SD 96.3 seconds), and learning about accelerometers and pedometers in the *Exercise* section (mean 89 seconds, SD 71.2 seconds).

Table 2. SurvivorSHINE key components accessed by participants (N=1905).

SurvivorSHINE website component	Total page views (page views ÷ total page views; N=1905), n (%)	User page views (page views ÷ total users; N=35), n (%)	Page views per user, median (IQR)	Total time spent on component in seconds, n	Average time spent on component per user in seconds (total time on component ÷ total users; N=35), mean (SD)
User profile	344 (18)	32 (97)	8 (9)	62,912	183 (173)
Home	301 (16)	31 (94)	9 (7)	9237	31 (33.1)
Healthy eating	196 (10)	32 (97)	4 (6)	8972	46 (44.6)
Vegetables and fruits	51 (3)	15 (46)	1 (0)	861	32 (21.2)
Benefits of a healthy diet	26 (1)	16 (49)	2 (0)	1135	44 (41.1)
Fats and fast food	24 (1)	17 (52)	1 (1)	1304	54 (16)
Sugar	21 (1)	19 (58)	1 (1)	1084	52 (24.2)
Meat	14 (0.7)	12 (36)	1 (0)	685	49 (34.3)
Whole grains	10 (0.5)	9 (27)	1 (0)	418	42 (34)
Alcohol	11 (0.6)	10 (30)	1 (0)	234	21 (12.1)
Healthy weight	192 (10)	32 (97)	6 (5)	6383	33 (68.1)
Calorie calculator	54 (3)	25 (76)	2 (1)	2160	40 (24.4)
BMI calculator	47 (3)	25 (76)	2 (1)	4142	86 (199)
Sample meal plans	39 (2)	18 (55)	3 (1)	4286	110 (96.3)
Benefits of healthy weight	31 (2)	16 (49)	2 (0)	1276	41 (28)
Common questions	24 (1)	14 (42)	1 (0)	1420	59 (54)
Exercise	124 (7)	32 (97)	3 (4)	6717	54 (63)
Resistance training exercises	38 (2)	23 (70)	2 (1)	1399	37 (23)
Benefits of exercise	29 (2)	18 (55)	1 (1)	1276	41 (28)
Calorie burning guide	28 (2)	19 (58)	1 (1)	980	35 (14)
Accelerometers and pedometers	28 (2)	21 (64)	1 (1)	2490	89 (71.2)
SMART[a] goals	27 (1)	20 (61)	1 (1)	873	32 (14)
News you can use	91 (5)	32 (97)	2 (3)	6984	77 (67.1)

[a]SMART: specific, measurable, attainable, relevant, and timely.

Behavior and Use Patterns

Survivors who are men and non-Hispanic White spent relatively more time on the website than survivors who are women and non-Hispanic Black. However, the time spent on SurvivorSHINE and website page views did not significantly differ by race or gender (Table 3). The frequency of log-ins was also not significantly different by subgroup; however, survivors who are non-Hispanic White had somewhat more log-ins than survivors who are non-Hispanic Black (P=.06).

The total time spent on the website was positively and significantly correlated with changes in self-reported and measured physical activity (Table 4). The frequency of log-ins was correlated with changes in measured physical activity but not self-reported. The total time on the website, frequency of log-ins, and page views were not correlated with changes in healthy lifestyle knowledge, weight, and diet.

Table 3. Mean rank log-ins, page views, and total time spent on SurvivorSHINE website by race and gender.

Variables	Female (n=17)	Male (n=16)	Mann–Whitney *U* test		Black (n=13)	White (n=20)	Mann–Whitney *U* test	
			Z value	*P* value			*Z* value	*P* value
Page views	16.5	17.6	0.32	.76	14.2	18.8	1.35	.18
Log-ins	16	18.1	0.65	.53	13.1	19.6	1.94	.06
Total time (minutes)	16.8	17.2	0.11	.93	14.7	18.5	1.12	.27

Table 4. The correlation between website use and posttest diet, exercise, and healthy lifestyle knowledge measures.

Measures	Total time		Log-ins		Page views	
	r_s	*P* value	r_s	*P* value	r_s	*P* value
Physical activity (measured)	0.73	.02	0.73	.02	.48	.19
Physical activity (self-reported)	−0.35	.049	−0.15	.41	−0.12	.51
Knowledge	0.04	.81	0.08	.66	−0.02	.91
Weight	0.15	.39	−0.01	.93	0.11	.54
Diet						
Total kilocalories	0.03	.90	0.11	.64	−0.03	.90
Saturated fat	0.07	.76	0.16	.49	−0.09	.68
Meat	−0.11	.62	0.10	.65	0.22	.34
Alcohol	−0.06	.80	−0.07	.75	0.02	.92
Fruit	0.17	.46	0.16	.49	0.16	.49
Vegetables	−0.001	.99	−0.21	.37	−0.27	.23

Discussion

Principal Findings

Overall, these findings indicate that survivors of cancer are willing to use and devote time to a healthy lifestyle website intervention. During the intervention, participants logged onto the SurvivorSHINE intervention website an average of 3.2 times and spent an average of 94 minutes viewing the website content over the 2-week intervention period. The *User Profile* and *Home* sections were the most visited components of the website. Our results found a positive correlation between the total time on the website and improvements in accelerometer-measured physical activity and self-reported physical activity. No other associations were observed between website use and improvements in knowledge related to healthy lifestyles, body weight, or dietary intake.

The frequency of website log-ins (an average of 3 times over 2 weeks, 1.5 log-ins per week) is consistent with other web-based lifestyle interventions among survivors of cancer and is comparable with the 5.3 times over 6 weeks reported by Bantum et al [4] and 10.3 times over 9 weeks reported by Forbes et al [5] (ie, approximately 1 log-in per week). The study participants also spent an average of 31 minutes per log-in while visiting the SurvivorSHINE website, which is more than the 12 and 11.3 minutes per log-in reported by 2 other healthy lifestyle interventions for survivors of breast cancer [13,22]. The survivors of cancer in this study were highly motivated to use the SurvivorSHINE website likely for its tailored, easy-to-read, and readily accessible healthy lifestyle information and resources [32].

This study also highlights web content on healthy lifestyle behaviors that may be most appealing to the survivors. The user profile was the most visited (number of page views and time spent) component of the website. We speculate that this was the most visited component as participants were encouraged to frequently update their user profile to ensure relevant tailored feedback and information on diet, exercise, and weight loss were received [32]. Comparably, Chen et al [48] reported the most frequently visited page among survivors of breast cancer was *setting goals* where users received personalized feedback based on individual goals and activities. Of the *Healthy Eating*, *Healthy Weight*, and *Exercise* sections, participants spent the most time in the *Healthy Eating* section. Moreover, the *Healthy Eating* section was the most visited lifestyle section of the website; however, changes in diet, as noted in a previous report, were fairly minimal [32], as were changes in weight. However, it must be borne in mind that the study period was only 2 weeks in duration; thus, the −0.3 kg change in body weight noted in the previous report may be clinically significant but did not produce a statistically significant correlation with website use. Similarly, changes in dietary intake may not correlate highly with website use, as the 2-week study period may not have allowed enough time to purchase healthier foods, such as fruits and vegetables, and procure substitutes for red meat. However, despite the *Exercise* section being the least visited of the 3 main health behavior change sections of the website, significant

improvements in physical activity were noted as well as significant correlations with website use.

The duration and frequency of the log-ins were positively associated with increased physical activity. Therefore, survivors of cancer who frequently logged on and spent substantial time on SurvivorSHINE were more physically active. Similar associations between website use and increased physical activity have also been reported in previous studies of older adults and Latinas. For example, a previous randomized controlled trial examining a web-based intervention to promote physical activity in sedentary older adults [49] found that higher levels of program use (website visits and total time) were associated with greater changes in physical activity [33,49]. Moreover, a recent web-based computer-tailored physical activity trial in older survivors of prostate cancer reported that increased engagement in the physical activity web component was associated with a higher level of moderate to vigorous intensity physical activity at follow-up [50]. Although this study corroborates and extends past findings to survivors of cancer, further investigation of website components that promote healthy lifestyle behaviors among survivors of cancer is needed in a larger population to learn more about the roles of diet and weight components.

This study also sheds light on the subgroups of survivors of cancer who might be more engaged in web-based approaches. The frequency of log-ins, time on the website, and page views did not significantly differ by race and gender, which supports the accessibility and utility of eHealth tools for diverse populations and allays concerns regarding digital divide. However, although a past review found that samples for past web-based lifestyle intervention studies for survivors of cancer were comprised mostly of non-Hispanic White women [21], trends in the current data suggested slightly greater time spent on website and page views in this group; such programs may be particularly appealing to non-Hispanic White women. Although there is limited research on web-based physical activity intervention usability in relation to race and gender, a review of web-based lifestyle interventions among survivors of cancer found that most of the participants in studies that recruited various cancer types were non-Hispanic White and women [21]. This is similar to the usability trend found in this study, suggesting that non-Hispanic White women frequently participate in web-based interventions and may be particularly interested in such programs. Thus, further investigation on how to make the website more relevant to survivors of cancer who are men and non-White may be necessary, especially given the association between website use and behavior change.

Strengths and Limitations

Although minority representation among survivors of cancer was a strength of this study, the sample was still predominately affluent and well educated, which limits the generalizability of these findings to other survivor groups. Another primary weakness was the number of tests conducted and the high probability of a type 1 error. An additional limitation was the possibility that participants could have created more than one account. In fact, 1 participant had 2 accounts (one where they registered with their work email and another account registered with their personal email) but was taken into account, which also could lead to inaccuracies in website use recordings. Moreover, inaccuracies in metrics of time spent on the website could have occurred in situations where users walked away from the computer or navigated to other websites while still logged into their SurvivorSHINE account. However, the available analytics were helpful in providing a rough estimate of website use for this study. Finally, this study was brief in duration and had a small sample size.

As for future directions, findings from this study need to be replicated and further examined in larger, longer-term studies. Qualitative research should be conducted to examine factors that influence the number of log-ins, time spent on the overall website, and specific features of web-based lifestyle interventions for survivors of cancer (eg, determining why the *Healthy Eating* section was visited more than the *Exercise* section).

Conclusions

The findings of this study provide insights into how survivors of cancer use web-based healthy lifestyle interventions. Diet content was popular, whereas exercise was less so, but perhaps more helpful, as analyses suggest that internet-delivered lifestyle program use was positively associated with increases in physical activity within this population. Website use and patterns were similar across demographic factors, which is promising for the potential reach of web-based lifestyle interventions for various segments of survivors of cancer.

Acknowledgments

This research was supported by the American Cancer Society (CRP-14-111-01-CPPB), National Cancer Institute of the National Institutes of Health (R01 CA106919), and the Cancer Prevention and Control Training Program (T32-CA047888). The authors are exceptionally grateful for the efforts of the 35 survivors of cancer who participated in this study.

Conflicts of Interest

None declared.

References

1. Lee M, Yun Y, Park H, Lee E, Jung K, Noh DY. A web-based self-management exercise and diet intervention for breast cancer survivors: pilot randomized controlled trial. Int J Nurs Stud 2014 Dec;51(12):1557-1567. [doi: 10.1016/j.ijnurstu.2014.04.012] [Medline: 24856854]

2. Kanera IM, Bolman CA, Willems RA, Mesters I, Lechner L. Lifestyle-related effects of the web-based Kanker Nazorg Wijzer (Cancer Aftercare Guide) intervention for cancer survivors: a randomized controlled trial. J Cancer Surviv 2016 Mar 17:883-897. [doi: 10.1007/s11764-016-0535-6] [Medline: 26984534]
3. Paxton RJ, Hajek R, Newcomb P, Dobhal M, Borra S, Taylor WC, et al. A lifestyle intervention via email in minority breast cancer survivors: randomized parallel-group feasibility study. JMIR Cancer 2017 Sep 21;3(2):e13 [FREE Full text] [doi: 10.2196/cancer.7495] [Medline: 28935620]
4. Bantum EO, Albright CL, White KK, Berenberg JL, Layi G, Ritter PL, et al. Surviving and thriving with cancer using a Web-based health behavior change intervention: randomized controlled trial. J Med Internet Res 2014;16(2):e54 [FREE Full text] [doi: 10.2196/jmir.3020] [Medline: 24566820]
5. Forbes CC, Blanchard CM, Mummery WK, Courneya KS. Feasibility and preliminary efficacy of an online intervention to increase physical activity in Nova Scotian cancer survivors: a randomized controlled trial. JMIR Cancer 2015 Nov 23;1(2):e12 [FREE Full text] [doi: 10.2196/cancer.4586] [Medline: 28410166]
6. Frensham LJ, Parfitt G, Dollman J. Effect of a 12-week online walking intervention on health and quality of life in cancer survivors: a quasi-randomized controlled trial. Int J Environ Res Public Health 2018 Sep 21;15(10) [FREE Full text] [doi: 10.3390/ijerph15102081] [Medline: 30248943]
7. Golsteijn RH, Bolman C, Volders E, Peels DA, de Vries H, Lechner L. Short-term efficacy of a computer-tailored physical activity intervention for prostate and colorectal cancer patients and survivors: a randomized controlled trial. Int J Behav Nutr Phys Act 2018 Oct 30;15(1):106 [FREE Full text] [doi: 10.1186/s12966-018-0734-9] [Medline: 30376857]
8. Hatchett A, Hallam JS, Ford MA. Evaluation of a social cognitive theory-based email intervention designed to influence the physical activity of survivors of breast cancer. Psychooncology 2013 Apr;22(4):829-836. [doi: 10.1002/pon.3082] [Medline: 22573338]
9. Hong YA, Goldberg D, Ory MG, Towne SD, Forjuoh SN, Kellstedt D, et al. Efficacy of a mobile-enabled web app (iCanFit) in promoting physical activity among older cancer survivors: a pilot study. JMIR Cancer 2015 Jun 26;1(1):e7 [FREE Full text] [doi: 10.2196/cancer.4389] [Medline: 28410158]
10. McCarroll ML, Armbruster S, Pohle-Krauza RJ, Lyzen AM, Min S, Nash DW, et al. Feasibility of a lifestyle intervention for overweight/obese endometrial and breast cancer survivors using an interactive mobile application. Gynecol Oncol 2015 Jun;137(3):508-515 [FREE Full text] [doi: 10.1016/j.ygyno.2014.12.025] [Medline: 25681782]
11. Puszkiewicz P, Roberts AL, Smith L, Wardle J, Fisher A. Assessment of cancer survivors' experiences of using a publicly available physical activity mobile application. JMIR Cancer 2016 May 31;2(1):e7 [FREE Full text] [doi: 10.2196/cancer.5380] [Medline: 28410168]
12. Rabin C, Dunsiger S, Ness KK, Marcus BH. Internet-based physical activity intervention targeting young adult cancer survivors. J Adolesc Young Adult Oncol 2011 Dec;1(4):188-194 [FREE Full text] [doi: 10.1089/jayao.2011.0040] [Medline: 23610737]
13. Short CE, Rebar A, James EL, Duncan MJ, Courneya KS, Plotnikoff RC, et al. How do different delivery schedules of tailored web-based physical activity advice for breast cancer survivors influence intervention use and efficacy? J Cancer Surviv 2017 Feb;11(1):80-91. [doi: 10.1007/s11764-016-0565-0] [Medline: 27498099]
14. Trinh L, Arbour-Nicitopoulos KP, Sabiston CM, Berry SR, Loblaw A, Alibhai SM, et al. RiseTx: testing the feasibility of a web application for reducing sedentary behavior among prostate cancer survivors receiving androgen deprivation therapy. Int J Behav Nutr Phys Act 2018 Jun 07;15(1):49. [doi: 10.1186/s12966-018-0686-0] [Medline: 29880049]
15. Uhm KE, Yoo JS, Chung SH, Lee JD, Lee I, Kim JI, et al. Effects of exercise intervention in breast cancer patients: is mobile health (mHealth) with pedometer more effective than conventional program using brochure? Breast Cancer Res Treat 2017 Feb;161(3):443-452. [doi: 10.1007/s10549-016-4065-8] [Medline: 27933450]
16. Valle CG, Tate DF, Mayer DK, Allicock M, Cai J. A randomized trial of a Facebook-based physical activity intervention for young adult cancer survivors. J Cancer Surviv 2013 Sep;7(3):355-368 [FREE Full text] [doi: 10.1007/s11764-013-0279-5] [Medline: 23532799]
17. Lynch SM, Stricker CT, Brown JC, Berardi JM, Vaughn D, Domchek S, et al. Evaluation of a web-based weight loss intervention in overweight cancer survivors aged 50 years and younger. Obes Sci Pract 2017 Mar 24;3(1):83-94 [FREE Full text] [doi: 10.1002/osp4.98] [Medline: 28392934]
18. Chatterjee A, Prinz A, Gerdes M, Martinez S. Digital interventions on healthy lifestyle management: systematic review. J Med Internet Res 2021 Nov 17;23(11):e26931 [FREE Full text] [doi: 10.2196/26931] [Medline: 34787575]
19. Haberlin C, O'Dwyer T, Mockler D, Moran J, O'Donnell DM, Broderick J. The use of eHealth to promote physical activity in cancer survivors: a systematic review. Support Care Cancer 2018 Jun 16. [doi: 10.1007/s00520-018-4305-z] [Medline: 29909476]
20. Roberts AL, Fisher A, Smith L, Heinrich M, Potts HW. Digital health behaviour change interventions targeting physical activity and diet in cancer survivors: a systematic review and meta-analysis. J Cancer Surviv 2017 Dec;11(6):704-719 [FREE Full text] [doi: 10.1007/s11764-017-0632-1] [Medline: 28779220]
21. Williams V, Brown N, Becks A, Pekmezi D, Demark-Wahnefried W. Narrative review of web-based healthy lifestyle interventions for cancer survivors. Ann Rev Res 2020 Mar 10;5(4) [FREE Full text] [doi: 10.19080/arr.2020.05.555670] [Medline: 33294850]

22. Kuijpers W, Groen WG, Oldenburg HS, Wouters MW, Aaronson NK, van Harten WH. eHealth for breast cancer survivors: use, feasibility and impact of an interactive portal. JMIR Cancer 2016 May 10;2(1):e3 [FREE Full text] [doi: 10.2196/cancer.5456] [Medline: 28410178]
23. Galiano-Castillo N, Cantarero-Villanueva I, Fernández-Lao C, Ariza-García A, Díaz-Rodríguez L, Del-Moral-Ávila R, et al. Telehealth system: a randomized controlled trial evaluating the impact of an internet-based exercise intervention on quality of life, pain, muscle strength, and fatigue in breast cancer survivors. Cancer 2016 Oct 15;122(20):3166-3174 [FREE Full text] [doi: 10.1002/cncr.30172] [Medline: 27332968]
24. Goode AD, Lawler SP, Brakenridge CL, Reeves MM, Eakin EG. Telephone, print, and web-based interventions for physical activity, diet, and weight control among cancer survivors: a systematic review. J Cancer Surviv 2015 Mar 11. [doi: 10.1007/s11764-015-0442-2] [Medline: 25757733]
25. Wang EY, Graff RE, Chan JM, Langlais CS, Broering JM, Ramsdill JW, et al. Web-based lifestyle interventions for prostate cancer survivors: qualitative study. JMIR Cancer 2020 Nov 10;6(2):e19362 [FREE Full text] [doi: 10.2196/19362] [Medline: 33170126]
26. Kuijpers W, Groen WG, Aaronson NK, van Harten WH. A systematic review of web-based interventions for patient empowerment and physical activity in chronic diseases: relevance for cancer survivors. J Med Internet Res 2013;15(2):e37 [FREE Full text] [doi: 10.2196/jmir.2281] [Medline: 23425685]
27. Joseph RP, Ainsworth BE, Vega-López S, Adams MA, Hollingshead K, Hooker SP, et al. Rationale and design of Smart Walk: a randomized controlled pilot trial of a smartphone-delivered physical activity and cardiometabolic risk reduction intervention for African American women. Contemp Clin Trials 2019 Feb;77:46-60. [doi: 10.1016/j.cct.2018.12.011] [Medline: 30576840]
28. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer Statistics, 2021. CA Cancer J Clin 2021 Jan;71(1):7-33 [FREE Full text] [doi: 10.3322/caac.21654] [Medline: 33433946]
29. Bluethmann SM, Mariotto AB, Rowland JH. Anticipating the "Silver Tsunami": prevalence trajectories and comorbidity burden among older cancer survivors in the United States. Cancer Epidemiol Biomarkers Prev 2016 Jul;25(7):1029-1036 [FREE Full text] [doi: 10.1158/1055-9965.EPI-16-0133] [Medline: 27371756]
30. Anderson M, Perrin A. Tech adoption climbs among older adults. Pew Research Center. 2017. URL: https://www.pewresearch.org/internet/2017/05/17/tech-adoption-climbs-among-older-adults/ [accessed 2022-02-04]
31. Hunsaker A, Hargittai E. A review of internet use among older adults. New Media Society 2018 Jul 16;20(10):3937-3954. [doi: 10.1177/1461444818787348]
32. Williams VA, Brown NI, Johnson R, Ainsworth MC, Farrell D, Barnes M, et al. A web-based lifestyle intervention for cancer survivors: feasibility and acceptability of SurvivorSHINE. J Cancer Educ 2021 Jun 01. [doi: 10.1007/s13187-021-02026-x] [Medline: 34061334]
33. Linke SE, Dunsiger SI, Gans KM, Hartman SJ, Pekmezi D, Larsen BA, et al. Association between physical activity intervention website use and physical activity levels among Spanish-speaking Latinas: randomized controlled trial. J Med Internet Res 2019 Jul 24;21(7):e13063 [FREE Full text] [doi: 10.2196/13063] [Medline: 31342902]
34. Atske S, Perrin A. Home broadband adoption, computer ownership vary by race, ethnicity in the U.S. Pew Research Center. 2021. URL: https://www.pewresearch.org/fact-tank/2021/07/16/home-broadband-adoption-computer-ownership-vary-by-race-ethnicity-in-the-u-s/ [accessed 2022-02-04]
35. Internet use by race/ethnicity. Pew Research Center. URL: https://www.pewresearch.org/internet/chart/internet-use-by-race/ [accessed 2022-02-04]
36. Morey MC, Snyder DC, Sloane R, Cohen HJ, Peterson B, Hartman TJ, et al. Effects of home-based diet and exercise on functional outcomes among older, overweight long-term cancer survivors: RENEW: a randomized controlled trial. J Am Med Assoc 2009 May 13;301(18):1883-1891 [FREE Full text] [doi: 10.1001/jama.2009.643] [Medline: 19436015]
37. Rock CL, Doyle C, Demark-Wahnefried W, Meyerhardt J, Courneya KS, Schwartz AL, et al. Nutrition and physical activity guidelines for cancer survivors. CA: Cancer J Clin 2012 Apr 26;62(4):242-274. [doi: 10.3322/Caac.21142]
38. Doran GT. There's a S.M.A.R.T. way to write management's goals and objectives. Manag Rev 1981;70:35-36 [FREE Full text]
39. Godin G, Shephard RJ. A simple method to assess exercise behavior in the community. Can J Appl Sport Sci 1985 Sep;10(3):141-146. [Medline: 4053261]
40. Amireault S, Fong AJ, Sabiston CM. Promoting healthy eating and physical activity behaviors: a systematic review of multiple health behavior change interventions among cancer survivors. Am J Lifestyle Med 2018;12(3):184-199 [FREE Full text] [doi: 10.1177/1559827616661490] [Medline: 30202391]
41. Amireault S, Godin G, Lacombe J, Sabiston CM. Validation of the Godin-Shephard Leisure-Time Physical Activity Questionnaire classification coding system using accelerometer assessment among breast cancer survivors. J Cancer Surviv 2015 Sep;9(3):532-540. [doi: 10.1007/s11764-015-0430-6] [Medline: 25666749]
42. Boyle T, Lynch BM, Courneya KS, Vallance JK. Agreement between accelerometer-assessed and self-reported physical activity and sedentary time in colon cancer survivors. Support Care Cancer 2015 Apr;23(4):1121-1126. [doi: 10.1007/s00520-014-2453-3] [Medline: 25301224]

43. Janz KF. Validation of the CSA accelerometer for assessing children's physical activity. Med Sci Sports Exerc 1994 Mar;26(3):369-375. [Medline: 8183103]
44. Melanson EL, Freedson PS. Validity of the Computer Science and Applications, Inc. (CSA) activity monitor. Med Sci Sports Exerc 1995 Jun;27(6):934-940. [Medline: 7658958]
45. Freedson PS, Melanson E, Sirard J. Calibration of the Computer Science and Applications, Inc. accelerometer. Med Sci Sports Exerc 1998 May;30(5):777-781. [doi: 10.1097/00005768-199805000-00021] [Medline: 9588623]
46. Subar AF, Kirkpatrick SI, Mittl B, Zimmerman TP, Thompson FE, Bingley C, et al. The Automated Self-Administered 24-hour dietary recall (ASA24): a resource for researchers, clinicians, and educators from the National Cancer Institute. J Acad Nutr Diet 2012 Aug;112(8):1134-1137 [FREE Full text] [doi: 10.1016/j.jand.2012.04.016] [Medline: 22704899]
47. 2020 demographic analysis estimates press kit. United States Census Bureau. 2020. URL: https://www.census.gov/newsroom/press-kits/2020/2020-demographic-analysis.html [accessed 2022-02-04]
48. Chen KB, Yen T, Sun W, Tevaarwerk AJ, Wiegmann DA, Heidrich SM, et al. Usage of a web-based workplace and symptom self-management intervention tool to improve work ability for breast cancer survivors. J Cancer Educ 2021 Sep 03. [doi: 10.1007/s13187-021-02048-5] [Medline: 34476769]
49. Irvine AB, Gelatt VA, Seeley JR, Macfarlane P, Gau JM. Web-based intervention to promote physical activity by sedentary older adults: randomized controlled trial. J Med Internet Res 2013;15(2):e19 [FREE Full text] [doi: 10.2196/jmir.2158] [Medline: 23470322]
50. Finlay A, Evans H, Vincent A, Wittert G, Vandelanotte C, Short CE. Optimising web-based computer-tailored physical activity interventions for prostate cancer survivors: a randomised controlled trial examining the impact of website architecture on user engagement. Int J Environ Res Public Health 2020 Oct 28;17(21) [FREE Full text] [doi: 10.3390/ijerph17217920] [Medline: 33126692]

Abbreviations

SMART: specific, measurable, attainable, relevant, and timely

Edited by A Mavragani; submitted 04.06.21; peer-reviewed by C Valle, M Salimi; comments to author 28.07.21; revised version received 04.10.21; accepted 27.12.21; published 21.02.22.

Please cite as:
Williams V, Brown N, Moore JX, Farrell D, Perumean-Chaney S, Schleicher E, Fontaine K, Demark-Wahnefried W, Pekmezi D
Web-Based Lifestyle Interventions for Survivors of Cancer: Usability Study
JMIR Form Res 2022;6(2):e30974
URL: https://formative.jmir.org/2022/2/e30974
doi:10.2196/30974
PMID:35188468

Original Paper

Exploring Deeper Causes Linking Adolescents' Mental Disorders to Mobile Phone Use Problems: Grounded Theory Approach

Zeyuan Sun[1*], BM; Yue Zhou[2*], BA; Yinan Zhang[3], BA, MSc; Bing Gui[4*], BM; Zhenmi Liu[1*], PhD

[1]West China School of Public Health and West China Fourth Hospital, Sichuan University, Chengdu, China

[2]Harvard Law School, Harvard University, Cambridge, MA, United States

[3]Wuxi Mental Health Center, Nanjing Medical University, Wuxi, China

[4]Department of Psychiatry Clinic, Affiliated Dalian Municipal Central Hospital, Dalian Medical University, Dalian, China

[*]these authors contributed equally

Corresponding Author:
Zhenmi Liu, PhD
West China School of Public Health and West China Fourth Hospital
Sichuan University
No. 17, Section 3, Ren Min Nan Road
Chengdu, 610041
China
Phone: 86 18116636482
Email: zhenmiliu@scu.edu.cn

Abstract

Background: Evidence from a variety of studies link mobile phone use with an increase in mental health problems, with the situation being particularly prevalent in China and exacerbated by the COVID-19 quarantine.

Objective: This study aims to reveal underlying connections between mobile phone use and mental disorders of adolescents, and to develop a theory to help parents and counseling psychologists better understand and intervene in future cases.

Methods: A total of 37 teenagers having both mental health and mobile phone use problems, along with their parents, were included for individual interviews. These interviews were transcribed, coded, and analyzed using qualitative methods of grounded theory.

Results: The grades-ranking-first mentality is one of the main factors causing problems such as defective family bonding and peer influences, pushing teenagers with mental disorders to seek comfort in the virtual world through their cellphones.

Conclusions: The idea proposed in this study is not only inspiring for psychological counseling and therapy on adolescents with mental problems but also beneficial for school educators and parents to better understand the adolescents. The findings of the study are also particularly noteworthy in the postpandemic age, where parents whose work locations and schedules are substantially affected due to any emergencies should try to build a relaxing and cozy atmosphere at home to avoid possible conflicts with adolescents.

(JMIR Form Res 2022;6(2):e31089) doi:10.2196/31089

KEYWORDS

mobile phone use; adolescent health; mental disorder

Introduction

Mental health problems have become prevalent among children and adolescents globally, with research estimating 11.3% to 15.9% of those aged 6-18 years [1] and 31% of those aged 10-19 years [2] being affected. According to the World Health Organization [3], depression has become one of the leading causes of mental illness and disability among young people, accounting for 16% of all illnesses and injuries among adolescents. Research also suggests that half of adult mental health disorders begin at age 14 years, but most cases go undetected or untreated [4], casting a profound and longstanding impact on individuals and societies.

On the other hand, evidence from a variety of studies links mobile phone use with an increase in mental distress, self-injurious behaviors, and suicidality among teenagers [5,6]. The situation is particularly prevalent in China, due to large groups of adolescents with extensive mobile phone use [7,8]

compared with other Asian countries like Korea [9] and India [10]. Limited research speculates that the rate of smartphone addiction among adolescents is around 30% in China [11] and that the rate of mobile phone use problems may be even higher [12]. Studies have indicated that Chinese adolescents with attention problems are more likely to have problematic mobile phone use [13]; other characteristics of the population include irrational procrastination and less physical activities [14]. Despite an increasing number of research investigating mobile phone use–related adolescent mental illness cases [15,16], the underlying causes of the problem remain unclear. Possible causes include low reward dependence, low self-esteem, low family function [17], and necessary mobile phone use required by occupation [18].

On the other hand, the stigmas associated with mental disorders prevent children and adolescents from expressing their distress and seeking help directly from others [19]. Even when they do so, Chinese adolescents are more likely to seek help from nonprofessionals (relatives and family members) rather than paying a visit to a psychiatrist or therapist [20]. Moreover, high school students with depression and suicidal ideation mainly seek help from friends and parents, and the rate of seeking help from mental health professionals is very low (about 1%). In fact, 30% of students do not seek help at all when they encounter psychological problems [21]. Additionally, visitors to psychological clinics in China tend to be resistant to participating in studies for ethical and privacy concerns [22], such as the fear of having medical records indicating visits to a mental clinic or private information leaking. Thus, the existing research literature is quite limited.

Having online courses during China's COVID-19 lockdown increased teenager mobile phone use and time spent with parents significantly, which intensified the conflicts within the family [23,24]. The latest research has indicated the loneliness and desire to escape from reality have added to their cell phone and social media use [25]. COVID-19 also prompted additional factors that may have intensified family conflicts: limited physical exercises, restricted social activities, and significant environmental changes [26,27]. These factors have noticeably contributed to the increase of adolescent visits to psychiatric clinics.

This study, therefore, aims to reveal some underlying connections between mobile phone use and mental disorders of adolescents, and to develop a theory to help parents and counseling psychologists better understand and intervene in future cases.

Methods

Study Design

Given the exploratory nature of the inquiry and the limited existing evidence base, we adopted the grounded theory (GT) approach to analyze the social process of mental disorder development among adolescents with mobile phone use problems. The study was conducted from August to December 2020 at a psychological clinic in Dalian Municipal Central Hospital [28], a top teaching hospital with the largest psychological clinic, providing over half of the psychotherapy services in Dalian City. The GT techniques of the constant comparison method and theoretical sensitivity were used throughout the study process, ensuring that the developing codes and theories remain grounded in the data.

Ethics Approval

Approval of the study was obtained from the Ethical Committee of the Dalian Municipal Central Hospital.

Recruitment and Sampling

Purposive sampling was used to target a group of participants reflecting a range of adolescents' characteristics and backgrounds [29]. All visitors were given flyers about the study, and those interested were screened according to preset criteria. Those who met the criteria then gave consent that their conversations with the psychologists would be transcribed into verbatim transcripts for research purposes. Since the vast majority of participants were minors, they were usually accompanied by guardians. The guardians would also sign the informed consent, and then the parents would leave after giving a brief explanation of the adolescent's situation. In special circumstances (eg, when parents or adolescents expressed strong interest to do the interview together or if the adolescent was too young to articulate themselves fluently), the doctor would allow parents to join the interview.

The selection criteria are as follows. We included adolescents aged 10-19 years, with serious mobile phone use problems (defined as they spontaneously talked about their mobile phone use issues and complained about the conflicts with their parents on mobile phone use). We excluded the participant if the doctor considered them unsuitable for the study or the mobile phone use was not salient enough. We also excluded participants with schizophrenia, major depression associated with suicidal thoughts and behaviors, and encephalitis and other organic diseases accompanied by mental disorders.

Data Collection

After completing the consent form, participants were interviewed, followed by a formal psychotherapy 1 to 2 days later. Individual interviews were conducted using a structured guide, and consensus was obtained for each interview [30]. The guide was designed by the corresponding authors (BG and ZL) using the Delphi methods (see Multimedia Appendix 1). All disagreements and conflicts were resolved by the involvement of third-party experts in psychiatry and qualitative research. In actual practice, questions were not entirely constrained by the guide and could be modified as needed. Interviewers are experienced psychological therapists who had training in conducting structured interviews.

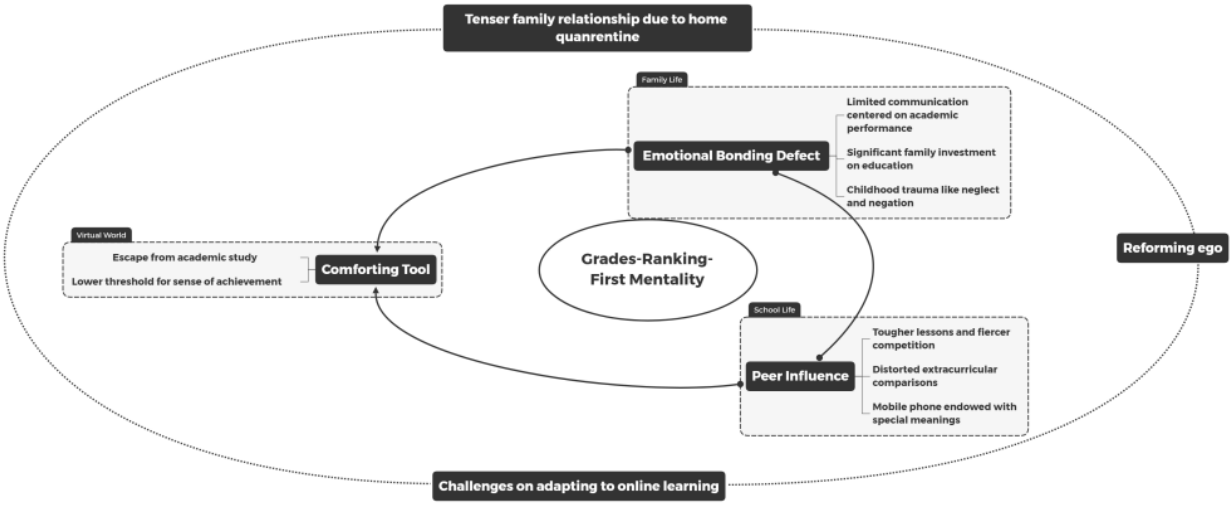
All interviews were conducted in Mandarin and each lasted about 45 minutes. After the interview, the researchers transcribed the interview into verbatim scripts immediately with necessary notes of facial expressions, movements, and emotional reactions that may assist the analysis. Parents accompanying participants were also interviewed, and their statements were recorded as supplementary data. After each verbatim script was

analyzed, the next interview was adjusted accordingly to reach thematic saturation.

Data Analysis

Data analysis was initiated as soon as the first verbatim transcription was completed to avoid recalling bias. The first authors (ZS and Y Zhou) completed initial line-by-line coding of interview transcripts using gerunds to identify facial expressions and other body languages of interviewees noted by the doctors (eg, crying or wiping tears). Two separate researchers (ZS and Y Zhang) completed the open coding, selective coding, and theoretical coding under the supervision of the corresponding authors (BG and ZL) to ensure rigor and interconnectedness among concepts, themes, and categories [31]. From there, we were able to develop a robust description of categories with nuanced properties (See Figure 1 and Textbox 1).

Figure 1. Depiction of the grades-ranking-first mentality system.



Textbox 1. Categories and themes of the participants with mobile phone use issues.

Grades-ranking-first mentality

- Emotional bonding defect (both parents and children)
  - Limited communication centered on academic performance
  - Significant family investment in education
  - Childhood trauma like neglect and negation
- Peer influence (children)
  - Lower threshold for a sense of achievement
  - Tougher lessons and fiercer competition
  - Distorted extracurricular comparisons
  - Mobile phone endowed with special meanings
- Comforting tool (children)
  - Escape from academic study
  - Lower threshold for a sense of achievement
  - Challenges on adapting to online learning
- Other (mixed, see below)
  - Tense family relationship (both parents and children)
  - Reforming ego (children)

Results

Participant Demographics

We had a sample of 37 participants, all living in Dalian City (see Table 1). All reported problems in emotion control or difficulty in an academic study recently and thus are seeking counseling and psychotherapy. Most participants were female (n=28, 76%), and the average age was 14.86 years, with nearly 90% coming from middle or high school.

Table 1. Demographic characteristics of participants.

	Participants (N=37)
Gender, n (%)	
Male	9 (24)
Female	28 (76)
Age (years), mean (SD)	14.86 (1.77)
Grade, n (%)	
Primary	1 (3)
Middle 1	4 (11)
Middle 2	5 (14)
Middle 3	9 (24)
Senior 1	4 (11)
Senior 2	7 (19)
Senior 3	4 (11)
Freshman	1 (3)
Sophomore	1 (3)
Drop out[a]	1 (3)
Reference to the clinic, n (%)	
Myself	5 (14)
Mother	21 (57)
Father	4 (11)
Teacher	6 (16)
Grandma	1 (11)
Symptom duration/month, n (%)	
<12 months	13 (35)
12-24 months	11 (30)
24-36 months	4 (11)
>36 months	9 (24)
Long-term absence of parent, n (%)	
Yes	26 (70)
No	11 (30)
Only child, n (%)	
Yes	25 (68)
No	12 (32)
School bullying, n (%)	
Yes	22 (59)
No	15 (41)
Corporal punishment, n (%)	
Yes	19 (51)
No	18 (49)
Self-harm, n (%)	
Yes	13 (35)
No	24 (65)

[a]The participant dropped out of school after middle school.

Core Category—Grades-Ranking-First Mentality

Mainstream cultural values and social environment have profoundly shaped the thoughts of parents and teachers, eventually influencing adolescents themselves. The idea of linking school performance with career and personal achievements forms an important part of Chinese traditional values. Every Chinese student is expected by their parents and teachers to get higher grades and superior rankings, as it is considered the only way for them to obtain higher education qualifications and seize more career development opportunities. Many Chinese adolescents are thus raised to study diligently and consider nothing but obtaining higher grades. However, when adolescents eventually encounter more real-life challenges, many became frustrated, self-doubted, and even developed mental disorders, which would bring them to counseling clinics.

Subcategory—Emotional Bonding Defect

We found that our participants have relatively weak bonding with their parents. There appeared to be little parental involvement during childhood for most of the participants, and some were taken care of entirely by their grandparents before school age, which leads to growing gaps in mutual understanding. Additionally, their neglect and frequent criticism during childhood diminished the children's self-esteem and confidence, and continued to influence them throughout adolescence.

Limited Communication Centered on Academic Performance

We found that most participants have busy parents who spent the most time at work. When asked about why they spent so little time with their children, the parents usually responded that they were busy earning money to support their family and pay for their children's education. A father said:

> *It is only for this family that I have worked so hard. I have spent all my time making money to pay for her education! That's why I'm so anxious when her grades drop.*

Another father said:

> *When I come home, he is usually asleep, or ready to go to sleep, so we barely talk. I don't know him well, so the only thing I can ask about is his school life.*

On the other hand, the adolescents found that communication with their parents to be difficult. A senior high school boy said:

> *When I get home, I just want to relax by myself, yet my parents wouldn't leave me alone. I am really sick of them asking "how was school today!"*

Significant Family Investment in Education

Another factor that further exacerbates the bonding defect is a considerable financial investment in education. Chinese parents consider investing in education one of the best ways to help their children improve their academic performance and take financial investment as an indicator of how much they value their children's education and care for the children. A father complained:

> *I have paid enough for her study. One 60-minute class of math costs 600 RMB, and she takes it weekly.*

A mother said:

> *We never hesitate to spend money on his study. We try our best to provide for him.*

However, for teenagers, such investments add to their pressure rather than signifying parental love. A senior high girl sobbed:

> *They signed up so many tutoring classes for me, but they wouldn't take me out to travel or buy me a new cellphone. I don't think they love me.*

Childhood Trauma Like Neglect and Negation

Low self-esteem is a typical characteristic for many participants. We believe it stems from specific childhood experiences, namely, neglect and negation of their value by their parents. For their parents, those children are more like "tools" that can bring them honor and pride with their grades than real human beings. A middle school boy said:

> *He said that getting good grades is my only mission. If I failed, I would be a loser and do not deserve to be his son...*

A middle school girl complained about her father:

> *He cares nothing but my grades. He often said that "Successful as I am, how did I give birth to such a fool like you." It hurt me so much!*

Subcategory—Peer Influence

Since school life gradually takes a larger proportion of their daily life, our participants were influenced by their peer groups profoundly. Apart from peer pressure on academic performance, peer comparisons have expanded to extracurricular areas like hobbies and mobile phone use freedom. Additionally, mobile phones have been given extra value and symbolism that goes far beyond their original functions and significance.

Tougher Lessons and Fiercer Competition

A number of changes take place in adolescence, pushing the participants into an ever more challenging and competitive environment. As high school is not part of compulsory education in China, there is naturally more peer pressure and more effort required, as most senior high school students are determined to enter university through the College Entrance Exam. Categorizing students according to their grades and rankings, and even adjusting the seats according to test scores have been common practices in many high schools and are used as an encouragement mechanism for students. Most participants have to study intensively in an atmosphere full of pressure and surrounded by hundreds of competitors. Complained by a senior high school girl:

> *Frankly speaking, I was the top student in middle school with very little effort, but I can't do that in high school. I have a lot of classmates who are smarter*

and more hard-working, so I struggled to keep my rankings not dropping.

Another senior high school student said:

Physics exam is too difficult. Even our teacher can't explain all of them. There is a question he tried to explain for half an hour, and became confused himself. Then the best student in our class went upstage to explain it to us.

Distorted Extracurricular Comparisons

Narrow-minded focus on academic study finally results in parents' restrictions on children pursuing hobbies that do not contribute to academic study and college recruitment. Some participants were encouraged by their parents to play specific sports or practice music instruments only because these skills will allow them to access special college or high school recruitment channels as athletes or art students. In contrast, parents would strictly prohibit hobbies that would not benefit their children with more advantaged positions in school recruitment. We describe these comparisons as "distorted" since these extracurricular comparisons were targeted as means to compete. A senior high school girl said:

I loved dancing when I was a kid, but then my parents stopped me because they asked me to do sports to enter high school as an athlete. Now it's clear that I won't be going to college as an athlete, but my parents still don't let me dance. They ask me to focus on my study.

Complained by a middle school boy:

When I read the extra-curricular books, they scolded me for doing useless things. But I need to read novels and magazines to catch up with the trends and to fit in with classmates.

Such comparisons led the adolescents to compare themselves with their peers in other aspects, such as expecting more freedom to manage their free time, more pocket money, or better at playing mobile phone games. A senior high school girl said:

I envy her (a childhood friend) so much. She can go wherever she wants and buy whatever she needs without restrictions.

Said by a middle school boy:

In fact, I sometimes play mobile phones just to prove that I can play games better than they do. If I can't be better than them in studying, I have to be better somewhere else.

Mobile Phone Endowed With Special Meanings

The mobile phone has become an important part of peer comparison. Although parents consider it a distraction for study, the adolescents consider it recognition from parents that they have sufficient self-control and permission of grown-up rights. On the other hand, it is prevalent for students to compare the brand and price of the mobile phone among their peers.

Sobbed by a senior high school girl:

All my classmates had cell phones in middle school or even earlier, so by now, they've surely played enough. But I didn't have a phone back then, I am just making up for the time I lost, but my parents still do not let me. It is not fair! They promised to compensate me with the latest iPhone, but they only bought me an ordinary brand of mobile phone.

Subcategory—Comforting Tool

With the heavy burden of school life that they are unwilling to face and impaired family relationships that failed to support them emotionally, adolescents seek to escape from study to the virtual world as an easier pathway to attain a sense of achievement.

Escape From Academic Study

To get higher grades, most students need to complete extra work at home, usually assigned by tutors of after school classes or directly from parents. Some started doing extra work as early as in primary school and thus become tired and resistant to such work. Indulging in the virtual world thus becomes an escape from academic study for them. A middle school boy said:

I don't want to do extra exercises assigned by my mom. They are a waste of time. As long as I procrastinate schoolwork a bit later, she would not let me do extra work, so I would deliberately slow down to avoid that.

A middle school boy said:

I do not write my homework immediately when I go home. I usually watch some funny videos for a while and start my homework when I am happy because studying is too boring.

A middle school girl said:

(What do you play with your phone?) Nothing specific, anything that attracts my interests. Just want to run away from studying.

Lower Threshold for Senses of Achievement

The majority of participants are characterized by their incapability to derive a sense of achievement through academic study while at the same time being obsessed with making accomplishments or achieving goals. Consequently, we found most of them addicted to mobile games and entertain themselves with winning a game or passing all levels of a game. A middle school boy said:

Playing games is much easier than getting good rankings. I am quite good at that, maybe I can become a professional player.

A middle school boy said proudly:

My technique is the best among my classmates. They always ask me to team them up and guide them in games.

Challenges on Adapting to Online Learning

Online learning was not a major part of Chinese education, but under the influences of the COVID-19 pandemic, it has become

routine for many students. However, not all students can easily adapt to online learning, which requires self-discipline for students to pay attention to the online videos and understand the studying materials by themselves at home. A middle school girl said:

> *I'm used to the traditional way of teaching, where teachers give us homework. There are regular tests so that I know my ranking, which gives me a target to fight for. But online learning has made some homework impossible to check, and tests much less frequent.*

A senior high school girl said:

> *When I took the online class, I knew it was not my type. As soon as the online classes end and the new semester started, my ranking dropped dramatically.*

Tense Family Relationship Due to Parents' Expectation and Increased Time Together

Some participants reported a tenser family relationship as the parents would wish their children to study instead of doing anything else, which usually leads to frequent reminders and even warnings. The situation was manageable as several hours at home on school days is insufficient for further conflict since high school students often spend over 12 hours at school. School teachers normally have weaker behavior control and monitoring of students than parents as well. However, home quarantine and online learning forced both the students and the parents to spend a much longer time at home. Parents' expectation for their children to be constantly studying manifests as monitoring the children's phone use and frequently checking on the children, especially when the children need to use a cell phone and other electronic devices to study. This leads to growing conflicts between parents and children about cell phone use time. A senior high school boy said:

> *My father stayed at home and supervised my study daily. He forbade me to play my cellphone and relax between classes. We fought over it several times.*

A middle school girl complained:

> *When I took online classes in my room, my mother often came in to see if I am behaving well. She always suspects me of secretly playing mobile games, which makes me feel very uncomfortable.*

Reforming Ego

To some extent, parents show understanding for their children's unusual behaviors and take them as normal changes of adolescence. Some expressed frustration at the adolescents' unwillingness to communicate with them and to defy their wishes, while others expressed anger. The mother of a senior high school girl said:

> *She was quite an obedient girl, but she has become rebellious since middle school. We understand it was due to the adolescence stage, so we tried not to fight with her and satisfied many of her requests, including her smartphone.*

The participants, however, expressed a strong preference for independence and cell phone ownership. A senior high school girl shouted:

> *Why should I listen to them? They don't understand me at all, I have my own feelings and thoughts and they should respect that.*

A middle school girl complained:

> *Mobile phone is a necessity for modern people, so I should have a mobile phone of my own. They have no right to decide how I use it, let alone checking my chatting history.*

Discussion

Summary of Findings

This study used qualitative methodology to explore underlying connections between mobile phone use and mental disorders of adolescents, and to develop a theory to better understand adolescents with mobile phone use issues based on both parent and adolescent perspectives. We found that adolescents' mobile phone use issues were strongly related to the current academic evaluation system of the grade-ranking-only judgment for Chinese teenagers, which manifested in the defective bonding with family members and overfierce competition among peers. Under this competitive and undependable environment, adolescents turn to the virtual world to seek comfort, mainly via their mobile phones. Thus, we theorize that the grades-ranking-first mentality is one of the main factors contributing to mobile phone use problems of adolescents with mental disorders.

Our study found that the grades-ranking-first mentality may serve as an important reason for mobile phone use in China. Initial researches conducted in other eastern and southeastern countries and regions have indicated causes like loneliness, low self-esteem, and low family functions [32,33], while later studies have noticed the roles of life satisfaction [34], family attachment, and peer influence [35]. These studies, which have all used scale measurements, revealed important insights on teenager mobile phone use. However, as quantitative studies, they are limited in demonstrating the deeper root causes and reflecting individual perspectives. As family interactions and school life emerged from the data, we are able to explain the phenomenon more comprehensively. Our discovery connected multi-aspects of the mobile phone use issues of the Chinese adolescents and qualitative exploration on the relationships between mobile phone use and mental problems of adolescents. It is not only inspiring for psychological counseling and therapy but also beneficial for school educators and parents to better understand the problems of adolescents.

The findings of the study are also particularly noteworthy during the current period. Recent studies have indicated the negative effects of COVID-19 quarantine on school performance among children and adolescents [36,37]. They proposed possible contributing factors like financial strains and the anxiety of parents [38,39]. Thus, we added related questions into the interview guide to further explore the contributing factors of the intensified parent-child conflicts during the pandemic.

Though no direct financial-related problems were discovered in our study, we did find evidence in parents' statements that income decline due to quarantine has caused anxiety. We found that the prolonged time spent together with children could make parents unconsciously pass their stress and anxiety to adolescents through increasing negative family interactions like scolding and corporal punishments. These mechanisms were also proposed by similar studies [40,41]. Other latest studies have shown the impact of COVID-19 and lockdown on the mental health of children and adolescents through increased family interaction time. Although there may be an initial increase of conflicts between parents and children, the situation can be improved by having open discussions and communications [42], and establishing a routine schedule like the one in school [43]. From this, we suggest that parents whose works are substantially affected by the pandemic and the quarantine should try to build a relaxing and cozy atmosphere at home to reduce the negative psychological impact on their children.

Limitations of This Study

Our study benefited from the inclusion of a larger sample than other studies, and we collected statements from both adolescents and their parents, which gave us a broader analytical perspective. However, some outside factors limited the time frame of participant recruitment. The clinic opens only on weekdays, so adolescents with strict school attendance policies, such as seniors who are under great pressure with exams, may take up a smaller proportion of the sample than estimated. Despite our efforts to expand the opening time of the clinic to obtain a more comprehensive sample, there is still unavoidable selection bias. Besides, the study may also have self-selection bias, as many teenagers were brought to the clinic by concerned parents who might be more willing to solve problems for their children in the first place.

Implication for Practice

The findings of the study are also reusable after the COVID-19 pandemic, as the pandemic is merely an environmental factor that exacerbates the problem of mental disorders caused by mobile phone use. Mediating factors of family conflict intensified by the pandemic (eg, reduced time of physical exercise or decrease in family income) can also be caused by other incidents in life, including unemployment, divorce [44,45], and sometimes natural disasters [46].

Implication for Research

We noticed that mobile phone use is simply one entry point to study adolescents with mental problems, as we found a series of other situations such as school bullying and corporal punishment that some of our participants were going through. The theory we developed may also help study adolescents with similar cultural and educational backgrounds in Asian countries like Korea and Singapore, where adolescents were also reported to have noticeable mobile phone use addiction and a high prevalence of mental disorders. Further qualitative studies could benefit from including more statements of family members and conducting more in-depth interviews with adolescents themselves, especially in Asian countries like China, where adolescents are sometimes too introverted to express their true feelings and opinions in time.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Guideline using the Delphi methods.
[DOCX File , 13 KB - formative_v6i2e31089_app1.docx]

References

1. 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]
2. Silva SA, Silva SU, Ronca DB, Gonçalves VSS, Dutra ES, Carvalho KMB. Common mental disorders prevalence in adolescents: a systematic review and meta-analyses. PLoS One 2020;15(4):e0232007 [FREE Full text] [doi: 10.1371/journal.pone.0232007] [Medline: 32324835]
3. Stansbury JP, Ried LD, Velozo CA. Unidimensionality and bandwidth in the Center for Epidemiologic Studies Depression (CES-D) Scale. J Pers Assess 2006 Feb;86(1):10-22. [doi: 10.1207/s15327752jpa8601_03] [Medline: 16436016]
4. 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]
5. Abi-Jaoude E, Naylor KT, Pignatiello A. Smartphones, social media use and youth mental health. CMAJ 2020 Feb 10;192(6):E136-E141. [doi: 10.1503/cmaj.190434] [Medline: 32041697]
6. Liu X, Luo Y, Liu Z, Yang Y, Liu J, Jia C. Prolonged mobile phone use is associated with poor academic performance in adolescents. Cyberpsychol Behav Soc Netw 2020 May;23(5):303-311. [doi: 10.1089/cyber.2019.0591] [Medline: 32191529]
7. Long J, Liu T, Liao Y, Qi C, He H, Chen S, et al. Prevalence and correlates of problematic smartphone use in a large random sample of Chinese undergraduates. BMC Psychiatry 2016 Nov 17;16(1):408 [FREE Full text] [doi: 10.1186/s12888-016-1083-3] [Medline: 27855666]

8. Wu J, Siu ACK. Problematic mobile phone use by Hong Kong adolescents. Front Psychol 2020;11:551804. [doi: 10.3389/fpsyg.2020.551804] [Medline: 33384636]
9. Cha S, Seo B. Smartphone use and smartphone addiction in middle school students in Korea: prevalence, social networking service, and game use. Health Psychol Open 2018;5(1):2055102918755046 [FREE Full text] [doi: 10.1177/2055102918755046] [Medline: 29435355]
10. Nikhita CS, Jadhav PR, Ajinkya SA. Prevalence of mobile phone dependence in secondary school adolescents. J Clin Diagn Res 2015 Nov;9(11):VC06-VC09 [FREE Full text] [doi: 10.7860/JCDR/2015/14396.6803] [Medline: 26672469]
11. Chen B, Liu F, Ding S, Ying X, Wang L, Wen Y. Gender differences in factors associated with smartphone addiction: a cross-sectional study among medical college students. BMC Psychiatry 2017 Oct 10;17(1):341 [FREE Full text] [doi: 10.1186/s12888-017-1503-z] [Medline: 29017482]
12. Lee PH, Tse ACY, Wu CST, Mak YW, Lee U. Validation of self-reported smartphone usage against objectively-measured smartphone usage in Hong Kong Chinese adolescents and young adults. Psychiatry Investig 2021 Feb;18(2):95-100 [FREE Full text] [doi: 10.30773/pi.2020.0197] [Medline: 33517618]
13. Zheng F, Gao P, He M, Li M, Wang C, Zeng Q, et al. Association between mobile phone use and inattention in 7102 Chinese adolescents: a population-based cross-sectional study. BMC Public Health 2014 Oct 01;14:1022 [FREE Full text] [doi: 10.1186/1471-2458-14-1022] [Medline: 25273315]
14. Shi M, Zhai X, Li S, Shi Y, Fan X. The relationship between physical activity, mobile phone addiction, and irrational procrastination in Chinese college students. Int J Environ Res Public Health 2021 May 17;18(10):5325 [FREE Full text] [doi: 10.3390/ijerph18105325] [Medline: 34067764]
15. Sohn SY, Rees P, Wildridge B, Kalk NJ, Carter B. Prevalence of problematic smartphone usage and associated mental health outcomes amongst children and young people: a systematic review, meta-analysis and GRADE of the evidence. BMC Psychiatry 2019 Nov 29;19(1):356 [FREE Full text] [doi: 10.1186/s12888-019-2350-x] [Medline: 31779637]
16. Kuss DJ, Kanjo E, Crook-Rumsey M, Kibowski F, Wang GY, Sumich A. Problematic mobile phone use and addiction across generations: the roles of psychopathological symptoms and smartphone use. J Technol Behav Sci 2018;3(3):141-149 [FREE Full text] [doi: 10.1007/s41347-017-0041-3] [Medline: 30238057]
17. Ko C, Yen J, Yen C, Lin H, Yang M. Factors predictive for incidence and remission of internet addiction in young adolescents: a prospective study. Cyberpsychol Behav 2007 Aug;10(4):545-551. [doi: 10.1089/cpb.2007.9992] [Medline: 17711363]
18. Aguaded Gómez MC, Márquez IM, Martín Vélez C. Disorders caused by the use and abuse of mobile phones in Spanish students. MOJ Addict Med Ther 2018 May 21;5(3):122-127. [doi: 10.15406/mojamt.2018.05.00104]
19. Chen SX, Mak WWS. Seeking professional help: etiology beliefs about mental illness across cultures. J Couns Psychol 2008 Oct;55(4):442-450. [doi: 10.1037/a0012898] [Medline: 22017551]
20. Zheng Y, Zheng X. Current state and recent developments of child psychiatry in China. Child Adolesc Psychiatry Ment Health 2015;9:10 [FREE Full text] [doi: 10.1186/s13034-015-0040-0] [Medline: 25972919]
21. Hesketh T, Ding QJ, Jenkins R. Suicide ideation in Chinese adolescents. Soc Psychiatry Psychiatr Epidemiol 2002 May;37(5):230-235. [doi: 10.1007/s00127-002-0536-9] [Medline: 12107715]
22. Zheng F, Gao P, He M, Li M, Tan J, Chen D, et al. Association between mobile phone use and self-reported well-being in children: a questionnaire-based cross-sectional study in Chongqing, China. BMJ Open 2015 May 11;5(5):e007302 [FREE Full text] [doi: 10.1136/bmjopen-2014-007302] [Medline: 25967996]
23. Lee J. Mental health effects of school closures during COVID-19. Lancet Child Adolesc Health 2020 Jun;4(6):421 [FREE Full text] [doi: 10.1016/S2352-4642(20)30109-7] [Medline: 32302537]
24. 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]
25. Li J, Zhan D, Zhou Y, Gao X. Loneliness and problematic mobile phone use among adolescents during the COVID-19 pandemic: the roles of escape motivation and self-control. Addict Behav 2021 Jul;118:106857 [FREE Full text] [doi: 10.1016/j.addbeh.2021.106857] [Medline: 33676160]
26. 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]
27. Shah K, Mann S, Singh R, Bangar R, Kulkarni R. Impact of COVID-19 on the mental health of children and adolescents. Cureus 2020 Aug 26;12(8):e10051 [FREE Full text] [doi: 10.7759/cureus.10051] [Medline: 32999774]
28. Dalian Municipal Central Hospital. URL: https://www.dlzxyy.com/pages/sys/web/home.html [accessed 2022-02-10]
29. Scott JC, Glaser BG. The discovery of grounded theory: strategies for qualitative research. Am Sociol Rev 1971 Apr;36(2):335. [doi: 10.2307/2094063]
30. Wimpenny P, Gass J. Interviewing in phenomenology and grounded theory: is there a difference? J Adv Nurs 2000 Jun;31(6):1485-1492. [doi: 10.1046/j.1365-2648.2000.01431.x] [Medline: 10849162]
31. Ellis C, Strauss A, Corbin J. Basics of qualitative research: grounded theory procedures and techniques. Contemp Sociol 1992 Jan;21(1):138. [doi: 10.2307/2074814]
32. Zulkefly SN, Baharudin R. Mobile phone use amongst students in a university in Malaysia: its correlates and relationship to psychological health. Eur J Sci Res 2009;37:18 [FREE Full text]

33. Hong F, Chiu S, Huang D. A model of the relationship between psychological characteristics, mobile phone addiction and use of mobile phones by Taiwanese university female students. Comput Hum Behav 2012 Nov;28(6):2152-2159. [doi: 10.1016/j.chb.2012.06.020]
34. Xie X, Zhao F, Xie J, Lei L. Symbolization of mobile phone and life satisfaction among adolescents in rural areas of China: mediating of school-related relationships. Comput Hum Behav 2016 Nov;64:694-702. [doi: 10.1016/j.chb.2016.07.053]
35. Xie X, Chen W, Zhu X, He D. Parents' phubbing increases adolescents' mobile phone addiction: roles of parent-child attachment, deviant peers, and gender. Child Youth Services Rev 2019 Oct;105:104426. [doi: 10.1016/j.childyouth.2019.104426]
36. Racine N, Cooke JE, Eirich R, Korczak DJ, McArthur B, Madigan S. Child and adolescent mental illness during COVID-19: a rapid review. Psychiatry Res 2020 Oct;292:113307 [FREE Full text] [doi: 10.1016/j.psychres.2020.113307] [Medline: 32707216]
37. Fegert JM, Vitiello B, Plener P, Clemens V. Challenges and burden of the Coronavirus 2019 (COVID-19) pandemic for child and adolescent mental health: a narrative review to highlight clinical and research needs in the acute phase and the long return to normality. Child Adolesc Psychiatry Ment Health 2020;14:20 [FREE Full text] [doi: 10.1186/s13034-020-00329-3] [Medline: 32419840]
38. Rosen Z, Weinberger-Litman SL, Rosenzweig C, Rosmarin DH, Muenning P, Carmody ER, et al. Anxiety and distress among the first community quarantined in the U.S due to COVID-19: psychological implications for the unfolding crisis. PsyArXiv Preprint posted online on April 14, 2020. [doi: 10.31234/osf.io/7eq8c]
39. Jiao WY, Wang LN, Liu J, Fang SF, Jiao FY, Pettoello-Mantovani M, et al. Behavioral and emotional disorders in children during the COVID-19 epidemic. J Pediatr 2020 Jun;221:264-266.e1 [FREE Full text] [doi: 10.1016/j.jpeds.2020.03.013] [Medline: 32248989]
40. Redmond D. The effect of video gaming on family communication and interaction. Iowa State University Digital Repository. 2010. URL: https://dr.lib.iastate.edu/entities/publication/22147d04-5260-467b-97ca-d7af01588f83 [accessed 2022-02-03]
41. Labella MH, Masten AS. Family influences on the development of aggression and violence. Curr Opin Psychol 2018 Feb;19:11-16. [doi: 10.1016/j.copsyc.2017.03.028] [Medline: 29279207]
42. Tang S, Xiang M, Cheung T, Xiang Y. Mental health and its correlates among children and adolescents during COVID-19 school closure: The importance of parent-child discussion. J Affect Disord 2021 Jan 15;279:353-360 [FREE Full text] [doi: 10.1016/j.jad.2020.10.016] [Medline: 33099049]
43. Zhao Y, Guo Y, Xiao Y, Zhu R, Sun W, Huang W, et al. The effects of online homeschooling on children, parents, and teachers of grades 1-9 during the COVID-19 pandemic. Med Sci Monit 2020 Sep 12;26:e925591 [FREE Full text] [doi: 10.12659/MSM.925591] [Medline: 32917849]
44. Nilsen SA, Breivik K, Wold B, Askeland KG, Sivertsen B, Hysing M, et al. Divorce and adolescent academic achievement: heterogeneity in the associations by parental education. PLoS One 2020;15(3):e0229183 [FREE Full text] [doi: 10.1371/journal.pone.0229183] [Medline: 32130231]
45. Brand J, Moore R, Song X, Xie Y. Why does parental divorce lower children's educational attainment? A causal mediation analysis. Sociol Sci 2019 Apr;6:264-292 [FREE Full text] [doi: 10.15195/v6.a11] [Medline: 31187049]
46. Cerna-Turoff I, Fischer H, Mansourian H, Mayhew S. The pathways between natural disasters and violence against children: a systematic review. BMC Public Health 2021 Jul 12;21(1):1249 [FREE Full text] [doi: 10.1186/s12889-021-11252-3] [Medline: 34247619]

Abbreviations

GT: grounded theory

Edited by A Mavragani; submitted 09.06.21; peer-reviewed by T Ntalindwa, G Mayer, CI Sartorão Filho; comments to author 12.08.21; revised version received 10.09.21; accepted 22.12.21; published 21.02.22.

Please cite as:
Sun Z, Zhou Y, Zhang Y, Gui B, Liu Z
Exploring Deeper Causes Linking Adolescents' Mental Disorders to Mobile Phone Use Problems: Grounded Theory Approach
JMIR Form Res 2022;6(2):e31089
URL: https://formative.jmir.org/2022/2/e31089
doi:10.2196/31089
PMID:35188470

Original Paper

Embedding and Integrating a Digital Patient Management Platform Into Everyday Primary Care Routines: Qualitative Case Study

Susanne Frennert[1], PhD; Gudbjörg Erlingsdóttir[1], PhD; Mirella Muhic[2], PhD; Christofer Rydenfält[1], PhD; Veronica Milos Nymberg[3], PhD, MD; Björn Ekman[3], PhD

[1]Department of Design Sciences, Lund University, Lund, Sweden

[2]Department of Informatics, Umeå University, Umeå, Sweden

[3]Department of Clinical Sciences, Malmö, Lund University, Malmö, Sweden

Corresponding Author:
Susanne Frennert, PhD
Department of Design Sciences
Lund University
Ingvar Kamprad Designcentrum
Sölvegatan 26
Lund, 223 62
Sweden
Phone: 46 2220195
Email: susanne.frennert@design.lth.se

Abstract

Background: Traditional primary care is characterized by patient consultations via phone and physical visits. However, the current development in Swedish primary care is to blend digital solutions with traditional solutions. This paper addresses this development by examining the normalization of embedding and integrating a digital health care platform into everyday care routines in a primary care clinic. The digital health care platform enables both synchronous (video calls) and asynchronous (chat) communication, as well as self-registration of patient data using automated questions and forms requiring the patient's input.

Objective: This study aims to explore the work that health care professionals (HCPs) have to undertake to implement and sustain a digital health care platform as part of their everyday work practice.

Methods: HCPs were observed and interviewed to assess their individual and collective engagement and the mechanisms involved in the implementation of the digital platform and its effects on everyday work routines. The normalization process theory (NPT) was used to frame the data analysis.

Results: The analysis identified several themes related to the four NPT constructs: coherence, cognitive participation, collective action, and reflexive monitoring. The use of these constructs enabled the analysis to identify ways of supporting implementation. For example, it showed the benefits of having implementation champions and scheduling work hours for HCPs to use the platform. The analysis also revealed a theme of *materiality* that deviated from the NPT constructs, as NPT gives ontological priority to human actors and social structures.

Conclusions: Digital health care platform implementation is a complex process. Our findings provide insights into how individual and collective actions can be supported to embed and integrate a digital platform into everyday care routines. Primary health care organizations need to involve HCPs throughout the implementation process by reorganizing work and providing frequent feedback loops. HCPs are more likely to engage with and commit to changing practices if they perceive the digital platform to be beneficial compared with the current practice. However, they also need resources (eg, time, training, and continuous support) to put the platform into practice. Patient engagement and appraisal are important elements in implementation. Unless patients are willing to use the platform, there is no motivation for HCPs to embed the digital platform into everyday care practice.

(JMIR Form Res 2022;6(2):e30527) doi:10.2196/30527

KEYWORDS

digital patient management platform; primary care; normalization process theory coherence; cognitive participation; collective action; reflexive monitoring

Introduction

Background

Sweden, similar to many other countries, is undergoing demographic changes, resulting in an increasing older population and a shortage of health care professionals (HCPs). Thus, the Swedish health care system is in transition, aiming to move from reactive to proactive care, with a strong focus on preventive care that is aimed at keeping people out of hospitals. The change from inpatient care to outpatient care means that the locus of care is increasingly shifting toward primary care and patients' homes.

Enhanced digitalization in primary care is seen as a way of meeting current and future challenges [1,2]. The digital transformation of health care is associated with the promise of an improved public sector, with self-care and greater self-management of the health of individuals, resulting in increased cost efficiency. Digitalization in primary care entails converting parts of the care process from analog to digital to meet the needs of citizens, patients, decision-makers, HCPs, and clinics. In Sweden, the digital transformation of primary care has involved several different kinds of digital solutions, including web-based physicians (general practitioners performing digital consultations) [3], self-monitoring mobile apps [4], patient access to their electronic medical records (OpenNotes) [5], digital reminders for medication [6], and digital platforms for patient management [7,8]. This paper concerns the normalization process of a digital platform for patient management. The platform, which will be referred to as Wolf in this paper, was developed by a Swedish company and is procured and used by both private and publicly run primary health care centers and other primary care operations. In Sweden, all types of primary care are tax funded. Both public and private primary care providers enjoy a large degree of economic autonomy. This means that capital investment decisions, such as the one studied here, are made by the owners of the primary care providers, who also finance the investments. Such investments have no direct effect on user fees and are made in the hopes of providing more efficient and less costly services.

The same digital platforms as studied in this paper have been studied in different primary care settings [7-9]. However, the findings show diverse results and have not explicitly focused on the normalization process. For example, Entezarjou et al [9] investigated how family medicine physicians and nurses at primary care centers experienced the implementation and use of digital communication in the form of automated patient interviewing software and chat-based patient–provider communication. Their findings indicated that the use of the digital platform streamlined consultations with patients and improved interdisciplinary communication. The study also found that digitalization was perceived as essential because of the expectations of patients and HCPs. However, their interviewees acknowledged that digital communication did not suit all patients and all kinds of patient requests, and thus, they perceived the platform as only a partial solution [9]. Cajander et al [7] studied a digital platform, which they called a digitalized and automated patient-centric service, in their research on how it affected work engagement among nurses in a national telephone nursing service. They found that nurses felt stressed by having to deal with many patients in parallel, knowing that patients were waiting in the queue to chat with them. As the chat function offered asynchronous communication, the nurses spent time waiting for patients to respond and checking over and over to see whether any patients in parallel, ongoing conversations had responded while, at the same time, a queue of new patients was piling up. Another disadvantage highlighted by the nurses was the lack of feedback from physicians and patients. This finding contrasts with the findings of Entezarjou et al [9], who indicated that the digital platform increased interdisciplinary communication and streamlined patient consultations. However, it is not clear whether the lack of feedback was because of the use of the digital platform or whether this was also common in the ordinary care routine before the implementation of the platform [7]. Another study [8] of the same platform, conducted at primary care centers, indicated that written communication with patients was perceived as more time consuming than communication by phone. It also produced large amounts of text, which made it more difficult to quickly scan information to get an overview of patient concerns. These findings are also somewhat at odds with the finding of Cajander et al [7] that nurses felt that written communication with patients was less emotionally demanding. The prime advantage of using the digital platform, according to Eldh et al [8], was the possibility for patients to upload photos, which reduced physical patient visits. Their findings indicated that the system of automated triage facilitated response and consultation with patients. They conclude that the HCPs in the study lacked the resources needed to benefit from the platform and settled upon different routines themselves, with varied results [8].

The findings of the aforementioned studies indicate that several intertwined factors rather than just the digital platform per se affect the outcome in care practice. Thus, technology adaption and implementation in primary care are not straightforward; rather, it is a complex process that is affected by multiple factors (technological, social, structural, historical, economic, and political), different actors (HCPs, patients, next-of-kin, management, and politicians), organizations, ecologies of technologies, and core ideas (eg, patient safety, patient access, and efficiency). These are interrelated and afford certain kinds of care and working conditions while restraining others [10-12]. Past research shows that the implementation of this type of digital platform is a complex process involving a wide range of actors who translate means, actions, and objectives into care practices in different ways [13] and do not always render the expected effects [7]. Furthermore, early research on digital implementation in health care has been accused of being rich in data but *information poor* [10]. However, several theoretical tools for comprehending and illuminating implementation failures or successes have been developed [10,13-16]. One such explanatory framework is the normalization process theory (NPT) [15]. NPT identifies and explains the important mechanisms that promote or inhibit the implementation process. It allows for a systematic exploration of how and why (or not) a digital health care platform becomes normalized and sustained

in health care practice. NPT "characterizes and explains implementation processes as interactions between 'emergent expressions of agency' (i.e. the things that people do to make something happen, and the ways they work with different components of a complex intervention to do so); and as 'dynamic elements of context' (the social-structural and social-cognitive resources that people draw on to realize that agency)" [17]. Other theoretical frameworks that have been used to better understand the implementation processes concerned with digital technologies in health care are the Rogers [18,19] diffusion of innovations theory and the technology acceptance model [20,21]. Although the latter is primarily concerned with the characteristics at the individual level that affect individuals' intentions to use the technology, the former attempts to connect the macro level to the micro level or context of use. However, although the Rogers [18,19] diffusion of innovation theory considers the context, it does not account for the dynamics of the context in the way that NPT does.

Aims and Objectives

This paper adds to the growing body of research on digital platforms in primary care and the understanding of how to facilitate implementation. We seek to explore whether and how a digital platform becomes part of the everyday primary care practice and the mechanisms involved in its implementation. NPT pays attention to what work people do, or need to do, for complex interventions to become a normal part of everyday work [15,16]. Therefore, in this paper, we explore whether and how the constructs of NPT are visible in the implementation of a digital platform in Swedish primary care and whether the use of NPT offers additional insights into more inductive approaches.

Methods

Overview

This study is part of a large project on the impacts of digitalization on the working environment of primary HCPs when transforming traditional primary care practices into combined digital and physical care practices. As part of the project, this qualitative study explores, through interviews and observations, the implementation process of a digital patient management platform in regard to NPT [15].

Theoretical Framework: NPT

NPT [15,17] is a middle-range theory that goes beyond identifying barriers to and facilitators of a specific implementation and focuses on the means by which people, individually and together, act and try to configure an implementation process. NPT describes four constructs that shape an implementation process: *coherence*, which is the extent to which an innovation is understood as meaningful, desirable, and practicable; *cognitive participation*, which is the enrollment of key individuals to drive the implementation of the innovation; *collective action*, which is the work that operationalizes the innovation into practice; and *reflexive monitoring*, which is the ongoing process of adjusting and appraising the innovation to sustain routines [17].

Our motive for using NPT is that the theory offers a way of understanding how HCPs individually and mutually engage in or foresee engagement with a digital platform in primary care and how specific organizational and social norms and values are being invented or reinvented in the interactions with the digital platform and other people [17,22,23]. As such, it allows for the anticipation of and reflection on the implications of technology implementation in practice to ensure informed implementation, use, and the governance of digital primary care solutions.

The Setting

The provision of health care in Sweden is the responsibility of the 21 autonomous health care regions, the largest of which are the region of Stockholm, region of Västra Götaland (including the city of Gothenburg), and region of Skåne (including the city of Malmö). The regions ensure the universal provision of health care to the citizens of the regions through the direct ownership of clinics and hospitals or by means of contracting with private providers. Most hospitals are publicly owned by the regions, whereas approximately 40% of primary care providers are private. Most private primary care providers are owned by larger companies that run chains of primary care clinics.

In 2019, Sweden spent approximately 11% of the Gross Domestic Product on health care (compared with the Organization for Economic Cooperation and Development average of 8%). Health care is funded primarily by means of regional income taxes (approximately 82% of total health care expenditure). Approximately 15% of the overall spending is in the form of out-of-pocket expenditures, largely because of user fees for health care and cost sharing with respect to pharmaceuticals. Only a small, albeit increasing, portion of the total health care spending goes to private health insurance.

Our study focuses on HCPs employed at Albra, a primary health care center in Sweden (the health care center, the digital platform, the company, and interviewees are anonymized. The names Albra, Wolf, and company D are thus fictitious). It has approximately 40 employees representing different health care professions. These physicians, nurses, psychologists, rehabilitation coordinators, and health care administrators provide care to approximately 10,000 patients, which makes Albra an average-sized Swedish primary care center. Albra has a team-based approach with close collaboration between different HCPs. Similar to many other primary care centers in Sweden, Albra uses a nurse-based triage of patient errands. This means that the first point of contact is a nurse, who, during the telephone conversation, assesses whether the patient needs to see a physician or another HCP in the team. Primary care nurses in Sweden are both registered nurses and nurses with specialist competence in primary care (district nurses). District nurses can prescribe some medications, such as topical steroids and antihistamines, usually after an in-person assessment. In addition to a varying number of remote patient contacts (usually by phone) and other patient-related administrative tasks, physicians usually meet between 10 and 20 patients daily, with visits varying between 15 and 45 minutes depending on the type of visit (emergency or planned follow-up of chronic diseases). These patients are often triaged by the first-contact nurses before

the booking. Primary and secondary care use different electronic medical records, and patients have limited access to medical information from their records by accessing a national platform called 1177 with a bank ID.

Albra was struggling with long patient queues and a stressful work environment, especially for nurses who were answering patients' phone calls. Therefore, in December 2019, the management and staff decided to implement a digital health care platform known as Wolf, in an attempt to address these problems. The purpose of the implementation was to increase patient accessibility, improve resource use, and decrease the workload for HCPs (primarily nurses).

The Platform

The digital platform, Wolf, was developed in Sweden in 2016 by people with backgrounds in health care to support work and patient management in primary and outpatient care. The platform has been widely implemented by private and public primary care centers in Sweden. It is also used in Norway, the Czech Republic, Poland, and England.

The purpose of the platform is to improve medical quality, resource use, and patient experience. In contrast to traditional Swedish primary care routines, the platform enables digital patient contact. Through the platform, patients can initiate contact with the health care center digitally instead of through phone calls. First, patients enter their symptoms and receive automated questions depending on their responses. The patient data are compiled as a medical report and communicated to the health care center via the platform. At the health care center, a nurse enters the platform, reads the medical report, and allocates the patient matter to appropriate health care personnel for attendance. Thereafter, communication and patient meetings take place either synchronously or asynchronously in the form of digital (video or chat) or physical meetings with different categories of HCPs.

Data Collection

The collected data covered the initial phase of implementation and its outcomes. The data were collected using the following settings:

- Two half-day observations during the HCP training sessions with the digital platform provider
- Observation at a formal workplace meeting in which the HCPs could comment and reflect on the platform together with colleagues and management
- Qualitative interviews with HCPs (N=11)
- A qualitative interview with one of the main initiators and developers of the platform

All 3 observations were conducted by a researcher who sat in the 2 training sessions and the workplace meeting. The observed personnel were aware of the researcher's identity and the reason for her presence. Notes were taken during all 3 observations. The notes were rewritten and added to the analysis.

The interviewees were recruited and interviewed by one of the authors. Data collection took place between January 2020 and June 2020. Each participant was interviewed individually by phone or video call because of the current COVID-19 pandemic. The interviews averaged 40 minutes and were digitally recorded and transcribed in full for in-depth analysis. The interview questions were directed at eliciting the views and experiences of Wolf and its implementation. They were asked to describe the implementation process and their role in it, how they learned about the digital platform and how to use it, their expectations about how the platform would affect their everyday practice, and how they then experienced the digital platform and its effect on their work.

Analysis

Our analytical emphasis was on the meanings our interviewees gave to the digital platform, its implementation, and its perceived effect on their work. The constructs of NPT (ie, coherence, collective participation, collective action, and reflexive monitoring) served as a framework for identifying the work and meanings that the HCPs associated with the embedding and integration of the digital platform into everyday care practice [15]. According to NPT, an innovation is likely to become successfully integrated if (1) those affected by it and its implementation understand and agree on its adaption and usefulness in practice (*coherence*), (2) those affected by it and its implementation engage and commit to using the new innovation (*cognitive participation*), (3) operational work is done to enact the innovation as part of the practice (*collective action*), and (4) those affected by it and its implementation positively and continuously appraise the utility and usefulness of the platform (*reflexive monitoring*) [15,17].

We reviewed all transcripts and highlighted all texts that appeared to describe important issues related to the platform and its implementation. All highlighted text was then coded using an inductive approach [24] to identify themes. MM, GE, and SF each coded the material independently, and the identified themes were then discussed and scrutinized until consensus was reached. Fifteen themes were identified: *beneficial for patients*, *assistance for nurses*, *as a supplement*, *close collaboration*, *implementation champions*, *patient engagement*, *familiarity*, *parallel work practices*, *new communication patterns*, *ease of use*, *enhancingteamwork*, *relief and less distress*, *flexibility*, *workarounds*, and *materiality*.

After the coding, data for each theme were examined to determine whether the categories of NPT described by May et al [15,17] were present.

Themes that could not be coded into one of these NPT categories were coded under another label that captured the essence of individual and collective actions and the mechanisms involved in the implementation. We compared the extent to which the data were supportive of the constructs described by the NPT versus how much they represented different mechanisms involved in the implementation of the digital platform.

Ethics

According to the Swedish Ethical Review Act (SFS 2003:460) [25], research that involves retrieval and handling of sensitive personal data or is likely to cause physical and psychological impact or in other ways harm the subjects is required to undergo ethical review. Data regarding participants' race, ethnic origin,

political opinion, religious conviction and the like, is considered as having a sensitive character.

In the study presented in this article, no ethical approval was required according to the Swedish Ethical Review Act (SFS 2003:460), as we did not ask interviewees about their own health or other sensitive topics (see above what is considered as sensitive data).

Participation was optional, and all participants provided written consent. The participants were informed that they could withdraw at any time, for any reason, or for no reason at all. To guarantee anonymity, no names or places are mentioned in the text. The researchers followed the guidelines on research ethics issued by the Swedish Research Council [26].

Results

Overview

The group of 11 interviewees comprised 1 of the main initiators and developers of the platform, 4 nurses, 3 physicians, 2 managers, 1 psychologist, and 1 rehabilitation coordinator from the health care center. All participants were female, except for 2 of the physicians and the developer. They all provided rich and often highly reflexive accounts of their experiences of, and views about, the digital platform and its implementation. To unpack these issues, we present their experiences and views through the four constructs (coherence, cognitive participation, collective action, and reflexive monitoring) described by May et al [15,17]. We also present the theme of *materiality*, which deviates from NPT constructs.

Coherence: Sense-Making Work HCPs Do When Implementing the Digital Platform

The first construct in NPT is coherence. Coherence or sense-making work refers to the process by which people give meaning and value to a new practice instigated by the utility of new technology and its implementation [15]. The sense-making process starts as soon as the ideas of a new implementation emerge. Thus, people not only make sense of an implementation based on their experience of it but also draw on promises, hopes, worries, and fears in regard to the new technology [27,28]. As a result, before the implementation starts, people already have ideas about what the technology is and how and whether it fits into their worldview, habits, and work. These ideas are not fixed and stable but are part of an iterative and collective sense-making process. The analysis revealed that the interviewees had a shared understanding of the aim of implementing the digital platform. Important coherence-related themes that emerged from the interview data were *beneficial for patients*, *assistance for nurses*, and *as a supplement* (Table 1).

Table 1. Overview: themes that adhered to the normalization process theory (NPT) constructs and the theme that deviated from the NPT.

NPT constructs and subthemes	Excerpts
Coherence	
Beneficial for patients	"That it becomes easily accessible for the patients, that they can contact us when it isn't...when we aren't open, but rather when it suits them...that it is a way for them to reach us" [rehab coordinator 1]
Assistance for nurses	"First, to make the nurses' situation easier—it is onerous for them to sit long days on the phone dealing with a lot of tasks. So it was about giving them good tools for triage and more variety in their work team." [physician 3]
As a supplement	"...the idea is that we should transfer as many tasks as possible, and those tasks that are appropriate...We would like to shift as many as possible over to Wolf, but not everyone can use digital contacts, not everyone has that option, for various reasons. So the telephone will always be with us, that's the way it is." [operations area manager]
Cognitive participation	
Close collaboration	"I think that the staff from Wolf who come to us, it is very...it feels very positive. They are deeply engaged and able to answer questions and so on." [psychologist 1]
Implementation champions	"We have had a physician here who has been positively disposed and sort of pulled us along. And I think that it has been much easier to get underway, given that it has actually been one of the physicians that has been the driving force." [nurse 1]
Patient engagement	"I think it works really well. It's actually been a positive surprise; I thought it would be much more difficult. But I think it has worked well. And the patients, sometimes they have doubts, but once they're used it, they feel very comfortable with it, and often come back to that approach." [nurse 3]
Familiarity	"I think it is very easy and very good. There are...the colours that are present are clear and good. And the dialogue boxes and how you should...one needs...It is easy to understand, it is easy to immerse yourself in it, so I think it's great." [nurse 2]
Collective action	
Parallel work practices	"You have to set aside time for it. It takes time. As long as its entered in our time sheets, which we make sure to do, then it goes well...Yes, but it has gotten more enjoyable in that it's a change from sitting talking on the phone." [nurse 4]
New communication patterns	"There are fewer disturbances for both nurses and physicians, given that asking advice from a colleague involves a push of a button rather than having to run up two flights of stairs and knock on a door, and then they would be busy and you'd have to do it again later. And a ton of similar bother." [physician 3]
Reflexive monitoring	
Ease of use	"Over the years I've been working, there have been good and bad digital systems, so to speak, and they don't always make things easier, but I think that Wolf is one of the systems that actually facilitates the work." [physician 1]
Enhancing teamwork	"Yes, but a little more collaboration, I would have to add. We collaborated very well among ourselves before as well, but it is a bit easier when you can write in the platform and you don't have to hunt for one another physically" [physician 2]
Relief and less distress	"When you are sitting talking on the phone, then...Wolf is not...It is not as burdensome in that way, because you don't have...You have the patients and patient contact, but you have it remotely. And in that way I find that my work environment has become easier, that we have days when you don't have quite so much patient contact, and that is actually restorative. Having them remotely, to be able in some way...Even though it's patient contact, it's not direct patient contact, and that's great for our recovery." [nurse 2]
Flexibility	"It has gotten a little easier compared to otherwise, when I am at home. Then it is indeed...Then I have been able to do some of my job tasks from home rather than having to put them off until I am back in the workplace, so it has made things a little easier for me." [psychologist 1]
Workarounds	"...we are accustomed to not being [laughs] synchronised with all the systems there are, but what happens is that we have to cope with this in a similar manner as we have done with certain systems before. So it's not actually a surprise to us, but it is clear that it adds an extra work step." [care manager]
Deviation from NPT constructs	
Materiality	"Yes, I receive more data. Because in Wolf it is the patients who determine completely what they want to reveal...or what you have on this questionnaire that I assume they are given for various symptoms they have, that's what I think. And then it's guided by the questions you have, so to speak, while with telephone contact I am the one who guides the actual questioning and zeroes in on the problem as quickly as I can, naturally." [nurse 3]

When asked about the expected benefits of using the digital platform, the interviewees commonly mentioned the possibilities of increased patient access, easier contact with and access to patients, less time spent on phone calls, and nurses spending less time trying to reach physicians for advice about patients and their symptoms. The physicians particularly highlighted the expected benefits for nurses: spending less time on the phone with patients, tools for triage, and increased flexibility. The nurses, on the other hand, particularly highlighted the expected benefits for patients—spending less time trying to reach the health care center by phone and having the option of contacting it through the digital platform. The interviewees also had very similar conceptualizations of how the digital platform and its utility differed from existing practices. NPT emphasizes that differentiation is part of the sense-making process, as it relates to understanding how an innovation and its use will produce a new practice and how the new practice differs from the existing practice [15]. The idea of new practices resulting from the use of the digital platform was welcomed by the interviewees. They perceived that the platform would enable them to *relocate* some of the work to patients, who would fill in their own symptoms and medical history, answer the follow-up questions provided automatically based on what the patients had filled in, and upload pictures. The envisaged benefits of using the digital platform include the digital platform collecting and providing detailed information about patients, thus providing a foundation for analysis and decision-making. This, together with the chat and video functionality, was seen as more efficient than current practices in which prospective patients generally contact the Albra primary care center by phone.

Although the value and benefits of using the digital platform were a positive fit with the interviewees' worldview and mindset, it is clear from their responses that they perceived the platform as a supplement to the current practice, not as a replacement for it. Many commented that the digital platform provided new communication possibilities that suited some patients but not all.

Cognitive Participation: Relationship Work That HCPs Do to Build and Sustain a New Practice

It is not enough that people understand and have a coherent view of a new practice. Practice will not change until people take personal responsibility and are committed to implementation. May et al [15] referred to this second construct of NPT as cognitive participation. It concerns individual and shared commitments and engagements in innovation and its implementation [15]. We identified four themes that were prominent for individual and mutual engagement and commitment in the implementation: *close collaboration*, *implementation champions*, *patient engagement*, and *familiarity*, as shown in Table 1.

It is clear from the data that the digital platform was implemented by the management in close collaboration with the HCPs employed at the health care center. The decision to implement it was taken by the management in close collaboration with one of the physicians and some of the nurses, who were all later trained to become *implementation champions*; that is, persons who drive the implementation and assist colleagues when they need help or encounter problems with the platform) [29]. Furthermore, the platform and its utility in practice were regularly discussed and scrutinized, both before and during the implementation, during formal meetings between the management and the employees, and during informal meetings between colleagues. Frequent meetings fostered commitment and engagement in the implementation. Moreover, the implementation was conducted in *close collaboration* with the platform provider. Initially, the platform provider presented a digital platform to the management and a small group of employees. Thereafter, the provider conducted half-day training sessions with the HCPs. Hence, implementation was not left to the management of the primary health care center or the HCPs themselves; instead, it was approached as a joint venture with the platform provider.

Interviewees commonly said that patient engagement with the platform was crucial. According to NPT, legitimation (ie, beliefs and ideas about whether HCPs feel that it is right for them to be involved and whether the new practice corresponds to their core values) is an important mechanism of cognitive participation. Another factor that affected the interviewees' engagement in the use of the platform was familiarity. The *familiarity* of the user interface, as well as prior experience with chat and videoconferencing, affirmed commitment to the implementation as the interviewees had encountered synchronous (video calls) and asynchronous (chat and photo uploads) chats before. The interface and functionalities were thus perceived as familiar and easy to use.

The familiarity of the functionalities, as well as the willingness of HCPs and patients to use the platform, appear to have acted as facilitators for cognitive participation. The interviewees, both individually and together, engaged in the implementation by taking part in the discourse on the positive effects of the digitalization of health care and in feedback loops regarding the platform, which were scheduled at regular meetings. The act of articulating and rearticulating the benefits of using the platform seemed to be an important source of cognitive participation and a possible contributor to creating and maintaining commitment and engagement in implementation.

Collective Action (Enacting Work): Operational Work That People Do to Enact a New Practice

According to NPT, collective action is the third construct of NPT [15]. Collective action refers to how a practice operationalizes (or does not) an innovation. It includes the actual work people do together and in relation to the implementation and the confidence HCPs have as individuals and as a team in the new practice. It also includes how work is outlined and allocated and how innovation is incorporated into institutional and social structures. The dominant themes related to collective action were *parallel work practices* and *new communication patterns* (Table 1).

When asked about how the digital platform was enacted in practice, the interviewees commonly said that there was time scheduled for nurses and physicians to work on the platform. During scheduled *platform time*, both nurses and physicians worked exclusively in the platform and did not take phone calls or meet patients in person. Thus, using the digital platform

became a parallel new work routine alongside the *existing routines*; it became the supplement and enhancement to current practice mentioned during the sense-making work. In most cases, it was perceived as a change for the better. It enabled variation in work tasks and provided a source of respite while meeting the needs of patients at the same time.

Although the interviewees saw the benefits of the new working routines, some also highlighted the disadvantage of parallel work practices as the digital platform added to the complexity of work tasks. However, concerns about the additional complexity because of having one more system to attend to did not seem to have a major negative effect on the use of the platform. The data revealed that the interviewees felt confident in using the platform and in enacting new routines. They felt that they had received proper training and that they could receive help, if needed, from colleagues (implementation champions). Owing to the digital platform, they developed new patterns of communication, both between colleagues and with patients.

The interviewees felt that communication had become more efficient not only between different kinds of HCPs but also with patients. Some patient requests could now be handled through the digital platform, and automated triage, chat, video, and photo uploads enabled HCPs to help patients without having to book physical visits. The ability to use written communication challenged the HCPs' habits, and they needed to develop or fine-tune their skills. This was a work in progress at the time of the interviews. Nevertheless, the interviews revealed that the platform had become embedded in everyday practice and ran in synergy with old routines. The different HCPs worked together with and through the platform. The transition from old routines to new routines using the platform was well-supported by the management, as the employees had scheduled time to use the platform. The HCPs had the same responsibilities as before (eg, the patient's initial contact was with a nurse, who decided whether he or she could attend to the patient's request or whether it should be forwarded to other HCPs). However, they perceived the new system as allowing more efficient communication with colleagues and patients.

Reflexive Monitoring: Appraisal Work HCPs Do to Evaluate and Understand How a New Practice Affects Them and Others

Reflexive monitoring refers to continuous judgments and beliefs regarding innovation and new practices, its impact on collaboration within the health care center, and its impact on service quality and value for individual HCPs and management. Reflexive monitoring also includes whether and how individuals alter or suggest changes to a digital platform to enhance its usefulness [15]. The dominant themes identified in the data that related to reflexive monitoring were *ease of use*, *enhancing teamwork*, *relief and less distress*, *flexibility*, and *workarounds* (Table 1).

Reflexive monitoring is an ongoing process of continuous interaction with other NPT constructs. It is not the *last stage* of the implementation but a continual process that reinforces sense making, commitment, and engagement, as well as how the implementation is operationalized. During the time we conducted the study, the interviewees had several positive appraisals of the digital platform and its effect on their working environment. A prominent theme was the *ease of use.* They did not need to invest much time to understand how to use the platform, which was perceived as easy to use with an intuitive interface. The interviewees not only perceived the platform as *easy to use* and understand but also praised it for enhancing teamwork. Several of them mentioned that the platform enabled and facilitated teamwork between different HCPs. The platform did not *configure* them or negatively alter professional relations with patients or other HCPs; it supported their professional work, social roles, and responsibilities. The narrative of *relief and less stress* was pervasive in nurses' accounts of using the platform.

Another positive appraisal regarding the platform was that it allowed *flexibility*. It facilitates accessibility without geographical boundaries (eg, HCPs can work from home, and patients can *see* HCPs through video calls without physically visiting the health care center). This type of accessibility was important as the implementation was initiated early in the COVID-19 pandemic. The platform was also perceived as enabling flexibility in terms of which patients were requested to handle and when. Nurses could go through the list of patient requests and decide which ones needed urgent attention; which they could handle themselves; and which needed to be forwarded to a physician, psychologist, or other HCP. Another aspect regarding flexibility was that the HCPs did not have to get hold of the patients but could communicate via the platform at any time, whereas the patient could read the message at a time convenient for them. Thus, the platform extended accessibility across time and space.

Although the general view of the platform was positive, the interviewees also emphasized that the new routines required some *workarounds*. For example, the digital platform was not compatible with patients' electronic health care records (EHRs). As a result, the nurses used two screens when working with the digital platform: 1 for the EHR and 1 for the digital platform. They needed to be able to see the patient's medical history in the patient health care record and compare it with the data in the digital platform; they also needed to copy and paste data from the digital platform into the patient health care records. This, in turn, raised questions about how much data could just be copied and pasted from the digital platform or whether summaries were needed of the data that patients contributed by answering automated questions and filling in forms. At the time of our study, most of the interviewees said that they simply copied and pasted data from the digital platform directly into the EHRs. They noted that this approach resulted in the scattering of patient data within patient health care records. The alternative to copying and pasting would be to extract the data collected in the digital platform; however, this would require more work and thus reduce efficiency. However, the copy-and-paste procedure may result in extra time having to be spent comprehending the data in the EHRs and thus may have a negative effect on patient safety in the long run. This was not something that the interviewees mentioned; instead, the workarounds were discussed during frequent feedback loops with colleagues and the management. All were under the

impression that this might be solved in the future or was something they would have to adjust to.

Materiality

The materiality (ie, the characteristics) of the platform and how it shaped and was shaped by the implementation cut across all NPT constructs. In our analysis, we decided to split the *materiality theme* from the NPT constructs as our understanding of NPT is that it gives ontological priority to humans and social structures (focusing on social mechanisms such as coherence, cognitive participation, enactment, and reflexive monitoring), whereas the materiality of the technology and the role and significance of the technology in the implementation are somewhat neglected. May et al [17] described the implementation process as follows:

> *Whenever some new way of thinking, acting, or organising is introduced into a social system of any kind, it is formed as a complex bundle—or better, an "ensemble"—of material and cognitive practices.*

As described by May et al [17], the social and the material are intertwined during the implementation, and it is hard to untangle the bundle. In our analysis, the materiality of the platform did not fit seamlessly into any of the constructs of NPT. Verbeek [30] argues that human-technology relations expand beyond mere functionality and use. To understand how humans and technology coshape experiences and practices, one has to think in terms of hybrids [30]. Here, *hybrid* refers to the way human experience, perception, and practices result from elements of technology and contexts (ie, technology mediates human experience and practices). In our study, the HCPs communicated with each other and with patients through the digital platform, which mediated a different kind of communication (written instead of oral and asynchronous instead of synchronous). The platform was also used to collect data from the patient's input, asking automated follow-up questions based on the patient's answers. Consequently, the platform directed what kind of questions the patient was given, resulting in automated triage. Thus, the HCPs' knowledge of the patients and what kind of care they need was mediated through the platform and its algorithms.

The materiality of the platform also affected how patients contacted the primary health care center. They were free to choose between different kinds of technologies. According to the interviewees, patients used different logic to achieve different goals; for example, patients with sexually transmitted diseases or mental illnesses preferred to seek help through the platform rather than by phone. Some patients made contact both by phone and via the digital platform, as the materiality of the technologies in practice enabled it. This, in turn, highlights that materiality is closely connected to usability and how design features are understood and used. However, Orlikowski [31] has shown in her work that the material and the social cannot be viewed separately as organizations, practices, people, social relations, power relations, digital systems, and so on and are interconnected [31]. In other words, materiality goes beyond focusing on material properties such as design features and interfaces. Furthermore, material properties are not stable entities but are subject to interpretations, reinterpretations, negotiations, and local adaptions [32]. Orlikowski [33] insists that it does not matter if a technology can do *X, Y, or Z*; if people do not perceive it as capable of doing *X, Y, or Z*, they will not use it, and as a result, the technology will not change the way they work [33].

What is interesting is that although the materiality of the digital platform supported certain actions while constraining others, the materiality of the platform and its mediating effect on practice was not questioned by the interviewees, as technological determinism reverberated in their accounts.

Discussion

Principal Findings

We used the constructs of NPT to make sense of our data and generalize the findings. The NPT constructs were visible in the data, and the theory helped us understand the important mechanisms involved while embedding and integrating the digital platform into everyday practice. NPT sheds light on the mechanisms and dynamics of the implementation process in a way that probably would not have been possible with frameworks such as the diffusion of innovations theory and the technology acceptance model or its derivatives [18,20,34]. Our findings indicate that the platform became normalized into everyday primary care routines. The work done before and during the implementation involved several actors and actions. The platform developers, who were clinicians, developed a platform to solve problems they faced when managing patients in everyday care practice. The development of the platform was iterative and involved HCPs throughout the development process. Thus, the problem and potential solutions were anchored in health care practices and easily translated by HCPs. Furthermore, the problem of inefficient patient management that the platform was expected to solve is well-known, both to HCPs and in society. As a result, it was readily accepted by patients as well.

Throughout the implementation, the management involved their employees, the HCPs (1) by initial presentations on and discussion about the digital platform and its potential in practice; (2) after the decision was made to implement the platform, training sessions were run by the platform provider; (3) key individuals were trained as implementation champions; (4) the HCPs had scheduled hours when they could work uninterrupted using the platform; and (5) there was ongoing discussion and reflection on the platform at formal meetings.

The HCPs had an equally important role in embedding the platform into everyday practice: (1) they interpreted digital communication with patients as a meaningful addition to current practice; (2) some took on the role of implementation champions and fostered the involvement and engagement of their colleagues by supporting them in their use of the platform and by an informal discussion of the platform; (3) they enacted the platform in current practice and communicated through the platform with colleagues and patients; (4) they reinforced organizational and social norms and values by praising the system for enhancing teamwork, increasing patient accessibility,

and being a source of respite; and (5) they found workarounds to fit the platform into practice.

We perceived NPT and its constructs as beneficial as it helped to deepen the analysis and move away from identifying barriers to and facilitators of understanding general means for supporting implementations. However, a problem we encountered with using the constructs of NPT was related to presenting the data. As the findings are presented sequentially, one can obtain the impression that the implementation is a clear-cut, linear process. However, it is an ongoing complex process, as pointed out by Nilsen [10]. The constructs of NPT are intertwined and interrelated, and the implementation is a continuous process that reinforces sense making, commitment and engagement, and operationalization and appraisal or configuration [15,17]. Translation studies show that translations are never completed but continuously negotiated and renegotiated. Thus, an initially successful implementation may turn into a failed or abandoned implementation [35]. As the study presented in this paper concerns only the first 6 months of the implementation, it is difficult to know whether the actors will sustain their involvement and engagement with the platform and whether the system will become a sustainable part of everyday care practice.

A perceived drawback with NPT, as we understand it, is that it gives ontological priority to humans and social structures while the materiality of the technology is relegated to the background. However, the materiality of technology should not be neglected as every digital health care innovation designs certain kinds of care and working conditions or at least certain ways of providing care [36]. In our study, the materiality of the platform enabled communication between HCPs and both patients and colleagues across time and space (asynchronous communication without geographical boundaries). The platform required patient input, and the patient was guided through automated questions and forms. As a result, the platform not only stores and transmits information but also transforms the patient data process by automated triaging. Consequently, the patient's experiences of health and sickness are enacted and reproduced through the platform and its algorithms. As such, the platform has performativity and affects the HCPs' understanding of the patient's medical status and, in turn, the care given. Thus, the platform, the patients, the HCPs, and the social contexts they are part of are entangled and coconstitutive of each other [31,37]. In other words, the platform and its materiality (eg, which actions and decisions it enables vs constrains) affect knowledge and power dynamics in primary care and, as a result, shape and mediate relations among HCPs, patients, and the health care center. This, in turn, raises ethical questions such as which norms and values are embedded in the platform and its algorithms. Further research is needed on the ethical dimensions of the platform and the effects of the platform's performativity.

Limitations of the Study

It is difficult to generalize about digital platform implementations in primary care and their effects on HCPs' working conditions from a small-scale exploratory study such as this. However, as the constructs of NPT were used to make sense of our data, we identified reasons why the digital platform became embedded and integrated into the practice we studied and what health care organizations can do to support the implementation of digital health care platforms. Long-term research at multiple sites may be valuable to test and validate our findings, as well as to identify mechanisms promoting or inhibiting sustainable implementation.

As the data collection and analysis were guided by the theoretical framework of NPT, we may have been blinded by the theory and thus failed to identify aspects of implementation that are not described by NPT. In an attempt to avoid this, we first analyzed the data using an inductive approach, searching for themes emerging from the data. After that, we deductively tested whether the NPT constructs were visible in the data.

Conclusions

Changing primary care practices from patient consultations by phone and physical visits to digital communication and automated patient assessments allows for increased patient accessibility across time and space. At the same time, digital solutions in primary care challenge current care practices and affect the working routines of HCPs. In this paper, we have shown how a digital health care platform became embedded and integrated into care practice as the HCPs perceived it as beneficial. They were also actively involved in the implementation, had scheduled hours for working with patients through the platform, and had ongoing support from management and implementation champions. Furthermore, the platform and its materiality (eg, which actions and decisions it enables vs constrains) shaped and mediated relations among HCPs, patients, and the health care center. The findings imply that implementation outcomes for changing work practices do not rely only on microlevel actions and engagements, as shown by the NPT analysis. Implementations also depend on interrelations with macrolevel conditions such as the current COVID-19 pandemic (eg, physical distancing) and the dominant discourse on digitalization as a prerequisite for efficiency and accessibility. This, in turn, circles back and influences microlevel actions and engagements.

Acknowledgments

AFA Insurance has provided funding for the research. The authors would like to thank the individuals and primary health care centers who generously shared their time and experience for this study.

Conflicts of Interest

The research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Penedo FJ, Oswald LB, Kronenfeld JP, Garcia SF, Cella D, Yanez B. The increasing value of eHealth in the delivery of patient-centred cancer care. Lancet Oncol 2020;21(5):e240-e251 [FREE Full text] [doi: 10.1016/S1470-2045(20)30021-8] [Medline: 32359500]
2. Moerenhout T, Devisch I, Cornelis GC. E-health beyond technology: analyzing the paradigm shift that lies beneath. Med Health Care Philos 2018;21(1):31-41. [doi: 10.1007/s11019-017-9780-3] [Medline: 28551772]
3. Westerlund T, Marklund B. Community pharmacy and primary health care in Sweden - at a crossroads. Pharm Pract (Granada) 2020;18(2):1927 [FREE Full text] [doi: 10.18549/PharmPract.2020.2.1927] [Medline: 32477436]
4. Bengtsson U, Kjellgren K, Hallberg I, Lindwall M, Taft C. Improved blood pressure control using an interactive mobile phone support system. J Clin Hypertens (Greenwich) 2016;18(2):101-108 [FREE Full text] [doi: 10.1111/jch.12682] [Medline: 26456490]
5. Erlingsdóttir G, Petersson L, Jonnergård K. A theoretical twist on the transparency of open notes: qualitative analysis of health care professionals' free-text answers. J Med Internet Res 2019;21(9):e14347 [FREE Full text] [doi: 10.2196/14347] [Medline: 31573905]
6. Andreassen M, Hemmingsson H, Boman IL, Danielsson H, Jaarsma T. Feasibility of an intervention for patients with cognitive impairment using an interactive digital calendar with mobile phone reminders (RemindMe) to improve the performance of activities in everyday life. Int J Environ Res Public Health 2020;17(7):2222 [FREE Full text] [doi: 10.3390/ijerph17072222] [Medline: 32224975]
7. Cajander Å, Larusdottir M, Hedström G. The effects of automation of a patient-centric service in primary care on the work engagement and exhaustion of nurses. Qual User Exp 2020;5(1):9. [doi: 10.1007/s41233-020-00038-x]
8. Eldh AC, Sverker A, Bendtsen P, Nilsson E. Health care professionals' experience of a digital tool for patient exchange, anamnesis, and triage in primary care: qualitative study. JMIR Hum Factors 2020;7(4):e21698. [doi: 10.2196/21698] [Medline: 33315014]
9. Entezarjou A, Bolmsjö BB, Calling S, Midlöv P, Milos Nymberg VM. Experiences of digital communication with automated patient interviews and asynchronous chat in Swedish primary care: a qualitative study. BMJ Open 2020;10(7):e036585 [FREE Full text] [doi: 10.1136/bmjopen-2019-036585] [Medline: 32709650]
10. Nilsen P. Making sense of implementation theories, models and frameworks. Implement Sci 2015;10:53 [FREE Full text] [doi: 10.1186/s13012-015-0242-0] [Medline: 25895742]
11. Best A, Greenhalgh T, Lewis S, Saul JE, Carroll S, Bitz J. Large-system transformation in health care: a realist review. Milbank Q 2012;90(3):421-456 [FREE Full text] [doi: 10.1111/j.1468-0009.2012.00670.x] [Medline: 22985277]
12. Greenhalgh T, Papoutsi C. Studying complexity in health services research: desperately seeking an overdue paradigm shift. BMC Med 2018;16(1):95 [FREE Full text] [doi: 10.1186/s12916-018-1089-4] [Medline: 29921272]
13. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci 2009;4:50 [FREE Full text] [doi: 10.1186/1748-5908-4-50] [Medline: 19664226]
14. Tabak RG, Khoong EC, Chambers DA, Brownson RC. Bridging research and practice: models for dissemination and implementation research. Am J Prev Med 2012;43(3):337-350 [FREE Full text] [doi: 10.1016/j.amepre.2012.05.024] [Medline: 22898128]
15. May C, Finch T. Implementing, embedding, and integrating practices: an outline of normalization process theory. Sociology 2009;43(3):535-554. [doi: 10.1177/0038038509103208]
16. May CR, Finch TL, Cornford J, Exley C, Gately C, Kirk S, et al. Integrating telecare for chronic disease management in the community: what needs to be done? BMC Health Serv Res 2011;11:131 [FREE Full text] [doi: 10.1186/1472-6963-11-131] [Medline: 21619596]
17. May C. Towards a general theory of implementation. Implement Sci 2013;8:18 [FREE Full text] [doi: 10.1186/1748-5908-8-18] [Medline: 23406398]
18. Rogers EM. Diffusion of Innovations. 4th Edition. New York: Simon and Schuster; 2010.
19. Putzer GJ, Park Y. Are physicians likely to adopt emerging mobile technologies? Attitudes and innovation factors affecting smartphone use in the Southeastern United States. Perspect Health Inf Manag 2012;9(Spring):1b [FREE Full text] [Medline: 22737094]
20. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989;13(3):319-340. [doi: 10.2307/249008]
21. Holden RJ, Karsh BT. The technology acceptance model: its past and its future in health care. J Biomed Inform 2010;43(1):159-172 [FREE Full text] [doi: 10.1016/j.jbi.2009.07.002] [Medline: 19615467]

22. Murray E, Treweek S, Pope C, MacFarlane A, Ballini L, Dowrick C, et al. Normalisation process theory: a framework for developing, evaluating and implementing complex interventions. BMC Med 2010;8:63 [FREE Full text] [doi: 10.1186/1741-7015-8-63] [Medline: 20961442]
23. O'Reilly P, Lee SH, O'Sullivan M, Cullen W, Kennedy C, MacFarlane A. Assessing the facilitators and barriers of interdisciplinary team working in primary care using normalisation process theory: an integrative review. PLoS One 2017;12(5):e0177026 [FREE Full text] [doi: 10.1371/journal.pone.0177026] [Medline: 28545038]
24. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today 2004;24(2):105-112. [doi: 10.1016/j.nedt.2003.10.001] [Medline: 14769454]
25. Act (2003: 460) on ethical review of research involving humans. Swedish Parliament. 2003. URL: https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/lag-2003460-om-etikprovning-av-forskning-som_sfs-2003-460 [accessed 2022-02-17]
26. Good research practice. Swedish Research Council. Stockholm: Swedish Research Council; 2017. URL: https://www.vr.se/english/analysis/reports/our-reports/2017-08-31-good-research-practice.html [accessed 2021-08-01]
27. Borup M, Brown N, Konrad K, Van Lente H. The sociology of expectations in science and technology. Technol Anal Strateg Manag 2006;18(3-4):285-298. [doi: 10.1080/09537320600777002]
28. Verbeek PP. Moralizing Technology: Understanding and Designing the Morality of Things. Chicago: University of Chicago Press; 2011.
29. Miech EJ, Rattray NA, Flanagan ME, Damschroder L, Schmid AA, Damush TM. Inside help: an integrative review of champions in healthcare-related implementation. SAGE Open Med 2018;6:2050312118773261 [FREE Full text] [doi: 10.1177/2050312118773261] [Medline: 29796266]
30. Verbeek PP. Cover story: beyond interaction: a short introduction to mediation theory. Interactions 2015;22(3):26-31. [doi: 10.1145/2751314]
31. Orlikowski WJ. The sociomateriality of organisational life: considering technology in management research. Cambridge J Econ 2010;34(1):125-141. [doi: 10.1093/cje/bep058]
32. Mol A, Moser I, Pols J. Care: putting practice into theory. In: Mol A, Moser I, Pols J, editors. Care in Practice: On Tinkering in Clinics, Homes and Farms. New York: Columbia University Press; 2010:7-26.
33. Orlikowski WJ. Using technology and constituting structures: a practice lens for studying technology in organizations. Organ Sci 2000;11(4):404-428. [doi: 10.1287/orsc.11.4.404.14600]
34. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425-478. [doi: 10.2307/30036540]
35. Callon M. Some elements of a sociology of translation: domestication of the scallops and the fishermen of St Brieuc bay. Sociol Rev 1984;32(1_suppl):196-233. [doi: 10.1111/j.1467-954x.1984.tb00113.x]
36. Bijker WE, Hughes TP, Pinch T. The Social Construction of Technological Systems, Anniversary Edition: New Directions in the Sociology and History of Technology. Cambridge: MIT Press; 2012.
37. Barad K. Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning. Durham: Duke University Press; 2007.

Abbreviations

EHR: electronic health care record
HCP: health care professional
NPT: normalization process theory

Edited by A Mavragani; submitted 18.05.21; peer-reviewed by C Callahan, H van Marwijk; comments to author 02.10.21; revised version received 21.10.21; accepted 06.01.22; published 22.02.22.

Please cite as:
Frennert S, Erlingsdóttir G, Muhic M, Rydenfält C, Milos Nymberg V, Ekman B
Embedding and Integrating a Digital Patient Management Platform Into Everyday Primary Care Routines: Qualitative Case Study
JMIR Form Res 2022;6(2):e30527
URL: https://formative.jmir.org/2022/2/e30527
doi:10.2196/30527
PMID:35191845

Original Paper

Importance of Getting Enough Sleep and Daily Activity Data to Assess Variability: Longitudinal Observational Study

María Óskarsdóttir[1,2*], PhD; Anna Sigridur Islind[1,2*], PhD; Elias August[2,3], PhD; Erna Sif Arnardóttir[1,2,3,4], PhD; François Patou[5,6], PhD; Anja M Maier[5,7], PhD

[1]Department of Computer Science, Reykjavík University, Reykjavík, Iceland

[2]Reykjavík University Sleep Institute, School of Technology, Reykjavík University, Reykjavík, Iceland

[3]Department of Engineering, Reykjavík University, Reykjavík, Iceland

[4]Internal Medicine Services, Landspitali University Hospital, Reykjavík, Iceland

[5]Department of Technology, Management and Economics, DTU-Technical University of Denmark, Copenhagen, Denmark

[6]Oticon Medical, Copenhagen, Denmark

[7]Department of Design, Manufacturing and Engineering Management, Faculty of Engineering, University of Strathclyde, Glasgow, United Kingdom

[*]these authors contributed equally

Corresponding Author:
María Óskarsdóttir, PhD
Department of Computer Science
Reykjavík University
Menntavegur 1
Reykjavík, 102
Iceland
Phone: 354 5996326
Email: mariaoskars@ru.is

Abstract

Background: The gold standard measurement for recording sleep is polysomnography performed in a hospital environment for 1 night. This requires individuals to sleep with a device and several sensors attached to their face, scalp, and body, which is both cumbersome and expensive. Self-trackers, such as wearable sensors (eg, smartwatch) and nearable sensors (eg, sleep mattress), can measure a broad range of physiological parameters related to free-living sleep conditions; however, the optimal duration of such a self-tracker measurement is not known. For such free-living sleep studies with actigraphy, 3 to 14 days of data collection are typically used.

Objective: The primary goal of this study is to investigate if 3 to 14 days of sleep data collection is sufficient while using self-trackers. The secondary goal is to investigate whether there is a relationship among sleep quality, physical activity, and heart rate. Specifically, we study whether individuals who exhibit similar activity can be clustered together and to what extent the sleep patterns of individuals in relation to seasonality vary.

Methods: Data on sleep, physical activity, and heart rate were collected over 6 months from 54 individuals aged 52 to 86 years. The Withings Aura sleep mattress (nearable; Withings Inc) and Withings Steel HR smartwatch (wearable; Withings Inc) were used. At the individual level, we investigated the consistency of various physical activities and sleep metrics over different time spans to illustrate how sensor data from self-trackers can be used to illuminate trends. We used exploratory data analysis and unsupervised machine learning at both the cohort and individual levels.

Results: Significant variability in standard metrics of sleep quality was found between different periods throughout the study. We showed specifically that to obtain more robust individual assessments of sleep and physical activity patterns through self-trackers, an evaluation period of >3 to 14 days is necessary. In addition, we found seasonal patterns in sleep data related to the changing of the clock for daylight saving time.

Conclusions: We demonstrate that >2 months' worth of self-tracking data are needed to provide a representative summary of daily activity and sleep patterns. By doing so, we challenge the current standard of 3 to 14 days for sleep quality assessment and call for the rethinking of standards when collecting data for research purposes. Seasonal patterns and daylight saving time clock change are also important aspects that need to be taken into consideration when choosing a period for collecting data and designing

studies on sleep. Furthermore, we suggest using self-trackers (wearable and nearable ones) to support longer-term evaluations of sleep and physical activity for research purposes and, possibly, clinical purposes in the future.

(JMIR Form Res 2022;6(2):e31807) doi:10.2196/31807

KEYWORDS

wearable technology; nearable technology; internet of health care things; sleep; Withings; study duration; establishing standards; seasonality; mHealth; digital health

Introduction

Background

Sleep disorders and short sleep durations are some of the main health challenges in current times. Obstructive sleep apnea is one such disorder and is estimated to affect 1 billion adults worldwide [1]. Insomnia, defined as difficulties in initiating or maintaining sleep, outlines another common sleep disorder [2,3]. Short sleep duration, although not a sleep disorder, is also a major risk factor for adverse health effects and death [3-5]. The gold standard measurement setting for clinical assessment of sleep quality and sleep disturbances is the use of polysomnography for 1 night, typically performed in a hospital environment [6]. Polysomnography is expensive and requires individuals to sleep with several sensors attached to their face, scalp, and body, which is cumbersome [7,8]. Furthermore, data from such a study gives no indication on important routine aspects of sleep quality such as the average total sleep time (TST) of individuals, when they normally go to bed and wake up, whether they are affected by seasonal changes, or whether they have insomnia [9]. Instead, to assess free-living sleep conditions, multiple night recordings in the home environment need to be performed [9]. In the medical field, this is typically accomplished by using wrist actigraphy, which involves a small watch-like device with an embedded accelerometer that often also records ambient light conditions and skin temperature [9]. The use of actigraphy is accompanied by a subjective sleep log or a sleep diary. Clinical guidelines recommend that the individual wears the actigraphy for 7 to 14 days; however, 72 hours of recording is generally sufficient to bill for testing in the United States [9]. For research purposes, 5 to 7 days of actigraphy measurements are often used to assess sleep behavior [10]. These data are used to assess, for example, average sleep duration, chronotype (morningness vs eveningness, commonly referred to as *A-type* vs *B-type*), and similar sleep parameters of interest. This type of data can also be used to facilitate the analysis of individual sleep patterns and for clustering purposes to show trends at the group level [11].

Consumer-Grade Self-Tracking Technologies

More recently, consumer-grade self-tracking technologies that facilitate sleep data collection over longer periods have emerged [12]. Wearable technology (wearables) is an umbrella term for body-worn connected sensors [8]. Smartwatches are an example of such wearables and can capture information similar to actigraphy. Often, they collect even a wider range of physiological signals, such as heart rate, skin temperature, and oxygen saturation [13-15]. Other self-tracking technologies are nearable technologies (nearables), which can also be used to monitor physiological signals by close approximation to the body. These are increasingly used in conjunction with wearables in health-related research studies [14-16]. For instance, and relevant to our study, they include connected mattresses to monitor sleep patterns in more detail [17]. In most cases, consumer-grade self-trackers are designed for the general purpose of activity tracking. However, their ability to monitor a broad range of physiological parameters means that they are now seriously being considered as alternatives to medical-grade technology for the monitoring of various clinical conditions [18,19]. In addition, the portability and affordability of these trackers open up opportunities for pursuing clinical research on larger cohorts of participants and for rethinking the implementation of remote monitoring care models in specific patient populations [20].

Recent years have seen a surge of research on sleep with consumer-grade self-trackers. Most of these studies focus on relating measurement from the wearable device to either mental or physical health and sometimes both [21,22]. In a few cases, the duration of data collection varies from days and weeks [23-26] to months and years [27,28]. In addition, large sample sizes obtained from the vast number of people who wear self-trackers in the general population have been leveraged to study and compare sleep patterns by age, gender, and BMI worldwide, as in the work presented by Jonasdottir et al [12]. In terms of duration of data collection, a similar study associated shorter sleep duration and greater variability of sleep duration with increased BMI [28]. Furthermore, the large amount of data collected with self-trackers has encouraged the use of advanced machine learning techniques and deep learning to predict clinical outcomes more robustly [29,31]. Although some studies have taken on the task of observing participants over a longer time span than the gold standard for clinical assessment of sleep quality and sleep disturbances, to our knowledge, only a single study has compared data collected over 1 week with data collected over 2 weeks, concluding that the shorter period is sufficient [24]. Although the sleep research community acknowledges the need for longer periods of data collection with wearable and nearable (nonwearable that is placed near the body) self-trackers, the question of whether the participants should wear self-trackers for a longer time than the gold standard to generate a more insightful portrait of their sleep patterns remains unanswered [11,31].

Aims and Overview

The primary goal of this paper is to investigate whether the time span of 3 to 14 days is sufficient for data collection when performing sleep measurements at home using wearable and nearable sensors. We address the primary goal through the following research question: is 3 to 14 days of data collection sufficient to capture the sleep habits and fluctuations in sleep

patterns of an individual in a reliable way for research purposes? Our secondary goals are to investigate whether there is a relationship between sleep quality, physical activity, and heart rate and whether individuals who exhibit similar activity and sleep patterns in general and in relation to seasonality can be clustered together. We address the secondary goals through the following three research questions:

1. Is there a relationship between sleep quality, physical activity, and heart rate?
2. Can individuals who exhibit similar activity be clustered together in an insightful manner?
3. Are there significant differences between sleep patterns of individuals that are affected by seasonality and daylight saving time (DST) clock changes?

Our a priori hypothesis is that 3 to 14 days' worth of data are neither sufficient to capture a person's sleep habits nor sufficient to observe fluctuations in sleep patterns that might be important for research purposes.

Methods

Data Collection

This study was proposed in the context of the Stanford Medicine X–Digital Health Challenge [32]. It was executed under an ethical waiver from the central Danish National Committee on Health Research Ethics. The participants were recruited through advertisements in 2 local newspapers (*Søndagsavisen Vestegnen* and *Villabyerne*) distributed within Greater Copenhagen in Denmark. A total of 82 adults aged >50 years were screened. The first screening was conducted over the phone. Candidates were then scheduled for a home visit, during which the Montreal Cognitive Assessment test was administered by a trained neuropsychologist. The Montreal Cognitive Assessment scores were collected but are not reported in this paper as it was outside the scope of this study (see [33] for details). Of the 82 individuals, 54 (66%) (aged 52-86 years; male: 35/54, 65%; female: 19/54, 35%) fulfilled the inclusion criteria of the study. All participants signed informed consent to join the study and agreed to share their data. At a second home visit, the participants were equipped with the wearable Withings Steel HR smartwatch (Withings Inc), tracking the number of steps and heart rate on a per-minute basis. Participants were also equipped with the nearable Withings Aura sleep mattress (Withings Inc), tracking the various phases of sleep (sleep onset latency, wake, light sleep, deep sleep, rapid eye movement [REM] sleep, and waking up times) on a per-minute basis [34]. The first day of data logging for the participants spanned from June 7, 2017, to September 25, 2017. Data logging stopped for all participants on December 28, 2017. Figure 1 shows an overview of the days for which data were acquired for all participants in the study. In addition, the participants' age, height, weight, and gender were noted upon entry into the study.

The study is based on data from the abovementioned devices—smartwatch (wearable) and sleep mattress (nearable)—and specifically the variables listed in Table 1. Most of the variables in Table 1 are either measured directly or calculated by the smartwatch during the day or sleep mattress during the night. In addition to those, we derived 2 commonly used variables in sleep research, namely, TST, which is the time in hours from falling asleep until final wake up, and sleep midtime, which is the midtime between falling asleep and final wake up.

Figure 1. Overview of data collection for the participants in the study. The dots indicate dates with measurements. Blue represents periods with complete data from the smartwatch during the day, and red represents data from the sleep mattress during the night.

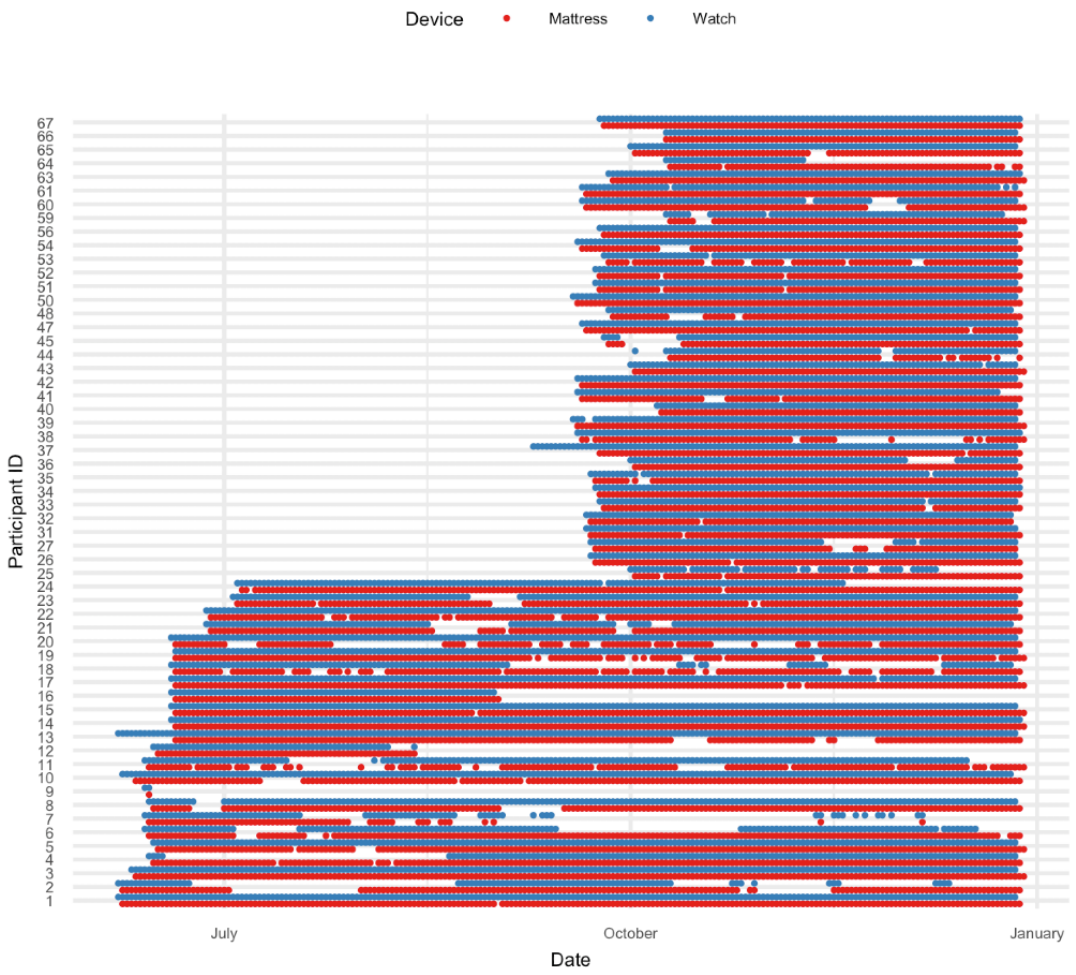


Table 1. Overview and definitions of the variables used in our analysis and the self-trackers used to collect them.

Name	Description	Device
Daily step count	Number of steps during the day	Smartwatch
Diurnal heart rate–average	Mean heart rate during the day	Smartwatch
Nocturnal heart rate–average	Mean heart rate during the night	Sleep mattress
Total duration in bed	Time in hours from going to bed until getting out of bed	Sleep mattress
Total sleep time	Time in hours from falling asleep until final wake up	Sleep mattress
Sleep onset latency	Time in minutes from going to bed until falling asleep	Sleep mattress
Number of times awake	Count of how often the individual woke up during the night	Sleep mattress
Deep sleep duration	Time in hours spent in deep sleep	Sleep mattress
REM[a] sleep duration	Time in hours spent in REM sleep	Sleep mattress
Light sleep duration	Time in hours spent in light sleep	Sleep mattress
Sleep midtime	Midtime between falling asleep and final wake up	Sleep mattress

[a]REM: rapid eye movement.

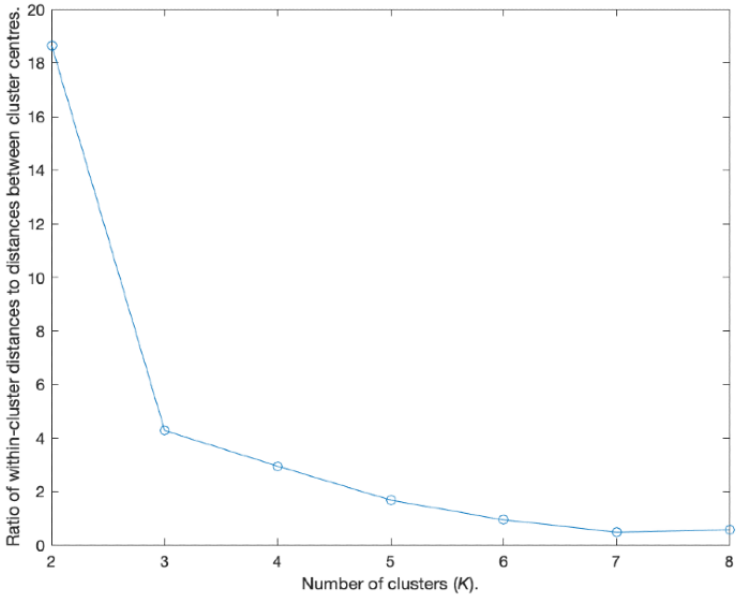
Group-Level Analyses

To inspect long-term changes and variability at the cohort level, we considered the participants for whom data were collected the longest, starting in June 2017 until December 2017. For this part of the cohort (25/54, 46%), we calculated the daily means and SDs of the measurements.

To inspect whether there are discernable patterns in the day-to-day activities of the participants, we used *K*-means clustering that aims to group together similar numerical data, where similarity is defined through the Euclidean distance, particularly, to partition *N* observations into *K* clusters [35]. First, we applied the method using only data from the last week of the trial and then from the last 2 weeks and so on. We decided to do so because of the different starting points and to avoid seasonality effects.

To determine the number of clusters, we used the elbow method [36]. The number of clusters was chosen such that adding an additional one does not increase (much) the information gained. Specifically, we recorded the ratio of within-cluster distances of all clusters to distances between cluster centers and used Figure 2 to determine when it ceases to change (much). This created an *elbow* in the graph at $K=3$, after which not much change occurs. Note that to investigate which variables differed with statistical significance between clusters, we used the 2-sampled Student 2-tailed *t* test with $P<.05$ significance level.

Figure 2. Elbow test.



Individual-Level Analyses

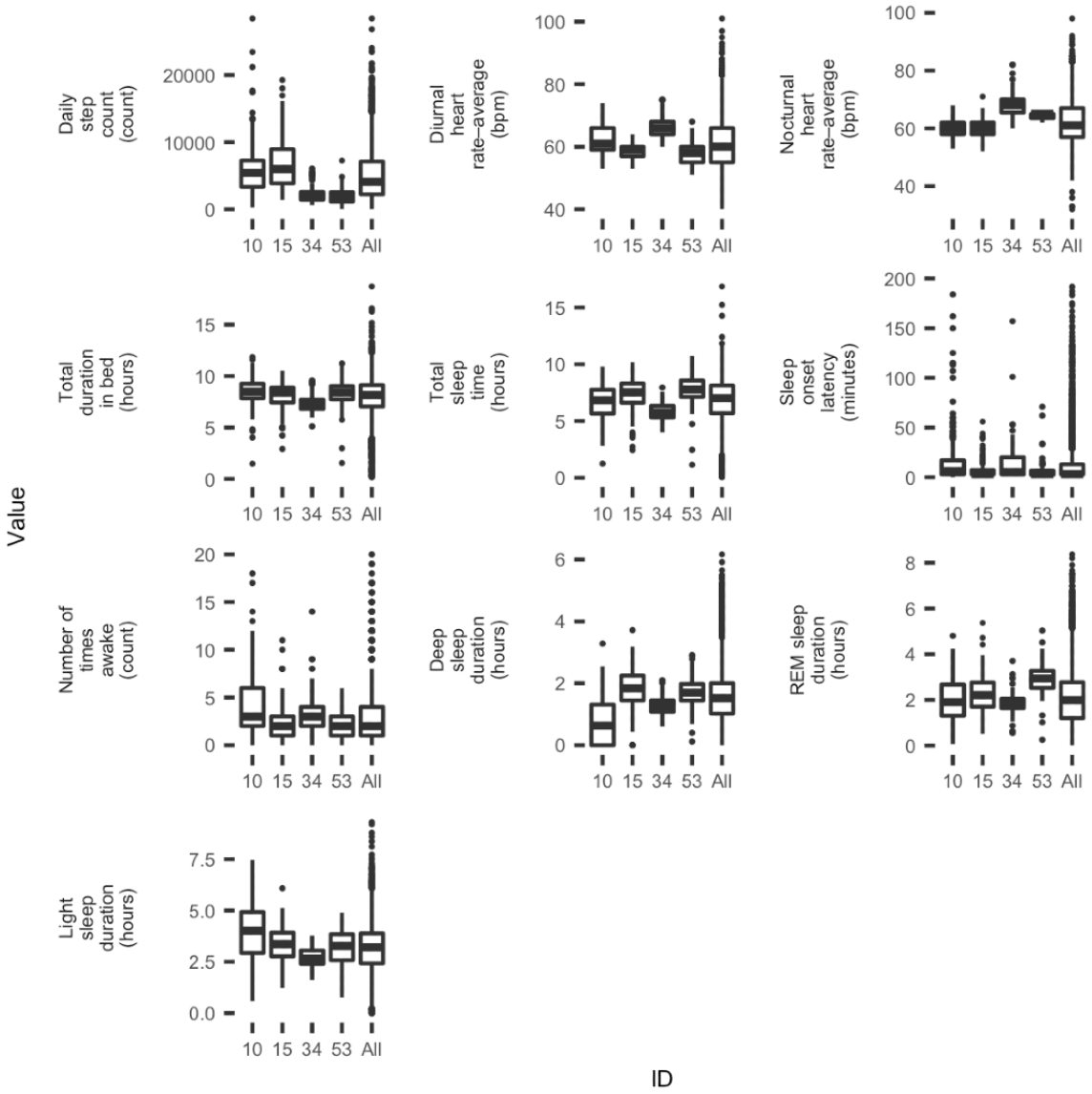
To demonstrate the variability in sleep and daily activity at the individual level, 7% (4/54) of the participants were selected at random and studied in depth. Following the selection, their values were compared with those of other participants. Figure 3 shows the variable distributions of these four participants, which fell within the same range as that of the entire population. These participants were not meant to be a representative sample of the cohort, and the rationale behind our choice to show only data from 7% (4/54) of participants was to clearly demonstrate the variety in measurement patterns among participants without compromising the readability of the figures.

We considered 3 perspectives in the individual-level analyses. First, we assessed the day-to-day values for variables associated with sleep quality for a span of 1 week. The week was chosen at random. Subsequently, we calculated the weekly mean and SD for the different variables for a span of 10 weeks for the same 4 participants, where we normalized variable values for each participant by dividing by the largest value measured in the collection period. These values showed how sleep and daily activity changed from day to day and week to week.

Second, we looked at the evolution of the SD of sleep and activity measurements. We calculated a rolling SD over 7 days with a 1-day moving window from the first week of October 2017 until the end of December 2017. Moreover, starting with the first 3 days of October 2017, we calculated the SD of each participant's measurements. Then, we added the next day and performed the calculation again. We repeated the procedure until 80 days had been added to the original 3 days. Thus, we obtained a sequence of SD values that described the variability in each participant's measurement.

Finally, to investigate seasonal effects and, in particular, the impact of the DST clock change on October 29, 2017, we used a 2-tailed *t* test to evaluate whether differences in the values of each of the 11 variables before and after the DST clock change were significant using a .05 significance level. For this, we considered 3 periods: (1) short-term: 15 days before and 15 days after October 29, 2017; (2) midterm: 30 days before and 30 days after October 29, 2017; and (3) long-term: 60 days before and 60 days after October 29, 2017.

Figure 3. Boxplots showing the distributions for the variables considered (daily step count, diurnal heart rate–average, nocturnal heart rate–average, total duration in bed, total sleep time, sleep onset latency, number of times awake, deep sleep duration, rapid eye movement sleep duration, and light sleep duration) for the four participants and the whole population. As sleep midtime is a circular variable, it is not considered in this figure. bpm: beats per minute; REM: rapid eye movement.



Results

Group-Level Analyses

Some of the participants in the study wore self-trackers for 6 months. This allowed us to look at trends over a longer period and assess seasonal patterns. Figure 4 shows daily means and SDs for TST, total duration in bed, sleep onset latency, sleep midtime, and daily step count. These variables showed the most evidence of seasonal effects. Major trends in the data indicate that total duration in bed increased, albeit the TST remained similar. The sleep onset latency leaped at the end of October 2017, when DST stops in Europe and the clock is set back by 1 hour. Clearly, the participants in this study were affected by this change, as shown by the increased time they took to fall asleep in the weeks after the change of the clock. We also saw a downward trend in the number of steps throughout the 6-month period and fluctuations in the sleep midtime.

Clustering analysis resulted in the suggestion of 2 distinct cohorts of approximately the same size (25 participants each) and a third one that we neglected for its small size when >4 weeks of data were used (Figure 5); the third cluster with 6 participants was omitted. Table 2 shows the mean and SD of all the variables in the two cohorts. The 2-tailed *t* test results showed that only differences in *number of times awake* were statistically significant. On average, the difference between the cohorts in the number of times awake was 1 time. On average, most participants woke up <4 times per night, whereas 6% (3/54) of study participants woke up >5 times per night. We did not find statistically significant differences in gender, age, or BMI between the two clusters.

Figure 4. Seasonal differences in (A) daily step count, (B) total duration in bed, (C) total sleep time, (D) sleep onset latency, and (E) sleep midtime. The blue line indicates the local polynomial regression fit, and the red dashed line indicates the start of daylight saving time on October 29, 2017.

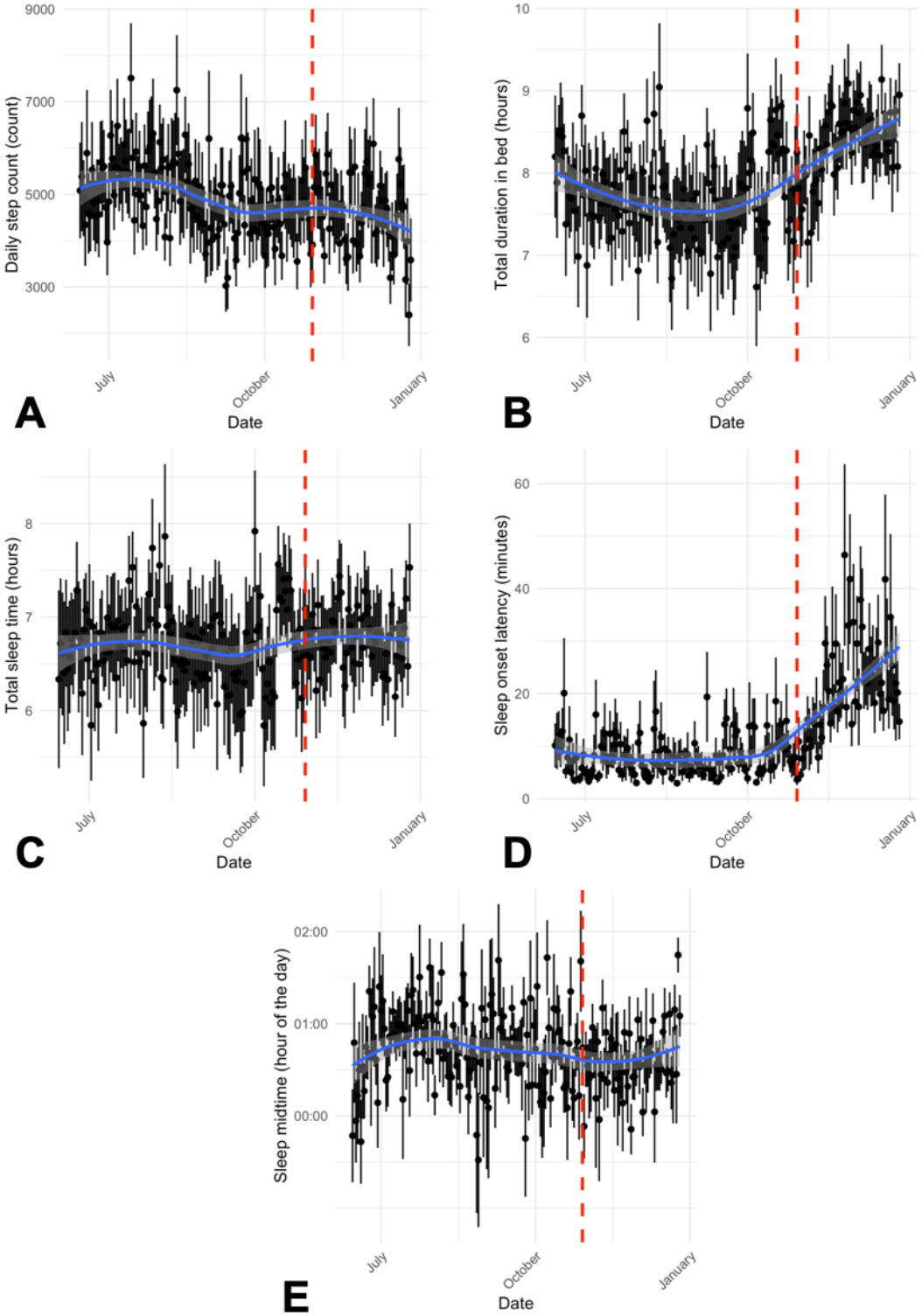


Figure 5. Cluster analysis revealed 2 cohorts (red and blue in the figure). The figure shows a difference in the variable number of times awake between the 2 cohorts. Other variables were less distinctive. REM: rapid eye movement.

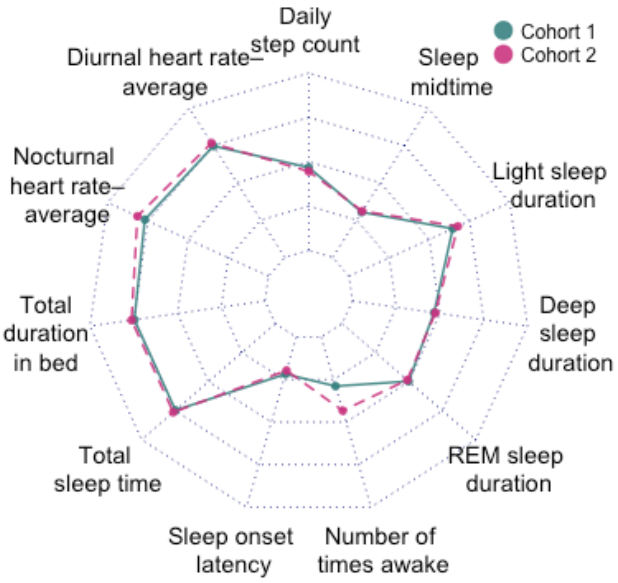


Table 2. Mean and SD of the 11 variables in the 2 cohorts.

Variable	Cohort 1, mean (SD)	Cohort 2, mean (SD)
Daily step count (count)	4895.97 (2772.25)	4996.31 (2230.61)
Diurnal heart rate–average (bpm)	58.89 (5.91)	61.85 (6.67)
Nocturnal heart rate–average (bpm)	61.10 (6.02)	62.66 (6.85)
Total duration in bed (h)	7.81 (0.82)	8.06 (1.24)
Total sleep time (h)	6.86 (0.99)	7.08 (1.29)
Sleep onset latency (min)	10.22 (5.25)	8.70 (4.05)
Number of times awake (count)	2.03 (0.93)[a]	3.06 (1.27)[a]
Duration of REM [b] sleep (h)	1.54 (0.34)	1.53 (0.63)
Duration of deep sleep (h)	2.18 (0.68)	2.23 (0.77)
Duration of light sleep (h)	3.13 (0.56)	3.32 (0.68)
Sleep midtime (time)	11,249.91 (2645.61)	11,493.10 (3465.17)

[a]Indicates that the difference was statistically significant at the .05 confidence level.

[b]REM: rapid eye movement.

Individual-Level Analyses

Figure 6 shows time-series data for 6 variables (total duration in bed, TST, sleep onset latency, sleep midtime, deep sleep duration, and REM sleep duration) collected for 1 week for 7% (4/54) of participants on a night-to-night basis. The remaining 5 variables (daily step count, diurnal heart rate–average, nocturnal heart rate–average, number of times awake, and light sleep duration) can be seen in Figure S1 in Multimedia Appendix 1. The figures show a clear difference for each participant on a day-to-day basis and among the four of them. The relationship between the duration of REM and deep sleep differed for the participants considered here. For participant 34, they were in sync, but not for the remaining participants. Finally, sleep onset latency appeared regular for all 4 participants, and of them, 2 (50%) had days where it peaked.

Figure 7 shows weekly averages over a period of 10 weeks for the same 7% (4/54) of participants and 6 of the variables (total duration in bed, TST, sleep onset latency, sleep midtime, deep sleep duration, and REM sleep duration). The remaining 5

variables (daily step count, diurnal heart rate–average, nocturnal heart rate–average, number of times awake, and light sleep duration) can be seen in Figure S2 in Multimedia Appendix 1. Here, we see that the variation in measurements was even greater. For example, the sleep onset latency of participant 10 was gradually increasing, a pattern that can also be discerned in the total duration in bed and sleep midtime plots. In some weeks, the SD was large, which indicated that the values in those weeks spanned a wide range. The measurements of participant 53 showed stark fluctuation during the 10-week period.

Figure 8 shows the correlation between the weekly averages in Figure 7 for the 7% (4/54) of participants, which varied greatly. Participants 10, 15, and 34 had some positive and negative correlations between their variables. For example, for participant 10, there was a positive correlation between light sleep duration and TST and a negative correlation between light sleep duration and sleep onset latency. In contrast, for participant 53, most of the correlations were strongly positive or negative, showing great synergy. Only sleep onset latency showed little correlation with the other variables.

Figure 6. Daily parameters over a period of 1 week for the 4 participants. Each column and color represent one of the participants. REM: rapid eye movement.

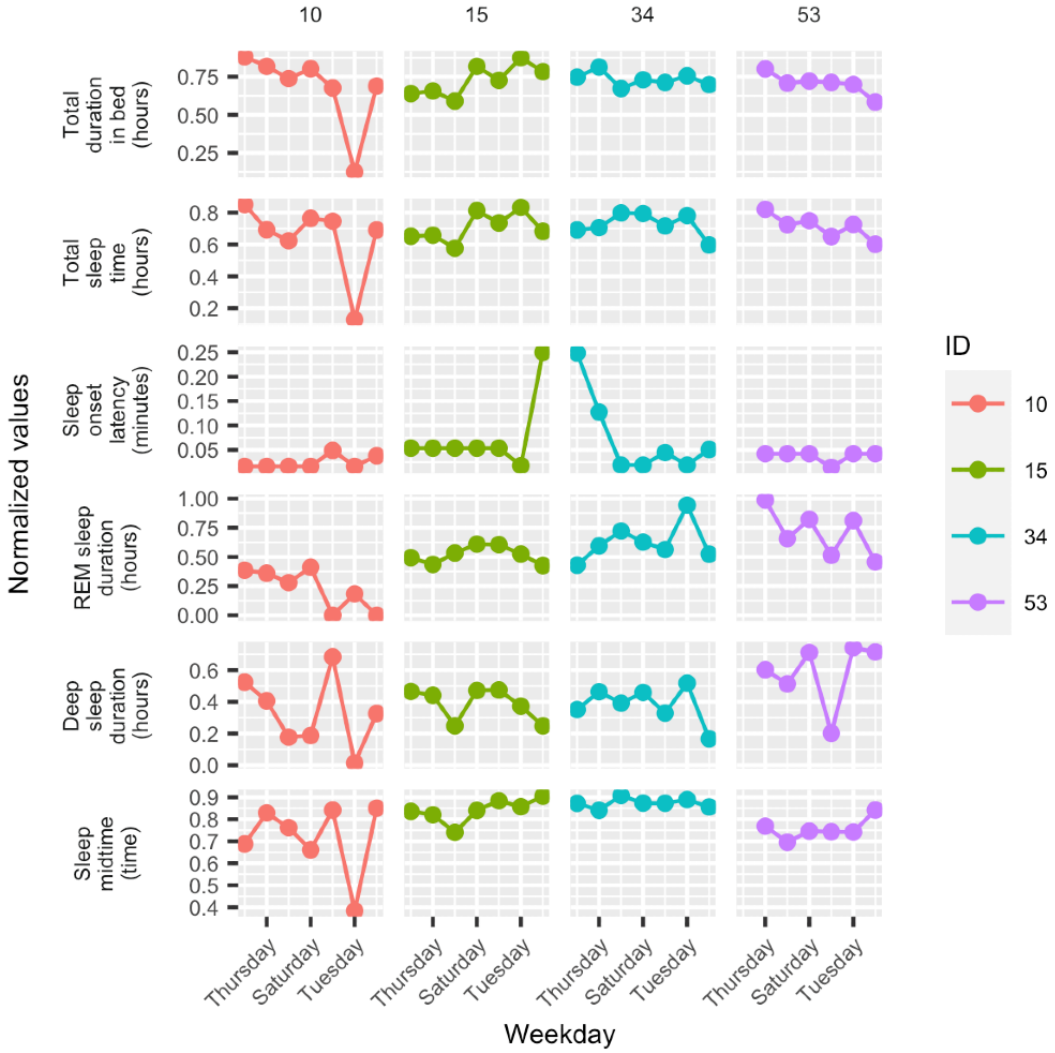


Figure 7. Average activity by week over a 10-week period for the 4 participants. The bars denote the SD within each week. REM: rapid eye movement.

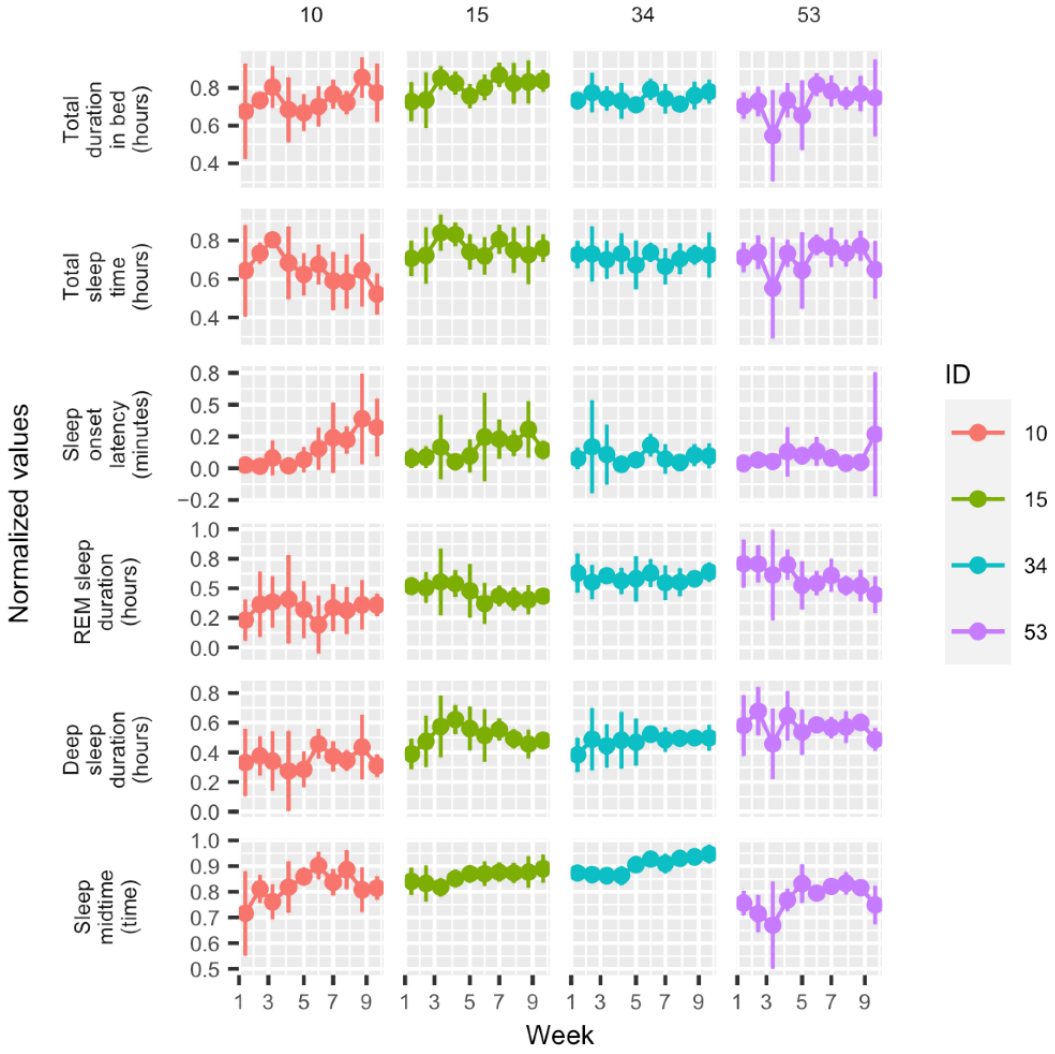
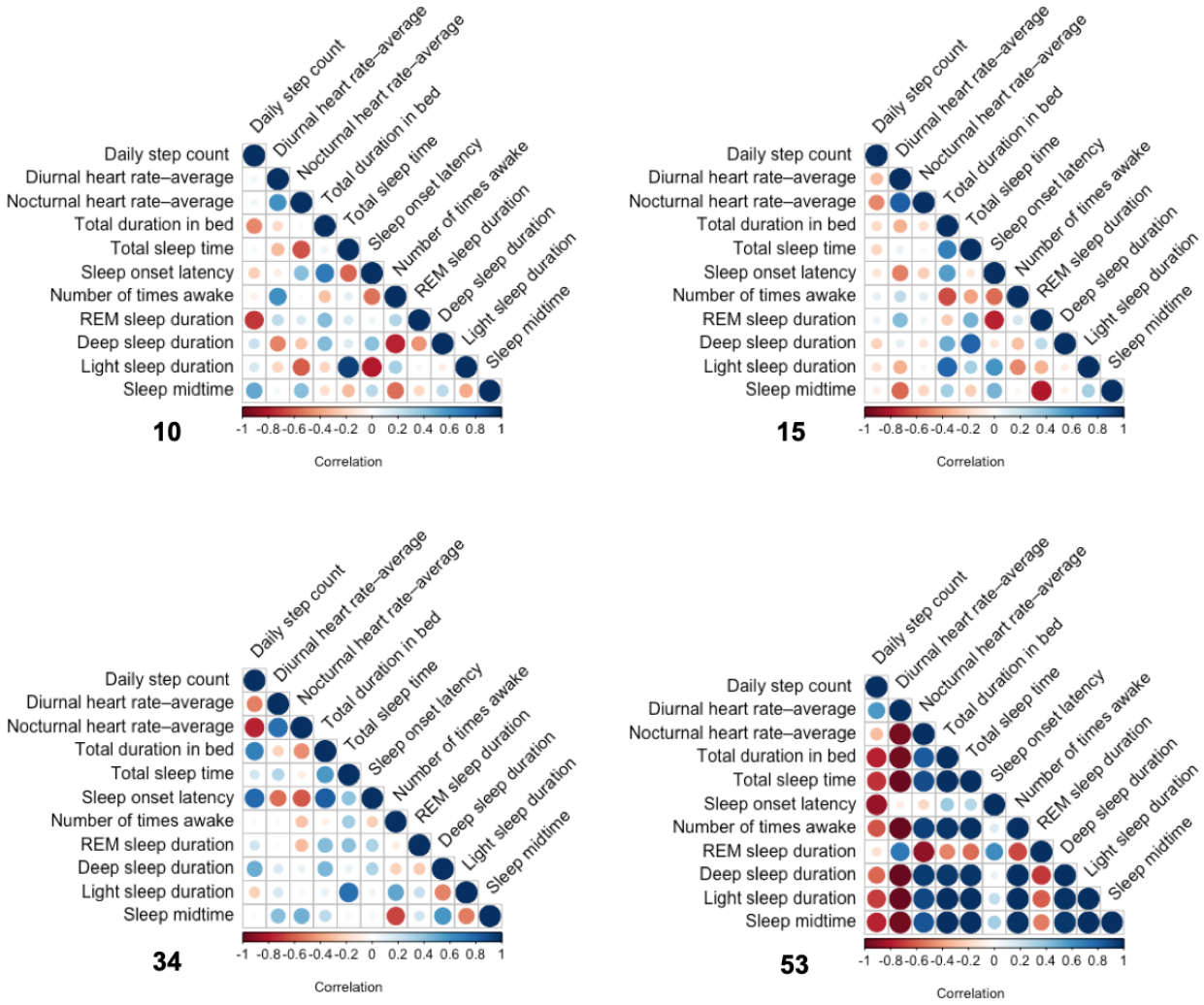


Figure 8. Correlation between the 11 variables for the 4 participants. REM: rapid eye movement.



Next, we investigated the variability in different variables by focusing on the changes in SD over time. Firstly, Figure 9 shows the rolling SD computed over 1 week with a rolling window of 1 day for 6 of the variables (TST, total duration in bed, sleep midtime, deep sleep duration, REM sleep duration, and sleep onset latency). The remaining 5 variables (daily step count, diurnal heart rate–average, nocturnal heart rate–average, number of times awake, and light sleep duration) can be seen in Figure S3 in Multimedia Appendix 1. From the figure, we see that the SD changed greatly throughout time for all participants and for all measurements. Participant 34 had little variation in measurements. The SD of TST and total duration in bed remains within 1 hour. However, in the first weeks, the SD of the sleep onset latency went up to 60 minutes. The other participants had greater fluctuations throughout the period, with SD of TST reaching 3 hours for participants 10 and 53. In addition, the variability in deep and REM sleep duration decreased over time. The data also shows that the variability in sleep onset latency had an increasing trend in the 10-week period.

Finally, Figure 10 shows the cumulative SD of the 7% (4/54) of participants for the same 6 variables. The remaining 5 variables can be seen in Figure S4 in Multimedia Appendix 1. These plots give a sense of the participants' overall variability over time and how it stabilized as more days were added to the data collection. The plots show that 1 week is not representative of someone's sleep behavior as it can change drastically from week to week.

We now assessed the effects of seasonality on sleep and sleep quality. More precisely, we investigated which, if any, of the variables were significantly different before and after the DST clock change when looking at short-term (15 days before and after October 29, 2017), midterm (30 days before and after October 29, 2017), and long-term (60 days before and after October 29, 2017) periods for the 7% (4/54) of participants. Table 3 and Figure 11 show which variables had a significant difference before and after October 29, 2017. The difference in sleep midtime was almost always significant. Also, long-term changes were the most statistically significant, and before the changing of the clock, the participants spent more time in REM sleep, the midtime of their sleep was earlier, and they fell asleep faster.

Figure 9. Rolling SDs over a 10-week period. These were calculated over 7 days with a 1-day rolling window. REM: rapid eye movement.

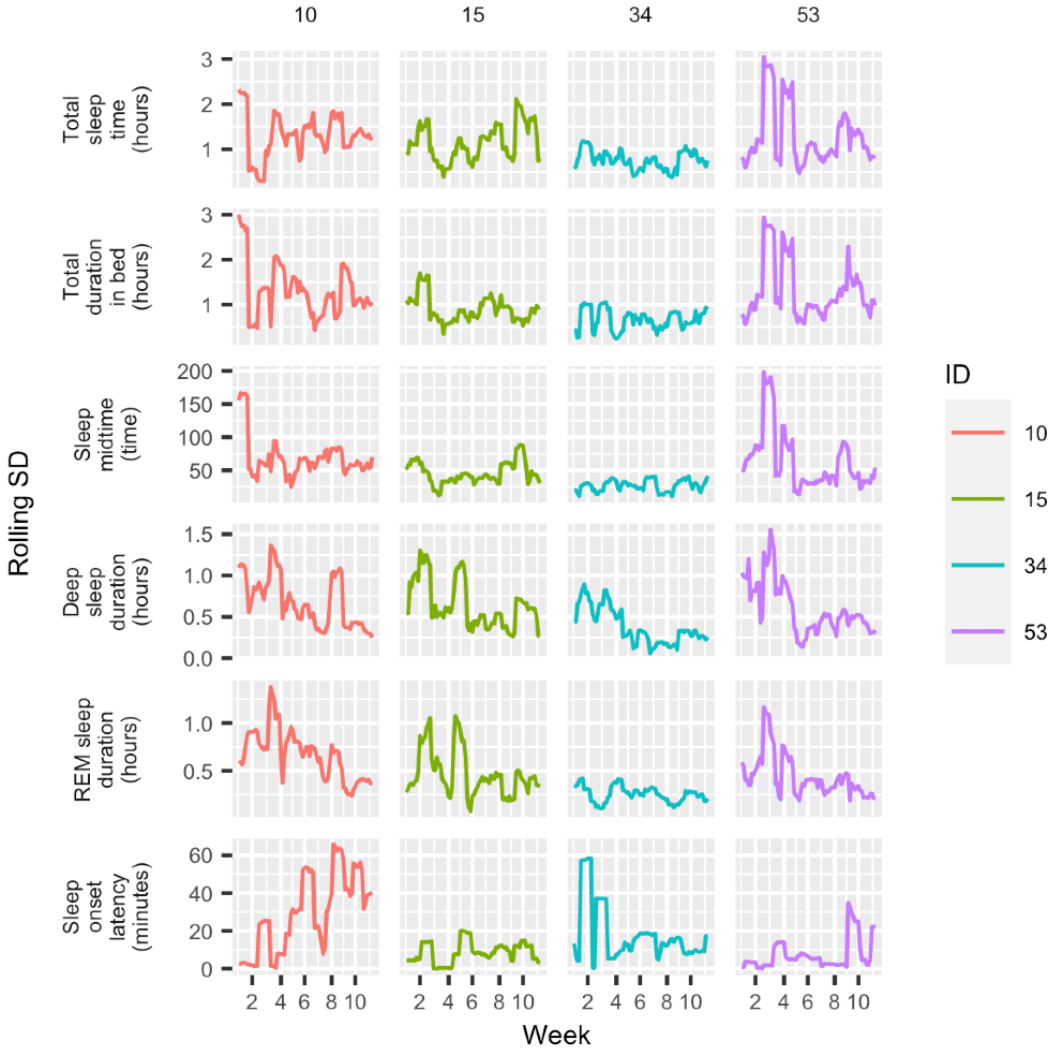


Figure 10. Cumulative SD over a 10-week period, adding 1 day at a time. REM: rapid eye movement.

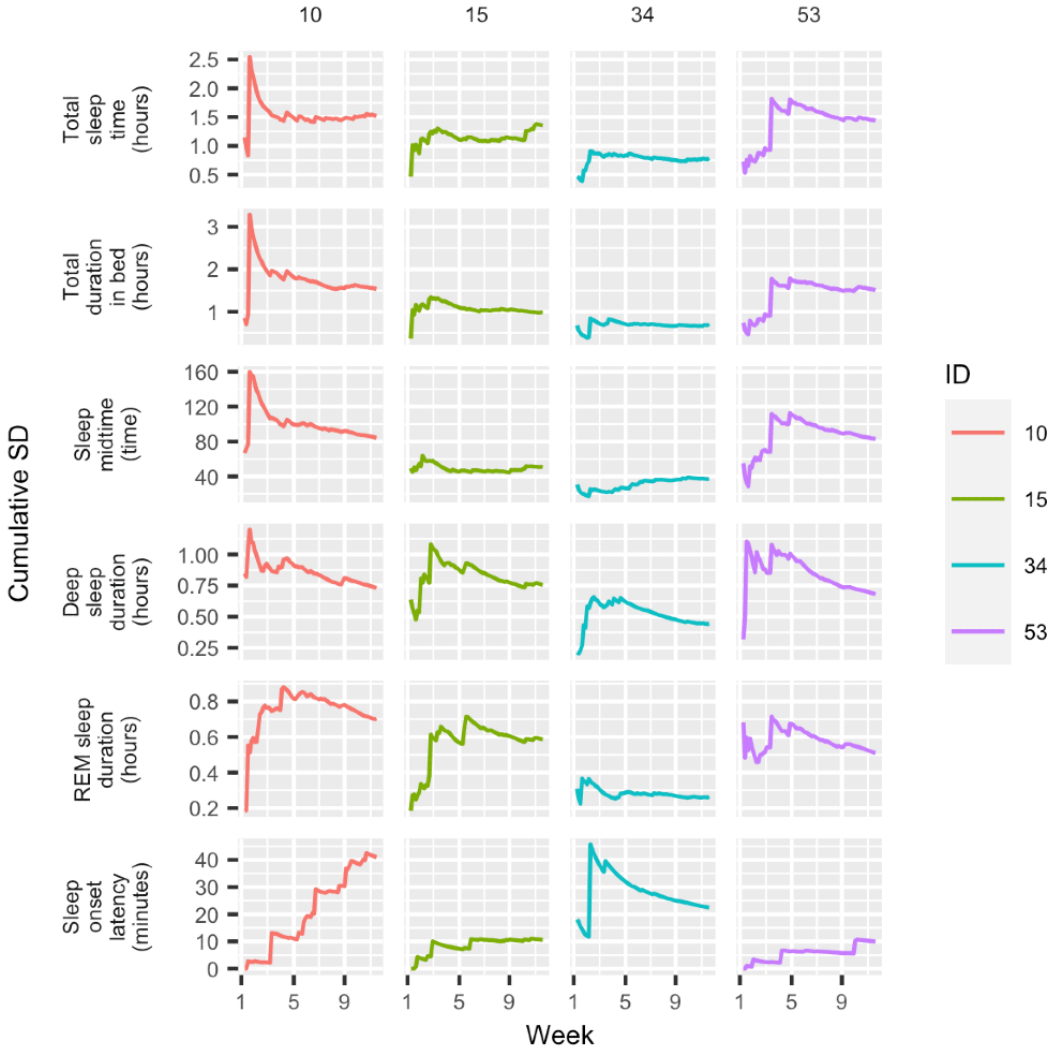


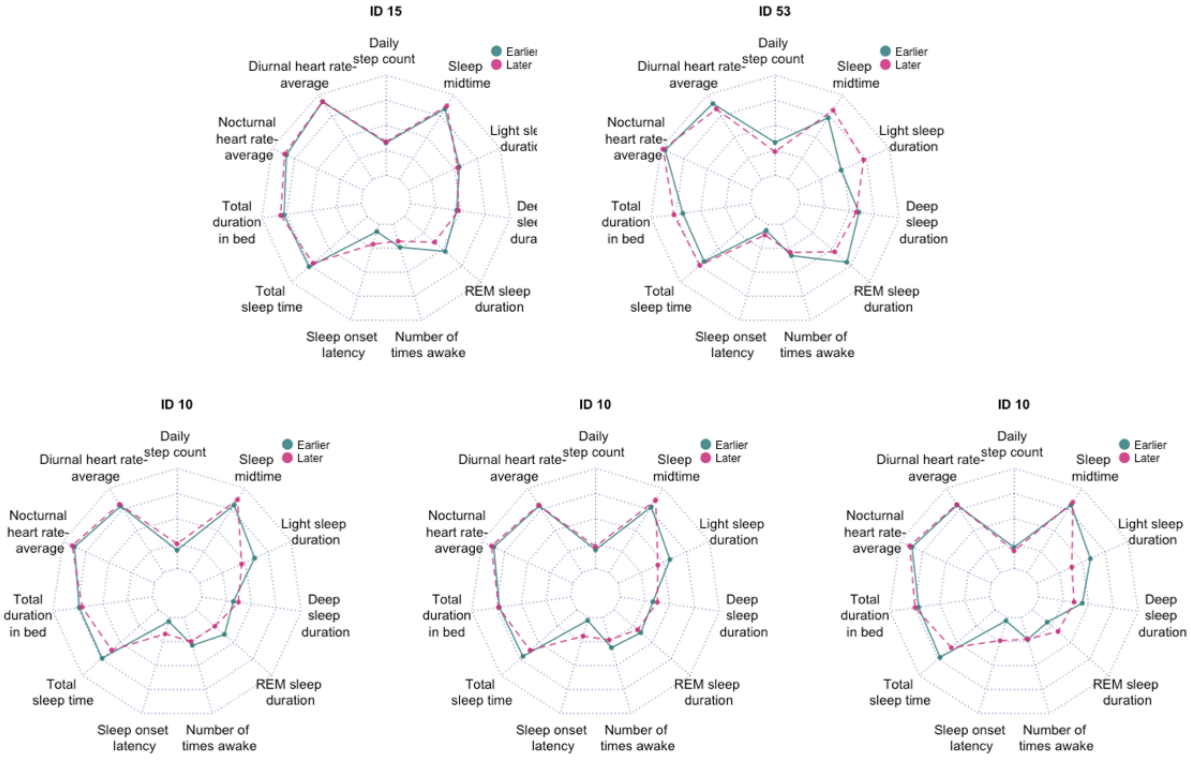
Table 3. Statistical significance of the difference in variables before and after October 29, 2017, for each of the four participants during the three periods.

Variables	Short term				Midterm				Long term			
	ID 10	ID 15	ID 34	ID 53	ID 10	ID 15	ID 34	ID 53	ID 10	ID 15	ID 34	ID 53
Daily step count												✓[a]
Diurnal heart rate–average						✓		✓				✓
Nocturnal heart rate–average					✓				✓	✓		
Total duration in bed												✓
Total sleep time	✓	✓			✓				✓	✓		
Sleep onset latency	✓				✓	✓			✓	✓		
Number of times awake					✓		✓			✓	✓	
REM[b] sleep duration						✓		✓	✓	✓		✓
Deep sleep duration										✓		
Light sleep duration	✓			✓	✓		✓	✓	✓			✓
Sleep midtime	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓

[a]Indicates a statistical significance at the .05 confidence level.

[b]REM: rapid eye movement.

Figure 11. Mean values for the variables before and after October 29, 2017. Top row: long-term patterns for participants 15 and 53. Bottom row: short-term, midterm, and long-term patterns, which represent 15, 30, and 60 days before and after October 29, 2017, respectively, for participant 10. Both participants 15 (P<.0001) and 53 (P<.001) had significantly longer REM sleep before the change. Participant 10 fell asleep sooner, slept longer, and spent more time in light sleep before the change. REM sleep duration also changed from short term to long term. REM: rapid eye movement.



Discussion

Principal Findings

Although the sleep research community welcomes the advancement of consumer-grade self-trackers, including wearables and nearables, they are also widely aware of the numerous challenges that remain, especially regarding the need to validate the devices to ensure their accuracy and reliability [8]. Although many of these barriers are for wearable and nearable technology companies to solve, it is the responsibility of the sleep research community and the medical informatics community to make a collective effort and decide upon necessary and sufficient requirements for validating the devices [31]. As we show in this paper, the duration of the validation period is an essential but grossly overlooked factor. Guillodo et al [11] acknowledged the need for long-term sleep studies, which could help identify connections between sleep quality and health outcomes. However, few attempts have been made with data from wearables and nearables. Although data collected over a longer period is essential, it is also important to make a clear distinction between group-level and individual-level approaches when it comes to research goals, clinical value, and data analysis. Although data collected over an extended period in a large cohort can reveal interesting insights about sleep patterns of the general population [10], there is much potential in using wearables and nearable devices for individualized medicine approaches as well. The approach used in this paper, where we studied individual patterns, has been fruitful for understanding sleep patterns over time.

Another shortcoming in the sleep literature is that it views and analyzes individual nights instead of analyzing time series, where trends, seasonality, and other long terms patterns can be discovered. On that basis, we showed that sleep patterns vary highly from person to person, and, because of that, an individualized approach may be more appropriate than pooling the data per night for several individuals, as is common in the literature. Moreover, we can see that the type of wearable or nearable is not the main value; instead, the main value is in comparing data from the same device, for the same individual, over an extended period. It has been acknowledged that clinical practices should embrace the unique characteristics of individual patients and their patterns and seek to individualize patient care; clearly, the same should hold for sleep [37].

In this paper, our primary research goal was to investigate whether the gold standard, the traditional time span of up to 2 weeks, is sufficient for obtaining reliable data to assess sleep duration and sleep quality of an individual when performing sleep measurements at home using wearable and nearable sensors. Our answer to this question is no. Specifically, we showed that there is much variability in the self-tracker measurements for individual participants across time. Furthermore, in our cohort analysis, we observed a clear distinction in the empirical data only when using sufficient data (>30 days) and could show the emergence of clusters that are robust to changes in the amount of data and the specific dates chosen for the analysis. However, when following individual behavior, an even longer period is needed, and we recommend >2 months.

The secondary research goal of this paper was to investigate whether there was a relationship among sleep quality, physical activity, and heart rate and whether within-group patterns in clusters of individuals exhibit similar activity and sleep patterns, both in general and in relation to seasonality. Our results show a seasonal effect on sleep patterns is related to the changing of the clock. This could both be because of overall seasonal changes and affected by the DST change, which has a significant effect on sleep patterns. This has been acknowledged previously, for instance, by international sleep and biological rhythm societies [38]. We show that there is much variability in the self-tracker measurements and apparent correlation between variables among participants [38].

Conclusions

In conclusion, analysis and exploration of time-series data have given new insights about collecting and analyzing data from self-trackers. The findings in this paper show that it is important to get enough sleep data when attempting to understand sleep patterns from self-trackers in depth. First, the gold standard is less useful as there is much variation in the measures, both on a day-to-day basis and a week-to-week basis. This means that when collecting data on individuals, we recommend a longer period to capture as much of this variability as possible. Second, the variation in the patterns in the data is high from person to person. Although cluster analysis indicates that some patterns seem common among groups of people, our individual observations indicate that the analysis should be conducted on a person-by-person basis by training algorithms to learn individual patterns. Thus, further analysis is needed to investigate the number of days suitable for data collection with self-trackers and whether these patterns and correlations observed are common among groups of people, particularly as our analyses were only based on data from self-trackers and additional information such as illnesses, exercise plans, medication, or medical history were not included. Further limitations of the study include recruitment bias as participants were not randomly selected but were included from a homogeneous sample, and the sample size of 54 individuals affects robust conclusions. The novel finding and call to action of this paper is to reconsider the gold standard in sleep research from 14 days to >3 months. The proposition of this paper is that wearables and nearables make this possible and appear promising for clinical research under free-living sleep conditions, such as at home.

Acknowledgments

The authors thank Withings Inc for providing the health devices—the smartwatch and the sleep mattress. The authors thank all the study participants and collaborators of the Stanford Medicine X–Digital Health Challenge.

Conflicts of Interest

ESA discloses lecture fees from Nox Medical, Philips, and ResMed, outside the scope of the current manuscript.

Multimedia Appendix 1
Additional figures showing trends in daily and weekly averages and weekly trends in cumulative and rolling SD for the five variables not included in the manuscript.
[DOCX File , 978 KB - formative_v6i2e31807_app1.docx]

References

1. Benjafield AV, Ayas NT, Eastwood PR, Heinzer R, Ip MS, Morrell MJ, et al. Estimation of the global prevalence and burden of obstructive sleep apnoea: a literature-based analysis. Lancet Respir Med 2019 Aug;7(8):687-698 [FREE Full text] [doi: 10.1016/S2213-2600(19)30198-5] [Medline: 31300334]
2. Zeng L, Zong Q, Yang Y, Zhang L, Xiang Y, Ng CH, et al. Gender difference in the prevalence of insomnia: a meta-analysis of observational studies. Front Psychiatry 2020;11:577429 [FREE Full text] [doi: 10.3389/fpsyt.2020.577429] [Medline: 33329116]
3. Kocevska D, Lysen TS, Dotinga A, Koopman-Verhoeff ME, Luijk MP, Antypa N, et al. Sleep characteristics across the lifespan in 1.1 million people from the Netherlands, United Kingdom and United States: a systematic review and meta-analysis. Nat Hum Behav 2021 Jan;5(1):113-122. [doi: 10.1038/s41562-020-00965-x] [Medline: 33199855]
4. Cappuccio FP, Cooper D, D'Elia L, Strazzullo P, Miller MA. Sleep duration predicts cardiovascular outcomes: a systematic review and meta-analysis of prospective studies. Eur Heart J 2011 Jun;32(12):1484-1492. [doi: 10.1093/eurheartj/ehr007] [Medline: 21300732]
5. Pienaar PR, Kolbe-Alexander TL, van Mechelen W, Boot CR, Roden LC, Lambert EV, et al. Associations between self-reported sleep duration and mortality in employed individuals: systematic review and meta-analysis. Am J Health Promot 2021 Jul;35(6):853-865. [doi: 10.1177/0890117121992288] [Medline: 33567861]
6. The AASM manual for the scoring of sleep and associated events. American Academy of Sleep Medicine. URL: https://aasm.org/clinical-resources/scoring-manual/ [accessed 2021-12-31]
7. Fischer J, Dogas Z, Bassetti CL, Berg S, Grote L, Jennum P, Executive Committee (EC) of the Assembly of the National Sleep Societies (ANSS), Board of the European Sleep Research Society (ESRS), Regensburg, Germany. Standard procedures for adults in accredited sleep medicine centres in Europe. J Sleep Res 2012 Aug;21(4):357-368 [FREE Full text] [doi: 10.1111/j.1365-2869.2011.00987.x] [Medline: 22133127]
8. Arnardottir ES, Islind AS, Óskarsdóttir M. The future of sleep measurements: a review and perspective. Sleep Med Clin 2021 Sep;16(3):447-464. [doi: 10.1016/j.jsmc.2021.05.004] [Medline: 34325822]
9. Ancoli-Israel S, Martin JL, Blackwell T, Buenaver L, Liu L, Meltzer LJ, et al. The SBSM guide to actigraphy monitoring: clinical and research applications. Behav Sleep Med 2015;13 Suppl 1:S4-38. [doi: 10.1080/15402002.2015.1046356] [Medline: 26273913]
10. Rognvaldsdottir V, Gudmundsdottir SL, Brychta RJ, Hrafnkelsdottir SM, Gestsdottir S, Arngrimsson SA, et al. Sleep deficiency on school days in Icelandic youth, as assessed by wrist accelerometry. Sleep Med 2017 May;33:103-108 [FREE Full text] [doi: 10.1016/j.sleep.2016.12.028] [Medline: 28449887]
11. Guillodo E, Lemey C, Simonnet M, Walter M, Baca-García E, Masetti V, HUGOPSY Network, et al. Clinical applications of mobile health wearable-based sleep monitoring: systematic review. JMIR Mhealth Uhealth 2020 Apr 01;8(4):e10733 [FREE Full text] [doi: 10.2196/10733] [Medline: 32234707]
12. Jonasdottir SS, Minor K, Lehmann S. Gender differences in nighttime sleep patterns and variability across the adult lifespan: a global-scale wearables study. Sleep 2021 Feb 12;44(2):zsaa169. [doi: 10.1093/sleep/zsaa169] [Medline: 32886772]
13. Fox G, Connolly R. Mobile health technology adoption across generations: narrowing the digital divide. Info Systems J 2018 Jan 29;28(6):995-1019. [doi: 10.1111/isj.12179]
14. Mettler T, Wulf J. Physiolytics at the workplace: affordances and constraints of wearables use from an employee's perspective. Info Systems J 2018 Jul 06;29(1):245-273. [doi: 10.1111/isj.12205]
15. Swan M. The quantified self: fundamental disruption in big data science and biological discovery. Big Data 2013 Jun;1(2):85-99. [doi: 10.1089/big.2012.0002] [Medline: 27442063]
16. Lindroth T, Islind AS, Steineck G, Lundin J. From narratives to numbers: data work and patient-generated health data in consultations. Stud Health Technol Inform 2018;247:491-495. [Medline: 29678009]
17. Depner CM, Cheng PC, Devine JK, Khosla S, de Zambotti M, Robillard R, et al. Wearable technologies for developing sleep and circadian biomarkers: a summary of workshop discussions. Sleep 2020 Feb 13;43(2):zsz254 [FREE Full text] [doi: 10.1093/sleep/zsz254] [Medline: 31641776]
18. Hwang Y, Kim MG, Rho J. Understanding Internet of Things (IoT) diffusion: focusing on value configuration of RFID and sensors in business cases (2008–2012). Inf Develop 2016 Jul 09;32(4):969-985. [doi: 10.1177/0266666915578201]
19. Islind AS. Platformization: co-designing digital platforms in practice. EfterCancern. 2018. URL: https://www.researchgate.net/publication/327175455_Platformization_Co-Designing_Digital_Platforms_in_Practice [accessed 2021-12-31]

20. Saganowski S, Kazienko P, Dziezyc M, Jakimów P, Komoszyńska J, Michalska W, et al. Review of consumer wearables in emotion, stress, meditation, sleep, and activity detection and analysis. Emotion Recognition Using Wearables. 2020. URL: https://www.researchgate.net/publication/341118509_Review_of_Consumer_Wearables_in_Emotion_Stress_Meditation_Sleep_and_Activity_Detection_and_Analysis [accessed 2021-12-31]
21. Thorpe JR, Forchhammer BH, Maier AM. Development of a sensor-based behavioral monitoring solution to support dementia care. JMIR Mhealth Uhealth 2019 May 30;7(6):e12013 [FREE Full text] [doi: 10.2196/12013] [Medline: 31199304]
22. Thorpe J, Forchhammer BH, Maier AM. Adapting mobile and wearable technology to provide support and monitoring in rehabilitation for dementia: feasibility case series. JMIR Form Res 2019 Oct 17;3(4):e12346 [FREE Full text] [doi: 10.2196/12346] [Medline: 31625951]
23. Berryhill S, Morton CJ, Dean A, Berryhill A, Provencio-Dean N, Patel SI, et al. Effect of wearables on sleep in healthy individuals: a randomized crossover trial and validation study. J Clin Sleep Med 2020 May 15;16(5):775-783 [FREE Full text] [doi: 10.5664/jcsm.8356] [Medline: 32043961]
24. Briscoe S, Hardy E, Pengo MF, Kosky C, Williams AJ, Hart N, et al. Comparison of 7 versus 14 days wrist actigraphy monitoring in a sleep disorders clinic population. Chronobiol Int 2014 Apr;31(3):356-362. [doi: 10.3109/07420528.2013.858163] [Medline: 24304408]
25. Bradshaw DA, Yanagi MA, Pak ES, Peery TS, Ruff GA. Nightly sleep duration in the 2-week period preceding multiple sleep latency testing. J Clin Sleep Med 2007 Oct 15;3(6):613-619 [FREE Full text] [Medline: 17993043]
26. Melton BF, Buman MP, Vogel RL, Harris BS, Bigham LE. Wearable devices to improve physical activity and sleep: a randomized controlled trial of college-aged African American women. J Black Stud 2016 Jul 27;47(6):610-625. [doi: 10.1177/0021934716653349]
27. Kalmbach DA, Fang Y, Arnedt JT, Cochran AL, Deldin PJ, Kaplin AI, et al. Effects of sleep, physical activity, and shift work on daily mood: a prospective mobile monitoring study of medical interns. J Gen Intern Med 2018 Jun;33(6):914-920 [FREE Full text] [doi: 10.1007/s11606-018-4373-2] [Medline: 29542006]
28. Jaiswal SJ, Quer G, Galarnyk M, Steinhubl SR, Topol EJ, Owens RL. Association of sleep duration and variability with body mass index: sleep measurements in a large US population of wearable sensor users. JAMA Intern Med 2020 Dec 01;180(12):1694-1696 [FREE Full text] [doi: 10.1001/jamainternmed.2020.2834] [Medline: 32926073]
29. Sathyanarayana A, Joty S, Fernandez-Luque L, Ofli F, Srivastava J, Elmagarmid A, et al. Sleep quality prediction from wearable data using deep learning. JMIR Mhealth Uhealth 2016 Nov 04;4(4):e125. [doi: 10.2196/mhealth.6562]
30. Zhai B, Perez-Pozuelo I, Clifton EA, Palotti J, Guan Y. Making sense of sleep: multimodal sleep stage classification in a large, diverse population using movement and cardiac sensing. Proc ACM Interact Mob Wearable Ubiquitous Technol 2020 Jun;4(2):1-33. [doi: 10.1145/3397325]
31. de Zambotti M, Godino JG, Baker FC, Cheung JC, Patrick K, Colrain IM. The boom in wearable technology: cause for alarm or just what is needed to better understand sleep? Sleep 2016 Sep 01;39(9):1761-1762 [FREE Full text] [doi: 10.5665/sleep.6108] [Medline: 27397564]
32. Withings precision research challenge. Stanford Medicine X. URL: https://medicinex.stanford.edu/precision-research-challenge/ [accessed 2021-12-31]
33. Maier AM, Bang MM, Özkil A, Forchhammer BH. Remember to remember: a feasibility study adapting wearable technology to the needs of people aged 65 and over with Mild Cognitive Impairment (MCI) and Alzheimer's dementia. In: Proceedings of the International Conference on Engineering Design. 2015 Presented at: International Conference on Engineering Design; Jul 27–30, 2015; Milan, Italy URL: https://tinyurl.com/2p84k54s
34. Edouard P, Campo D, Bartet P, Yang R, Bruyneel M, Roisman G, et al. Validation of the Withings Sleep Analyzer, an under-the-mattress device for the detection of moderate-severe sleep apnea syndrome. J Clin Sleep Med 2021 Jun 01;17(6):1217-1227. [doi: 10.5664/jcsm.9168] [Medline: 33590821]
35. Bishop CM. Pattern Recognition and Machine Learning: All "just the Facts 101" Material. New Delhi, India: Springer; 2006.
36. Claude J. Morphometrics with R. New York: Springer; 2008.
37. Lillie EO, Patay B, Diamant J, Issell B, Topol EJ, Schork NJ. The n-of-1 clinical trial: the ultimate strategy for individualizing medicine? Per Med 2011 Mar;8(2):161-173 [FREE Full text] [doi: 10.2217/pme.11.7] [Medline: 21695041]
38. To the EU Commission on DST. European Sleep Research Society. URL: https://esrs.eu/wp-content/uploads/2019/03/To_the_EU_Commission_on_DST.pdf [accessed 2021-12-31]

Abbreviations

DST: daylight saving time
REM: rapid eye movement
TST: total sleep time

Edited by A Mavragani; submitted 06.07.21; peer-reviewed by M de Zambotti, M Salimi, Y Zhang; comments to author 25.08.21; revised version received 17.10.21; accepted 28.11.21; published 22.02.22.

Please cite as:
Óskarsdóttir M, Islind AS, August E, Arnardóttir ES, Patou F, Maier AM
Importance of Getting Enough Sleep and Daily Activity Data to Assess Variability: Longitudinal Observational Study
JMIR Form Res 2022;6(2):e31807
URL: https://formative.jmir.org/2022/2/e31807
doi:10.2196/31807
PMID:35191850

Original Paper

Persuasive Technologies and Social Interactions in Professional Environments: Embedded Qualitative Case Study

Barb Marcolin[1*], BCom, PhD; Chad Saunders[2*], PhD; Benoit Aubert[3*], PhD

[1]Faculty of Management, The University of British Columbia, Okanagan Campus, Kelowna, BC, Canada

[2]Haskayne School of Business, University of Calgary, Calgary, AB, Canada

[3]Department of Information Technologies, HEC Montreal, Montreal, QC, Canada

[*]all authors contributed equally

Corresponding Author:
Barb Marcolin, BCom, PhD
Faculty of Management
The University of British Columbia, Okanagan Campus
1137 Alumni Avenue, EME4113
Kelowna, BC, V1V 1V7
Canada
Phone: 1 250 300 9731
Fax: 1 250 807 8533
Email: barb.marcolin@ubc.ca

Abstract

Background: Although previous studies have highlighted the impact of interactions on the web in the context of patient–health care professional (HCP) dyads, this paper extends that context to a triad that includes the role of employers and associated settings with social groups.

Objective: This study aims to evaluate how the interactions between individuals and the social use of the platform affect individuals' use of persuasive technology and, in turn, their work environment actions and responses, by implementing a persuasive technology health and wellness platform in a work environment.

Methods: For 8 months, we deployed a persuasive technology platform with different combinations of health-related features and content in 1 embedded case design with 8 fire stations for a small Canadian city (total number of participant firefighters, n=141) assigned to 1 of 2 treatments—interactive or static webpages. We used text-based content analysis techniques for outcome measures, drawn from a total of 29 participant exit interviews. In addition, medical assessments were conducted at baseline, midpoint, and end point by 7 HCPs and 1 researcher (BM), who also served as the data steward and managed the study.

Results: Our results reveal that group, social, and work influences introduce new elements to the use of persuasive technology, which interact to foster higher levels of individual success. The platform in our study served as part of a larger social system, providing information that facilitated new behaviors at work and home. The 8-month group programs centered on exercise, nutrition, and smoking cessation. Groups of participants coached by certified professionals showed significant increases in sodium awareness, levels of actual exercise, and consistency of activities. As a result of the study, of 141 people, 15 (10.6%) were notified of serious medical health issues and 29 (20.6%) underwent blood work assessments and a privacy shield (protected by federal law) was enacted to protect employees from losing their employment based on any health concerns disclosed.

Conclusions: The persuasive technology platform, in combination with self-management and professional management and social interactions, significantly altered work management behaviors. Interactions among individual outcomes, group influences, and social situations strongly influenced individuals' behaviors in their work and home environments. Three things further improved the positive results that we observed: privacy shields (which allowed employees to reveal health concerns without fear of professional consequences), individual private activities aligned with group activities, and integration between HCP work with localized, organizational work roles.

(JMIR Form Res 2022;6(2):e32613) doi:10.2196/32613

KEYWORDS

persuasive technology; patient experience platforms; group effects; professional work management; services co-design; self-management; health and wellness outcomes; social environments; work influence

Introduction

Background

A wide array of health initiatives rely on the use of information technology to improve health outcomes [1-4]. One area where these technologies show continued promise is facilitating self-monitoring and feedback at work and home with the use of social systems.

Although individual use of persuasive technology platforms has been extensively explored in a health care setting and well summarized in a recent exemplar study [1], far less is known about the use of persuasive technology (1) within a broader work context and (2) in integration with social and professional medical care teams [1,5-8].

Persuasive systems and technologies are generally designed to aid and motivate people to adopt good behaviors [5], while avoiding behaviors that are detrimental to themselves or their community [6]. Other related work can be seen in a recent literature review [9] where many define persuasive systems and technologies as the set of "information technology systems and services developed to change [these] users' attitudes and behaviours or both" [10] for people to use better in organizations. In order for persuasive technology [9] to be most effective and foster healthy lifestyles, users must be in the right mood, time, and place to receive feedback [11].

In outlining underlying principles of their framework for designing persuasive systems, Oinas-Kukkonen and Harjumaa [10] introduced a social component, as well as individual use elements that organizations implement in practice. Similar to Oinas-Kukkonen and Harjumaa [10], we suggest that groups and social interactions play a key role in creating this ideal context for persuasive systems within work practices [12,13].

However, research continues and still focuses mainly on the individual [1]. Accordingly, we aim to address this gap in the literature by expanding the underexplored social aspect of persuasive technology. To do so, we rely on social cognitive theory [14,15] and a sociotechnical perspective to discuss the interconnections between social and technical aspects of improving users' health with persuasive technology in organizational practice [11,16].

A recent publication with experiences from the field of practice [17] notes that people generally perceive benefits and costs when setting specific, measurable, attainable, realistic, and time-bound (SMART) goals in the context of home, work, and social influences. People equate individual level with better health outcomes, where some goals get aligned, and some goals do not. According to Burton-Jones and Grange [18], using a tool or system effectively requires cycling through the concepts of actions, consequences, disturbances, perceptions, goals, comparisons, and feedback, which is consistent with SMART goals; however, it provides more concepts in the iterations to help explain different patterns [19].

Previous research has primarily considered dyads of technology-mediated interactions within practice settings between patients and health care professionals (HCPs) [1]. There are other relevant interactions at play beyond the traditional patient–HCP dyad, notably when a platform is introduced within work environments. In such a context, the patient is interacting not only with HCPs but also with a range of other stakeholders, including insurers, supervisors, colleagues, family, friends, and mentors, where, arguably, platform use can influence all these interactions. We know that information technology can change how people interact at work [20], and when everyone in a work environment can access the same persuasive technology platform [9], it can generate conversations beyond the traditional dyads. The influence of these interactions can either reinforce or discourage system use in practice. We anticipate that these work-related factors will be especially salient when good health is an employment requirement.

Setting

A prime example of a work environment that prioritizes good health is a fire station, which is a practical setting. We use this environment as our research setting for several reasons. First, firefighting requires good physical and mental health, and one might assume that firefighters are generally in excellent health to meet these demands of the work. Firefighters' individual perceptions of benefits and effective use are well described in the literature [18]. However, cardiac deaths account for the largest proportion of deaths of firefighters each year [21], and rates of mental health problems are elevated for Canadian firefighters [22]. Moreover, although fitness and wellness standards have been recommended, only 30% of the US fire departments implement programs targeting these standards. Within a Canadian context, firefighters have higher than average body strength but are comparable with the general population in terms of aerobic capacity [23]. As such, firefighters would likely benefit significantly from using a health platform to improve their aerobic capacity. Perhaps most startling is the fact that firefighters have a 50% chance of developing cancer in their lifetime, possibly because of compounds within the smoke they breathe [19]. Thus, within these work environments, there is a strong need to focus on wellness that is integrated with medical care, family, work, and social systems, which we do in this paper.

Methods

Ethics Approval

The study was approved by the National Research Council – Industrial Research Assistance Program - Research Ethics Board.

Recruitment

We collected data using an embedded qualitative case study design. The embedded case study facilitates the simultaneous consideration for multiple subunits of analysis, which for our

purposes were the individual, group, and organizational. Specifically, we considered 8 fire station organizations, each composed of groups and individuals. Stations were randomly assigned to 2 treatment groups of either interactive webpages or static webpages [24,25]. Data consisted of document review and interview data presented with quotes [24,25]. We gained participation from 94.6% (141/149) of the firefighters.

Firefighters from a small Canadian city were selected because, in an initial meeting, they confirmed their need for increased physical activity and general health improvements, would participate in the study, and would help tailor content to their firefighter terminology; these met our selection criteria. Participation in the study required completion of wellness interventions at various points, willingness to use the platform, and openness to exploring collaborative relationships between leaders (eg, fire chiefs, human resource (HR) managers, union representatives, and mentors) and workers using the persuasive technology platform.

Site Selection

The 1 embedded case study in a city with 8 fire stations allowed assignment of the stations to 2 treatment conditions, which were interactive persuasive technology webpages or static webpages, with half the stations (A, B, C, and D) randomly given interactive pages and the other half the stations (E, F, G, and H) given static pages. Fire stations received different combinations of nutrition, exercise, lifestyle, and medical goal-setting and advice on the persuasive technology platform over the 8-month study period. The research design data shown in Table 1 are best practices seen in the stations using interactive pages and absent in the stations using static pages.

Table 1. Research design for the study treatment program.

Description	Fire stations A, B, C, and D	Fire stations E, F, G, and H
Information sessions, account setup, and chili kickoff meeting (N=149)	77	72
Study treatment	Interactive webpages condition	Static webpages condition
Start baseline medical assessments (n=141)	73	68
End medical assessments (N=110)	57 using the webpages	53 using the webpages
If requested, received baseline prescription (N=53)	26 entered trackers at the beginning	27 entered trackers at the beginning
Interactive dashboard on webpages (N=73)	Interactive SMART[a] goals, to-do action items, encouragements, blogs, RSS[b] latest information feed, individual self-management, and extended access to HCPs[c]	None
Noninteractive dashboards on webpages (N=68)	None	Basic static content, basic self-tracking on paper, and basic information about interacting with persuasive technology platform
Nutrition and sodium awareness[d]	Interactive	Basic
Exercise step[d] (N=19)	Interactive	Basic
Exercise bike[d] (N=19)	Interactive	Basic
Dietitian[d]	Interactive—Canada Food Guide and DASH[e] diet and salt-reduced diet	Basic—Canada Food Guide and DASH diet and salt-reduced diet
Behavior change	Interactive	Basic
Smoking cessation[d] (N=21)	Interactive	Basic
Weight management	Interactive	Basic
Exit interviews (N=29)	Yes	Yes

[a]SMART: specific, measurable, attainable, realistic, and time-bound.

[b]RSS: Really Simple Syndication (web-based XML feed).

[c]HCP: health care professional.

[d]Group activity.

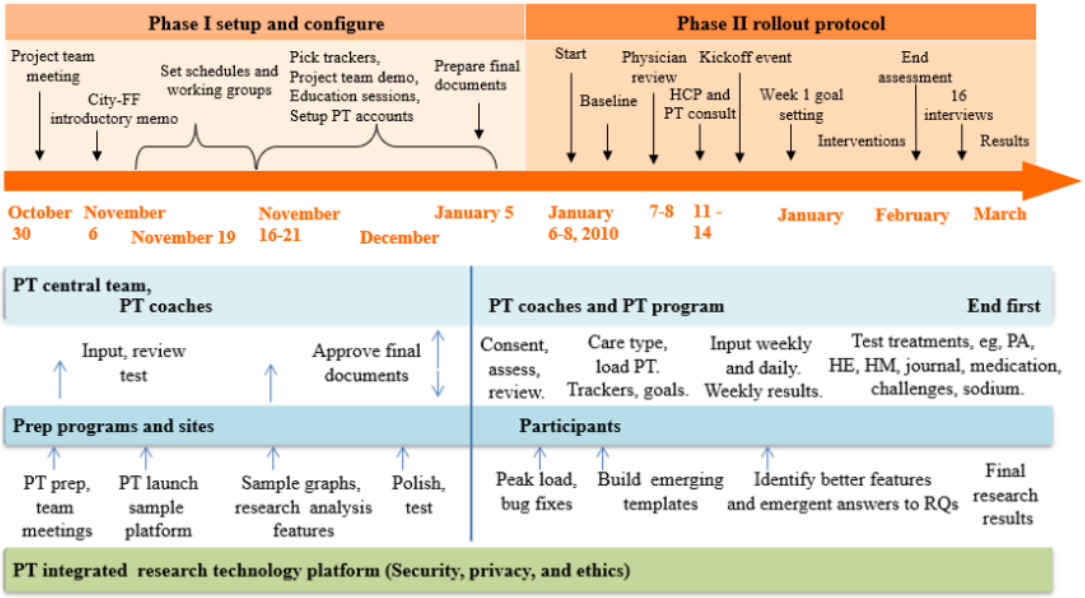
[e]DASH: Dietary Approaches to Stop Hypertension (for diabetics; reduced-sodium).

Timeline and Procedures

The entire study period of 8 months was divided as follows. First, the phase I setup took 4 months, and phase II rollout proceeded over 4 months. It was followed by a maintenance (phase III) for 2 years after the study. As shown in Figure 1, the phase I setup had adaptable planning and approval for clinical practice guidelines that occurred in training and rehearsal of the framework proposed by Oinas-Kukkonen and Harjumaa [10]. The nature of firefighters' tasks allowed us to apply the design

principles identified by Oinas-Kukkonen and Harjumaa [10] to the professionals' tasks in professional roles, which included HCPs, organizational HR professional, firefighters' health expert, firefighters exercise expert, privacy officer, and stop smoking expert. The reduction of concepts (principle 1) fit because professional judgments were reduced to key points and directions based on pretesting, then the tunneling of points was strengthened to focus on ignoring other messages and tunneling to messages that professionals gave (principle 2), and the tailoring (principle 3) and personalization (principle 4) of messages for patients allowed professionals the freedom to be responsible for their decisions, to adapt the content and to create draft program page or pages, documents for weekly issue management, and frequently asked question help sheets in development [10]. The research team participated in pretesting of instruments in a formal rehearsal and reviewed the stations assigned to experimental study combinations.

Figure 1. Timeline of research design events. FF: firefighter; HCP: health care professional; HE: healthy eating; HM: health management; PA: physical activity; PT: persuasive technology; RQ: research question.



The kickoff event for firefighters and HCPs had a 94.6% (141/149 firefighters participated) participation rate and engaged all participants from all 8 fire stations to follow at least one of the persuasive technology platform activities.

Medical assessments generally involved the physician overseeing the nurse capturing medical baselines for all participants, the other HCPs interacting with the firefighters, and the technology team (3/8, 38% to 3/15, 20%) answering questions and getting firefighters signed into their platform accounts.

To conduct the study, in addition to recruiting research participants, we recruited a 25- to 36-person *study team* that included 7 HCPs: 1 emergency room physician, 1 heart program nurse, 2 dietitians, 2 certified personal trainers, and 1 lifestyle mentor. Some of the other professional roles on the study team fell under an organization team (N=10-14), which included 1 union representative, 1 HR representative, and 8 to 12 organizational governance authorities, depending on the volume of activities with firefighters. The study team further included a technology team (N=8-15) of 5 to 12 technology development roles (depending on program pages being loaded), 1 data steward, and 2 HR consultants qualified to deliver the smoking cessation, stress, and general wellness guidelines.

All 141 participating firefighters were offered possible individual goals on a private webpage that individuals would only see if they chose to use the individual page through the persuasive technology platform. Half the stations (4/8, 50%) had interactive pages (A, B, C and D), and half (4/8, 50%) had static basic pages (E, F, G, and H). In the interactive rollout, goals followed the SMART guidelines (ie, they were specific, measurable, attainable, realistic, and time-bound) [22] that were designed to increase behavior change for both individuals and groups. Nutrition information was based on either the reduced-sodium Dietary Approaches to Stop Hypertension *(DASH)* diet for hypertension [26-28] or the Canada Food Guide. The DASH diet is popular for sodium awareness given by the dietitian and heart disease advice given by the nurse. The static basic pages had general messages to stay healthy and eat well. Figure 2 illustrates the order of activities in the study.

Figure 2. Process of goal-setting in phase II (rollout protocol timeline).

All participants (N=141 at the beginning of the study and N=110 at the end of the study, using the webpages) were also offered possible group goals for group exercise step challenges, group exercise bike challenges, general nutrition (eg, reducing sugary drink intake), group sodium awareness, dietitian-recommended meal plans, weight management, and smoking cessation programs. The HCPs and organizational professionals helped firefighters establish overall group goal targets and group weekly wellness tasks via ongoing monitoring and blogging on webpages.

Analysis Framework

We used qualitative case-based analysis [24,25], which included simple tallies in addition to thematic analysis. The data analysis generated evidence about complex interactions among individuals, groups, and work and social settings alongside HCPs, systems experts, and health measures using fun activities.

Thematic coding was implemented using NVivo software (QSR International) [29] to develop themes and identify patterns. The qualitative analysis involved multiple iterations of coding and comparing and compiling results. We created tallies and tables to demonstrate changes. Coding was done by 1 systems researcher and then checked by the technology and organization teams.

Ultimately, we compiled transcribed notes, compared patterns, and determined how many participants had succeeded (or failed) in reaching goals. We used the 11 factors identified by Orji and Moffatt [6] in their persuasive technology classification and analysis coding scheme to structure analysis, as shown in Table 2. Other persuasive technology studies presented these types of evidence.

Table 2. Persuasive technology classification scheme by Orji and Moffatt [6].

Factor	Identification	Description
1	Targeted (health) domain	Outcome changes for physical activity, healthy eating, medical issues, sleep, and hydration
2	Technology platform	Persuasive technology platform integrated web, mobile phone, weight scales, wearables, sensors, and watches
3	Duration of evaluation	Hours, days, weeks, months, and 2 years (a few people)
4	Behavior theories	Social cognitive theories with confidence and ability to do work tasks, goal-setting theory with interventions, self-management theory, performance evaluation theories, and processes improvement change theories
5	Motivation strategies	SMART[a] goals, professional protocols, and intervention strategies; start anywhere, make changes for life, and any improvement matters
6	Evaluation method	Mixed methods comparing qualitative text-based, quantitative small sample size, log analysis, website feature use, and interviews
7	Targeted age group	≥18 years
8	Number of participants	8 fire stations with 141 people
9	Study country	East Coast, Canada
10	Targeted behavioral or psychological outcome	Integration of social cognitive abilities, behaviors, attitudes, adherence, mentoring behaviors, and social interactions
11	Findings or results	Whether successful overall and smaller changes during the paths that the users took

[a]SMART: specific, measurable, attainable, realistic, and time-bound.

Given that this research study relied on health assessments (taken by health professionals) and interview data from participants, we used appropriate methods to analyze the data, notably the qualitative data [30]. Specifically, we used analytic induction, reflexivity, triangulation, and member checks to convey aspects of trustworthiness of the research in terms of dependability, credibility, confirmability, and transferability as ways to assess the quality of qualitative research [31]. As shown in Figure 3, the dependability of the study is based on the logical, traceable, and documented nature of the embedded case design. The credibility of the research is based on using a multidisciplinary research team, drawing on multiple data sources and methods to triangulate the findings, while reaching theoretical saturation for our key themes. We also used member

checks with participants to reflect our interpretations of the situation and whether they were reasonable assessments of what happened. The conformability is an indication of the link between the findings and interpretations to the data, which we have synthesized in Tables 3-6, whereby researcher reflexivity was used to identify themes and relationships among the themes and demonstrate better design principles and practices in persuasive systems and technology, adding to a growing body of related work. The transferability is illustrated by demonstrating similarities and differences between this study and previous work and theoretical and practical contributions made.

Figure 3. Mechanisms implemented to ensure research quality of data and analysis. HCP: health care professional.

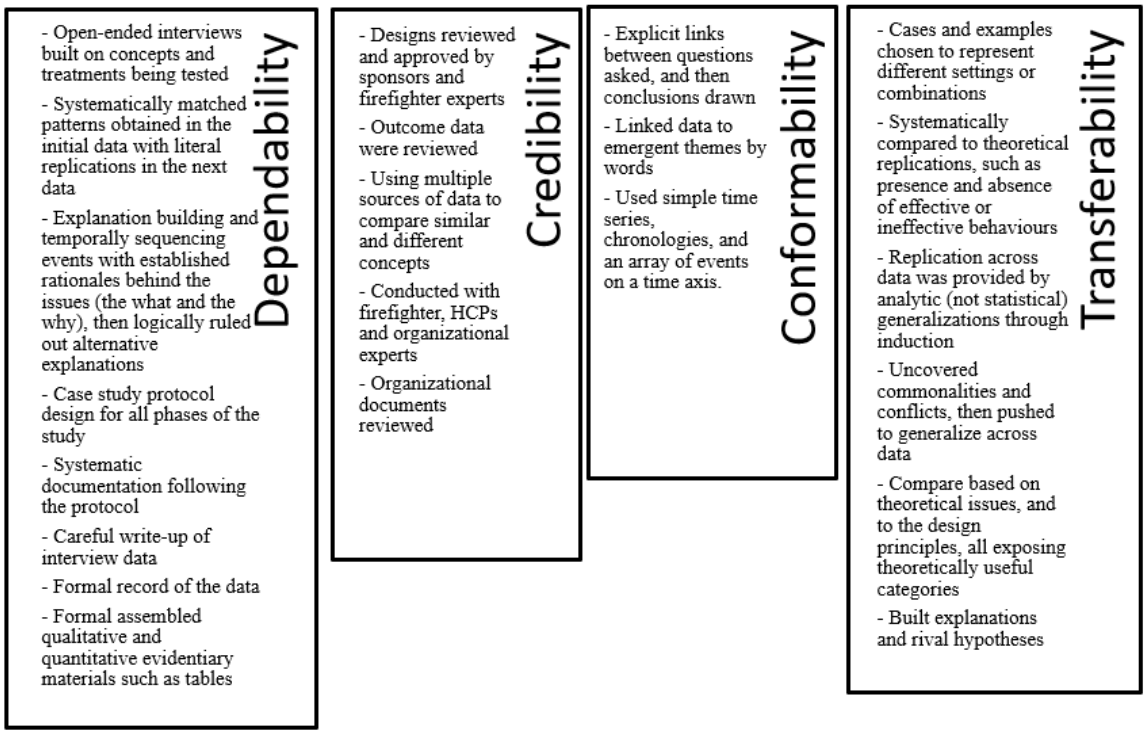


Table 3. Exemplars: operationalization of constructs for primary task support (design principles 1 to 7).

Design principle	Implementation of design principle in this study
Reduction: a system that reduces complex behavior into simple tasks helps users perform the target behavior, and it may increase the benefit-cost ratio of a behavior (principle 1).	• Goal-setting for a wide range of key goals, including nutrition, exercise, and lifestyle changes (sleep, hydration, smoking cessation, weight loss, etc). • Overall project outcomes set at the beginning to have fun, manage medical benefits costs better, improve health and wellness, manage cancer risks from exposure to fires, and gather evidence for work safety. • Provide a private environment for personal discussions to occur. • Access to resources and interactions that firefighters cannot easily coordinate themselves, such as interactive advice with HCPs[a].
Tunneling: using the system to guide users through a process or experience provides opportunities to persuade along the way (principle 2).	• Care plans developed as a result of a medical assessment by a physician and nurse and embedded in the persuasive technology platform to ensure the firefighters received appropriate care plans according to their medical conditions. • The nurse identified firefighters with heart stents and updated the platform and coaches to alter the exercise suggestions.
Tailoring: information provided by the system will be more persuasive if it is tailored to the potential needs, interests, personality, use context, or other factors relevant to a user group (principle 3).	• The programs were tailored by HCPs and other organizational professionals to fit firefighter needs.
Personalization: a system that offers personalized content or services has a greater capability for persuasion (principle 4).	• A comprehensive set of measures were included for individuals to select exercise, nutrition, medical, and lifestyle trackers. The personalized measures could be chosen from exercise (light, moderate, and vigorous converted by the platform), walking, steps, running, strength training (sets, reps, and weight), fruits and vegetables, grain products, meat and alternatives, milk products, total sodium, total fiber, water, sleep, cigarettes per day, blood pressure, 4 cholesterol groups (total cholesterol, LDL[b], HDL[c], and ratio of TC[d]/HDL) and triglycerides, blood glucose, hip and waist measurements, height, weight, and a few other trackers that were not used in this paper (well-being, positive attitude, and medications taken). • Any personalized measure could be refused or turned on at any time. • The platform offered personalized advice based on their specific situation and progress. • Individuals could release their trackers on their individual private webpage or release their trackers into a group challenge through a release feature. Users released with anonymous numbers or full names and could hide or rerelease data at any time in the future when they wanted. They had individual control.
Self-monitoring: a system that helps track one's own performance or status supports in achieving goals (principle 5).	• Self-management controlled by individual firefighters with their own unique username and password; they had individual pages with personal summary information of their trackers (measures), goals set with medical or persuasive technology staff, group challenges to join, or lifestyle changes to make. Individuals could enter data at any time and choose to share their data or keep it private.
Simulation: systems that provide simulations can persuade by enabling them to observe immediately the link between the cause and its effect (principle 6).	• None.
Rehearsal: a system providing means to rehearse a behavior can enable people to change their attitudes or behavior in the real world (principle 7).	• Rehearsal with experts within the firefighters' organizations (eg, union representatives, station mentors, and chiefs) ensured approval and quality before the start, as well as consistency and thoroughness. The researcher prepared and pretested protocols with the professional team to prove the accuracy of the platform operation. • Rehearsal for firefighters occurred with account log-in demonstrations and privacy options during the first interaction—the medical assessment day.

[a]HCP: health care professional.

[b]LDL: low-density lipoprotein.

[c]HDL: high-density lipoprotein.

[d]TC: total cholesterol.

Table 4. Exemplars: operationalization of constructs for dialogue support (design principles 8 to 14).

Design principle	Implementation of design principle in this study
Praise: by offering praise, a system can make users more open to persuasion (principle 8).	• There were weekly check-ins via the persuasive technology platform with HCPs[a] and other organizational professionals who offered praise for accomplishments and encouragement to increase goals. • The platform facilitated "walking buddies" (pairs) who could encourage regular walks and wellness discussions.
Rewards: systems that reward targets may have great persuasive powers (principle 9).	• The platform gave digital rewards, such as trophy icons, on the fly through HCPs (eg, chats and email). Winning teams were listed on leaderboards. • Learning from 1 healthy breakfast offered through the platform to persuade other stations. Dietitians offered improvements around healthy pancake breakfast for a first station and went to stations to illustrate the changes in their cupboards, fridges, and group breakfasts.
Reminders: if a system reminds users of their target behavior, the users will more likely achieve their goals (principle 10).	• To their individual private webpage, users received weekly results and reminders of their individual trackers (measures), goals, and interactions (summarized from users' primary tasks, which they were given by HCPs). • The platform presented group averages, which individual firefighters could review whenever they wanted.
Suggestion: systems offering suggestions at opportune moments will have greater persuasive powers (principle 11).	• The platform produced a professional dashboard with graphs where HCPs reviewed graphs to identify which users needed new suggestions (eg, a personal trainer could post new tailored exercises). • Other suggestions were planned for firefighters who plateaued and stopped progressing as suggested in care plans. In weekly reviews, the HCPs devised new suggestions at the exact time they saw an issue on dashboard graphs. Hence, interventions happened quickly. • HCPs also watched for gaps in the evidence and the absence of questions from firefighters to offer more suggestions. A lack of questions could indicate plateauing, and advisers watched to probe if firefighters wanted more suggestions.
Similarity: people are more readily persuaded through systems that remind (them of) themselves in some meaningful way (principle 12).	• Content tailored to the firefighters' wellness documents, their union's wellness terminology, and experts' suggestions. HCPs and other organizational experts connected terminology to other national standards (eg, hypertension and privacy standards) to the content on webpages.
Liking: a system that is visually attractive for its users is likely to be more persuasive (principle 13).	• The look and feel of the platform were designed to be visually appealing and easy to navigate and used gamification principles where possible.
Social role: if a system adopts a social role, users will more likely use it for persuasive purposes (principle 14).	• Stations mentors, coaches, personal trainers, persuasive technology account helpers, and HCPs took on social roles (principle 14) to advise other firefighters and begin dialogues. Once dialogues started, then the social dialogue management moved into suggesting principle 11.

[a]HCP: health care professional.

Table 5. Exemplars: operationalization of constructs for system credibility support (design principles 15 to 21).

Design principle	Implementation of design principle in this study
Trustworthiness: a system that is viewed as trustworthy (truthful, fair, and unbiased) will have increased powers of persuasion (principle 15).	• Curated content was written by professionals to fit national protocols and approved by firefighter-sanctioned expert groups (ie, nutrition, exercise, and union wellness standards).
Expertise: a system that is viewed as incorporating expertise (knowledge, experience, and competence) will have increased powers of persuasion (principle 16).	• Professional certifications were important to ensure firefighters followed the advice and engaged in the activities. The HCPs[a] also worked primary jobs that lent credibility to their expertise, which included the physician who worked in the local hospital emergency room, the nurse who worked in heart program clinics, the personal trainers who owned popular local gyms, and the dietitians who worked in local health system practices. • The platform was built by the technology company, primarily led by the research director who designed pleasing interfaces, and the persuasive technology was then managed by the developers for back-end technical infrastructure, database, and security. Privacy and confidentiality were managed by the research director. The professionals' content was given to the tech company, which staged the content, rehearsed with professionals, and integrated with study wellness goals. The physician and nurse medical protocols were demonstrated in development-type use cases to be compliant with Health Level 7 tracker and terminology standards.
Surface credibility: people make initial assessments of the system credibility based on a firsthand inspection (principle 17).	• The platform pages were designed to allow trials and demonstrations. The interface, flow, and navigation shown in the initial information session received positive feedback.
Real-world feel: a system that highlights people or organizations behind its content or services will have more credibility (principle 18).	• The platform connected trackers for work and home lifestyle changes to manage the more difficult behavior changes (eg, lack of sleep, overeating, substance abuse, and mental health issues). The platform adapted to requests for individuals or groups to work on something, which included trying to foster family lifestyle changes.
Authority: a system that leverages roles of authority will have enhanced powers of persuasion (principle 19).	• All content and pages were approved by station chiefs, union, city government, and sponsors for funding of the study. Sponsors only received aggregated results and had no access to individual trackers (measures), groups, or comments. Reports were summarized into anonymous personas or group averages.
Third-party endorsements: third-party endorsements, especially from well-known and respected sources, boost perceptions of system credibility (principle 20).	• Platform trackers (measures) and group challenges fit the union representative's expectations, and specific sections in the international standard wellness document were noted on the bottom of webpages.
Verifiability: credibility perceptions will be enhanced if a system makes it easy to verify the accuracy of site content via outside sources (principle 21).	• Presented comparable information to the Canada Food Guide for general nutrition advice and DASH[b] salt-reduced diet suggestions for reduced-sodium diets when required (eg, for patients with heart disease or high-risk participants).

[a]HCP: health care professional.

[b]DASH: Dietary Approaches to Stop Hypertension (for diabetics; reduced-sodium).

Table 6. Exemplars: operationalization of constructs for social support (design principles 22 to 28).

Design principle	Implementation of design principle in this study
Social learning: a person will be more motivated to perform a target behavior if he or she can use a system to observe others performing the behavior (principle 22).	• The persuasive technology platform presented chats on the landing page (eg, chat with a dietitian or whichever professional was logged in at the time). It also presented social walls such as Facebook functionality, which were optional and internal to the software that were used to create a social environment and friend one another, discuss similar interests, and get to know users and their family information.
Social comparison system: users will have a greater motivation to perform the target behavior if they can compare their performance with the performance of others (principle 23).	• The platform presented group averages for step challenge and bike challenge, where comparisons could be made privately on individual viewing pages or shared in group challenge pages where users' data would be visible under the name they chose (real full name, real short name, pseudonym fake name, or anonymous).
Normative influence: a system can leverage normative influence or peer pressure to increase the likelihood that a person will adopt a target behavior (principle 24).	• The platform presented national medical norms within a table on users' individual page, where their personal individual medical assessment and trackers were compared with national norms (by age, gender, and race) for weight, blood pressure, cholesterol, and so on. • Medical assessments were done at baseline and the middle and end of the study. • Individuals could track their progress against averages for the entire sample.
Social facilitation: system users are more likely to perform target behavior if they discern via the system that others are performing the behavior along with them (principle 25).	• The platform presented a "trash talk" discussion board for the bike exercise challenge where firefighters exchanged fun discussion and egged each other on. • Group challenges were done with group workouts in stations.
Cooperation: a system can motivate users to adopt a target attitude or behavior by leveraging human beings' natural drive to cooperate (principle 26).	• All programs were designed to build social comparisons into them with friendly encouragement in cooperative ways and not too competitive. The smoking cessation program paired smokers with a nonsmoking buddy to alter at-work smoking behavior and offer social interaction when they wanted to smoke. This social support worked because smokers indicated they wanted social interaction at work; the evidence shows that smoking is not good for them, but many smoke, anyway.
Competition: a system can motivate users to adopt a target attitude or behavior by leveraging human beings' natural drive to compete (principle 27).	• Six challenges were offered (ie, step, bike, sodium awareness, nutrition generally, smoking cessation, and weight management).
Recognition: by offering public recognition (for an individual or a group), a system can increase the likelihood that a person or group will adopt a target attitude or behavior (principle 28).	• Group challenge winners were celebrated with digital rewards and persuasive technology platform webpage announcements. • Kickoff celebration offered healthy chili recipe.

Results

In this section, we review the results that interact across multiple levels, starting with a brief overview of individual results and then focusing on study results that extend the current understanding by examining groups of people using a platform (ie, examining social and work contexts).

Individual Integrated Self-management Behaviors

Our results for individuals achieved many positive health and wellness impacts similar to those of other studies [1], where individuals improved medical, exercise, nutrition, and lifestyle trackers. In the final interviews (N=29), 51% (15/29) of participants stated that they viewed the persuasive technology platform as a positive catalyst for their wellness behavior change and 37% (11/29) individuals lost an average of 7.3 kg. Although many stated they already had good wellness behaviors, it was unclear what that meant in practice. Notably, the platform had a stronger influence on those who did not have an established routine (ie, an average of 60% score on behavior change measure; N=29) compared with those who did (ie, an average of 44% score on behavior change measure; N=29).

User Journeys

Overview

For each user story, results are generally organized by level of analysis, starting with the individual, then the group, and finally the organization level. Within each level, we first consider key themes and link those to important outcomes that each user story illustrates.

Patterns began to emerge that illustrated differences in individuals' holistic outcomes—that is, outcomes captured at every level of an individual's life (N=141 at the beginning of the study and N=110 at the end of the study, using the webpages). For example, it cannot be considered a positive move if someone stops smoking at work only to take up drinking to manage the stress of quitting cigarettes. Thus, we sought to capture combinations that accounted for such possibilities. In order to convey those fast and flexible combinations, we mapped the 28 design principles identified by Oinas-Kukkonen and Harjumaa [10] onto our constructs, as shown in Tables 3-6.

Among these various user journeys, several achieved better patterns using interactive webpages but also showed that integrated change compared with the design principles needs

to be made carefully with much adaptability to local situations (N=141 at the beginning of the study and N=110 at the end of the study, using the webpages). Qualitative case evidence shown in Tables 3-6 provide data to support these claims that demonstrate the better practices seen under interactive webpages (ie, stations A, B, C, and D), which were absent when compared with patterns under static webpages (ie, stations E, F, G, and H). The evidence data are exemplars of better design principles; they are not exhaustive. Overall, our results showed that systems and platforms yield far better results when integrated with support from society, families, colleagues, social settings, and organizations, which begins some evidence of this integration.

In the user journeys, evaluation of platform use patterns (along the principles described in Tables 3-6) are presented along with outcome evaluations by professionals. All evaluations were performed on a none, poor, fair, good, great, or excellent scale. These show linkages between use of features following the design principles of persuasive technology and outcomes.

Individual User Journey 1: Tom (Poor Platform Use Pattern, Leading to Fair Health Outcomes; Influencers: Vague Wellness Objectives and Few Good Success Concepts; Orthotics Barriers)

One person that the persuasive technology platform helped was Tom, a firefighter in his mid-40s, paid by the city, who had been dealing with some physical ailments that created challenges with exercise. Tom was a prime example of a user who set vague wellness objectives, not as specific as other firefighters who set SMART goals and who achieved higher outcomes. Tom had high computer skills but did not push himself very hard to improve. Sleep apnea was preventing him from getting the proper rest he needed to fuel his 2-job lifestyle (he had a second job in construction on weekends). Time was also a barrier for exercise; however, he remained active at least three days a week. Tom used his pedometer to diligently track and record his own steps, deliberately challenging himself. He watched the group challenge, without actively participating, by logging in to the system as a spectator but was hesitant to join the actual activities. Tom needed new orthotics, and to him, that was a substantial barrier to lower body training and his exercise goal of gaining muscle; this made it impossible for him to track strength training efforts. Because exercise caused joint discomfort, he only walked each day, missing much of the moderate exercise asked of him. He reviewed program software pages and read educational material but did not engage much further when others began to adopt new behaviors. Tom zeroed in on general nutrition information in the software and got a push in the right direction, managing to lose 10 lbs. The biggest barriers for Tom were lack of time and complex physical and medical barriers. Orthotics were ordered in the hope that this would address his lack of exercise at the moderate levels asked by the HCPs.

We coded the user journey of Tom as *poor* according to the primary task support principles (principles 1 to 7) identified by Oinas-Kukkonen and Harjumaa [10] because he did not reduce the concepts well, had little pathway tunneling, had little tailoring (general vague wellness objectives, not specific as SMART goals suggest), had little personalization (focused on broader, general program pages), had some self-monitoring for his own steps but little else (eg, step challenge, sodium reduction, and strength exercises), had no simulation or trying new things, and had little rehearsal with others in the group [10]. Hence, overall, Tom underperformed, even though any positive change could be viewed as a success and was celebrated with him, when comparing across patterns, he missed these many behaviors that others adopted to reach higher or better outcomes as seen in the next journeys.

Group Follower User Journey 2: Harvey (Lurker Not a Participant, Fair Platform Use Pattern, Leading to Fair Health Outcomes: Typical Individual Wellness Patterns but Lacking Significant Improvement; Influencers: Avoiding Group Platform Activities)

The persuasive technology platform also helped Harvey, a user who typified those who followed groups but never participated in group activities. The introduction of groups allowed individuals to join organizational groups or to ask work friends to join them. In this study, they chose to place themselves together in groups (eg, firefighters, union representatives, chiefs, and HR department) [32]. These emergent groups could then challenge other groups and join other groups.

Similar to many other participants, Harvey had a great general nutrition and exercise regime, wanted to get blood work done, and addressed a previous high cholesterol issue. His nutrition had improved before the program when a family member discovered they were hypertensive and was maintained in the average range throughout the program. He was concerned about family wellness and lack of time.

Harvey followed the bike challenge by monitoring the group results on the program page but did not participate. However, he reported not being able to find his medical results, which were on the same program page. Harvey was also looking for additional ways to destress. He had a good grasp on his balanced lifestyle but wanted new ways to incrementally improve his approach. He achieved this by making a small shift in his nutrition, exercise, and sodium awareness, hence leading to minor improvements. He confirmed that his cholesterol levels had shifted within the normal range. We coded the user journey of Harvey overall as *fair* according to the primary task support principles (principles 1 to 7) around broader, general wellness goals identified by Oinas-Kukkonen and Harjumaa [10] because he was only watching the social activities, not participating in or performing the other behaviors, which was principle 22 (social learning), and was mostly not well engaged in activities such as the others were.

Group Challenge User Journey 3: Allen (Excellent Platform Use Pattern, Leading to Excellent Health Outcomes; Influencers: Joined Too Many Groups, but Fun Groups Were His Organizing Mechanism)

Fun group challenges available through the persuasive technology platform included excellent participants such as Allen in the sodium reduction, steps, bike exercise, and general nutrition programs (Table 6; all principles 22-28). Allen did the many group activities described here, as the exemplar excellent group participant. He liked to access HCPs' chats and blogs,

which were well received (Table 5; principles expertise and surface credibility). The group challenges were appreciated because they delivered integrated engaging social, group pages, while allowing individualized medical advice to be seen only on individual private screens (Table 3, primary task [all]; Table 4, rewards; Table 5, real-world feel, not a demo; and Table 6, normative influence and recognition). The personalized individual timing (Table 3; personalization) kept pace with public and fun group activities (Table 3; tailoring) [9].

The overall results reveal an immediate change in shopping behaviors and a 73% change in sodium awareness (N=29), as well as a 59% change in overall nutrition awareness which Allen achieved. Actual nutrition intake improved by 22% on average (N=29 interviewed of 94.6% (141/149) using webpages at the beginning and 73.8% (110/149) using webpages at the end. We purposely combined individual, group, and social concepts that fit the setting to capture the complex work and home interactions that existed.

The program further motivated firefighters such as Allen with a fun kickoff event game called *guess the salt*, which included access to content and features on the persuasive technology platform through direct knowledge from experts and appealing interactive pages. This approach was used to manage the user patterns systematically without too much drill-down that becomes fatiguing for HCPs. The firefighters indicated in the final interviews (N=29) that the game's well-designed graphics helped them change their behaviors and achieve their excellent results. After the kickoff event, the program in parallel conducted tailored and personalized baseline assessments for individuals (Table 3; tailoring and personalization) [33,34] in a private setting close to work activities (N=141 at the beginning of the study and N=110 at the end of the study, using the webpages). This approach helped employees, employers, and professionals improve more easily [7,9].

Group programs [35] were offered around achievable social challenges [3,14,16] chosen by firefighters such as Allen, who showed excellent platform use patterns (eg, exercise, nutrition, sodium awareness, and smoking cessation and led by certified professionals). These groups combined two or more people in to (1) engage the power of group check-in and persuasion, (2) identify a wellness challenge goal or objective that all group members wanted to meet [32-34,36], and (3) pick specific measures [33,36] that participants could record on paper and in the persuasive technology platform, which was then converted into the shared effort on the platform.

The program also had weekly behaviors [35] that firefighters, including Allen, recorded for weekly group results charting, chatting between people or leader competitions. It set end points 6-8 weeks from the start to give closure and motivation. End celebrations [36] included digital rewards or physical prizes from the organization. The maintenance phase III continued for 2 years to identify sustainable behaviors for life [10,17] beyond the typical 2-month falling-off point when participants stop good behaviors and go back to their starting point. In parallel and in an integrated manner, the program managed professional oversight dashboards [6] to deliver program information [5] and identify when an intervention was needed to the few exceptions by the professionals' known patterns of participation.

Group nutrition behaviors were shaped by program design options such as the fun sodium and nutrition challenges, which involved all of the participants attending a dietitian-led interactive sodium awareness lecture that Allen enjoyed (N=149), followed by discussion, label reading introduction, and 2 months with features on the persuasive technology platform (N=141 at the beginning of the study and N=110 at the end of the study, using the webpages). The dietitians advised compliance with the Canada Food Guide (69/141, 48.9%) and Dietary Approach to Stop Hypertension low-sodium diet (72/141, 51.1%; numbers provided by a dietitian) with identified actionable behaviors such as *your next trip to the grocery store* and immediate ways to remove unwanted sodium.

In general, group exercise awareness increased by 55% (average 5.5/10 ratings for 19/29, 63%), and the most frequent comment from participants such as Allen was that they were "more consistent in filling the periods of inactivity with exercise" (quote from the interview). Among those who participated in the group step and bike challenges, 65% (19/29) of the individuals reported results from their participation [23].

We observed evidence of exercise avoidance among the moderate to low exercisers who did investigate exercise and nutrition information via the platform but did not push themselves hard or consistently. We defined low to moderate exercisers as those who exercised, on average, 3 days a week, for 1.10 hours at a time, at a 4.46 intensity level (lower), and within more general workouts. These are much lower than the intense exercisers, who we defined as those who exercised, on average, 4.23 days per week, for 1.09 hours at a time, and at a 7.32 intensity level (high). Our findings suggest that persuasive technology platforms should offer a wide variety of choices to sway moderate to low exercisers.

Distrusted Organization User Journey 4: David (Good Platform Use Pattern—Once Trust Issues Were Addressed and There Were Good Health Outcomes, He Stopped Smoking; Influencers: Required Privacy Shield to Continue in the Smoking Cessation Program Group Because He Distrusted the Organization Work Environment for Demonstrated Reasons; Well Organized by Nonsmoking Buddy at His Fire Station Work Environment)

Overview

The *stop smoking* program enlisted 21 firefighters, including David, at the initial assessment (evidence counted), which resulted in good outcomes for David but tempered by distrust of the work organization. These participants chose to conduct the programs on the platform as a group social challenge. However, when it came time to participate in the stop smoking program, only 62% (13/21) of the firefighters confirmed their participation. David was one of the participants who might have backed out, so we use him as an example of distrust toward the organization and to explore what it took to get him through.

Distrust in the Work Environment

Many firefighters, including David, hesitated to use employer wellness resources and medical benefits that would reveal that they smoked, as they believed that would allow their employer to classify them as individuals who engaged in unhealthy behaviors and even fire them. The firefighters reported stories about other firefighters losing a promotion or being let go for smoking.

As related evidence, no firefighter took up the offer from the HR manager to cover CAD $150 (US $118.5) of nicotine patches as a medical benefits payment. The distrust of HR by David resulted in concealing unhealthy behaviors from his employer, as he believed, such as many of the firefighters, that this exposure would threaten his career. Thus, one perceived downside of incorporating health platforms at work is a worry that health conditions could be misinterpreted by employers as weaknesses on the job, hindering or halting career development.

Privacy Shield Design Element

The technology company quickly developed a *privacy shield* to address the distrust by David and to allow all firefighters confidential access to the smoking cessation resources. The firefighters also confirmed the need for anonymity as important.

Credible Book or Information Source

As a consequence of all these forces, the smoking cessation program needed credible content behind the privacy shield to be successful (Table 5; principle of surface credibility), which David appreciated. The first component was the *Easy Way to Stop Smoking* book, chosen by a third-party consultant to ensure it was appropriate for an organizational environment. The persuasive technology platform anonymously averaged the number of cigarettes not smoked for the group.

Nonsmoking Social Buddy System

The other component that contributed to the success of the smoking cessation program was introducing each user to a nonsmoking coworker (ie, social nonsmoking buddy), which we propose further reinforced the social group effect and a displacement effect with healthier wellness behaviors. Several firefighters reported they stopped smoking on their own or with the help of the family physician after discussions that took place in the program. Notably, however, they went directly to a physician, not using the patch offered by HR—an indication of their ongoing distrust of the organizational work environment.

Valued Discussions in the Workplace User Journey 5: Jack (Fair Platform Use Pattern, Leading to Great Health Outcomes; Influencers: Access to HCP Advice but Not Group Activities)

Overview

Rich and complex work-related discussions occurred that highlighted the challenges associated with the pursuit of health initiatives in work environments (N=141 at the beginning of the study and N=110 at the end of the study, using the webpages; N=29 interviewed at the end of the study). In some cases, when health issues are directly related to the work environment, even the narrow use of the platform can lead to significant improvements. One particular work-integrated medical assessment outcome discussion between the nurse and one of the firefighters, Jack, is used as an illustration. We use the term *great work-related discussion* to encompass those rich settings beyond the initial, typical discussions between the HCPs and participants (detailed further in the next section). In this way, the persuasive technology platform and program had the potential to reveal information or behaviors unknown to the organization, which entailed work implications beyond other influences such as the details provided in the next section.

Effective Work-Integrated Assessment Outcomes

Firefighters considered the blood work screenings to be work-related activities because of their aforementioned increased occupational risk of cancer. This danger motivated many firefighters, including Jack, to undertake the platform challenges and individual activities to get markers of clean blood or evidence of contaminated blood. To them, tasks such as blood work represented long-term, effective, work-related evidence for occupational hazards and a way to manage their work-related health concerns.

For Jack, the middle and end medical assessments showed improved medical outcomes, where he specifically valued medical outcomes evaluated in the professional work context that improved for him and many others. Firefighters such Jack felt that it was improved workplace access to HCPs (such as the nurse), who were more aware of specific work-related options available to the firefighters, that motivated them to achieve these great results, which were comparable with recommended national norms (N=141 at the beginning of the study and N=110 at the end of the study, using the webpages).

Work-Related Professional User Journey 6: Nurse (Excellent at the Start but Failing Toward the End, Platform Use Pattern; Influencers: Excellent HCP Work but Fell Back to Silos)

Although firefighters perceived medical assessments as meaningful for individual medical outcomes or group averages, HCPs perceive medical assessments as a normal, typical part of professional health care work [5]; this created different perceptions that influenced the integrated persuasive technology platform use. Hence, we examine how HCP work roles fall back into silos if they are not reworked into holistic team-based care along with the principles (Table 5).

Excellent Work-Related Medical Assessment Process and Primary Tasks in the Work Setting of Practice (Nurse)

The nurse completed 53 blood work requisitions for the 141 people at the baseline assessment, which appear as typical HCP work with physical measurements of participants, and then assessed medical and family histories. Considered by itself, the medical baseline, middle, and end look similar to excellent but regular nurses' work.

Slippage From Integrated Work Roles Back to Siloed Work Roles

The persuasive technology platform systematically delivered blood work results to each firefighter and provided an internal

email from the physician and nurse for medical assessment onto a private individual webpage.

The study integrated HCPs' work well for the work setting where the physician's and nurse's approved protocols and prescriptions were coordinated and individualized blood work options, relevant to the medical history of Jack and clinical conditions that met both professions' protocols [5].

Organization and Program Manager Results With Reasonable Costs

The user stories illustrate several interesting patterns, which suggest that everyone values the services that was done at a reasonable cost. Statements about values included "oversight of entire program progress gave greater comfort, ease of program setup eased the burden of preparing, consent and messages gave less to manage and destroy, engaging fun group activities, improving wellness overall" (ie, increase exercise, nutrition, and sodium awareness), and "implementing joint union-mandated wellness programs." These quotes match social and technical concepts that build on and improve the quality of the persuasive technology platform.

The reasonable costs were achieved at a fraction of the cost of other technologies and activities. The persuasive technology wellness challenge was licensed at approximately CAD $12,000 (US $9480) in the first year to operate the challenge and avoided CAD $17,613 (US $13,914.27) in other general wellness and disease costs for this small municipal local government organization. The total cost savings include CAD $6913 (US $5461.27) of general wellness cost saving calculated by improvements for wellness program outcomes for the number of firefighters multiplied by savings per year identified in the literature. Outcomes for the increase in physical activity, smoking cessation, weight loss and maintenance, and improvements in nutrition were multiplied by appropriate cost numbers noted in literature and then added to form a total. The rest of the cost savings was CAD $10,700 (US $8453) for disease cost savings calculated in a similar way from acute conditions better managed (eg, high blood glucose, back pain, muscle pain, sleep apnea, hypothyroidism, stroke, high cholesterol, high blood pressure, gastrointestinal ailments, asthma, colitis, kidney issues, chronic obstructive pulmonary disease, and depression) [8,28].

Discussion

Principal Findings

The results of this study confirm recent findings [1] that persuasive technology can alter individuals' outcomes and behaviors through frequent interaction beyond the initial contact points with the technology. By studying the use of a persuasive technology platform in a work environment, we found that the combined, integrated influences of work, group, social, and individual outcomes reinforce one another through this platform. In this way, we illustrate an interconnected social and technical system—but it is one that needs far more work to implement holistically in team-based care settings for life.

Individual Outcomes

As expected, the persuasive technology platform [1] improved firefighters' individual outcomes on several fronts. There was dramatic improvement in sodium awareness (29/141, 20.6%) with a reported 73% improvement on our awareness scale, 37% (11/29) of individuals lost an average of 7.3 kg, and a medical baseline (N=141) was established with medical issues identified for blood pressure, total cholesterol, low-density lipoprotein cholesterol, blood glucose, and other conditions.

The firefighters in our study were motivated by these processes and perceived benefits. They interacted with programs as individuals and as colleagues in groups, whereas professionals created the material to make the program happen. Consistent with previous findings [24], we observed that using feedback to compare consequences with goals generally led to more effective use of the platform and perceived benefits. Individuals engaged more with the persuasive technology platform's features when HCPs' work tasks were integrated into the use of those features through the platform and the firefighters' work and social networks (N=141 at the beginning of the study and N=110 at the end of the study, using the webpages). This finding was confirmed during final interviews. All design and execution elements of the study had to focus on benefits for successful behavior change to achieve the great number of improvements seen in the study.

Organizational Work Outcomes

The persuasive technology platform created an environment that motivated firefighters to have more health- and wellness-related conversations at their organizational work, even though it did not always lead to exercise-tracking behavior. We suggest that a repeatable persuasive technology platform with sustainable programs, deployed in an organizational work setting, produced these combined health improvements. The platform changed work behaviors, which station mentors observed as higher numbers of participants discussing exercise around the stations and asking more specific wellness questions. Work discussions with friends, coworkers, and HCPs increased greatly but had to fit within the organizational workplace certification requirements and be adapted to meet union guidelines. In some settings, peer and supervisor influences can prevent technology use if they do not like that technology. Given that firefighters eat together while on the job, there was much discussion with dietitians about how to change station group meals. Firefighters developed a stronger understanding of their own health through talking directly to dietitians at work.

The studied organizational work environment—such as any work environment—was not neutral, and this hindered participation. There was a level of distrust between the employer and employees in this organization and certain information that employees did not want shared with their employer. The smoking cessation program revealed the need for a secure *privacy shield* for these in-depth organizational health and wellness programs to succeed. The fact that some behaviors or conditions could eventually be made visible to the employer was a significant concern for participants.

At the same time, as there are many common elements in the individual participants' conditions, the organizational work environment could be very well suited to offer support and well-targeted interventions by HCPs. In our study, professional assessments were adjusted to consider the unique features of the professional work environment. The study helped align multiple health care protocols and checked for possible work-related diseases. For some participants, significant medical issues were uncovered that would not have been detected as early without the program and without the persuasive technology platform in the organizational work setting (29/53, 55% of the prescriptions were out of range, where 15/29, 52% were serious medical issues). Firefighters were not visiting their physicians regularly, and this was made visible by the persuasive technology platform.

It is puzzling that even in an organizational work environment where poor health could lead to loss of employment, we still found that participants struggled to develop healthy behaviors and observed unproductive workplace wellness messages. In this sense, firefighters represent a critical case, and the observations from this environment can serve as important examples for other organizations.

Groups Offer Fun, Social Distractions at Work Around Confidential, Serious Work-Related Health Issues

Group components, which represented the HCPs' tasks in the programs they ran, were designed to support participants' work processes and became meaningful tasks for the individual firefighters. Goals were set for individuals and for groups, which increased motivation and changed behaviors. Individuals were motivated by the desire to not disappoint the group. This motivation was especially salient with the exercise, sodium reduction, and smoking components. Ultimately, the study's comprehensive program achieved many outcomes around exercise improvements, smoking cessation, management of serious medical issues, and improvements in general nutrition (N=29 final interviews; N=141 at the beginning of the study and N=110 at the end of the study, using webpages).

In this way, group influence went beyond the motivational component to embed in individuals' tasks and organizational work processes. Discussions with work colleagues around health are likely to contribute to a better understanding of the data presented in a persuasive system and of the expected consequences of actions. This increased understanding is likely to lead to appropriate action [18]. Evidence from the discussions among the participants supports this pattern. Hence, we argue that integrating social components of groups in organizations with persuasive technology and systems increases the potential benefits of these platforms.

Better Integrating the Social, Work, and Technical Environments

The interplay among these persuasive technology platforms, work, and personal interactions enabled the success of the programs and fostered the firefighters' willingness to continue after the study ended: at the end of the 8-month study period, the firefighters asked the systems researcher to establish additional studies through the HR department and emergency room physician, which she designed and handed to the organization. These participants acknowledged that they could not have figured out the design or executed it at the high level that was achieved in the study.

Platform use helped users gain control over their health conditions throughout iterations, leading to improved health and wellness after several weeks. The persuasive technology platform made efforts visible to the group and reminded individuals that their respective efforts were contributing to a larger group or social goal. Seeing the progression of results helped participants persevere toward their goals with the added motivation, as noted, of not disappointing their peers. This result is consistent with research in other contexts where interfaces provided feedback about performance and led to increased group performance through social reinforcement [35].

The repeatable persuasive technology platform gave participants integrated, personalized medical advice and personalized intervention oversight on a large, systematic scale. The study's professionals had to work within organizational settings, which included HCPs in team-based care that was facilitated by the technology, through fast and flexible approaches. The end result provided firefighters with a comprehensive approach that accounted for the complexity of the local situation, integrating individual, group, work, social, home, and professional influences.

Path Forward

This paper argues that health and wellness improvements occurred in this study because the organization and participants involved used a persuasive technology platform that was designed to reflect its use in work, group, home, and social systems beyond initial interactions. The persuasive technology platform enabled employees to reveal behaviors such as smoking, drinking, or reckless driving behind a privacy shield, which probably would not have occurred without this protection, and then change those behaviors. In alignment with the 28 principles identified by Oinas-Kukkonen and Harjumaa [10] for designing persuasive systems, we allowed people to disclose sensitive data to subsequently improve their situation. Tables 3-6 provide a useful illustration of how the features of an implementation reflect 96% (27/28) of the design principles [10].

We present a vision for personal health record platforms that considers the wider improvements that could be made in understanding health decisions, addressing participants' motivation, and facilitating meaningful use of the technology in the individual users' local work setting. These persuasive technology platforms fundamentally offer a different approach to encourage new social interaction through work and home environments beyond the initial interactions. Given that a significant portion of the provincial government budgets in Canada already pay for the health system and many people still lack primary health care [37-39], these holistic platforms and team-based HCP care at a reasonable cost should be embedded in many organizational work settings.

Limitations and Future Research

The fact that our data were drawn from 1 Canadian city limits our ability to directly compare our context with others with different characteristics. For instance, large urban areas may make it harder to create strong groups if the social fabric is not as tightly woven as it is in smaller communities. Because of data aggregation, we are unable to provide findings by smaller groupings, which limits our results. Furthermore, we intentionally conducted this study in a context where good health is an essential job requirement. However, it would be interesting to see whether these results are replicated in the same manner in settings where health is not seen as essential for the work performed; this could be a boundary condition for the findings observed. Our findings illustrate that workplace health and wellness are inextricably linked with what happens with the individual and the group both on and off the job. Any sustainable persuasive technology that links workers with professional care teams must navigate the work versus personal life combined with individual versus group dynamics.

Acknowledgments

The authors thank Ms Mirna Hennawy, who worked as a research assistant.

Conflicts of Interest

None declared.

References

1. Shang L, Zuo M, Ma D, Yu Q. The antecedents and consequences of health care professional-patient online interactions: systematic review. J Med Internet Res 2019;21(9):e13940 [FREE Full text] [doi: 10.2196/13940] [Medline: 31573908]
2. Coelhoso CC, Tobo PR, Lacerda SS, Lima AH, Barrichello CR, Amaro Jr E, et al. A new mental health mobile app for well-being and stress reduction in working women: randomized controlled trial. J Med Internet Res 2019;21(11):e14269 [FREE Full text] [doi: 10.2196/14269] [Medline: 31697244]
3. 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;21(11):e13687 [FREE Full text] [doi: 10.2196/13687] [Medline: 31719028]
4. Orji R, Lomotey R, Oyibo K, Orji F, Blustein J, Shahid S. Tracking feels oppressive and 'punishy': exploring the costs and benefits of self-monitoring for health and wellness. Digit Health 2018;4:2055207618797554 [FREE Full text] [doi: 10.1177/2055207618797554] [Medline: 30202544]
5. Volkoff O, Strong DM, Elmes MB. Technological embeddedness and organizational change. Organ Sci 2007;18(5):832-848 [FREE Full text] [doi: 10.1287/orsc.1070.0288]
6. Orji R, Moffatt K. Persuasive technology for health and wellness: state-of-the-art and emerging trends. Health Informatics J 2016;24(1):66-91. [doi: 10.1177/1460458216650979] [Medline: 27245673]
7. Slight SP, Berner ES, Galanter W, Huff S, Lambert BL, Lannon C, et al. Meaningful use of electronic health records: experiences from the field and future opportunities. JMIR Med Inform 2015;3(3):e30 [FREE Full text] [doi: 10.2196/medinform.4457] [Medline: 26385598]
8. Kendall PR. Investing in prevention: improving health and creating sustainability. Office of the Provincial Health Officer. 2010. URL: https://www.health.gov.bc.ca/library/publications/year/2010/Investing_in_prevention_improving_health_and_creating_sustainability.pdf [accessed 2021-11-30]
9. Yun JJ, Won D, Park K, Yang J, Zhao X. Growth of a platform business model as an entrepreneurial ecosystem and its effects on regional development. Eur Plan Stud 2017;25(5):805-826 [FREE Full text] [doi: 10.1080/09654313.2017.1282082]
10. Oinas-Kukkonen H, Harjumaa M. A systematic framework for designing and evaluating persuasive systems. In: Oinas-Kukkonen H, Hasle P, Harjumaa M, Segerståhl K, Øhrstrøm P, editors. Persuasive technology. PERSUASIVE 2008. Lecture notes in computer science, vol 5033. Heidelberg: Springer; 2008:164-176.
11. Spanakis EG, Santana S, Tsiknakis M, Marias K, Sakkalis V, Teixeira A, et al. Technology-based innovations to foster personalized healthy lifestyles and well-being: a targeted review. J Med Internet Res 2016;18(6):e128 [FREE Full text] [doi: 10.2196/jmir.4863] [Medline: 27342137]
12. Matthews J, Win KT, Oinas-Kukkonen H, Freeman M. Persuasive technology in mobile applications promoting physical activity: a systematic review. J Med Syst 2016;40(3):72. [doi: 10.1007/s10916-015-0425-x] [Medline: 26748792]
13. Yoganathan D, Kajanan S. Persuasive technology for smartphone fitness apps. In: Proceedings of 2013 Pacific Asia Conference on Information Systems. 2013 Presented at: PACIS'13; June 18-22, 2013; Jeju Island p. 185 URL: http://aisel.aisnet.org/pacis2013/185
14. Avgerou C. Social mechanisms for causal explanation in social theory based IS research. J Assoc Inf Syst 2013;14(8):399-419 [FREE Full text] [doi: 10.17705/1jais.00341]
15. Burton-Jones A, Volkoff O. How can we develop contextualized theories of effective use? A demonstration in the context of community-care electronic health records. Inf Syst Res 2017;28(3):468-489 [FREE Full text] [doi: 10.1287/isre.2017.0702]

16. Shaw J, Shaw S, Wherton J, Hughes G, Greenhalgh T. Studying scale-up and spread as social practice: theoretical introduction and empirical case study. J Med Internet Res 2017;19(7):e244 [FREE Full text] [doi: 10.2196/jmir.7482] [Medline: 28687532]
17. Physical activity guidelines for adults 2nd edition. U.S. Department of Health and Human Services. 2018. URL: https://health.gov/sites/default/files/2019-09/Physical_Activity_Guidelines_2nd_edition.pdf [accessed 2021-11-30]
18. Burton-Jones A, Grange C. From use to effective use: a representation theory perspective. Inf Syst Res 2013;24(3):632-658 [FREE Full text] [doi: 10.1287/isre.1120.0444]
19. Baker R. Cancer top cause of firefighter fatal workplace claims, study finds. CBC News. 2018. URL: https://www.cbc.ca/news/canada/british-columbia/cancer-tops-list-of-firefighter-fatal-workplace-claims-1.4599348 [accessed 2021-11-30]
20. Lau T, Wong YH, Chan KF, Law M. Information technology and the work environment - does IT change the way people interact at work? Hum Syst Manag 2001;20(3):267-279 [FREE Full text] [doi: 10.3233/hsm-2001-20309]
21. Storer TW, Dolezal BA, Abrazado ML, Smith DL, Batalin MA, Tseng CH, PHASER Study Group. Firefighter health and fitness assessment: a call to action. J Strength Cond Res 2014;28(3):661-671 [FREE Full text] [doi: 10.1519/jsc.0b013e31829b54da] [Medline: 24566608]
22. Carleton RN, Afifi TO, Turner S, Taillieu T, Duranceau S, LeBouthillier DM, et al. Mental disorder symptoms among public safety personnel in canada. Can J Psychiatry 2018;63(1):54-64 [FREE Full text] [doi: 10.1177/0706743717723825] [Medline: 28845686]
23. Nazari G, MacDermid JC, Sinden KE, Overend TJ. Comparison of Canadian firefighters and healthy controls based on submaximal fitness testing and strength considering age and gender. Int J Occup Saf Ergon 2019;25(1):1-7. [doi: 10.1080/10803548.2017.1372086] [Medline: 28877646]
24. Eisenhardt KM, Graebner ME. Theory building from cases: opportunities and challenges. Acad Manage J 2007;50(1):25-32 [FREE Full text] [doi: 10.5465/amj.2007.24160888]
25. Yin RK. Case Study Research and Application: Design and Methods. 6th Edition. Thousand Oaks: SAGE Publications; 2017.
26. 2015 - 2020 dietary guidelines for Americans. 8th Edition. U.S. Department of Health and Human Services and U.S. Department of Agriculture. 2015. URL: https://health.gov/our-work/food-nutrition/previous-dietary-guidelines/2015 [accessed 2021-11-30]
27. U.S. Department of Agriculture, U.S. Department of Health and Human Services. Dietary guidelines for Americans 2010. 7th edition. U.S. Department of Health and Human Services. Washington DC: U.S. Government Printing Office; 2010. URL: https://web.archive.org/web/20160901170759/https:/health.gov/dietaryguidelines/dga2010/dietaryguidelines2010.pdf [accessed 2021-11-21]
28. Your guide to lowering your blood pressure with DASH. U.S. Department of Health and Human Services. 2006. URL: https://www.nhlbi.nih.gov/files/docs/public/heart/new_dash.pdf [accessed 2021-11-30]
29. NVivo software. QSR International. 2019. URL: https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home [accessed 2021-10-15]
30. Eriksson P, Kovalainen A. Qualitative Methods in Business Research. 2nd Edition. Thousand Oaks: SAGE Publications; 2016.
31. Lincoln YS, Guba EG. Naturalistic Inquiry. Thousand Oaks: SAGE Publications; 1985.
32. Lunenburg FC. Goal-setting theory of motivation. Int J manag Bus Adm 2011;15(1):1-6 [FREE Full text]
33. Canadian intellectual property office. Government of Canada. 2019. URL: https://www.ic.gc.ca/eic/site/cipointernet-internetopic.nsf/eng/home [accessed 2019-11-19]
34. Office of the privacy commissioner of Canada. Office of the Privacy Commissioner of Canada. 2019. URL: https://www.priv.gc.ca/en/ [accessed 2019-11-19]
35. Jung JH, Schneider C, Valacich J. Enhancing the motivational affordance of information systems: the effects of real-time performance feedback and goal setting in group collaboration environments. Manage Sci 2010;56(4):724-742 [FREE Full text] [doi: 10.1287/mnsc.1090.1129]
36. MacLeod L. Making SMART goals smarter. Physician Exec 2012;38(2):68-72 [FREE Full text] [Medline: 23971403]
37. Krueger H, Rasali D, Fong D. The economic burden of risk factors in British Columbia, 2015: excess weight, tobacco smoking, alcohol use, physical inactivity and low fruit and vegetable consumption. BC Centre for Disease Control. 2018. URL: http://www.bccdc.ca/pop-public-health/Documents/economic_burden_five_risk_factors_BC_2015.pdf [accessed 2021-11-30]
38. Primary health care. World Health Organization. 2019. URL: https://www.who.int/news-room/fact-sheets/detail/primary-health-care [accessed 2021-11-30]
39. 50% of the world lack primary health care. World Health Organization. 2019. URL: https://www.who.int/news-room/fact-sheets/detail/primary-health-care [accessed 2022-01-27]

Abbreviations

DASH: Dietary Approaches to Stop Hypertension
HCP: health care professional

HR: human resource
SMART: specific, measurable, attainable, realistic, and time-bound

Edited by A Mavragani; submitted 03.08.21; peer-reviewed by H McCall, L Shang, K Oyibo; comments to author 14.09.21; revised version received 22.10.21; accepted 21.12.21; published 23.02.22.

Please cite as:
Marcolin B, Saunders C, Aubert B
Persuasive Technologies and Social Interactions in Professional Environments: Embedded Qualitative Case Study
JMIR Form Res 2022;6(2):e32613
URL: https://formative.jmir.org/2022/2/e32613
doi:10.2196/32613
PMID:35195527

Original Paper

Iterative Development of a Mobile Phone App to Support Community Health Volunteers During Cervical Cancer Screening in Western Kenya: Qualitative Study

Jacob Stocks[1*], MSc; Yujung Choi[1*], MSc; Saduma Ibrahim[2*], BSc; Megan Huchko[1,3*], MPH, MD

[1]Center for Global Reproductive Health, Duke Global Health Institute, Durham, NC, United States

[2]Kenya Medical Research Institute, Nairobi, Kenya

[3]Department of Obstetrics and Gynecology, Duke University School of Medicine, Durham, NC, United States

[*]all authors contributed equally

Corresponding Author:
Jacob Stocks, MSc
Center for Global Reproductive Health
Duke Global Health Institute
310 Trent Drive
Durham, NC, 27710
United States
Phone: 1 919 681 7760
Email: jacob_stocks@med.unc.edu

Abstract

Background: To achieve the World Health Organization targets for cervical cancer elimination, low- and middle-income countries will need to develop innovative strategies to provide human papillomavirus (HPV)–based screening at a population level. Although mobile health (mHealth) interventions may help realize these goals by filling gaps in electronic specimen tracking and patient education, effective implementation of mHealth interventions is dependent upon context-specific development that is acceptable and usable by the target population. Detailed feedback should be gathered at the design and development stages to yield final products that reflect the needs, desires, and capabilities of target users.

Objective: The aim of this study is to develop an mHealth app (mSaada) to support HPV-based screening in partnership with community health volunteers (CHVs) and program planners in western Kenya.

Methods: A team of student programmers developed a prototype to meet previously identified gaps in screening: patient education, protocol support, data capture, and specimen tracking. The prototype was iteratively developed through 2 waves of in-person working sessions with quantitative (survey) and qualitative (in-depth interview) feedback. Research staff engaged key stakeholders from both urban and rural locations and with varying levels of experience in delivering screening services. During the sessions, participants completed simulation exercises and role-play activities to become familiar with the platform. Once feedback was gathered and synthesized after each wave of in-person data collection, developers implemented changes to improve mSaada functionality.

Results: A total of 18 CHVs and clinicians participated in the in-person sessions. Participants found mSaada useful, easy to use, and would meet the needs of CHVs to provide HPV-based cervical cancer screening (electronic data capture, client education resources, and specimen tracking). They provided key feedback to enhance user experience, workflow, and sustainability. Key changes included altering the appearance of the wireframes, adding translation in additional local languages, changing potentially insensitive figures, alphabetizing lengthy dropdown menus, adding clinically relevant logic checks when entering data, and incorporating the ability to make real time edits to client records. They also made recommendations for additional features that might enhance mSaada's impact at the facility and health system levels, specifically the inclusion of a report-generating tool consistent with the Ministry of Health standards.

Conclusions: Using a process of iterative feedback with key stakeholders and rapid response from developers, we have developed a mobile app ready for pilot testing in HPV-based screening programs led by CHVs.

(JMIR Form Res 2022;6(2):e27501) doi:10.2196/27501

KEYWORDS

mHealth; cervical cancer screening; Kenya; HPV testing; user-testing; community health volunteers; mobile phone

Introduction

Background

According to projections from the World Health Organization (WHO), incident cases of and annual deaths due to cervical cancer could increase by 37% and 47% to 777,000 and 459,000, respectively, by 2040 [1]. The vast majority of this disease burden is observed in low- and middle-income countries in sub-Saharan Africa and Southeast Asia [1,2]. To reduce rates of morbidity and mortality due to cervical cancer, screening and vaccination programs that target human papillomavirus (HPV), the causative agent in almost all cervical cancers, must be able to reach the target population in a cost-effective and culturally appropriate manner [3]. Such programs are nascent in many low- and middle-income countries and, despite the availability of simplified protocols, face infrastructural and financial challenges [4,5]. Kenya's Ministry of Health has developed national guidelines for cervical cancer screening, which align with the WHO-endorsed framework for HPV-based screening and treatment. Although this may drive greater access to services throughout the country, there are significant barriers to implementation [3]. One such barrier is the use of nonformally trained lay providers, such as community health volunteers (CHVs), to deliver care. Although these individuals can reduce the clinical load on a strained health system, they require adequate supervision and support to carry out these services.

Poor adherence to clinical guidelines has been cited as a challenge to using CHV-led programs and a risk of delivering substandard care. An initiative to use lay health providers in Siaya district, Kenya, showed poor adherence to clinical guidelines by trained community health workers [6]. Further training, on-the-job supervision, and technical support have been shown to increase the quality of service delivery. A study conducted in the Morogoro region of Tanzania found that lay health workers "value supervision and appreciate the sense of legitimacy that arises when supervisors visit them in the village" [7]. However, it is important that when supervision is offered, it occurs in a respectful, nonjudgmental way, or it could lead to decreased self-efficacy and lowered desire to complete assigned tasks. The provision of job aids, such as pamphlets, flipcharts, mobile apps, and handbooks, has been cited as effective at increasing adherence and may provide a means to avoid overly critical supervision that still yields the necessary support [6,8].

Mobile health (mHealth) has been defined by the WHO as "medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, personal digital assistants, and other wireless devices" [9]. mHealth approaches, which typically use information sharing, education, communication, or data collection strategies to meet specified goals [10], have been used to deliver service reminders, promote behavior change, enhance medication adherence, provide health education, and collect and store patient data [11-15]. There is a large body of evidence supporting the feasibility, acceptability, and usability of a variety of mHealth interventions that target a wide range of health conditions such as depression, diabetes, HIV/AIDS, and cancer [16-22]. Although many mHealth interventions are patient sided, there are a growing number of provider-sided apps that seek to support and improve the delivery of health services [10,16]. These novel approaches provide a platform for the effective delivery of evidence-based practices targeting a range of health outcomes in often unreached or hard-to-reach populations and can be used to support task shifting of health delivery to CHVs or other lay health workers [23].

Effective implementation of mHealth interventions is reliant upon context-specific development that is appropriate, useable by the target population, and reflective of mobile phone ownership and rates of use [24,25]. Such considerations must be incorporated in the design and development stages of mobile apps and other interventions to yield final products that reflect the needs, desires, and capabilities of target users [25]. Otherwise, poor development of mobile apps could result in low levels of usefulness and user uptake [26]. A recent study by Huchko et al [27] showed that CHVs delivering cervical cancer screening in community- and facility-based settings in western Kenya desired more protocol and decision support tools to effectively complete screening. In addition, the study identified a clear need for continued education regarding the cause, risk, transmission, and prevention of HPV [28-30]. On the basis of high reported mobile phone ownership within Kenya and past research documenting success with SMS text messaging–based delivery of screening results in western Kenya [31,32], the introduction of a mobile app–based intervention to address gaps in education and provider support appears feasible in this context.

Study Objective

The aim of this study is to iteratively develop and refine the mSaada mobile app in consultation with key stakeholders before small-scale pilot testing in health facilities in western Kenya.

Methods

Study Setting

This 2-wave qualitative study sought perspectives on the functionality of and user experience with a newly developed mobile phone app. Data collection was conducted in Migori and Kisumu, Kenya. Both locations, one rural and one urban, offer cervical cancer screening programs within local health facilities and were being considered for government-supported implementation of HPV-based screening. We chose these locations based on the target end user of the app, CHVs, who are commonly employed in both places.

App Platform

mSaada, meaning *support* in Swahili, is a counseling and decision support tool designed for use by CHVs. As an android-based mobile phone app developed by students as part of a computer science course at Duke University, mSaada was

designed to address logistical and educational gaps in HPV-based cervical cancer screening in western Kenya. The app features were developed based on prior research in Migori and Kisumu counties [27,28], and through web-based communication with Kenyan key stakeholders, who provided feedback on the overall functionality, wireframes, and pilot app. As described in Table 1, the app includes four main features to guide the CHVs through the entire process of screening, including counseling and decision-making, answers to frequently asked questions, data collection, and specimen tracking. Feature headings were defined based on stakeholder feedback and use more common, abbreviated language (*Add New Client*, *Screening Info*, *Questions*, and *Search Client*).

Table 1. Description of mSaada mobile app features.

Feature heading and purpose	Components
Add New Client	
Collect relevant client data during screening	A 34-item clinical questionnaire collecting demographic and health history information
Link client data to laboratory specimens for tracking and results notification	Barcode scanner for linkage of client records with laboratory specimens
Screening Info	
Aid in education and counseling of clients	Cervical Cancer Education Module
Aid in explanation of self-collection steps of HPV[a]-based cervical cancer screening	Kenya Ministry of Health Cervical Cancer Prevention Protocol
__[b]	Explanatory figures and diagrams
Questions	
Support client education	A total of 65 questions and answers addressing myths and misperceptions of HPV and cervical cancer, grouped by relevant topics including screening, treatment, and transmission
Search Client	
Access, review, and edit client records	Searchable client record database

[a]HPV: human papillomavirus.

[b]Component applies to both purposes.

Study Sample

To gather input and perspectives regarding mSaada and its features, we recruited individuals in three categories: experts (n=6), end users (n=6), and lay persons (n=6). Expert study participants were individuals with direct experience providing cervical cancer-related services, such as clinicians or study staff working in related research. End user study participants were individuals who had worked or were currently working as CHVs and had performed cervical cancer screening using HPV testing via self-collection. Lay person study participants were individuals who had no formal training or experience in cervical cancer screening. Research staff identified potential study participants based on prior engagement with the individuals or by recommendation from partner facility staff. Study staff recruited participants by phone or in-person before the beginning of the study period.

Study Design

The iterative development period lasted 8 weeks in total and consisted of alternating waves of data collection and integration of feedback into the mSaada platform. During data collection waves, participants attended day-long feedback sessions that were categorized by participant group (ie, experts, end users, and lay persons). To obtain a combination of novel and continuous feedback on mSaada, we asked 2 representative experts, end users, and lay persons (6/18, 33%) to participate in both waves of data collection activities, completing an in-depth interview during both waves. All other study participants (12/18, 67%) were asked to attend only 1 day of feedback sessions, completing only 1 in-depth interview. This resulted in 24 in-depth interviews among the 18 participants, with an average length of 45 minutes.

Feedback sessions (n=6) began with a description of study aims, intended methods of data collection, and completion of written informed consent. Following this discussion, researchers provided detailed, screen-by-screen explanations of the mSaada app and its features, allowing participants to follow along using study phones loaded with the platform during the demonstration. After initial hands-on familiarization in a group setting, participants completed simulation activities in pairs, including CHV and client role-plays and example scenarios to gain more experience with the platform [33]. Explanation, familiarization, and role-plays with mSaada lasted approximately 3 hours, on average. Thereafter, we conducted individual in-depth interviews to gather detailed feedback about the platform. Both the group exploration and familiarization sessions and individual in-depth interviews were conducted in English and audio-recorded for transcription. We completed transcription using Otter.ai, a free, open-source software. Study staff transcribed recordings using the software, and reviewed and edited resulting transcripts for accuracy. Researchers followed the same protocol for both waves of data collection. All study activities occurred within

the Kisumu Office of the Duke Center for Global Reproductive Health and the Migori County Hospital. Group feedback sessions were conducted within the conference area of the office, whereas individual in-depth interviews were conducted within private rooms. Research staff and app developers convened to discuss and integrate key participant feedback after it was gathered.

Communication With App Developers

Researchers consolidated data from the first wave of interviews into a master list of recommended changes and updates. The resulting 12-page document was shared with Kisumu- and Durham-based researchers to focus and prioritize revision efforts. Although most of the recommended changes were considered for integration before the second wave of interviews, some participants suggested changes that were not feasible to undertake between the 2 feedback waves. We addressed these changes, including a large expansion in capabilities of the existing *Search Client* feature, after the iterative development period.

After agreed-upon prioritization of changes on the master list, researchers provided participant feedback to the Nairobi-based app developer for integration. To track and discuss progress during the 3-week period of app refinement, the app developer supplied Kisumu-based researchers with intermittent versions of mSaada, which were downloaded to study phones and reviewed for accuracy. This back-and-forth process helped facilitate effective communication between team members and allowed for successful completion of app refinement.

Qualitative Analysis

We developed a 2-part interview guide that was used for all participants during both waves of data collection (Multimedia Appendix 1). First, we asked participants to reflect on each of the 4 features of the app. For each feature, participants were asked their opinion on usability, user control, aesthetics, comfort, ease of use, and any challenges they faced. The interview concluded with questions regarding overall impressions of the platform and thoughts about implementation and use of mSaada within a Kenyan context. We asked participants for any recommended changes or updates to the app's layout and features, and any concerns they had about the app's use within facilities.

We developed an initial codebook with deductive codes that mirrored the domains described above in the interview guide (usability, user control, aesthetics, comfort, and ease of use) and further refined it with inductive codes as participant feedback was gathered. We analyzed the qualitative data using thematic analysis and a 4-stage process [34]. Analysis was aided by NVivo (version 12). First, researchers reviewed all transcripts and created document memos to summarize key points from each participant and to get a strong sense of the collected data. Second, we identified deductive structural codes based on the two sections of the interview guide (ie, feature-specific feedback and overall impressions). Third, inductive thematic codes were identified via thorough review and rereview of participant transcripts. As thematic codes were identified, researchers added them to the codebook and recoded transcripts based on observed themes. Fourth, after completion of thematic coding, we wrote analytic memos for each identified theme detailing the similarities and differences in feedback between features of mSaada to broadly summarize the gathered information. Interim analysis, completed between waves 1 and 2 of data collection, used a similar process but was not aided by NVivo (version 12). To accelerate the procedure, researchers completed a consolidated version of the aforementioned 4-stage process.

Ethics Approval

We obtained ethical approval from Duke University Campus Institutional Review Board (2019-0650) and the Kenya Medical Research Institute's Scientific and Ethics Review Unit (2918). All participants provided written informed consent before initiation of study activities. Implementation of the study followed all local human participant research policies.

Results

Overview

Feedback resulted in changes to all 4 areas of the mSaada app functionality and appearance. Changes to the app focused on improving user experience and enhancing the visual appeal of the platform. Figure 1 depicts the wireframes of the final version of mSaada after iterative development. Table 2 outlines key updates made to each feature of mSaada. Qualitative data gathered during the iterative development process were classified into four categories: ease of use, accessibility of information, anticipated workflow considerations, and acceptability.

Figure 1. Example wireframes of mSaada: (A) home page depicting 4 main features; (B) page showing option to select language preference; (C) self-collection screening instructions within the Screening Info feature; (D) patient data entry page within the Add New Client feature; (E) searchable frequently asked questions with corresponding answer; and (F) Search Client feature illustrating client record lookup. The client-facing language used in wireframes C, D, and E is reflective of participant recommendations (see the Anticipated Workflow Considerations section). mSaada is a provider-sided app. DOB: date of birth; HPV: human papillomavirus; LU: last updated; SA: survey administrator; VIA/VILI: Visual Inspection with Acetic Acid and Lugol's Iodine.

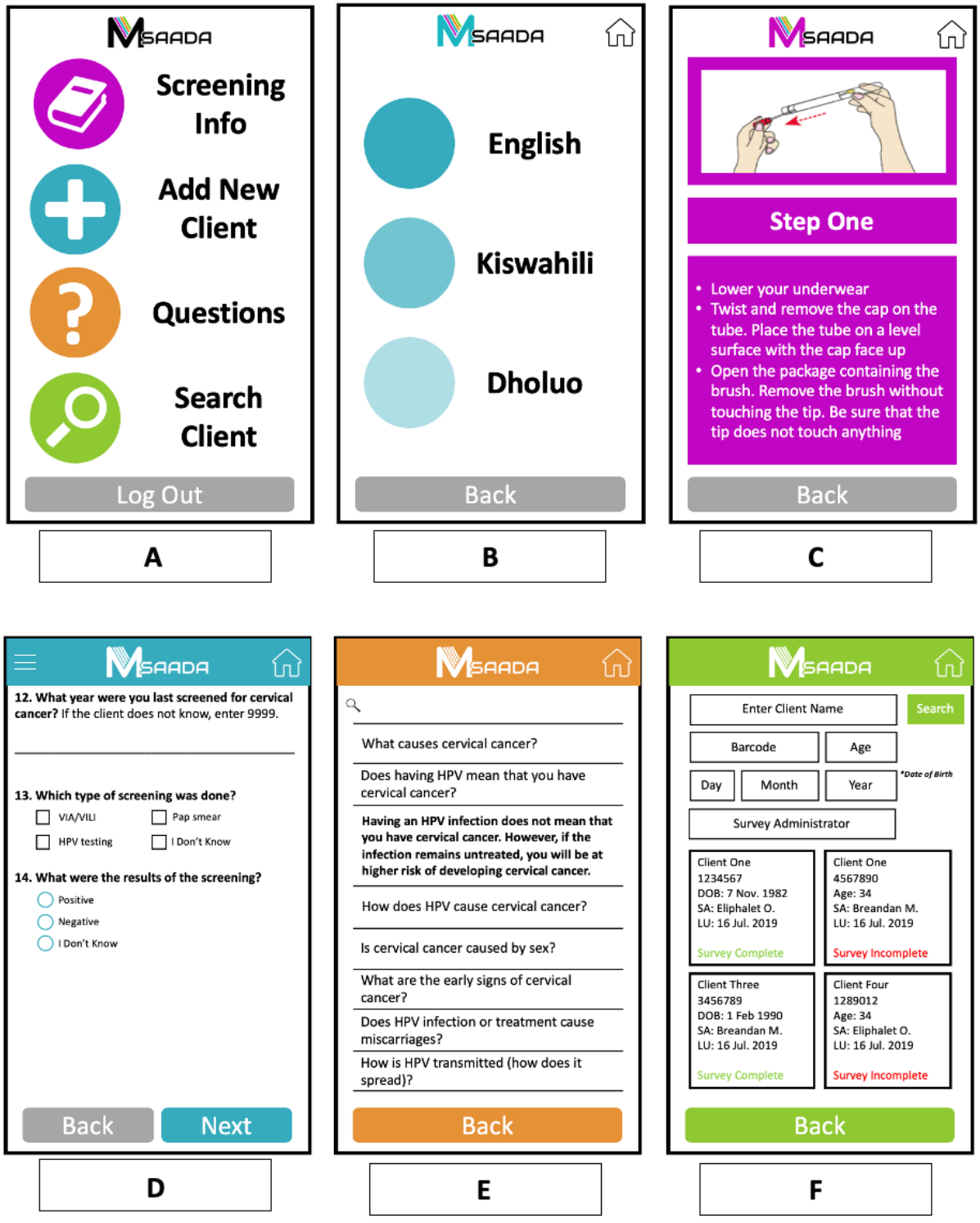


Table 2. Feature-specific feedback and the resulting revisions to mSaada.

App feature and weakness identified	Changes or update made
Add New Client	
Lack of clarity in questionnaire items	Refined phrasing in consultation with Kisumu-based staff
Lack of questionnaire items gathering administrative details	Added items regarding screening location and provider name
Questionnaire items presented only in English	Translated feature content into Swahili and Luo
Screening Info	
Lack of sufficiently detailed graphics and instructions for self-collection	Updated step-by-step instructions to include more detail
Need for additional images or graphics to aid in counseling	Included Kenya Ministry of Health Cervical Cancer Prevention Protocol for additional support
Content presented only in English	Translated feature content into Swahili and Luo
Questions	
Difficulty using built-in keyword search query	Broaden keyword search query to include synonymous terms
Educational content presented only in English	Translated into Swahili and Luo
Search Client	
Unable to effectively edit client records	Incorporated editing capabilities within feature[a]
Search fields of database query returning broad group of client records	Added additional search fields to aid in narrowing of client records

[a]Not completed before pilot testing.

Ease of Use

Overall, participants were comfortable using the main features of mSaada. Participants cited previous mobile phone use, specifically smartphone use, as a main reason for their comfort with the app:

> *The practitioners here in Kisumu, I think most of them are used to these phones, so I don't think they will have many challenges, maybe it is just a matter of familiarizing with the app. It's quite different, but almost everyone has smartphones, so they're used to them.* [Expert 4]

When asked about their use of specific features of mSaada, participants reported ease with all features except the *Questions* feature. They liked the logical flow from screen to screen within each feature, allowing for overall easy use of the app. For example, participants liked the use of the *Back* and *Next* buttons for navigation within the app and felt the scrolling and swiping aspects of the platform were effective, responsive, and straightforward. Regarding the *Questions* feature, participants reported difficulty in identifying the correct keyword necessary to locate the needed information. When asked about their challenges with the *Questions* feature, participants believed this difficulty was due to a lack of familiarity and experience with mSaada, not a weakness in the design or functionality of the platform. Participants felt that the use of the *Questions* feature would improve with repeated exposure and did not recommend any changes:

> *I think the feature helps a lot in replying to questions. All we have to do is master the concept, so that during the [keyword search] you already have an insight. That way when a client asks you a question, you know what to type and then get the correct information.* [End User 4]

Participants concluded that the most effective way to become well versed with the app was through role-play and other simulation activities. They recommended detailed and engaging trainings for CHVs to facilitate successful use before going to the field.

Accessibility of Information

Accessibility of information emerged as both a strength and weakness of mSaada. Overall, participants commended the use of very simple, direct, nonmedical jargon within the app. Participants suggested that this would likely increase the usefulness of mSaada, as CHVs would be able to effectively communicate necessary information with clients at a contextually appropriate level of understanding, regardless of their lack of formal medical training. In addition, the format of educational content, specifically within the Cervical Cancer Education Module, was well received, as participants found the scrollable nature of the electronic information easier to use in comparison with the bulky flip charts currently used within facilities.

Although strengths in the presentation of information were identified, all participants highly recommended translation of app content into Swahili, a national language of Kenya, as well as Luo, a local language in western Kenya, so that CHVs could effectively convey important information to clients of any language preference. It was mentioned that, by presenting only an English version of the app, CHVs were responsible for translation of information on the fly to non-English–speaking clients, which could allow for inconsistencies or inaccuracies in the transmission of information from app to CHV to client. Thus, to lower the burden of responsibility on CHVs, participants recommended that all information intended for

client consumption be uniformly translated to both local languages.

In addition to the inclusion of local languages, participants highlighted sentences and phrases included within the app's *Add New Client* and *Screening Info* features that were not well understood or did not appropriately convey the intended meaning of the statement within a Kenyan context. To ensure effective use of mSaada and reduce varied or unintended interpretation of information, participants placed emphasis on the need to assure that phrasing of statements was nuanced and reflective of the speech patterns of those in the target area.

Finally, participants were concerned with the lack of ability to access or edit completed client records once information was entered into the app. They felt it was important to include an editing aspect within the *Search Client* feature as well as a way of uniquely identifying client records to get "all of the info about [a] person." Participants cited examples from past experience where small typos had been made or where there were questions about a certain client record that needed further evaluation.

Anticipated Workflow Considerations

Overall, participants believed that mSaada would help accelerate the screening and client data collection processes. Errors and inefficiencies within the app were highlighted, and changes were recommended, especially within the *Add New Client* feature. Participants emphasized the need for the incorporation of clinically relevant logic checks throughout the clinical questionnaire. Recommended checks included the following: age, to ensure clients are within the recommended target age range for screening; history of hysterectomy, to ensure no contraindication of screening; and pregnancy status, which is of clinical importance during the treatment of HPV infection or cervical cancer. In addition, participants recommended refinement of response options, specifically the alphabetizing of lengthy dropdown menus, within the clinical questionnaire to make the data collection process more efficient.

Participants also emphasized the need to cater the presentation of information within mSaada toward end users (CHVs) for successful and efficient use within facilities. Participants felt that the wording of items within the clinical questionnaire should be constructed so that CHVs could read the text directly from the app as if they were talking to a client, rather than having to reframe the question after reading. During role-play activities, participants often stumbled over questionnaire items not written in this way.

Finally, participants identified a data persistence issue within the *Add New Client* feature, which they felt would negatively impact the CHV workflow within clinical settings. The possibility of data loss and the need for re-entry of client data concerned participants.

> *The issue...was that when you go back [to a previous screen using the "Back" button], now you have to again key in the same set of information and this might be a problem if you have a number of patients who are on the line because we may make errors and we may also want to change [answers] without doing away with the whole information.* [Lay Person 2]

Acceptability

Participants believed that mSaada included all components and features necessary to aid CHVs in the successful screening of clients. They found the Cervical Cancer Education Module within the *Screening Info* feature to be comprehensive but concise. In addition, although not exhaustive, participants found the *Questions* feature content to be extremely useful for client education and believed the information covered many of the main topics about which a client might inquire. Participants did, however, recommend that the 4 main features on the home screen be reordered. Participants believed that presenting the features in chronological order of a client's screening visit would help reduce confusion of app users (CHVs) and help facilitate the visit.

> *The "Add New Client" is what we'll be using mostly for our new clients. It ["Screening Info"] should be the second because after adding a patient is when you go to the "Screening Info" and the education module. And then you go to "Questions". And lastly the "Search Client" should be the last because maybe you could use that later, after the interview.* [End User 4]

Finally, participants recommended the addition of a fifth feature for administrative purposes. The additional feature, a report-generating tool, was recommended to enhance mSaada's usefulness at the facility and health system levels. Participants cited a need for up-to-date records on clinical outcomes and services rendered, which is requested by the Ministry of Health.

Discussion

Principal Findings

To our knowledge, we are the first to describe the development of an app designed for health workers offering HPV-based cervical cancer screening services in Kenya. The iterative development process, whereby relevant stakeholders with a diversity of perspectives provided input in a cyclical manner, proved to be effective at creating a viable, functioning app for pilot testing. The stakeholder-engaged, iterative process, which has been used in many studies [18-20,35,36], yielded critical insight that could not have been gathered otherwise. For example, many of the word choices and sentence structures included within the platform were considered to be unclear and not likely to be understood by clients during use, even though this language originated from materials provided by Kenyan research staff. Although the app was designed to improve the quality of counseling, having end users test the app revealed additional concerns about the possibility of miscommunication. These lessons learned call to attention the need for highly engaged, locally driven processes of intervention development, even outside of mHealth approaches.

Comparison With Prior Work

Many mHealth interventions are developed with a singular intent and, therefore, use only 1 approach to achieve their purpose, through communication, education, data collection, or information sharing [10]. This method was countered in this study, as participants applauded mSaada's multifaceted nature,

offering a comprehensive solution to many of the challenges experienced in the delivery of screening services. This stands as an example of the importance of multidimensional approaches for future mHealth interventions.

A key strength of this study is the feedback cycle. The mSaada platform underwent multiple rounds of testing and refinement in the 2-month study period, and this process involved a variety of individuals of varying engagement in cervical cancer screening. In addition, key stakeholder and end user feedback was integral to decision-making and revisions of the platform. A study by Fishbein et al [19] showed that the inclusion of a broad range of stakeholders and perspectives proved beneficial for the development of the app, and the sense of cocreation likely contributed to the resulting acceptability of the app. By striving to produce a product that was context specific and relevant to its target audience, those with the most firsthand knowledge and deepest insight into the success and failure of current screening efforts were able to drive mSaada's development, likely resulting in a better and more useful final product. Another strength of this study was the use of qualitative methods to provide detailed feedback, opinions, and perspective on the app's development and use. Qualitative methods, if executed well, can produce a plethora of actionable information for use in intervention development and further revision. Many usability testing frameworks emphasize the use of qualitative methods when developing and evaluating mHealth interventions [33].

Limitations

There were also a few limitations to this study. First, although we encouraged feedback from a variety of individuals, our small sample size could have missed important feedback. Given the novel nature of the app within this setting, this study was focused on feedback from app end users. However, feedback from women being screened by providers using the app is critical to its implementation and was missing from our study. In addition, although consulted informally, hospital and district health administrators were not included within our study sample, therefore, insight into how mSaada might integrate into facilities from a macro level was not gathered. Finally, participants were given only 3 hours, on average, to use the app before providing feedback. This limited interaction may not have provided sufficient time to fully explore and identify an issue with the platform, and functionality in clinical settings should be tested in a pilot study. Future mHealth development studies should strive to gather feedback from a sufficient number of stakeholders at all levels of the health system and provide ample interaction with their proposed platforms.

Conclusions

This study demonstrates the usefulness of iterative approaches to mHealth development. By engaging a variety of key stakeholders, we were able to quickly develop a mobile app that would be well received, have ownership among end users, and ensure readiness for small-scale pilot testing. In this study, we show a process for an iterative approach to app development that builds on context-specific preliminary work to further improve the functionality before introduction in a clinical setting.

Acknowledgments

The work presented in this paper was supported by the National Cancer Institute of the National Institutes of Health under award number R01CA188248 (80%) and a grant from the Duke Global Cancer Program (20%). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest

None declared.

Multimedia Appendix 1
In-depth interview guide.
[DOCX File , 20 KB - formative_v6i2e27501_app1.docx]

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018 Nov;68(6):394-424 [FREE Full text] [doi: 10.3322/caac.21492] [Medline: 30207593]
2. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 2015 Mar 01;136(5):E359-E386 [FREE Full text] [doi: 10.1002/ijc.29210] [Medline: 25220842]
3. WHO Guidelines for Screening and Treatment of Precancerous Lesions for Cervical Cancer Prevention. Geneva: World Health Organization; 2013.
4. Morema EN, Atieli HE, Onyango RO, Omondi JH, Ouma C. Determinants of cervical screening services uptake among 18-49 year old women seeking services at the Jaramogi Oginga Odinga Teaching and Referral Hospital, Kisumu, Kenya. BMC Health Serv Res 2014 Aug 06;14:335 [FREE Full text] [doi: 10.1186/1472-6963-14-335] [Medline: 25100298]

5. Abdikarim IK, Carole Atieno WM, Habtu M. Prevalence and associated factors of cervical cancer screening among Somali women in an urban settlement in Kenya. J Comm Pub Health Nurs 2017;03(01). [doi: 10.4172/2471-9846.1000159]
6. Rowe SY, Kelly JM, Olewe MA, Kleinbaum DG, McGowan JE, McFarland DA, et al. Effect of multiple interventions on community health workers' adherence to clinical guidelines in Siaya district, Kenya. Trans R Soc Trop Med Hyg 2007 Feb;101(2):188-202. [doi: 10.1016/j.trstmh.2006.02.023] [Medline: 17064747]
7. Roberton T, Applegate J, Lefevre AE, Mosha I, Cooper CM, Silverman M, et al. Initial experiences and innovations in supervising community health workers for maternal, newborn, and child health in Morogoro region, Tanzania. Hum Resour Health 2015 Apr 09;13:19 [FREE Full text] [doi: 10.1186/s12960-015-0010-x] [Medline: 25880459]
8. Scott K, Beckham SW, Gross M, Pariyo G, Rao KD, Cometto G, et al. What do we know about community-based health worker programs? A systematic review of existing reviews on community health workers. Hum Resour Health 2018 Aug 16;16(1):39 [FREE Full text] [doi: 10.1186/s12960-018-0304-x] [Medline: 30115074]
9. mHealth: New Horizons for Health Through Mobile Technologies: Second Global Survey on eHealth. Geneva: World Health Organization; 2011.
10. Amoakoh-Coleman M, Borgstein AB, Sondaal SF, Grobbee DE, Miltenburg AS, Verwijs M, et al. Effectiveness of mHealth interventions targeting health care workers to improve pregnancy outcomes in low- and middle-income countries: a systematic review. J Med Internet Res 2016 Aug 19;18(8):e226 [FREE Full text] [doi: 10.2196/jmir.5533] [Medline: 27543152]
11. Zhao J, Freeman B, Li M. Can mobile phone apps influence people's 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]
12. Fry JP, Neff RA. Periodic prompts and reminders in health promotion and health behavior interventions: systematic review. J Med Internet Res 2009 May 14;11(2):e16 [FREE Full text] [doi: 10.2196/jmir.1138] [Medline: 19632970]
13. Gurman TA, Rubin SE, Roess AA. Effectiveness of mHealth behavior change communication interventions in developing countries: a systematic review of the literature. J Health Commun 2012;17 Suppl 1:82-104. [doi: 10.1080/10810730.2011.649160] [Medline: 22548603]
14. Ashar R, Lewis S, Blazes DL, Chretien JP. Applying information and communications technologies to collect health data from remote settings: a systematic assessment of current technologies. J Biomed Inform 2010 Apr;43(2):332-341 [FREE Full text] [doi: 10.1016/j.jbi.2009.11.009] [Medline: 19961957]
15. Shirima K, Mukasa O, Schellenberg JA, Manzi F, John D, Mushi A, et al. The use of personal digital assistants for data entry at the point of collection in a large household survey in southern Tanzania. Emerg Themes Epidemiol 2007 Jun 01;4:5 [FREE Full text] [doi: 10.1186/1742-7622-4-5] [Medline: 17543099]
16. van Heerden A, Sen D, Desmond C, Louw J, Richter L. App-supported promotion of child growth and development by community health workers in Kenya: feasibility and acceptability study. JMIR Mhealth Uhealth 2017 Dec 05;5(12):e182 [FREE Full text] [doi: 10.2196/mhealth.6911] [Medline: 29208588]
17. Griffin L, Lee D, Jaisle A, Carek P, George T, Laber E, et al. Creating an mHealth app for colorectal cancer screening: user-centered design approach. JMIR Hum Factors 2019 May 08;6(2):e12700 [FREE Full text] [doi: 10.2196/12700] [Medline: 31066688]
18. Laurence C, Wispelwey E, Flickinger TE, Grabowski M, Waldman AL, Plews-Ogan E, et al. Development of PositiveLinks: a mobile phone app to promote linkage and retention in care for people with HIV. JMIR Form Res 2019 Mar 20;3(1):e11578 [FREE Full text] [doi: 10.2196/11578] [Medline: 30892269]
19. Fishbein JN, Nisotel LE, MacDonald JJ, Amoyal Pensak N, Jacobs JM, Flanagan C, et al. Mobile application to promote adherence to oral chemotherapy and symptom management: a protocol for design and development. JMIR Res Protoc 2017 Apr 20;6(4):e62 [FREE Full text] [doi: 10.2196/resprot.6198] [Medline: 28428158]
20. Fjeldsoe BS, Miller YD, O'Brien JL, Marshall AL. Iterative development of MobileMums: a physical activity intervention for women with young children. Int J Behav Nutr Phys Act 2012 Dec 20;9:151 [FREE Full text] [doi: 10.1186/1479-5868-9-151] [Medline: 23256730]
21. Whitehead L, Seaton P. The effectiveness of self-management mobile phone and tablet apps in long-term condition management: a systematic. J Med Internet Res 2016 May 16;18(5):e97 [FREE Full text] [doi: 10.2196/jmir.4883] [Medline: 27185295]
22. Cafazzo JA, Casselman M, Hamming N, Katzman DK, Palmert MR. Design of an mHealth app for the self-management of adolescent type 1 diabetes: a pilot study. J Med Internet Res 2012 May 08;14(3):e70 [FREE Full text] [doi: 10.2196/jmir.2058] [Medline: 22564332]
23. Royston G, Hagar C, Long L, McMahon D, Pakenham-Walsh N, Wadhwani N. Mobile health-care information for all: a global challenge. Lancet Global Health 2015 Jul;3(7):e356-e357. [doi: 10.1016/S2214-109X(15)00054-6]
24. 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;Suppl:49-56. [doi: 10.2345/0899-8205-46.s2.49] [Medline: 23039777]
25. Schnall R, Rojas M, Bakken S, Brown W, Carballo-Dieguez A, Carry M, et al. A user-centered model for designing consumer mobile health (mHealth) applications (apps). J Biomed Inform 2016 Apr;60:243-251 [FREE Full text] [doi: 10.1016/j.jbi.2016.02.002] [Medline: 26903153]

26. Brown W, Yen P, Rojas M, Schnall R. Assessment of the health IT usability evaluation model (health-ITUEM) for evaluating mobile health (mHealth) technology. J Biomed Inform 2013 Dec;46(6):1080-1087 [FREE Full text] [doi: 10.1016/j.jbi.2013.08.001] [Medline: 23973872]
27. Huchko MJ, Ibrahim S, Blat C, Cohen CR, Smith JS, Hiatt RA, et al. Cervical cancer screening through human papillomavirus testing in community health campaigns versus health facilities in rural western Kenya. Int J Gynaecol Obstet 2018 Apr;141(1):63-69 [FREE Full text] [doi: 10.1002/ijgo.12415] [Medline: 29197067]
28. Rosser JI, Njoroge B, Huchko MJ. Changing knowledge, attitudes, and behaviors regarding cervical cancer screening: the effects of an educational intervention in rural Kenya. Patient Educ Couns 2015 Jul;98(7):884-889 [FREE Full text] [doi: 10.1016/j.pec.2015.03.017] [Medline: 25858634]
29. Rosser JI, Njoroge B, Huchko MJ. Cervical cancer screening knowledge and behavior among women attending an urban HIV clinic in Western Kenya. J Cancer Educ 2015 Sep;30(3):567-572 [FREE Full text] [doi: 10.1007/s13187-014-0787-7] [Medline: 25595965]
30. Rosser JI, Njoroge B, Huchko MJ. Knowledge about cervical cancer screening and perception of risk among women attending outpatient clinics in rural Kenya. Int J Gynaecol Obstet 2015 Mar;128(3):211-215 [FREE Full text] [doi: 10.1016/j.ijgo.2014.09.006] [Medline: 25467908]
31. Aker JC, Mbiti IM. Mobile phones and economic development in Africa. J Econ Perspect 2010 Aug 01;24(3):207-232. [doi: 10.1257/jep.24.3.207]
32. Huchko MJ, Saduma I, Blat C, Oketch S, Bukusi EA. How providing cervical cancer screening results via cell phone affects patient follow-up rates in western Kenya. J Glob Oncol 2019 Jun;5:1-8 [FREE Full text] [doi: 10.1200/JGO.18.00264] [Medline: 31246553]
33. Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: a scoping review. Int J Med Inform 2019 Jun;126:95-104. [doi: 10.1016/j.ijmedinf.2019.03.018] [Medline: 31029270]
34. Introduction to applied thematic analysis. In: Applied Thematic Analysis. Thousand Oaks, California: SAGE Publications; 2012.
35. Pereira-Azevedo N, Osório L, Fraga A, Roobol MJ. Rotterdam prostate cancer risk calculator: development and usability testing of the mobile phone app. JMIR Cancer 2017 Jan 06;3(1):e1 [FREE Full text] [doi: 10.2196/cancer.6750] [Medline: 28410180]
36. Simons D, De Bourdeaudhuij I, Clarys P, De Cocker K, Vandelanotte C, Deforche B. A smartphone app to promote an active lifestyle in lower-educated working young adults: development, usability, acceptability, and feasibility study. JMIR Mhealth Uhealth 2018 Feb 20;6(2):e44 [FREE Full text] [doi: 10.2196/mhealth.8287] [Medline: 29463491]

Abbreviations

CHV: community health volunteer
HPV: human papillomavirus
mHealth: mobile health
WHO: World Health Organization

Edited by T Leung; submitted 17.03.21; peer-reviewed by L Nisotel, L Guo; comments to author 05.08.21; revised version received 29.09.21; accepted 29.12.21; published 24.02.22.

Please cite as:
Stocks J, Choi Y, Ibrahim S, Huchko M
Iterative Development of a Mobile Phone App to Support Community Health Volunteers During Cervical Cancer Screening in Western Kenya: Qualitative Study
JMIR Form Res 2022;6(2):e27501
URL: https://formative.jmir.org/2022/2/e27501
doi:10.2196/27501
PMID:35200151

Original Paper

Understanding the Needs of a Mobile Phone–Based Telemonitoring Program for Pregnant Women at High Risk for Pre-Eclampsia: Interpretive Qualitative Description Study

Anam Shahil Feroz[1,2], BSc, MSc; Kristina De Vera[1], BSc, MHI; Nadia D Bragagnolo[3], BSc, MSc; Sarah Saleem[2], MBBS, MSc; Zulfiqar Bhutta[4,5,6], MBBS, PhD; Emily Seto[1,7], BAS, MSc, PhD

[1]Institute of Health Policy, Management, and Evaluation, University of Toronto, Toronto, ON, Canada

[2]Community Health Sciences Department, Aga Khan University, Karachi, Pakistan

[3]Physical Sciences, Sunnybrook Health Sciences Centre, Toronto, ON, Canada

[4]Centre for Global Child Health, SickKids, Toronto, ON, Canada

[5]Center of Excellence in Women and Child Health, Aga Khan University, Karachi, Pakistan

[6]Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

[7]Centre for Global eHealth Innovation, Techna Institute, University Health Network, Toronto, ON, Canada

Corresponding Author:
Anam Shahil Feroz, BSc, MSc
Institute of Health Policy, Management, and Evaluation
University of Toronto
155 College Street 4th Floor
Toronto, ON, M5T 3M6
Canada
Phone: 1 6478676702
Email: anam.shahil@mail.utoronto.ca

Abstract

Background: Lack of early risk detection, diagnosis, and treatment of pregnant women at high risk for pre-eclampsia can result in high maternal mortality and morbidity not only in Pakistan but also in other low- to middle-income countries (LMICs). A potential tool for supporting pregnant women at high risk for pre-eclampsia for early detection is telemonitoring (TM). However, there is a limited body of evidence on end-user needs and preferences to inform the design of the TM programs for pregnant women at high risk for pre-eclampsia, specifically in LMICs such as Pakistan.

Objective: This study aims to explore the needs of TM for pregnant women at high risk for pre-eclampsia in Karachi, Pakistan, to inform a potential future feasibility trial of a mobile phone–based TM program.

Methods: An interpretive qualitative description approach was used to conduct and analyze 36 semistructured interviews with 15 (42%) pregnant women and 21 (58%) key informants, including clinicians; nurses; maternal, neonatal, and child health specialists; and digital health experts to explore the perspectives, needs, and preferences of a mobile phone–based TM program to support pregnant women at high risk for pre-eclampsia. Pregnant women were identified through heterogeneous sampling, whereas key informants were selected through purposive sampling. The interview transcripts were analyzed using a conventional content analysis technique.

Results: The following four themes emerged from the analysis of the transcripts: poor use of antenatal care during pregnancy, the value of a TM program in high-risk pregnancy, barriers influencing the adoption of TM programs and potential strategies, and considerations for implementing TM programs. The pregnant women and health care providers were willing to use a TM program as they perceived many benefits, including early identification of pregnancy complications, prompt treatment, convenience, cost-effectiveness, increased sense of empowerment for one's health care, improved care continuity, and reduced clinical workload. However, some providers and pregnant women mentioned some concerns regarding the adoption of a TM program, including malfunctioning and safety concerns, potential inaccuracy of blood pressure machines, increased clinical workload, and resistance to learning new technology. Our study recommends building the capacity of patients and providers on TM program use, sensitizing the community and family members on the usefulness of the TM program, using an approach incorporating user-centered design and phased implementation to determine the clinical workload and whether additional staff for the TM program is required, and ensuring greater levels of co-design and the engagement of consumer representatives.

Conclusions: Our findings highlight the perceived feasibility of a mobile phone–based TM program for pregnant women at high risk for pre-eclampsia and provide insights that can be directly used for the design of future TM programs with the aim of reducing mortality and morbidity from pre-eclampsia and eclampsia in LMICs.

(JMIR Form Res 2022;6(2):e32428) doi:10.2196/32428

KEYWORDS

telemonitoring; pre-eclampsia; qualitative study; Pakistan; pregnant women at high risk; low- to middle-income country; pregnant; pregnancy; women; mobile phone

Introduction

Background

In Pakistan, approximately one-third (34%) of maternal mortality in tertiary-level facilities is attributable to pre-eclampsia and eclampsia [1]. High maternal mortality from pre-eclampsia and eclampsia results from the lack of early risk detection, diagnosis, and treatment of pregnant women at high risk for pre-eclampsia [1]. To meet the United Nations Sustainable Developmental Goal target of 3.1 (maternal mortality ratio <70/100,000 live births) by 2030 [2], innovations are required to help decrease pre-eclampsia and eclampsia related mortality. There is substantial empirical evidence on the use of telemonitoring (TM) to support pregnant women at high risk for pre-eclampsia by remotely monitoring their blood pressure readings at home [3-11]. TM is a promising tool in which individuals who are pregnant take blood pressure measurements and record symptoms at home, and these readings and self-reported symptoms are sent to their health care providers in real time [9]. Van Den Heuvel et al [10] found that pregnant women at high risk in the Netherlands could highly benefit from TM as it facilitates better blood pressure control, early risk identification and treatment, fewer hospital visits, and cost savings [10]. In Pakistan, TM has been implemented by community health workers as part of the Community-Level Interventions for Pre-eclampsia (CLIP) trial [9] and Control of Blood Pressure and Risk Attenuation–Bangladesh, Pakistan, and Sri Lanka studies [12]. In the CLIP trial, the Piers on the Move mobile health app directed community health workers to first observe women to rule out emergency conditions that would warrant immediate referral to a facility. The CLIP Piers on the Move tool facilitated the stratification of pregnant women by community health workers into 1 of 3 care pathways: usual antenatal and postnatal care, nonurgent referral, and urgent referral to a higher facility. The CLIP trial was well received by families; however, it did not have a significant impact on either the composite outcome of maternal, fetal, and newborn mortality and severe morbidity or individual components thereof [13].

Although the literature exploring the use of TM for supporting pregnancy care is expanding [10,14-16], there are limitations in the data collected from TM systems, such as monitoring of only a few gestation parameters, which makes it hard or impossible for health professionals to provide holistic assistance to the pregnant women and fetuses [16]. The review by Eysenbach et al [17] on the effectiveness of TM in obstetrics concluded that TM could be tentatively recommended for pregnant women at risk for preterm delivery, given the high methodological risk of bias among the included studies [18]. In addition, very few studies have focused on understanding patient needs for the design and development of TM platforms. The limited body of qualitative evidence on end user needs and preferences to inform the design of TM programs could be a barrier to the successful development and implementation of more applicable, effective, and user-centric TM platforms [19-22]. A qualitative study conducted at Vanderbilt University Medical Center explored the practices, health needs, and strategies related to pregnancy care for pregnant women and caregivers to inform the development and implementation of health information technologies [23]. Most expectant mothers in the study encountered everyday problems with mobility and household management and desired more assistance from caregivers, who often did not know how to help. The study identified technological innovations, including health-tracking watches to take basic vital measurements, virtual assistants, and cellular apps, to connect fellow pregnant women with others in their region to support expectant families [22].

Most reported TM programs had been implemented in high-income countries (eg, United Kingdom, Canada, United States, and Belgium) [6,7,9,10], with a paucity of evidence on the use of TM to support pregnant women at high risk in low- to middle-income countries (LMICs). In our scoping review on the use of digital health interventions for pregnant women at high risk for pre-eclampsia in LMICs, we identified only 9 unique digital health interventions from mainly South Asia and sub-Saharan Africa. Of these interventions, 2 served the purpose of predicting risk for adverse maternal health outcomes, whereas 7 focused on monitoring pregnant women at high risk, for managing pre-eclampsia and eclampsia (publication in review). The review identified only 1 TM intervention for monitoring pre-eclampsia and eclampsia in an LMIC. This was the Pre-eclampsia Integrated Estimate of RiSk on the Move app, which was used in the CLIP trials in India, Pakistan, and Mozambique, conducted from 2014 to 2017.

Objective

TM is a complex intervention and is sensitive to the context in which it is applied, including sociodemographic and sociocultural considerations, financial constraints, clinical workflows, and health system systems [24]. Thus, it cannot be assumed that the TM needs of pregnant women in high-income countries will necessarily apply to pregnant women in LMICs. We propose to assess the feasibility of a mobile phone–based TM program to support pregnant women at high risk for pre-eclampsia at Jinnah Post Graduate Medical Center (JPMC) in Karachi, which is the largest city in Pakistan. As the first step of this project, this study aims to explore the needs of TM for

pregnant women at high risk for pre-eclampsia in Karachi, Pakistan, to inform a potential future feasibility trial of a mobile phone–based TM.

Methods

Research Design and Setting Overview

Given our intention to understand the local context and needs of the intended users, an interpretive descriptive design was used [25] to explore the perspectives, needs, and preferences of pregnant women at high risk for pre-eclampsia for mobile phone–based TM through semistructured interviews with patient participants and key informants. Patient participants included pregnant women at high risk for pre-eclampsia who were the intended beneficiaries of the TM program, whereas key informants included the diverse group of stakeholders who provide or supply health care in different capacities, including nurses; clinicians; maternal, neonatal, and child health (MNCH) specialists; and digital health experts. This study was conducted at the JPMC, a 1650-bed tertiary-level public sector hospital in Karachi, which provides hospital care to >1 million people of low socioeconomic status coming from Karachi, Interior Sindh, Baluchistan, and other remote areas [1]. This research focused on the outpatient area of the JPMC obstetrics and gynecology (OB-GYN) department, which serves the vast majority of low-income women with high-risk pregnancies.

Further details about the study sites and protocol have been previously published [19]. The study was approved by the Aga Khan University (AKU) ethical review committee (2020-2153-8519), the JPMC institutional review board (44379), and the University of Toronto research ethics board (30635).

Proposed Mobile Phone–Based TM Program

During the interviews, patient participants were provided with an overview of the proposed mobile phone-based TM program. The proposed TM program includes the use of a Bluetooth-enabled home blood pressure device that is validated for use during pregnancy and a mobile app (in the Urdu language). The TM program will enable pregnant women to take their blood pressure reading every morning at home and answer symptom questions using the mobile app. All enrolled women would receive real-time automated instructions based on their readings, such as taking additional blood pressure readings, calling the medical officer (ie, a trained physician in Pakistan), or visiting the OB-GYN emergency department. The real-time automated instructions would be delivered in Urdu text on the mobile app. The medical officer would receive alerts from the TM system if their patient's blood pressure values were out of the target range. The medical officer would act as a central point person to communicate with the patients (phone calls or using the asynchronous app chat feature) and with the rest of the participant's care team as needed.

Participant Recruitment and Eligibility Criteria

The interpretive description approach used a maximum variation sampling technique [26,27] to purposively recruit a range of pregnant women at high risk for pre-eclampsia who had differing needs and preferences for the mobile phone–based TM program. Textbox 1 provides the definition of pregnant women at high risk for pre-eclampsia as per the National Institute for Health and Care Excellence guidelines. Owing to the COVID-19 pandemic–related restrictions, nurses at the hospital were asked to support the identification and recruitment of eligible pregnant women at high risk for pre-eclampsia for in-person interviews. Nurses identified the eligible pregnant women from the JPMC outpatient department, where women wait for several hours to be seen by their health care providers. The nurses contacted the research staff once a potential participant was identified, and then, the research staff made an immediate visit to the clinic to conduct in-person patient interviews at the hospital. The key informants such as clinicians, including specialists in OB-GYN, nurses, MNCH specialists, and digital health experts, were purposively recruited from the JPMC OB-GYN department, AKU, and other leading health care institutions in Karachi, Pakistan. Textbox 2 provides a list of the eligibility criteria for the patient participants and the key informants.

Textbox 1. Definition of pregnant women at high risk for pre-eclampsia as per the National Institute for Clinical Excellence guidelines.

Definition

National Institute for Health and Care Excellence guidelines define pregnant women at high risk for pre-eclampsia as those who have 1 high-risk factor or >1 moderate risk factor for pre-eclampsia.

High-risk factors:

- Hypertensive disease in a previous pregnancy
- Chronic kidney disease
- Autoimmune diseases, such as systemic lupus erythematosus or antiphospholipid syndrome
- Type 1 or type 2 diabetes
- Chronic hypertension

Moderate risk factors:

- First pregnancy
- Aged ≥40 years
- Pregnancy interval of >10 years
- BMI of ≥35 kg/m^2 at the first visit
- Family history of pre-eclampsia
- Multifetal pregnancy

Textbox 2. Eligibility criteria of patient participants and key informants.

Pregnant women at high risk for pre-eclampsia

Inclusion criteria:

- Pregnant women at high risk for pre-eclampsia who were visiting the Jinnah Post Graduate Medical Center outpatient department for antenatal visits
- Pregnant women at high risk for pre-eclampsia who met the National Institute for Health and Care Excellence guidelines [27] definition of high risk for pre-eclampsia
- Pregnant women who were diagnosed as high risk for pre-eclampsia for at least 3 weeks and have had time to reflect on the disease condition, associated difficulties, and needs for telemonitoring
- Pregnant women with the ability to speak English, Urdu, or Sindhi language

Exclusion criteria:

- Pregnant women at high risk for pre-eclampsia who were admitted to the inpatient wards and emergency care for treatment purposes
- Pregnant women at high risk for pre-eclampsia who refuse to consent for participating in the needs assessment study

Key informants

Inclusion criteria:

- Key informants must be clinicians; nurses; maternal, neonatal, and child health specialists; or digital health experts who were directly or indirectly involved in the care of pregnant women at high risk for pre-eclampsia
- Key informants with the ability to speak Urdu or English language

Exclusion criteria:

- Key informants who refuse to consent for participating in the needs assessment study

Data Collection Methods

Interviews were conducted between March 2020 and August 2020. The research team and study participants had access to all necessary personal protective equipment to help prevent the risk of spreading the COVID-19 virus during data collection. Midwives and nurses at JPMC used the National Institute for Health and Care Excellence guidelines [28] to identify pregnant women at high risk for pre-eclampsia from the outpatient department of the JPMC hospital. The identified pregnant women were informed about the study's purpose and procedures (including the recording of interviews), and their willingness

to participate in the study was ascertained by the primary researcher (ASF). If the participant was unable to read the consent form, the primary researcher explained the consent form verbally in their local language. Pregnant women at high risk for pre-eclampsia who were unable to write their names were asked to provide a thumbprint to mark their consent to participate. The interviews were conducted on the same day by the primary researcher when eligible pregnant women were identified. To avoid any disruption during the interview and ensure confidentiality, the consenting pregnant women were asked to move to a separate private room for the interview. We anticipated conducting and recording 10 to 15 interviews with pregnant women at high risk, to reach data saturation. The intent was to allow caregivers (husband, mother-in-law, or any family member) to accompany the pregnant women during the interviews if preferred by the patient. However, caregiver involvement during the interview was not possible because of COVID-19–related restrictions. After the interviews, the patient participants were given a 10-minute illustrative presentation to provide them with an understanding of the pre-eclampsia condition, its consequences, and simple ways to manage the condition through regular follow-ups. Key informants such as clinicians, nurses, MNCH specialists, and digital health experts were interviewed to understand their perspectives, preferences, and needs regarding the use of TM for pregnant women at high risk for pre-eclampsia. The key informants were identified from AKU, JPMC, People's Primary Healthcare Initiative, Sehat Kahani, Digital Care, eHealth Association of Pakistan, Commission on Science and Technology for Sustainable Development in the South, Shaheed Zulfikar Ali Bhutto Institute of Science and Technology, Tech4Life Enterprises, the Aga Khan Development Network Digital Health Resource Center, and other relevant institutions. All the key informants were invited to participate in the qualitative study via email, and key informants were requested to sign informed consent forms before the interview began. Most interviews with key informants were conducted through the web via Zoom (Zoom Video Communications, Inc) in either Urdu or English, whereas a few interviews with clinicians and nurses were conducted face to face at the JPMC OB-GYN department. We anticipated conducting 13 to 15 interviews with key informants to reach data saturation. The interviews lasted between 40 and 60 minutes.

The approach by Spradley [29] was used to design 2 semistructured interview guides for pregnant women at high risk for pre-eclampsia and key informants. This approach emphasizes the importance of having *grand tour* questions to allow for the free flow of rich and deep information. The interview guide for pregnant women at high risk for pre-eclampsia involved a general discussion about pre-eclampsia, causes of pre-eclampsia, perceptions toward the use of TM for pregnant women at high risk for pre-eclampsia, perceived benefits of TM, potential limitations or concerns related to TM for pre-eclampsia, and feasibility of smartphone-based TM. The interview guide for key informants included grand questions on causes of pre-eclampsia, routine obstetric care for pre-eclampsia, use of TM for supporting pregnant women at high risk for pre-eclampsia, and perceived facilitators of and barriers to the implementation of TM. The guide was pilot-tested with 2 pregnant women at high risk for pre-eclampsia and 2 key informants who shared the same traits as the study sample [30,31].

Data Analysis

The audio recordings from the interviews were professionally transcribed and translated into the English language, with no identifying characteristics included in the transcriptions. The anonymized transcripts were uploaded to NVivo (version 12 Plus; QSR International) to enable easy and organized retrieval of data for analysis. The conventional content analysis approach [32] was used to inductively analyze all the interview transcripts. The primary researcher (ASF) independently coded all the transcripts as the primary reviewer, whereas KDV and NDB independently coded key informants' and patients' interviews, respectively, as the second reviewers. The interviews of key informants and patient participants were analyzed separately by the researchers, and later, the codes and themes were compared across the 2 groups to identify overarching themes and subthemes. The main themes and subthemes were identified independently by the primary researcher and second reviewers for the 2 groups of respondents and then discussed in the larger research group until agreement on the themes was achieved. To gain a more complete understanding of the perspectives, preferences, and needs of TM for women at high risk for pre-eclampsia, the subthemes from all the interviews were compared and contrasted by multiple researchers (ASF, KDV, and NDB) to seek convergence and corroboration through data triangulation between the patient and key informant interviews [27,33]. The four main themes, which were consistent between the patient and key informant groups, and subthemes were finalized once a consensus was achieved after a total of four meetings: 2 meetings with the group of reviewers who analyzed the transcripts and 2 meetings with the larger research group.

Results

Overview

A total of 36 semistructured interviews were conducted to explore the perspectives, needs, and preferences of a mobile phone–based TM program to support pregnant women at high risk for pre-eclampsia. Of the 36 interviews, 15 (42%) interviews were conducted with pregnant women at high risk for pre-eclampsia, whereas 21 (58%) were conducted with various key informants, including 8 (38%) clinicians, 3 (14%) nurses, 3 (14%) MNCH specialists, and 7 (33%) digital health experts. Each of the interviews lasted between 30 and 50 minutes. All the participants (36/36, 100%) who were approached by the study team agreed to participate. The demographic information for all the key informants and patient participants is illustrated in Tables 1 and 2, respectively.

On the basis of the inductive analysis, four overarching themes were identified: (1) poor use of antenatal care during pregnancy, (2) value of a TM program in high-risk pregnancy, (3) barriers influencing adoption of TM and potential strategies, and (4) considerations for implementing TM programs. These themes and their subthemes are summarized in Textbox 3.

Table 1. Characteristics of key informants (N=21).

Characteristics and category	Values
Gender, n (%)	
Female	16 (76)
Male	5 (24)
Age (years)	
Values, mean (SD)	46.21 (11)
Values, median (range)	47 (28-65)
Role, n (%)	
Nurses	3 (14)
Clinicians (OB-GYN[a])	8 (38)
MNCH[b] specialists and digital health experts	10 (48)
Specialty, n (%)	
OB-GYN	11 (52)
MNCH	3 (14)
Digital health	7 (33)
Experience (years)	
Values, mean (SD)	16.52 (10)
Values, median (range)	20 (4-40)

[a]OB-GYN: obstetrics and gynecology.

[b]MNCH: maternal, neonatal, and child health.

Table 2. Characteristics of patient participants (N=15).

Characteristics of pregnant women at high risk for pre-eclampsia and category	Values
Gender, n (%)	
Female	15 (100)
Age (years)	
Values, mean (SD)	28.26 (5.284)
Values, median (range)	28 (20-38)
Educational level, n (%)	
No education	3 (20)
Less than high school	3 (20)
High school	8 (53)
College or university	1 (7)
Occupation, n (%)	
Housewife	11 (73)
Professional	4 (27)
History of pre-eclampsia, n (%)	
Yes	9 (60)
No	6 (40)
Pregnancy (weeks)	
Values, mean (SD)	31.67 (6.06)
Values, median (range)	33 (12-37)
Frequency of blood pressure measurement, n (%)	
Daily	5 (33)
Once or twice a week	4 (27)
Thrice a week	2 (13)
As per need	3 (20)
Never	1 (7)
Access to a personal home blood pressure machine, n (%)	
Yes	2 (13)
No	13 (87)
Gravida[a], n (%)	
Primigravida	4 (27)
Multigravida	11 (73)
Parity[b], n (%)	
Nulliparity	4 (27)
Multiparity	9 (60)
Grand parity	2 (13)
Access to a mobile phone, n (%)	
Basic mobile phone	4 (27)
Smartphone	11 (73)
Personal or shared access to a mobile phone, n (%)	
Individual access	9 (60)
Shared access	6 (40)

Characteristics of pregnant women at high risk for pre-eclampsia and category	Values
Access to the internet, n (%)	
Yes	8 (53)
No	7 (47)

[a]Total number of pregnancies.

[b]Live births and stillbirths.

Textbox 3. Themes and categories.

Poor use of antenatal care during pregnancy

- Inadequate access to quality maternal health care services
- Poor awareness and self-management during a high-risk pregnancy

Value of a telemonitoring (TM) program in high-risk pregnancy

- Early identification of pregnancy complications and prompt treatment
- Impact on physician's workload
- Convenient and cost-effective
- Sense of empowerment in own health care

Barriers influencing adoption of TM and potential strategies

- Lack of willingness by pregnant women and health care providers to use TM program
- Weak technological literacy to use TM program
- Lack of technological infrastructure to use TM program
- Sociocultural factors impacting TM program use

Considerations for implementing the TM program

- Features and ease of use of TM program
- Handling, maintenance, and sustainability of the TM program

Poor Use of Antenatal Care During Pregnancy

Inadequate Access to Quality Maternal Health Care Services

When asked about the use of antenatal care during current and previous pregnancies, some pregnant women described the accessibility issues they face in using antenatal care, such as the long commute to the public hospital. They reiterated that traveling is time consuming and costly, given that the public hospitals are few and quite far from their residential areas:

> *Every time I had to spend Rs. 1000 for JPMC visit. Obviously, this [per visit costs for travel] is very hard for us to afford.* [Woman with pregnancy 03]

Most pregnant women at high risk for pre-eclampsia mentioned that they usually visit small private clinics near their residential areas to receive care in the first few months of their pregnancy. However, these clinics are not well-equipped and are capable of managing only normal pregnancies. For this reason, pregnant women at high risk are referred to a tertiary-level public hospital such as JPMC by these private clinics for management and hospital delivery:

> *Before this [JPMC], I was going to a private hospital at Korangi's side for checkups. They [Doctor at the private hospital] told me that maybe there is some issue with the umbilical cord, so, maybe you would need to be operated...and referred me to Jinnah. So, I made my antenatal card over here [JPMC].* [Woman with pregnancy 12]

Most pregnant women at high risk described an unpleasant environment at the public hospital because of the excessive environmental noise, crowds, and long waiting times:

> *Actually, we have to wait a lot here and the atmosphere is a little supportive...the people [pregnant women] who come here [for antenatal care] have to wait for most of the time...and the noise here [due to crowd] makes us go into more depression. So, this is a bit of an issue.* [Woman with pregnancy 05]

Key informants highlighted the impact of high physician workload on the quality of patient care and health education, especially in public sector hospitals. Owing to the increased volume of patients and time constraints, providers often found it difficult to educate pregnant women at high risk about self-management such as dietary advice, regular blood pressure

monitoring, regular antenatal visits, and the importance of physical activity and exercise during pregnancy. In addition, the increased workload of health care workers was thought to not enable providers to counsel pregnant women about the adverse outcomes of the disease condition:

> *You know our physicians are overwhelmed, they do not have time or even the realization that they need to educate their patients. So, it begins from there.* [Digital health expert 14]

Poor Awareness and Self-management During a High-Risk Pregnancy

The interviews with pregnant women and key informants revealed that pregnant women have poor awareness and knowledge about pre-eclampsia and eclampsia, its symptoms, and pregnancy care in general. A few pregnant women stated that they knew about the disease symptoms but only through their current and previous pregnancy experiences. Key informants highlighted that pregnant women generally do not understand their disease condition because of low literacy levels, and when they begin to realize and report the disease symptoms, they get tagged as normal pregnancy symptoms:

> *Women do not even know what their symptoms are, so they are generally not educated enough to understand what they are going through...So, lots of these women do not even know what is happening, and by the time they do start realizing, it is routinely discarded off just as the symptoms of the pregnancy, it is not investigated further.* [Digital health expert 02]

The interviews revealed that inadequate awareness about the disease condition influences the self-care behaviors of pregnant women, such as dietary precautions, regular exercise, and blood pressure monitoring among pregnant women at high risk. When asked about self-care behaviors during pregnancy, most pregnant women described the different types of dietary precautions they take during pregnancy. However, none of them reported performing exercise and physical activity during pregnancy. In addition, pregnant women reported measuring their blood pressure during pregnancy either through a home blood pressure machine or visiting clinics. The frequency of blood pressure measurement varied among all pregnant women. A minority of pregnant women verbalized that daily blood pressure monitoring caused them stress, and therefore, they tend to avoid measuring blood pressure unless it is required.

In addition, pregnant women at high risk for pre-eclampsia and key informants described several sociocultural and financial factors leading to poor disease awareness and self-management. The pregnant women in the study received different levels of support from their husbands, in-laws, extended family members, and colleagues at the workplace for pregnancy care. Most pregnant women at high risk for pre-eclampsia acknowledged the support they received from their husbands and in-laws during pregnancy, whereas some pregnant women seemed to have trouble finding support for their responsibilities, especially childcare. Therefore, they did not have time to attend antenatal visits or get their blood pressure checked at a nearby clinic. Key informants further mentioned that women do not receive enough attention and care during the pregnancy period:

> *They [Pregnant women] have a great busy schedule at home. They cannot afford the luxury of enjoying downtime during their pregnancy they work really hard...their husbands and in-laws do not accept that these pregnant mothers need to have better hygiene, better nutrition, and rest hours. So, all of these factors contribute to poor pregnancy care and we all are aware of these factors, and there is nothing hidden.* [Digital health expert 14]

Key informants believed that cultural norms influence the empowerment of pregnant women to make decisions about their pregnancy care. They stated that women do not consider their health as a priority over other things, largely because of cultural factors. One such instance that key informants highlighted was that pregnant women do not make use of the health card provided to them by the public sector hospital upon registration of their pregnancy. The intent of the health card is for them to receive antenatal services; instead, the women would only visit the hospital and use their health card at the time of delivery or if there was an emergency:

> *So, they only have the card with them, and if you see, then there is only one entry on it...and on it is mentioned that you have to get all these tests done regularly but they never come. They only come at the time of delivery.* [Clinician 03]

Key informants described that pregnant women are empowered by health care workers during regular clinic visits to improve self-care behaviors. However, health care workers forget that the household decision-making power is with husbands and mothers-in-law. Thus, they suggested having husbands and family members be present along with pregnant women during health education sessions to have a significant impact on the uptake of maternal health services:

> *So, most of the time we try to empower the pregnant woman, and we forget that she does not have the decision-making power in that set. So, it is very important for us, to understand the dynamics, so the husband is very important. He needs to be on board and educated when it comes to the self-care of his wife. A mother-in-law is very important because in most cases she is the decision-maker. So, she needs to be very much aware of the situation and, also aware of what needs to be done and If that is not done, what will happen.* [Digital health expert 11]

In terms of financial constraints, some women expressed that they could not afford to buy nutritious food, medicine for their pre-eclampsia condition, or equipment for the regular monitoring of blood pressure at home. Most pregnant women acknowledged the importance of regular blood pressure monitoring. However, they mentioned that they could not get regular blood pressure measurements during their pregnancy as their husbands had to accompany them to the clinic, which often required their husbands to take time off from work and for the pregnant women to work around their husbands' schedules.

Value of a TM Program in High-Risk Pregnancy

Early Identification of Pregnancy Complications and Prompt Treatment

Approximately all key informants and most pregnant women at high risk for pre-eclampsia believed that the TM program would aid in the early identification of alarming signs through regular monitoring of blood pressure and disease symptoms. A few clinicians and MNCH specialists expressed that the TM program is highly valuable for reducing maternal and child morbidity and mortality rates:

> *Its benefit is that if even 90% of it is being done then it would be very beneficial. Meaning that the patients with us having eclampsia would not be dying. Because when they come to us in tertiary care. They come in bad to worse conditions...The mortality rate increases and morbidity itself is increased in those patients. And because of this our mortality and morbidity can be decreased.* [Clinician 04]

Key informants further mentioned that the early identification of danger signs would help in the timely referral of pregnant women at high risk to secondary health care facilities where their condition could be managed through either medications or immediate delivery. Although some key informants expressed that TM should be complementary to antenatal visits as clinicians would like to see pregnant women at high risk in clinics to monitor signs that cannot be captured by TM programs, an expert described an emergency where health care providers can immediately intervene with the help of the TM program:

> *So, if there are no kicks [fetal movement] for the last eight hours or ten hours we [healthcare providers] can immediately intervene and have the patients come to the hospital, in an emergency, do the ultrasound and do some kind of intervention to save the life of the baby. So, that why it is important. It is very, very important, especially for precious pregnancies.* [Digital health expert 14]

Pregnant women articulated that blood pressure readings would be more accurate when taken at home as they would not have to travel and wait for long hours in the stressful environment of JPMC to be seen by health care providers. In addition, pregnant women felt that there would be improved care continuity through this TM program as they will remain connected with their respective physicians through mobile phones without needing to go out of their homes:

> *There are many benefits like most doctors say that due to blood pressure anything can happen to the child and mother. Both can be saved. Through the TM program doctor will remain connected to us through the machine, phone message, or call...so it is much better. Then I would not have to go out of my home, I can stay at home and talk to my doctor.* [Woman with pregnancy 02]

Impact on Physician's Workload

Key informants, including clinicians and nurses, believed that the TM program could help in reducing physicians' workload and crowds in public hospitals, given that face-to-face visits by pregnant women would be reduced to a large extent. Although some key informants thought that this is a potential benefit of implementing a TM program, others argued that it might also be detrimental to clinicians' time as clinicians would need to remotely monitor and track the health conditions of pregnant women through the TM program. In addition, key informants stated that being able to remotely monitor the health conditions of pregnant women would empower health care providers and facilitate better clinical decision-making for managing high-risk pregnancies. The experts thought that the TM program would simplify their work as the blood pressure monitoring data will flow regularly to health providers through the smartphones of pregnant women, which would help in clinical decision-making:

> *It simplifies their (healthcare provider) work it facilitates in monitoring their women on regular basis...it becomes very, very important to give these devices to women so that the data can flow regularly to these health providers...to help them make a better decision about the patients or the women that they are taking care of.* [Digital health expert 14]

Convenient and Cost-effective

Most pregnant women and key informants felt that the TM would be convenient as patients will be able to easily monitor their blood pressures at their homes. They further iterated that this would save them time and the costs they incur when visiting the public hospital to receive pregnancy care. Most importantly, pregnant women believed that there would be less disruption to their daily routine, and they would not have to depend on their husbands or in-laws for traveling to the hospital and for childcare responsibilities:

> *It would be quite beneficial as I have two children with me at home, whom I keep with someone when I come here (JPMC). Because my own mother cannot take care of them because she has a job, and my mother-in-law would not take care of them. So, for me, a blood pressure machine is beneficial in this way that I can give my record while staying at home and can be with my children too.* [Woman with pregnancy 02]

Sense of Empowerment in Own Health Care

Pregnant women at high risk for pre-eclampsia felt that they would be more confident about their health through the use of the TM program and would gain a sense of empowerment in their own health care decisions. Several key informants also articulated that the TM program might prove essential for increasing personal health responsibility, given that the program would encourage self-care behaviors among pregnant women. The act of monitoring, reporting blood pressures via phone, and receiving alerts and messages on appropriate self-care behaviors during high-risk pregnancy would, in turn, empower pregnant women at high risk:

> *Most important is women's empowerment. Anything that we do in the communities to have the women monitor their own health, send information about their own self or receive information about how to better care about themselves...do[es] empower them within the families and in the communities and I think that is really important as well.* [Digital health expert 09]

Barriers Influencing Adoption of TM and Potential Strategies

Lack of Willingness of Pregnant Women and Health Care Professionals to Use the TM Program

Some health care providers, including clinicians and nurses, were unwilling to use the TM program, given that it would require providers to perform additional tasks, whereas other health care providers showed a willingness to use the TM program and suggested implementing TM to leverage the high adoption of virtual care during the pandemic. Key informants expressed that health care providers would not readily accept the TM program, given that it will require them to invest time and effort to learn and use it:

> *So, I think the biggest problem is going to be that it will increase work for all the nurses and doctors...they already have so much work on their hands...trying any new gadget will increase time and will increase monitoring as well. So, I think one barrier will be, that they will not accept it readily because of the time being consumed in doing this.* [Digital health expert 02]

Although some key informants thought that implementing TM would lessen physician workload, most clinicians and nurses were reluctant to take complete responsibility for running the TM program at public sector hospitals, including regular and timely monitoring of blood pressure readings of pregnant women at high risk. Health care providers iterated that they are already overwhelmed with the substantial patient influx at the public hospital. Therefore, they suggested that the implementation of a new TM program would require the recruitment of separate staff to support TM program functioning:

> *Over here one thing is, that the patient burden is too much. So, maybe the doctor would not be able to see the daily readings or the data, which is received of the woman, or maybe she [the doctor] could not directly get in touch with the patient. Yes, but if there is someone specific [Clinician or Nurse], who will be overseeing these patients then that would be easier as they will have to see only those patients, who are already booked with them and with whom they would be doing telemedicine or this TM.* [Clinician 04]

Key informants believed that the senior health care providers who are accustomed to seeing patients in clinics might resist change and show reluctance in learning new technology. A clinician highlighted that the health care providers who are comfortable in seeing patients during in-person visits might believe that the use of the TM program may cause issues in identifying important signs that could be monitored during physical examination:

> *Because they are so much used to the manual part and they are so much used to the physical presence of the patient that they will be reluctant to use new technology...they are used to seeing patients in their clinics; thus, they will think that the TM may miss the physical examination sorts of things of the patient.* [Clinician 04]

Regarding willingness for TM program use among pregnant women, most pregnant women stated desire and trust in using the TM program to share their blood pressure readings with their providers regularly, whereas a few pregnant women were unable to recognize the usefulness of the TM program and showed a lack of willingness to use it because of anticipated trust issues and harms associated with the use of the TM program. Pregnant women were particularly concerned about the accuracy of blood pressure readings as the blood pressures would be measured using electric machines:

> *The disadvantage can be this that it is an electric thing and there is no reliability of it. It can give wrong readings...even if you take a new blood pressure machine.* [Woman with pregnancy 06]

A few pregnant women showed a lack of willingness to use the TM program as they were concerned that blood pressure machines and smartphones might emit radiation that could be harmful to their fetus:

> *It [TM program] is a bit risky because the child might get rays...like the usage of microwaves.* [Woman with pregnancy 05]

To address trust issues and safety concerns among pregnant women, key informants suggested that it is necessary to fully understand the needs and concerns of patients and providers before implementing a TM program. Furthermore, key informants suggested the provision of incentives to pregnant women and health care providers to promote and increase TM program use, such as paying for transportation costs:

> *In my opinion, in this, you shall need to give some incentive to pregnant women and providers at some stage. You need to give them some support to kick-start the project...Like the women be given some incentive like you place some transportation cost.* [Digital health expert 13]

Weak Technological Literacy to Use TM Program

The interviews also revealed concerns regarding the technological literacy of pregnant women and health care providers in using the TM program. Key informants believed that senior health care providers might find it difficult to use the TM program, whereas many young health care providers may find it easy to use. A digital health expert mentioned that most health care providers usually wish to continue their regular conventional clinical practices and stay limited in terms of the scope of their work:

> *I think it is mainly the interests that they would have in such kind of a thing, many health providers here*

> *want to stay limited in what they are doing, they just want to do regular practice and that is it.* [Digital health expert 09]

All study participants unanimously believed that the young, educated, and financially secure pregnant women would be more able to learn and use the TM program as opposed to older women who are uneducated and belong to a low socioeconomic class. Nonetheless, it was mentioned that such women would be able to receive help from immediate family members to participate in a TM program:

> *Those who would be educated would be able to use it, and those who would not be educated would not be able to use it.* [Woman with pregnancy 01]

To address concerns related to technological illiteracy, key informants and pregnant women suggested building the capacity of health care providers and pregnant women through adequate training and demonstrations. In addition, key informants and pregnant women highlighted the need to train people in support networks such as husbands, in-laws, daughters, and sisters on the use of the TM program to increase TM adoption among older women:

> *I mean what you have told me right now, theoretically, the idea sounds doable, what resources it involves I do not really know, but one important factor that needs to be kept in mind is that women and their family members have to be trained and you need to make sure that they do it the right way.* [Digital health expert 12]

Lack of Technological Infrastructure for Implementation of TM Program

Key informants mentioned that the weak technological infrastructure, including poor access to smartphones, the internet, and blood pressure machines, might hinder TM program implementation. On the other hand, some key informants verbalized that cell phone penetration has increased significantly as a growing number of pregnant women coming to public hospitals have either personal or shared access to basic mobile phones. To address concerns related to weak technological infrastructure, digital health experts highlighted the need to use basic mobile phones and their functions such as SMS text messaging and phone calls to engage and empower more women for TM program use. A digital health expert highlighted the importance of using the Global System for Mobile Communications to leverage basic mobile phones' SMS text messaging function, given that the vast majority have access to basic mobile phones:

> *I think that GSM is still better. There are many people in our communities, who are still dependent on non-smartphones and that is where GSM plays a very important role. So, if that data can very simply be entered or directly transferred from the machine to your GSM device, I would say that is a better way. This is what we did in our monitor that the data is connected, and it is sent through the text messaging to their own phones and to the phones of health providers rather than using any kind of internet because that is another limitation. Because I doubt, the most that they would be able to use the Bluetooth-enabled technologies and smartphones themselves. We have to make sure that it is very easy to use, very simple.* [Digital health expert 09]

Sociocultural Factors Affecting Program Use

When asked about anticipated sociocultural factors influencing TM program use, most women expressed that they might face restrictions on the use of the TM program from mothers-in-law and husbands. Most women mentioned that they would require permission from their husbands and in-laws before opting into the TM program. This perception was echoed by some key informants, who also thought that family members might discourage pregnant women from using the TM program and buying blood pressure machines. An MNCH specialist expressed that women are not allowed to use phones because of cultural reasons, and, therefore, these women can be reached through their husbands' phones:

> *We are getting some families who do not allow us to use phones because of cultural barriers. They [pregnant women] are not allowed to keep the mobile phones or use the internet and you know, messaging and everything on phones. So, their husbands had the phones, and these women are contacted through their husband's phones. The women are not directly allowed to keep the phone. So, this is a challenge too, the cultural one.* [MNCH specialist 10]

However, other pregnant women at high risk for pre-eclampsia and key informants verbalized that, in most cases, there would be no cultural restrictions, and women would be able to receive support from their immediate family members for TM program use. A digital health expert mentioned that it is only an assumption that the TM program will not be accepted in the community, and there will be cultural restrictions on the use of the TM program:

> *I am saying that it is an assumption that people are not going to acknowledge it in the communities, my experience is that they are extremely receptive if the person who is giving that technology is known to them. So, the nurse is familiar to them they will be excited, they will be interested. Everyone wants to try new gadgets it is a human phenomenon, and they are no different. I do not see a cultural barrier in trying a new instrument.* [Digital health expert 02]

To respond to sociocultural factors associated with the use of the TM program, pregnant women and key informants, including clinicians, nurses, and digital health experts, suggested sensitizing the community and family members to encourage TM program use for supporting pregnant women at high risk. Digital health experts suggested having a team to work around the sensitization component. The main component of sensitization and advocacy was deemed to be the involvement of families in the decision-making process at the very onset of recruitment into the program for improving the uptake of the TM program among pregnant women as family members would be made responsible for sending the information to providers:

> *As I said that whatever these sorts of things need to be initiated. If at the first go it should be like the decision including the family, then acceptance will be very good. But if we just tell the woman and she goes back and tells, then I do not think there will be acceptance. They [family] may think that they are using machines maybe it will impact their privacy. And these sorts of things. So, I always include one of the family members and make them responsible, instead of the pregnant woman. So, that works.* [Clinician 04]

Considerations for Implementing the TM Program

Features and Ease of Use of TM Program

Clinicians and pregnant women suggested that the TM program should be designed in the local languages (Urdu or Sindhi), keeping in mind that health care providers, pregnant women, and people from their support network are proficient in at least one of these two languages. In addition, pregnant women and clinicians suggested a range of relevant maternal and fetal parameters to be included in the TM program, including the weight, hemoglobin level, complete blood count, calcium and vitamin D levels, and pre-eclampsia disease symptoms of pregnant women, as well as fetus weight and heartbeat. Some digital health experts, clinicians, and a few pregnant women suggested establishing an algorithm in the TM program, which could serve as a reminder and warning sign alert system to empower pregnant women, facilitate the patient's role in decision-making, and serve as a real-time clinical decision support system for health care providers:

> *And I do not know how this device or instrument would look like or how it will monitor and give the results but if it is in the form of color...maybe it can show red, yellow, and green color bars. And that actually shows the mother, like she is okay, she is a bit in the not okay or maybe in danger.* [MNCH specialist 03]

Pregnant women and digital health experts emphasized the inclusion of an educational component in the TM program to address the issue of poor awareness and knowledge about disease conditions and their management. A digital health expert stressed that the educational component should serve the purpose of educating pregnant women through SMS text messages even if the device monitoring is stopped after the intervention period.

Handling, Maintenance, and Sustainability of TM Program

When asked about handling and maintenance of the TM program, pregnant women assured that blood pressure machines would be gently handled and safely stored. However, clinicians and nurses anticipated issues with the handling and maintenance of the machines. Clinicians and nurses at JPMC expressed that there might be issues with retrieving smartphones and blood pressure machines from pregnant women because of the high loss to follow-up rate among pregnant women. A clinician iterated that even if the equipment is retrieved, it may get damaged because of poor handling by pregnant women and family members. Through past experiences, JPMC nurses and clinicians suggested either charging a minimum fee for equipment use or keeping their official documents such as national identity cards and marriage certificates until equipment is returned to create accountability on the part of pregnant women:

> *There should be something that can guarantee the equipment return from the patient. I would suggest keeping an ID card or marriage certificate.* [Clinician 01]

Key informants suggested that the sustainability of the TM program would require a substantial commitment from the public health department for integrating the TM program into existing workflows. A key informant suggested developing a sustainability plan with associated documentation at the provincial level to ensure TM program sustainability through the fiscal year budget:

> *I think that monitoring would also be extremely important. I mean there is always the risk when you do such programs that the devices do not reach the right people. If it is made part of the system, right from the planners and their commitment down to the health providers, and the supervision level to make sure that these devices are used properly and goes to the right people.* [Digital health expert 09]

Discussion

Principal Findings

This study provides an in-depth investigation into the perspectives, needs, and preferences of a mobile phone-based TM program for pregnant women at high risk for pre-eclampsia in Karachi, Pakistan. The interviewed pregnant women at high risk, nurses, clinicians, MNCH specialists, and digital health experts perceived an opportunity to establish a TM program for early identification of pregnancy complications and prompt treatment, care continuity, self-management, and clinical decision-making. Although some providers thought that there were potential benefits of implementing a TM program, there was no consensus on whether it would increase or decrease the clinical workload. The study identified the need for having a dedicated clinician to help with the operationalization of the TM program. A cost–benefit evaluation could help determine whether the recruitment of additional staff is justified to support the implementation of a TM program. Pregnant women stated their willingness to use the TM program and thought that TM would be convenient and cost-effective and provide a sense of empowerment in their own health care. However, some pregnant women at high risk were apprehensive of TM program malfunctioning and safety concerns associated with the use of smartphones and automated blood pressure machines. The study data indicated varied capacity to learn and use new TM programs among pregnant women because of sociocultural and financial factors and technological illiteracy. The study identified the need to provide technological support for pregnant women and providers and sensitize the community and family members to address sociocultural barriers. Pregnant women and clinicians suggested the establishment of a warning sign alert system as part of the TM program and the inclusion of some

maternal and fetal parameters to be monitored in the TM program. To ensure that TM equipment would be returned, the key informants suggested creating accountability on the part of pregnant women by charging a minimum fee for equipment use.

Comparison With Previous Research

Previous research has informed the needs for TM programs, mainly in high-income countries [10,34-37]. This study was the first to provide unique insights into the needs of the TM program for pregnant women at high risk in an LMIC such as Pakistan. Our study supported some findings from previous studies conducted in high-income countries regarding the convenience associated with TM. For instance, the study by Van Den Heuvel et al [10] conducted in the Netherlands reported that compared with the experiences of hospital admission in high-risk pregnancy, TM allowed pregnant women to be in a comforting and private environment during an anxious time in their lives. Consistent with the perceptions voiced in our study, another study by Van Den Heuvel et al [34] on TM for complicated pregnancies in the Netherlands highlighted the advantages of monitoring from home, such as reduced stress, increased rest periods for patients, reduction of admission, and possible reduction of costs. Similar to our study, the Primer and Provider Selection Guide on telehealth (2013), a whitepaper developed by the LeadingAge Center for Aging Services Technologies in a high-income country, highlighted that TM is a useful tool for empowering pregnant women at high risk by encouraging them to identify and report symptoms of exacerbation of their condition [35]. In addition, consistent with our study, the Primer and Provider Selection Guide on telehealth also recognized the value of TM in terms of supporting clinician decision-making and providing prompt treatment [35].

Our study reported some caveats to the willingness of pregnant women living in LMICs to use the TM program, such as TM program malfunctioning and safety concerns. Consistent with our findings, the Primer and Provider Selection Guide on telehealth emphasized the use of safe monitoring technologies to provide an enhanced sense of security, prolonged independence, and improved quality of life [35]. In our study, providers indicated reservations in using TM because of the anticipated increase in workload associated with responding to alerts generated through the TM program. This is consistent with the study by Anderson et al [35] on *unpacking TM work*, which confirmed that TM is time consuming and considered a burden on the clinical workload [36]. Their study highlighted that telephone calls have an important function in TM, such as supporting clinical decision-making and enabling the provision of patient-centric care. However, the telephone calls between patients and clinicians were found to increase the time spent in remote monitoring [36].

Digital health experts in our study highlighted the need to use basic mobile phones and SMS text messaging functions to address issues associated with weak technological infrastructure. The Centers for Disease Control and Prevention (health protection agency in the United States) Guide (2019) on *using technologies for data collection and management* also recommends that the choice of technology platforms should be driven by the existing technological infrastructure, goals of the investigation, and training and skills of available staff [37]. To address the issue of technological literacy for TM program use, our study findings suggested building the capacity of pregnant women and health care providers through training and demonstrations. This is consistent with the Primer and Provider Selection Guide, which emphasized the value of clearly defining the new model of care and preparing staff and end users through training and support plans before starting any new program [35].

In contrast to studies conducted in high-income countries, our study found some interesting findings on the sociocultural factors influencing TM program use in the LMIC context. Our study highlighted that pregnant women in LMICs might face restrictions from mothers-in-laws and husbands or require permission from them for using TM program. A similar finding was reported in the qualitative study by Qureshi et al [38] on health-seeking behaviors during pregnancy. The study concluded that maternal health care use is heavily influenced by social, economic, and cultural factors in rural Pakistani communities [38]. Their study revealed that principal decision makers for health care use are husbands and mothers-in-laws, and thus, women are expected to follow their decisions [38]. To respond to sociocultural factors associated with the use of the TM program in LMICs, our study identified the need to sensitize the community and family members on the usefulness of the TM program and to involve families in the decision-making process at the onset of recruitment into the TM program to help with its uptake.

Recommendations

Our findings suggest that TM could be successfully designed, implemented, and sustained in an LMIC such as Pakistan to support the early identification of pregnancy complications and prompt treatment. We offer the following recommendations for clinicians, MNCH specialists, digital health experts, and policy makers to consider in developing a TM program to address the needs of pregnant women at high risk with pre-eclampsia and eclampsia in LMICs:

1. Building the capacity of pregnant women and health care providers through training and demonstrations could improve technological literacy for TM program use.
2. There is a need to sensitize the community and family members on the usefulness of the TM program to address sociocultural factors affecting TM program use. This could be done at the very onset of recruitment into the TM program to improve its uptake.
3. Clinicians and nurses are in a unique position to use the TM program; however, this might increase their clinical workload associated with the TM of pregnant women. Therefore, an approach using user-centered design and phased implementation could help to determine the impact of TM program use on current clinical workload and establish the need for additional staff to ensure adequate adoption of the TM program.
4. Future work should consider greater levels of co-design and the engagement of consumer representatives to ensure the development of a context-specific TM program. For

instance, given that the vast majority of the population in LMICs such as Pakistan have access to basic mobile phones, basic mobile phones and SMS text messaging functions instead of smartphones should be considered for use in a TM program.

Strengths and Limitations

A strength of the study is the use of multiple data sources (pregnant women, digital health experts, MNCH specialists, clinicians, and health care providers), which assisted researchers in data triangulation and identification of converging and diverging lines of inquiry. Another strength was that the primary researcher maintained a reflexive journal during all stages of the research to recognize and acknowledge biases during the research process.

Limitations to the study included that it was conducted with a focus on TM within a single setting (JPMC), which could limit the transferability of the findings to other settings [39]. However, this study may provide insights into other similar public hospitals across Pakistan and in other LMICs that are interested in supporting pregnant women at high risk through TM. Second, caregivers were not allowed to be present during interviews with pregnant women because of the COVID-19 pandemic, which may have limited the exploration of shared narratives on the needs of the TM program. Third, the researchers were unable to conduct member checking with study participants as it would have been exceedingly difficult to contact patient participants after the initial interview. However, at the end of each interview, the primary researcher restated and summarized the information obtained during the interview with the aid of interview notes to ensure that the study data resonated with the participant's perspectives and experiences. Finally, some participants were interviewed via the web through Zoom technology. Hence, the researchers did not have the opportunity to build rapport with the respondents or obtain nonverbal cues during interviews.

Conclusions

The pregnant women at high risk for pre-eclampsia, clinicians, MNCH specialists, and digital health experts thought that a mobile phone–based TM program may be feasible to implement, largely as mobile phones are becoming increasingly pervasive in LMICs. The pregnant women and health care providers were willing to use a mobile phone–based TM program as they perceived many benefits, including early identification of pregnancy complications and prompt treatment, convenience, cost-effectiveness, increased sense of empowerment in their own health care, and improved care continuity. However, participants cited several caveats to their willingness to use the TM program: the monitoring system would have to be easy to use, the TM program must be safe, and the blood pressure measuring device must be reliable. Recommendations from this work include the following: (1) build the capacity of patients and providers on TM program use, (2) sensitize the community and family members on the usefulness of the TM program, (3) use an approach incorporating user-centered design and phased implementation to determine the clinical workload and establish the need for additional staff for TM program monitoring, and (4) ensure greater levels of co-design and the engagement of consumer representatives. The findings from this study highlight the perceived feasibility of a mobile phone–based TM program for pregnant women at high risk for pre-eclampsia and provide insights that can be directly used for the design of future TM programs with the aim of reducing mortality and morbidity from pre-eclampsia and eclampsia in LMICs. A future feasibility study will be conducted based on the findings of this study to determine if an effectiveness trial is warranted.

Conflicts of Interest

None declared.

References

1. Khowaja AR, Qureshi RN, Sheikh S, Zaidi S, Salam R, Sawchuck D, et al. Community's perceptions of pre-eclampsia and eclampsia in Sindh Pakistan: a qualitative study. Reprod Health 2016 Jun 08;13 Suppl 1(S1):36 [FREE Full text] [doi: 10.1186/s12978-016-0136-x] [Medline: 27357953]
2. Sustainable development goal 3: ensure healthy lives and promote well-being for all at all ages. United Nations. URL: https://sustainabledevelopment.un.org/sdg3 [accessed 2022-02-03]
3. Hinton L, Hodgkinson J, Tucker KL, Rozmovits L, Chappell L, Greenfield S, et al. Exploring the potential for introducing home monitoring of blood pressure during pregnancy into maternity care: current views and experiences of staff-a qualitative study. BMJ Open 2020 Dec 01;10(12):e037874 [FREE Full text] [doi: 10.1136/bmjopen-2020-037874] [Medline: 33262186]
4. Perry H, Sheehan E, Thilaganathan B, Khalil A. Home blood-pressure monitoring in a hypertensive pregnant population. Ultrasound Obstet Gynecol 2018 Apr 22;51(4):524-530 [FREE Full text] [doi: 10.1002/uog.19023] [Medline: 29468771]
5. Xydopoulos G, Perry H, Sheehan E, Thilaganathan B, Fordham R, Khalil A. Home blood-pressure monitoring in a hypertensive pregnant population: cost-minimization study. Ultrasound Obstet Gynecol 2019 Apr 12;53(4):496-502 [FREE Full text] [doi: 10.1002/uog.19041] [Medline: 29516615]
6. Aquino M, Munce S, Griffith J, Pakosh M, Munnery M, Seto E. Exploring the use of telemonitoring for patients at high risk for hypertensive disorders of pregnancy in the antepartum and postpartum periods: scoping review. JMIR Mhealth Uhealth 2020 Apr 17;8(4):e15095 [FREE Full text] [doi: 10.2196/15095] [Medline: 32301744]
7. Khalil A, Perry H, Lanssens D, Gyselaers W. Telemonitoring for hypertensive disease in pregnancy. Expert Rev Med Devices 2019 Aug 18;16(8):653-661. [doi: 10.1080/17434440.2019.1640116] [Medline: 31282755]

8. World Health Organization. WHO Recommendations on Antenatal Care for a Positive Pregnancy Experience. Geneva: World Health Organization; 2016.
9. van den Heuvel JF, Kariman SS, van Solinge WW, Franx A, Lely AT, Bekker MN. SAFE@HOME - Feasibility study of a telemonitoring platform combining blood pressure and preeclampsia symptoms in pregnancy care. Eur J Obstet Gynecol Reprod Biol 2019 Sep;240:226-231. [doi: 10.1016/j.ejogrb.2019.07.012] [Medline: 31330428]
10. van den Heuvel JF, Teunis CJ, Franx A, Crombag NM, Bekker MN. Home-based telemonitoring versus hospital admission in high risk pregnancies: a qualitative study on women's experiences. BMC Pregnancy Childbirth 2020 Feb 04;20(1):77 [FREE Full text] [doi: 10.1186/s12884-020-2779-4] [Medline: 32019499]
11. Lapidus AM. Effects of preeclampsia on the mother, fetus and child. Contemporary OB/GYN. 2010. URL: https://www.contemporaryobgyn.net/view/effects-preeclampsia-mother-fetus-and-child [accessed 2022-02-03]
12. Jafar TH, Gandhi M, de Silva HA, Jehan I, Naheed A, Finkelstein EA, et al. A community-based intervention for managing hypertension in rural South Asia. N Engl J Med 2020 Feb 20;382(8):717-726. [doi: 10.1056/nejmoa1911965]
13. Qureshi RN, Sheikh S, Hoodbhoy Z, Sharma S, Vidler M, Payne BA, CLIP Pakistan Working Group. Community-level interventions for pre-eclampsia (CLIP) in Pakistan: a cluster randomised controlled trial. Preg Hypertens 2020 Oct;22:109-118 [FREE Full text] [doi: 10.1016/j.preghy.2020.07.011] [Medline: 32777710]
14. Hinton L, Tucker KL, Greenfield SM, Hodgkinson JA, Mackillop L, McCourt C, et al. Blood pressure self-monitoring in pregnancy (BuMP) feasibility study; a qualitative analysis of women's experiences of self-monitoring. BMC Preg Childbirth 2017 Dec 19;17(1):427 [FREE Full text] [doi: 10.1186/s12884-017-1592-1] [Medline: 29258469]
15. O'Brien E, Rauf Z, Alfirevic Z, Lavender T. Women's experiences of outpatient induction of labour with remote continuous monitoring. Midwifery 2013 Apr;29(4):325-331. [doi: 10.1016/j.midw.2012.01.014] [Medline: 23159160]
16. Shorey S, Yang YY, Dennis C. A mobile health app-based postnatal educational program (Home-but not Alone): descriptive qualitative study. J Med Internet Res 2018 Apr 19;20(4):e119 [FREE Full text] [doi: 10.2196/jmir.9188] [Medline: 29674314]
17. Alves DS, Times VC, da Silva EM, Melo PS, Novaes MD. Advances in obstetric telemonitoring: a systematic review. Int J Med Inform 2020 Feb;134:104004. [doi: 10.1016/j.ijmedinf.2019.104004] [Medline: 31816495]
18. Lanssens D, Vandenberk T, Thijs IM, Grieten L, Gyselaers W. Effectiveness of telemonitoring in obstetrics: scoping review. J Med Internet Res 2017 Sep 27;19(9):e327 [FREE Full text] [doi: 10.2196/jmir.7266] [Medline: 28954715]
19. Feroz A, Saleem S, Seto E. Exploring perspectives, preferences and needs of a telemonitoring program for women at high risk for preeclampsia in a tertiary health facility of Karachi: a qualitative study protocol. Reprod Health 2020 Sep 15;17(1):135 [FREE Full text] [doi: 10.1186/s12978-020-00979-8] [Medline: 32928235]
20. Slevin P, Kessie T, Cullen J, Butler MW, Donnelly SC, Caulfield B. Exploring the barriers and facilitators for the use of digital health technologies for the management of COPD: a qualitative study of clinician perceptions. Quart J Med 2020 Mar 01;113(3):163-172. [doi: 10.1093/qjmed/hcz241] [Medline: 31545374]
21. Fishbein JN, Nisotel LE, MacDonald JJ, Amoyal Pensak N, Jacobs JM, Flanagan C, et al. Mobile application to promote adherence to oral chemotherapy and symptom management: a protocol for design and development. JMIR Res Protoc 2017 Apr 20;6(4):e62 [FREE Full text] [doi: 10.2196/resprot.6198] [Medline: 28428158]
22. Alberts NM, Badawy SM, Hodges J, Estepp JH, Nwosu C, Khan H, et al. Development of the incharge health mobile app to improve adherence to hydroxyurea in patients with sickle cell disease: user-centered design approach. JMIR Mhealth Uhealth 2020 May 08;8(5):e14884 [FREE Full text] [doi: 10.2196/14884] [Medline: 32383683]
23. Huth HB, Skeens R, Anders S, Herzberg S, Simpson C, Novak L, et al. Health management in the home: a qualitative study of pregnant women and their caregivers. J Patient Exp 2020 Dec 14;7(6):1227-1233 [FREE Full text] [doi: 10.1177/2374373520948442] [Medline: 33457569]
24. Boddy D, King G, Clark JS, Heaney D, Mair F. The influence of context and process when implementing e-health. BMC Med Inform Decis Mak 2009 Jan 30;9(1):9 [FREE Full text] [doi: 10.1186/1472-6947-9-9] [Medline: 19183479]
25. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000 Aug;23(4):334-340. [doi: 10.1002/1098-240x(200008)23:4<334::aid-nur9>3.0.co;2-g] [Medline: 10940958]
26. 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]
27. Creswell JW, Hanson WE, Clark Plano VL, Morales A. Qualitative research designs. Counsel Psychol 2016 Jun 30;35(2):236-264. [doi: 10.1177/0011000006287390]
28. Hypertension in pregnancy. National Institute for Health and Care Excellence (NICE) Guidelines. URL: https://www.nice.org.uk/guidance/qs35/chapter/Quality-statement-2-Antenatal-assessment-of-pre-eclampsia-risk [accessed 2022-02-03]
29. Spradley J. The Ethnographic Interview. Long Grove, Illinois: Waveland Press; 2016.
30. Creswell JW, Poth CN. Qualitative Inquiry and Research Design: Choosing Among Five Approaches. Thousand Oaks, CA: Sage Publications Inc; 2017.
31. Yin RK. Case Study Research and Applications: Design and Methods. Thousand Oaks, CA: SAGE Publications; 2017.
32. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov 01;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
33. Patton M. Enhancing the quality and credibility of qualitative analysis. Health Serv Res 1999 Dec;34(5 Pt 2):1189-1208 [FREE Full text] [Medline: 10591279]

34. van den Heuvel JF, Ayubi S, Franx A, Bekker MN. Home-based monitoring and telemonitoring of complicated pregnancies: nationwide cross-sectional survey of current practice in the Netherlands. JMIR Mhealth Uhealth 2020 Oct 28;8(10):e18966 [FREE Full text] [doi: 10.2196/18966] [Medline: 33112250]
35. Telehealth and remote patient monitoring for long-term and post-acute care: a primer and provider selection guide. A LeadingAge CAST Report. 2013. URL: https://www.mtelehealth.com/wp-content/uploads/2015/10/TelehealthandRemotePatientMonitoringRPMforLongTermandPostAcuteCareA PrimerandProviderSelectionGuide2013.pdf [accessed 2022-02-03]
36. Andersen TO, Nielsen KD, Moll J, Svendsen JH. Unpacking telemonitoring work: workload and telephone calls to patients in implanted cardiac device care. Int J Med Inform 2019 Sep;129:381-387 [FREE Full text] [doi: 10.1016/j.ijmedinf.2019.06.021] [Medline: 31445281]
37. Hamilton JJ, Hopkins RS. Using technologies for data collection and management. In: The CDC Field Epidemiology Manual. Oxford, United Kingdom: Oxford University Press; 2019.
38. Qureshi RN, Sheikh S, Khowaja AR, Hoodbhoy Z, Zaidi S, Sawchuck D, CLIP Working Group. Health care seeking behaviours in pregnancy in rural Sindh, Pakistan: a qualitative study. Reprod Health 2016 Jun 08;13 Suppl 1(S1):34 [FREE Full text] [doi: 10.1186/s12978-016-0140-1] [Medline: 27356863]
39. Donmoyer R. Generalizability and the single-case study. In: Case Study Method. Thousand Oaks, CA: SAGE Publications; 2009:45-68.

Abbreviations

AKU: Aga Khan University
CLIP: Community-Level Interventions for Pre-eclampsia
JPMC: Jinnah Post Graduate Medical Center
LMIC: low- to middle-income country
MNCH: maternal, neonatal, and child health
OB-GYN: obstetrics and gynecology
TM: telemonitoring

Edited by A Mavragani; submitted 28.07.21; peer-reviewed by T Turk, E Villalba-Mora; comments to author 27.12.21; revised version received 08.01.22; accepted 12.01.22; published 24.02.22.

Please cite as:
Feroz AS, De Vera K, D Bragagnolo N, Saleem S, Bhutta Z, Seto E
Understanding the Needs of a Mobile Phone–Based Telemonitoring Program for Pregnant Women at High Risk for Pre-Eclampsia: Interpretive Qualitative Description Study
JMIR Form Res 2022;6(2):e32428
URL: https://formative.jmir.org/2022/2/e32428
doi:10.2196/32428
PMID:35200152

Original Paper

A Smartphone-Based Information Communication Technology Solution for Primary Modifiable Risk Factors for Noncommunicable Diseases: Pilot and Feasibility Study in Norway

Inger Torhild Gram[1,2], MD, PhD; Guri Skeie[2], PhD; Sunday Oluwafemi Oyeyemi[2], MD, PhD; Kristin Benjaminsen Borch[2], PhD; Laila Arnesdatter Hopstock[2], PhD; Maja-Lisa Løchen[2], MD, PhD

[1]Norwegian Centre for E-health Research, University Hospital of North Norway, Tromsø, Norway

[2]Department of Community Medicine, Faculty of Health Sciences, UiT The Arctic University of Norway, Tromsø, Norway

Corresponding Author:
Inger Torhild Gram, MD, PhD
Norwegian Centre for E-health Research
University Hospital of North Norway
Forskningsparken i Breivika, 3rd floor
Sykehusvn. 23
Tromsø, 9019
Norway
Phone: 47 92401177
Email: inger.gram@ehealthresearch.no

Abstract

Background: Cardiovascular diseases, cancers, chronic respiratory diseases, and diabetes are the 4 main noncommunicable diseases. These noncommunicable diseases share 4 modifiable risk factors (tobacco use, harmful use of alcohol, physical inactivity, and unhealthy diet). Short smartphone surveys have the potential to identify modifiable risk factors for individuals to monitor trends.

Objective: We aimed to pilot a smartphone-based information communication technology solution to collect nationally representative data, annually, on 4 modifiable risk factors.

Methods: We developed an information communication technology solution with functionalities for capturing sensitive data from smartphones, receiving, and handling data in accordance with general data protection regulations. The main survey comprised 26 questions: 8 on socioeconomic factors, 17 on the 4 risk factors, and 1 about current or previous noncommunicable diseases. For answers to the continuous questions, a keyboard was displayed for entering numbers; there were preset upper and lower limits for acceptable response values. For categorical questions, pull-down menus with response options were displayed. The second survey comprised 9 yes-or-no questions. For both surveys, we used SMS text messaging. For the main survey, we invited 11,000 individuals, aged 16 to 69 years, selected randomly from the Norwegian National Population Registry (1000 from each of the 11 counties). For the second survey, we invited a random sample of 100 individuals from each county who had not responded to the main survey. All data, except county of residence, were self-reported. We calculated the distribution for socioeconomic background, tobacco use, diet, physical activity, and health condition factors overall and by sex.

Results: The response rate was 21.9% (2303/11,000; women: 1397/2263; 61.7%, men: 866/2263, 38.3%; missing: 40/2303, 1.7%). The median age for men was 52 years (IQR 40-61); the median age for women was 48 years (IQR 35-58). The main reported reason for nonparticipation in the main survey was that the sender of the initial SMS was unknown.

Conclusions: We successfully developed and piloted a smartphone-based information communication technology solution for collecting data on the 4 modifiable risk factors for the 4 main noncommunicable diseases. Approximately 1 in 5 invitees responded; thus, these data may not be nationally representative. The smartphone-based information communication technology solution should be further developed with the long-term goal to reduce premature mortality from the 4 main noncommunicable diseases.

(JMIR Form Res 2022;6(2):e33636) doi:10.2196/33636

KEYWORDS

eHealth; feasibility study; modifiable risk factor; noncommunicable disease; pilot study; smartphone-based information communication technology solution; short text message service; feasibility; risk; factor; information communication technology; smartphone; development; monitoring

Introduction

Cardiovascular diseases, cancers, chronic respiratory diseases, and diabetes are commonly grouped as the main noncommunicable diseases as they are the world's biggest killers [1-3]. These noncommunicable diseases share 4 modifiable risk factors (tobacco use, harmful use of alcohol, physical inactivity, and unhealthy diet). An important part of the United Nation's Sustainable Development Goal target 3.4 is to reduce premature mortality from the 4 main noncommunicable diseases by one-third relative to 2015 levels, by 2030 [4]. Encouraging reduced tobacco use, less harmful use of alcohol, increased physical activity, and healthy diet are simple and cost-effective measures to reduce premature death and disability from the 4 main noncommunicable diseases [5]. Surveillance of the 4 modifiable risk factors is crucial to be able to prevent and control premature death from the 4 main noncommunicable diseases according to the 2030 Sustainable Development Goal agenda [3].

In 2013, the World Health Assembly, the decision-making body of the World Health Organization, adopted a Global Monitoring Framework for noncommunicable diseases with 25 key indicators to track progress in prevention and control of noncommunicable diseases [6]. Before this, the World Health Organization had already introduced the STEPwise approach [7] for the surveillance of noncommunicable disease risk factors. Step 1 included self-reported demographic and behavioral risk factors as well as history of noncommunicable diseases and related conditions; step 2 included physical measurements; step 3 consisted of biochemical measurements.

Statistics Norway performs annual surveys, with representative samples, on tobacco use [8]. Furthermore, there have been several large population surveys [9-12] conducted in various regions and counties during the last fifty years, repeated at approximately 8-year intervals, which have collected some of the data included in the 3 steps of the STEPwise approach [11,12]. In addition, special surveys have been conducted, usually with 10-year intervals, to collect data on detailed dietary intake [9]. Special surveys on physical and sedentary activity have also been conducted with more than 5-year intervals [10]. In summary, there has been a lack of annual data on tobacco use, the harmful use of alcohol, physical inactivity, and unhealthy diet from a nationally representative sample. The Norwegian legislation on public health work [13] requires counties and municipalities to have an overview regarding risk factors, health conditions, and measures to promote health in their respective populations.

In Norway, more than 95% of individuals aged 16 to 54 years, and between 74% to 88% of those aged 55 to 74 years, have smartphones [14]. Short smartphone surveys have the potential to identify modifiable risk factors for individuals and monitor trends. Our main objective was to develop a smartphone-based information communication technology solution with functionalities for collecting data annually on the 4 modifiable risk factors. The secondary objective was to collect nationally representative data.

Methods

Study Design

This pilot study, which included a smartphone-based solution, a website, and 2 smartphone surveys, was developed over a 2-year period and conducted during fall 2019.

Development of the Smartphone-Based Information Communication Technology Solution

The details of the technical and architectural parts of the solution were developed by a private enterprise (Healthcom). For answers to the continuous questions, a keyboard was displayed for entering numbers (Figure 1); there were preset upper and lower limits for acceptable response values (Figure 2). For categorical questions, pull-down menus with response options were displayed (Figure 3).

The information communication technology solution was intended for capturing sensitive data from smartphones and was developed in accordance with general data protection regulations [15]. The cloud-based information communication technology solution automatically created a unique identification number for each respondent, when the initial SMS dialogue started. If the respondent clicked on the survey link later, the respondent's unique identifier was detected, and the participant could continue to fill in the answers. When the survey was submitted, the responses were anonymized. Subsequently, for analyses and storage, data were transferred to the Research Electronic Data Capture database hosted by the University Hospital of North Norway (Northern Norway Regional Health Authority server system) [16,17]. Communication with the database (over the internet and hospital intranet) was encrypted, and 2-factor authentication was required for researchers retrieving the data.

Figure 1. Screenshot of data entry keyboard.

Figure 2. Screenshot of smartphone-based survey.

Figure 3. Screenshot of pull-down menus.

Website

We developed a website to inform invitees and other interested parties about the study and survey (Figure 4). The website contained information about the background of the study (Figure 5), the study population, ethical assessments, survey results (Figure 6), the status of ongoing plans, study funding, and collaborating partners of the study.

Figure 4. Screenshot of information for invitees.

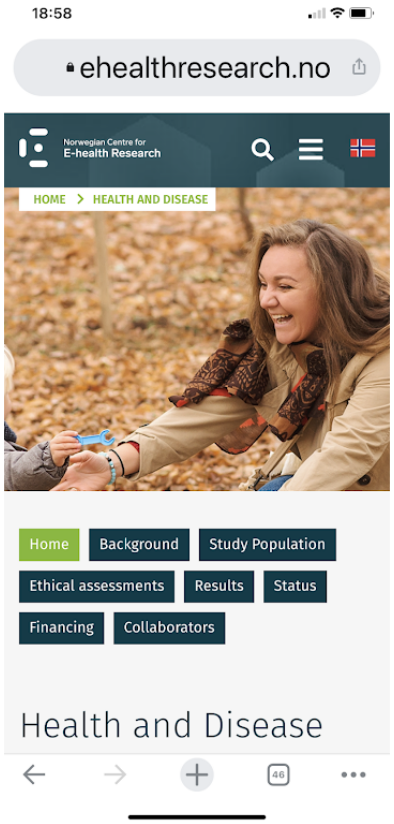


Figure 5. Screenshot of background information.

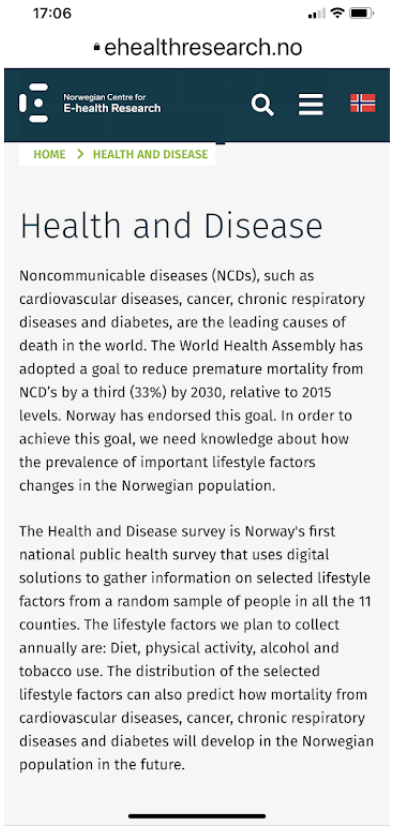
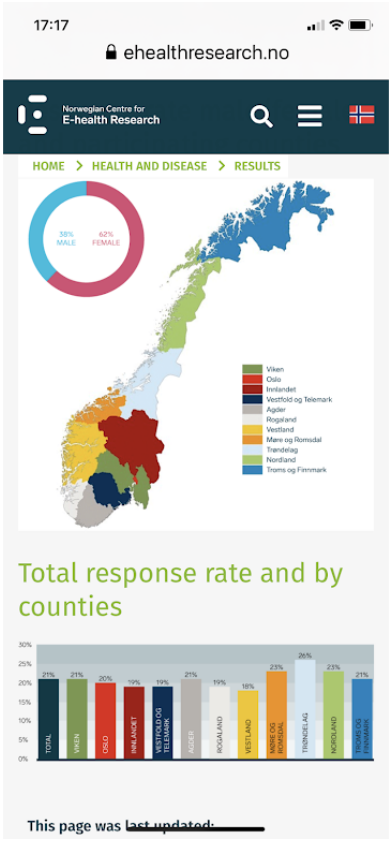


Figure 6. Screenshot of results presented to users.



Smartphone Surveys

Overview

We designed 2 mobile phone surveys—the first was sent to all invitees (main), and the second was sent only to a random sample of invitees who had not responded to the main survey. A total of 4 sets of SMS messages were sent: 1 request to participate in the main survey, 2 reminders, and 1 request to participate in the second survey (Multimedia Appendix 1). Invitees could obtain additional information when they accessed the survey and the link to the website (Multimedia Appendix 2).

Main Survey

The survey comprised 26 questions in Norwegian. There were 8 questions on socioeconomic background factors: year of birth, sex, education in years (≤10, 11-13,14-16, ≥17), marital status (single, married or cohabitating, divorced, widow or widower), number of persons in the household over and under 18 years (0,1,2,3, ≥4), gross household income in Norwegian Kroner (1 NOK equals approximately US $0.11) during the previous year (≤350,000 NOK, 351,000-550,000 NOK, 551,000-750,000 NOK, 751,000-1,000,000 NOK, ≥1,000,000 NOK), and occupational status (full time work, part-time work, student, retired, home-keeper, military service, and miscellaneous social benefits).

There were 17 questions on the 4 main risk factors. The questionnaire contained 3 questions about cigarette, snus, and e-cigarette use (never, former occasionally, former daily, current occasionally, current daily) and 1 question about alcohol consumption (yes, no). Consumers of alcohol were also asked the first 3 questions from the Alcohol Use Disorders Identification Test [18]—frequency (≤1 time per month, 2-4 times per month, 2-3 times per week, ≥4 times per week), number of units usually consumed (1-2, 3-4, 5-6, 7-9, ≥10), and frequency of occasions of consumption of ≥6 units of alcohol (never, less than monthly, monthly, weekly, daily or almost daily). There were 4 questions related to physical activity from the short version of the International Physical Activity Questionnaire [19]—number of days of strenuous physical activity in the last 7 days (0, 1-2, 3-4, 5-6, 7), number of days of moderate physical activity in the last 7 days (0, 1-2, 3-4, 5-6, 7), number of days of walking for ≥10 minutes in the last 7 days (0, 1-2, 3-4, 5-6, 7), hours spent sitting (excluding sleeping hours) on a regular weekday in the last 7 days (0-2, 3-5, 6-8, 9-11, 12-14)—and 7 questions on dietary intake frequency—servings per day of fruits and berries (0, 1, 2, 3, 4, ≥5), servings per day of lettuce and vegetables (0, 1, 2, 3, 4, ≥5), number of glasses per day sugar-sweetened drinks (0, 1, 2, 3, 4, 5, 6, ≥7), number of times fish and fish products are eaten per week (0, 1, 2, 3, 4, ≥5), number of times red meat is eaten per week (0, 1, 2, 3, 4, ≥5), number of times processed meat is eaten per week (0, 1, 2, 3, 4, ≥5), and how often extra salt is added to food before eating (never, occasionally, often, always).

There was 1 question asking if respondents have or have had any of the following conditions: high blood pressure, high cholesterol, atrial fibrillation, myocardial infarction, heart failure, stroke, chronic respiratory disease, cancer, diabetes, or none of these. In addition, there was an open box for comments.

The estimated time for completing the main survey was approximately 5 minutes.

Second Survey

The second survey comprised 9 statements that could be answered in the affirmative or negative: "I never answer such surveys"; "I did not want to answer several of the questions"; "I think the questions were difficult to answer"; "I was unsure who the SMS came from"; "I was expecting a login with BankID" (BankID is an electronic identification scheme in Norway for safe log in); "I was afraid that privacy was not taken care of"; "I had not heard of the Norwegian Centre for E-health Research"; "I had not heard of Healthcom"; "I thought the SMS I received looked like spam." The estimated time for completing this survey was approximately 1 minute. There was also an open box for comments.

The main survey was conducted on 11 consecutive days, with 2 reminders sent 5 and 10 days after the original invitation. The second survey was conducted 1 week after the main survey was completed.

Ethics

The Regional Committee for Medical and Health Research Ethics concluded that approval for this study was not necessary because the study fell outside the Norwegian Health Research Act. The study was approved by the Data Protection Section of the University Hospital of North Norway. All participants provided consented before answering any survey questions.

Recruitment

A total of 11,000 individuals (an equal number of men and women) aged 16 to 69 years, with 1000 from each of the 11 counties in Norway, were selected randomly from the Norwegian National Population Registry. All Norwegian residents have an 11-digit personal identification number in the National Population Registry. This registry contains the name, address, sex, age, and mobile phone number of each person. Persons who did not have a registered mobile phone number were replaced. For each person, name, personal identification number, phone number, sex, age, year of birth, and county of residence were retained. Subsequently, the system created a unique survey number for each participant. The Data Protection Section of the University Hospital of North Norway provided a server area specifically for this project where the linkage file could be stored. For the second survey, SMS text messages were sent to a random sample of 100 invitees who had not responded to the main survey from each of the 11 counties.

Dissemination

The communication group at the Norwegian Centre for E-health Research sent a standard press release to each county before the launch of the survey. The press release was followed up with phone contact to selected local media in each county. The communication group used social media (Facebook and Twitter) to promote the survey. Several municipalities informed their residents about the survey on municipality website. The project leader also gave interviews on the local radio and to newspapers in some of the counties. In total, the survey was covered in various media channels 127 times.

Statistical Analysis

All data used in the analyses, except county of residence, consisted of self-reported data. We calculated means (with standard deviations) or medians (with interquartile range) and percentages for each health variable, overall and by sex, using STATA (version 16.0; Stata Corp).

Results

Main Survey

Altogether, 25.2% (2769/11,000) participants opened the survey, while 21.0% (2305/11,000) submitted their responses. We excluded 2 participants reporting to be outside the targeted age groups. The remaining 2303 men and women constituted the respondents to the main survey.

The response rate to the main survey was 21.9% (2303/11,000; women: 1397/2303, 60.7%; men: 866/2303, 38.3%; missing: 40/2303, 1.7%). Among men and women, the response rates were 15.7% (866/5500) and 25.4% (1397/5500), respectively. The median age for men was 52 years (IQR 40-61), and the median age for women was 48 years (IQR 35-58). Trøndelag county had the highest response rate (26.0%, 260/1000), and Vestland county had the lowest response rates (18.1%, 181/1000).

Of the 2303 respondents, 1419 (61.6%) answered all the questions; 15 of 29 variables were missing ≤3% values. Each of the questions on socioeconomic background factors had missing values, which ranged from 10.3% (238/2303) for marital status to 20.6% (474/2303) for number of persons under 18 years old in household. Tobacco and alcohol consumption variables had ≤3% missing values, physical activity variables had 1.8% (41/2303) to 9.7% (224/2303) missing values, and food variables had 1.5% (35/2303) to 5.9% (136/2303) missing values (Table 1).

Most respondents were in the oldest (50-69 years) age group (1142/2274, 50.2%), in the highest educational category (graduate or postgraduate university education: 1088/1989, 55.1%), married (1497/2065, 73.1%,), and employed full-time (1158/2015, 57.8%) (Table 2).

Table 3 shows that 55.4% (1225/2209) of respondents reported being either overweight (794/2209, 35.9%) or obese (431/2209, 19.5%), with 63.9% (539/844) of men and 50.3% (686/1365) of women reporting being either overweight or obese.

Daily use of cigarettes (men: 100/855, 11.7%; women: 140/1380, 10.1%), snus (men: 153/850, 18.0%; women: 96/1360, 7.0%), and e-cigarettes (11/848, 1.3%; women: 19/1359, 1.4%) was reported by respondents, and 32.2% (237/735) of men reported drinking alcohol more than once a week, of whom 7.1% (51/724) reported consuming ≥6 units on the same occasion weekly, and 22.7% (254/1119) of women reported drinking alcohol more than once a week, of whom 1.7% (19/1103) reported consuming ≥6 units on the same occasion weekly. Overall, 17.2% (378/2199) reported 0 days with moderate physical activity during the last 7 days (Table 3).

Table 4 shows that 38.5% (875/2268) of respondents reported 1 serving per day of fruit and berries, 48.3% (1091/2255) of respondents reported 1 serving per day of lettuce and vegetables, 67.3% (1458/2167) of respondents reported 0 glasses of sugary drinks per day, 56.6% (1233/2178) of respondents reported ≥2 meals with fish or fish products per week, 67.6% (1465/2167) of respondents reported ≤2 meals of red meat per week, and 23.5% (517/2198) of respondents reported never adding extra salt before eating.

Only 44.5% (984/2209) of respondents were in line with national recommendations for BMI, and 42.0% (932/2219) of respondents were in line with national recommendations for tobacco or e-cigarette use [20]; 81.2% (1501/1848) were in line with national recommendations for alcohol consumption (ie, consuming ≥6 units of alcohol on one occasion less than monthly or never [21]). Only 34.2% (704/2056) of respondents met national recommendations for physical activity levels (ie, walking ≥10 minutes every day the last 7 days [22]), and only 22.9% (522/2281) of respondents met national recommendations for eating fruits and vegetables (ie, at least 5 servings per day [23]).

Table 1. Overall respondents and missing values to the main smartphone survey.

Variable	Respondents, N	Missing, n (%)
Age	2274	29 (1.3)
Sex	2263	40 (1.7)
Education	1989	314 (13.6)
Marital status	2065	238 (10.3)
Number of persons (>18 years old) in household	1914	389 (16.9)
Number of persons (<18 years old) in household	1829	474 (20.6)
Gross household income 2018	1946	357 (15.5)
Work life condition/occupation	2015	288 (12.5)
Chronic disease conditions (noncommunicable disease)	2245	58 (2.5)
Height	2270	33 (1.4)
Weight	2247	56 (2.4)
BMI[a]	2237	66 (2.9)
Smoking history	2261	42 (1.8)
Snus use history	2245	58 (2.5)
E-cigarette use history	2234	69 (3.0)
Alcohol consumption (yes/no)	2284	19 (0.8)
Frequency of alcohol consumption[b]	1876	29 (1.5)
Units of alcohol usually consumed[b]	1868	37 (1.9)
Frequency of occasions of consumption of ≥6 units of alcohol[b]	1848	57 (3.0)
Number of days of strenuous physical activity in the last 7 days	2262	41 (1.8)
Number of days of moderate physical activity in the last 7 days	2199	104 (4.5)
Number of days of walking for ≥10 minutes in the last 7 days	2079	224 (9.7)
Hours spent sitting on a regular weekday in the last 7 days (excluding sleeping hours)	2166	137 (6.0)
Fruits and berries intake (servings) per day	2268	35 (1.5)
Lettuce and vegetable intake (servings) per day	2255	48 (2.1)
Sugary drinks, number of glasses per day	2167	136 (5.9)
Fish and fish products, number of times eaten per week	2178	125 (5.4)
Red meat, number of times eaten per week	2167	136 (5.9)
Processed meat, number of times eaten per week	2167	136 (5.9)
Addition of extra salt to food	2198	105 (4.6)

[a]BMI: body mass index.

[b]Among 1905 participants who responded *yes* to alcohol consumption.

Table 2. Respondents' background characteristics (overall and by sex).

Characteristic	All, n (%)[a]	Men, n (%)[a]	Women, n (%)[a]
Age groups (years)			
16-29	347 (15.4)	121 (14.0)	226 (16.3)
30-49	760 (33.8)	264 (30.7)	496 (35.7)
50-69	1142 (50.8)	476 (55.3)	666 (48.0)
Education (number of years)			
≤10 (primary school)	219 (11.1)	104 (13.6)	115 (9.5)
11-13 (high school)	666 (33.8)	290 (37.8)	376 (31.2)
14-16 (graduate)	508 (25.7)	180 (23.5)	328 (27.2)
≥17 (postgraduate)	580 (29.4)	193 (25.2)	387 (32.1)
Marital status			
Single	389 (19.0)	155 (19.7)	234 (18.6)
Married or cohabitating	1497 (73.1)	575 (73.1)	922 (73.1)
Divorced	126 (6.1)	48 (6.1)	78 (6.2)
Widow or widower	36 (1.8)	9 (1.1)	27 (2.1)
Number of adults (age ≥18 years) in household			
0	151 (7.5)	57 (7.3)	94 (7.7)
1	474 (23.6)	184 (23.7)	290 (23.6)
2	1013 (50.5)	383 (49.3)	630 (51.3)
3	230 (11.5)	93 (12.0)	137 (11.2)
≥4	137 (6.8)	60 (7.7)	77 (6.3)
Number of children (age <18 years) in household			
0	1180 (61.6)	482 (65.1)	698 (59.4)
1	293 (15.3)	95 (12.8)	198 (16.8)
2	305 (15.9)	115 (15.5)	190 (16.2)
3	106 (5.5)	40 (5.4)	66 (5.6)
≥4	32 (1.7)	9 (1.2)	23 (2.0)
Gross household income in 2018 (in Norwegian Kroner)			
≤350,000	249 (12.9)	80 (10.7)	169 (14.3)
351,000-550,000	342 (17.7)	117 (15.6)	225 (19.0)
551,000-750,000	316 (16.4)	118 (15.8)	198 (16.7)
751,000-1,000,000	430 (22.2)	177 (23.7)	253 (21.4)
>1,000,000	595 (30.8)	256 (34.2)	339 (28.6)
Occupational status			
Full-time work	1158 (57.9)	524 (68.1)	634 (51.5)
Part-time work	234 (11.7)	37 (4.8)	197 (16.0)
Student	146 (7.3)	50 (6.5)	96 (7.8)
Retired	180 (9.0)	76 (9.9)	104 (8.5)
Other[b]	281 (14.1)	82 (10.7)	199 (16.2)

[a]Missing responses were not included in calculated percentages.

[b]Home-keeper, military service, and miscellaneous benefits (sick leave, unemployment, disabilities, social security).

Table 3. Respondents' lifestyle factors (overall and by sex).

Factor	All, n (%)[a]	Men, n (%)[a]	Women, n (%)[a]
BMI[b] (kg/m^2)			
<18.5	29 (1.3)	3 (0.4)	26 (1.9)
18.5-24.9	955 (43.2)	302 (35.8)	653 (47.8)
25.0-29.9	794 (35.9)	367 (43.5)	427 (31.3)
≥30.0	431 (19.5)	172 (20.4)	259 (19.0)
Smoking history			
Never	1087 (48.6)	392 (45.9)	695 (50.4)
Former occasionally	344 (15.4)	125 (14.6)	219 (15.9)
Former daily	469 (21.0)	201 (23.5)	268 (19.4)
Current occasionally	95 (4.3)	37 (4.3)	58 (4.2)
Current daily	240 (10.7)	100 (11.7)	140 (10.1)
Snus use history			
Never	1726 (77.8)	571 (67.2)	1155 (84.4)
Former occasionally	97 (4.4)	51 (6.0)	46 (3.4)
Former daily	80 (3.6)	51 (6.0)	29 (2.1)
Current occasionally	66 (3.0)	24 (2.8)	42 (3.1)
Current daily	249 (11.2)	153 (18.0)	96 (7.0)
E-cigarette use history			
Never	2069 (93.7)	786 (92.7)	1283 (94.4)
Former occasionally	66 (3.0)	27 (3.2)	39 (2.9)
Former daily	15 (0.7)	10 (1.2)	5 (0.4)
Current occasionally	27 (1.2)	14 (1.7)	13 (1.0)
Current daily	30 (1.4)	11 (1.3)	19 (1.4)
Alcohol consumption			
No	376 (16.7)	120 (13.9)	256 (18.4)
Yes	1880 (83.3)	743 (86.1)	1137 (81.6)
Frequency			
≤1 time per month	493 (26.6)	149 (20.3)	344 (30.7)
2-4 times per month	870 (46.9)	349 (47.5)	521 (46.6)
2-3 times per week	402 (21.7)	191 (26.0)	211 (18.9)
≥4 times per week	89 (4.8)	46 (6.3)	43 (3.8)
Units usually consumed			
1-2	994 (53.8)	366 (49.7)	628 (56.6)
3-4	538 (29.1)	215 (29.2)	323 (29.1)
5-6	208 (11.3)	86 (11.7)	122 (11.0)
7-9	84 (4.6)	56 (7.6)	28 (2.5)
≥10	22 (1.2)	14 (1.9)	8 (0.7)
Frequency of occasions with ≥6 units consumed			
Never	545 (29.5)	132 (18.2)	403 (36.5)
Less than monthly	956 (51.7)	385 (53.2)	561 (50.9)
Monthly	277 (15.0)	156 (21.6)	120 (10.9)

Factor	All, n (%)[a]	Men, n (%)[a]	Women, n (%)[a]
Weekly	68 (3.7)	49 (6.8)	19 (1.7)
Daily or almost daily	2 (0.1)	2 (0.3)	(0.0)
Days with strenuous physical activity in the last 7 days			
0	639 (28.6)	198 (23.2)	441 (32.0)
1-2	828 (37.1)	308 (36.0)	520 (37.7)
3-4	527 (23.6)	228 (26.7)	299 (21.7)
5-6	180 (8.0)	9 (10.5)	90 (6.5)
7	61 (2.7)	31 (3.6)	30 (2.2)
Days with moderate physical activity in the last 7 days			
0	378 (17.4)	133 (16.0)	245 (18.2)
1-2	927 (42.6)	350 (42.1)	577 (42.9)
3-4	510 (23.5)	187 (22.5)	323 (24.0)
5-6	227 (10.4)	108 (13.0)	119 (8.8)
7	133 (6.1)	53 (6.4)	80 (6.0)
Days walking for ≥10 minutes in the last 7 days			
0	99 (4.8)	55 (7.1)	44 (3.4)
1-2	373 (18.1)	154 (19.7)	219 (17.2)
3-4	480 (23.4)	189 (24.2)	291 (22.8)
5-6	400 (19.5)	153 (19.6)	247 (19.4)
7	704 (34.2)	229 (29.4)	475 (37.2)
Hours spent sitting on a regular weekday in the last 7 days (excluding sleeping hours)			
0-2	162 (7.6)	58 (7.1)	104 (7.9)
3-5	876 (40.9)	292 (35.5)	584 (44.3)
6-8	598 (27.9)	245 (29.8)	353 (26.7)
9-11	348 (16.3)	150 (18.2)	198 (15.0)
12-14	109 (5.1)	59 (7.2)	50 (3.8)
≥15	48 (2.2)	18 (2.2)	30 (2.3)

[a]Missing responses were not included in calculated percentages.

[b]BMI: body mass index.

Table 4. Respondents' dietary intake variables (overall and by sex).

Variable	All, n (%)[a]	Men, n (%)[a]	Women, n (%)[a]
Fruits and berries (servings per day)			
0	260 (11.6)	136 (15.8)	124 (9.0)
1	875 (39.1)	378 (44.1)	497 (35.9)
2	621 (27.7)	205 (23.9)	416 (30.1)
3	309 (13.8)	93 (10.8)	216 (15.6)
4	99 (4.4)	25 (2.9)	74 (5.4)
≥5	77 (3.4)	21 (2.5)	56 (4.0)
Lettuce and vegetables (servings per day)			
0	134 (6.0)	82 (9.6)	52 (3.8)
1	1091 (49.0)	494 (58.1)	597 (43.3)
2	643 (28.9)	190 (22.4)	453 (32.9)
3	225 (10.1)	50 (5.9)	175 (12.7)
4	72 (3.2)	19 (2.2)	53 (3.8)
≥5	63 (2.8)	15 (1.8)	48 (3.5)
Sugary drinks (glasses per day)			
0	1458 (68.1)	497 (60.7)	961 (72.7)
1	415 (19.4)	192 (23.4)	223 (16.9)
2	134 (6.3)	67 (8.2)	67 (5.1)
3	60 (2.8)	28 (3.4)	32 (2.4)
4	26 (1.2)	10 (1.2)	16 (1.2)
≥5	47 (2.2)	25 (3.1)	22 (1.7)
Fish and fish products (number of times eaten per week)			
0	151 (7.0)	48 (5.8)	103 (7.7)
1	769 (35.7)	293 (35.6)	476 (35.8)
2	789 (36.7)	312 (38.0)	477 (35.8)
3	339 (15.7)	136 (16.6)	203 (15.3)
4	82 (3.8)	28 (3.4)	54 (4.1)
≥5	23 (1.1)	5 (0.6)	18 (1.3)
Red meat (number of times eaten per week)			
0	172 (8.0)	37 (4.5)	135 (10.3)
1	664 (31.0)	241 (29.3)	423 (32.1)
2	629 (29.4)	255 (31.0)	374 (28.4)
3	369 (17.2)	163 (19.8)	206 (15.6)
4	179 (8.4)	78 (9.5)	101 (7.7)
≥5	127 (5.9)	49 (5.9)	78 (5.9)
Processed meat (number of times eaten per week)			
0	341 (15.9)	96 (11.7)	245 (18.6)
1	868 (40.6)	328 (39.9)	540 (40.9)
2	581 (27.2)	230 (28.0)	351 (26.6)
3	266 (12.4)	127 (15.5)	139 (10.5)
4	51 (2.4)	27 (3.3)	24 (1.8)
≥5	33 (1.5)	13 (1.6)	20 (1.5)

Variable	All, n (%)[a]	Men, n (%)[a]	Women, n (%)[a]
Addition of extra salt to food			
Never	517 (23.8)	192 (23.2)	325 (24.2)
Occasionally	1264 (58.2)	465 (56.1)	799 (59.5)
Often	306 (14.1)	127 (15.3)	179 (13.3)
Always	84 (3.9)	45 (5.4)	39 (2.9)

[a]Missing responses were not included in calculated percentages.

Second Survey

Altogether, 18.1% (199/1100) of the invitees to the second survey opened it, but 8.2% (90/1100) did not submit a response, while 9.9% (109/1100) replied to some or all and submitted their answers. The latter group constituted the respondents of the second survey. The most common reason for not participating (in the main survey) was "I had not heard of Healthcom" (yes: 43/109, 39.4%; no: 10/109, 9.2%; missing: 56/109, 49.5%), followed by both "I had not heard of the National Center for E-Health Research" and "I was unsure who the SMS came from" (yes: 37/109, 33.9%; no: 11/109, 10.1%; missing: 61/109, 56.0%), then "I thought the SMS I received looked like spam" (yes: 35/109, 32.1%; no: 10/109, 9.2%; missing: 64/109, 58.7%).

Discussion

Principal Results

This feasibility study is, to the best of our knowledge, the first successfully piloted smartphone-based information communication technology solution with technically, ethically, and regulatory functionalities for collecting annual data on the 4 main modifiable risk factors for the 4 main noncommunicable diseases. The primary reason reported for not participating in the initial survey was that the sender of the SMS was unknown. Due to the differences in response rates between men and women and between counties, our secondary objective to collect nationally representative data may not have been achieved. If the respondent sample has a similar distribution with respect to background characteristics, life style factors, and dietary intake variables, they may be nationally representative regardless of response rate.

Comparison With Past Work

We are not aware of any previous work with the same overall objectives as ours.

We chose to use SMS text messaging as the means of initial contact, instead of internet or email surveys, ordinary or computer-assisted telephone interviews, or interactive voice response because text messages are often perceived as personal forms of communication; they are more likely to be read quickly, understood, and responded to upon receipt [24]; and they are relatively cheap, with high reachability because mobile phones are ubiquitous. We anticipated that SMS text messaging would allow simple, low-commitment participation in the survey. We considered the intrusiveness (or push factor) to be an advantage; however, one disadvantage is that messages are limited to 160 characters, after which additional payments fare required for every 160 characters. Our request to participate had many more characters; therefore, we had to pay 3 times the ordinary amount for each message. Another disadvantage is that more and more commercial organizations have started using SMS text messages for customer satisfaction surveys. This can lead to an overload of text messages considered to be spam.

In 2017, Pariyo et al [25] suggested that mobile phone surveys have the potential to become a powerful data collection tool to address public health challenges, such as those arising from noncommunicable diseases, in low- and middle-income countries. Ethical considerations in global mobile phone-based surveys of noncommunicable diseases have also recently been discussed, and a need for a broad conceptual framework for the ethical, legal, and societal issues associated with mobile phone surveys for noncommunicable disease risk factors was identified [26].

Response Rate

In addition to randomized selection of participants, response rate is often considered to be an important factor in obtaining representative data. Our response rate (2305/11,000, 21.0%) was similar to those of 2 Norwegian eHealth research studies—a large population-based randomized controlled trial on smoking cessation with SMS text messaging or emails [27] and a web-based cross-sectional survey on diabetes [28]. We are satisfied that we achieved a response rate from the general population that was similar to that in these 2 surveys, which both had highly selected participants.

It is common for researchers to use monetary incentives (such as gift certificates or lottery tickets) to increase participation [29-31]. Instead of monetary incentives, we had intended to develop personalized feedback, based on national guidelines, for respondents. Though part of this feature was developed, due to time, monetary, and personnel constraints, we were not able to create detailed personalized responses for all possible combinations of answers to the survey. Since we did not complete this, the invitees were not informed about or given any options for personalized feedback. Most importantly, we were not able to examine if this would be an incentive to respond to the survey.

There is, as far as we know, no commonly accepted minimal response rate. In the UK Biobank cohort [32], there were close to 10 million invitees and the response rate was 5.45%. A response rate of 60% has been used as the threshold of acceptability for population-based face-to-face or postal surveys; however, this response rate can also cover up response bias [33]

if the characteristics of nonrespondents differ from those of the respondents. However, if respondents are perfectly representative of the source population, a low response rate is not a problem. [33]. Low participation, high dropout, or high loss to follow-up may be expected features of eHealth research and likely should not be looked upon as failures [34].

We assume that our response rate would have been higher if our university, a renowned entity, had been the sender. Instead, the Norwegian Centre for E-health Research, which was established in 2016 (a few years before the survey), and Healthcom both appeared as senders of the initial SMS. A systematic review [35] found that response rates for paper questionnaires were higher if they originated from universities rather than from other sources such as commercial organizations. We have previously found that more specific titles to otherwise identical questionnaires influenced the response rate of mailed surveys [36]; therefore, in addition, the name of the survey—*Health and Disease*—was too general.

Another Norwegian eHealth study [37], which examined recruitment to public health surveys with electronic forms on 2 different platforms, found that (1) sampling from the national health website (30 000 invitees) yielded a response rate of 40.8%, whereas sampling from the National Population Registry (36 000 invitees) yielded a response rate of 41.5% and (2) there were systematic and pronounced differences in the responses of the 2 samples. Skogen et al [37] concluded that limiting recruitment to users of *Helsenorge (Health Norway)* services resulted in further selection problems.

Our survey was conducted in 2019 (before the COVID-19 pandemic) in Norway, and approximately 1 in 5 respondents (431/2209, 19.5%) reported being obese. In fall 2020, after the first national lockdown, the Norwegian Institute of Public Health conducted a pilot study, collecting data on risk factors for the 4 main noncommunicable diseases and several other topics [38], which noted as a limitation, that most questions in the survey had not yet been validated, but that this would be done at a later stage. As we did, they randomly sampled from each of the 11 counties, but they invited twice as many from each county and used both email and SMS text messaging to contact the invitees. The response rate to this survey [38] was 38.1%, and 16% of respondents reported being obese, and 59% of the men and 47% of the women reported being overweight or obese. We do not have a good explanation why we found greater percentages—63.9% (539/844) of men and 50.3% (686/1365) of women reporting being either overweight or obese.

Similar to our findings, (1) the 3 previously described eHealth studies [27,28,38] found that more women than men responded, and (2) the 2020 national survey [37] found that Trøndelag county had the highest response rate. One reason may be that the people residing in Trøndelag have been invited 4 times to the Trøndelag Health Study [11], which collected health survey data from the same geographic population [39].

In our survey, 10.7% (240/2235) reported smoking daily. For 2019, Statistics Norway reported overall daily smoking to be 9% for both sexes [40]. The proportion of men who reported being daily snus use in our study (153/850, 18.0%) was a little lower than that found by Statistics Norway in 2019 (20% [40]), while the proportion of women found by Statistics Norway (7% [40]) was the same as that found in our study (96/1368, 7.0%).

In our survey, 32.3% (237/735) men reported drinking alcohol more than once a week; the corresponding figure for Statistics Norway survey was 34%. In our survey, 7% (51/724) of men reported having consumed more than 6 units on the same occasion weekly; the corresponding figure for the 2019 Statistics Norway survey was 5%. In our survey, 23% (254/1119) of women reported drinking alcohol more than once a week; the corresponding figure for the 2019 Statistics Norway survey was 28%. In our survey, less than 2% reported having consumed more than 6 units on the same occasion weekly; the corresponding figure for the 2019 Statistics Norway survey was 3%.

A recent review [41] found that pilot or feasibility studies are still poorly reported, and only 8.9% of the 90 studies led to subsequent main studies. This study will continue to be developed as part of a larger project [42].

Strengths

For a nationally representative study population, participants were sampled randomly from each of the 11 counties. Major assets were that our main survey could be answered in 5 minutes and comprised already validated questions.

Limitations

The main limitation of our study was the low response rate. Other limitations include the use of self-reported information and that only individuals who had a smartphone could participate; however, since smartphone ownership exceeds 80% in Norway [14], not being eligible due to not having a smartphone is a minor concern. This may be a concern in other countries.

Implications for Future Research

Short smartphone surveys have the potential to be used to monitor trends annually to identify high-risk groups for the 4 main noncommunicable diseases. This knowledge can subsequently be used for better targeting of interventions and in policy making, to meet the United Nations Sustainable Development Goal to reduce premature mortality from noncommunicable diseases by 33% by 2030 [4].

We encourage the further study of short mobile phone-based surveys regarding the modifiable risk factors for the 4 main noncommunicable diseases in high-, low-, and middle-income countries. Our study included only Norwegians, of whom the majority had more than high school level of education. Future studies should develop surveys that are easily recognized as research, have a well-known sender, and can be distinguished from spam. We recommend examining if an offer to the invitees for personalized feedback in response to their answers about the modifiable risk factors will increase overall participation.

Conclusions

We successfully developed and piloted a smartphone-based information communication technology solution for collecting data annually on the four modifiable risk factors for the 4 main noncommunicable disease from a random sample of the

Norwegian population; 1 in 5 responded, thus our secondary objective to collect nationally representative data may not have been achieved.

Acknowledgments

The project was funded in part by the Research Council of Norway (HELSE-EU-project 269882).

We would like to thank all the participants in the Health and Disease survey. We would also like to thank Niels Frost, at Healthcom, and Jarl-Stian Olsen, for designing the Norwegian and English websites and the poster, and Lene Lundberg and Randi Laukli, for their efforts to promote the main survey on social and other media, at the Norwegian Centre for E-health Research.

Authors' Contributions

ITG conceived the idea. ITG, GS, LAH, KBB, and MLL designed and conducted the study. SOO and ITG performed analyses. ITG drafted the manuscript. All authors contributed to interpreting the data, drafting the manuscript, and critically revising the manuscript for important intellectual content.

Conflicts of Interest

None declared.

Multimedia Appendix 1
SMS text message request to participate in main and second survey.
[DOCX File , 15 KB - formative_v6i2e33636_app1.docx]

Multimedia Appendix 2
Additional information given to the invitees after accessing the survey.
[DOCX File , 14 KB - formative_v6i2e33636_app2.docx]

References

1. Noncommunicable diseases 2018. World Health Organization (WHO) 2018. URL: https://www.who.int/health-topics/noncommunicable-diseases#tab=tab_1 [accessed 2021-02-05]
2. NCD Countdown 2030 collaborators. NCD Countdown 2030: worldwide trends in non-communicable disease mortality and progress towards Sustainable Development Goal target 3.4. Lancet 2018 Sep 22;392(10152):1072-1088. [doi: 10.1016/S0140-6736(18)31992-5] [Medline: 30264707]
3. NCD Countdown 2030 collaborators. NCD Countdown 2030: pathways to achieving sustainable development goal target 3.4. Lancet 2020 Sep 26;396(10255):918-934 [FREE Full text] [doi: 10.1016/S0140-6736(20)31761-X] [Medline: 32891217]
4. Transforming our world: the 2030 agenda for sustainable development 2020. United Nations. 2020. URL: https://sdgs.un.org/2030agenda [accessed 2021-02-05]
5. World cancer report: cancer research for cancer prevention 2020. International Agency for Research on Cancer. URL: https://www.iccp-portal.org/system/files/resources/IARC%20World%20Cancer%20Report%202020.pdf [accessed 2021-05-29]
6. Riley L, Guthold R, Cowan M, Savin S, Bhatti L, Armstrong T, et al. The World Health Organization STEPwise approach to noncommunicable disease risk-factor surveillance: methods, challenges, and opportunities. Am J Public Health 2016 Jan;106(1):74-78. [doi: 10.2105/AJPH.2015.302962] [Medline: 26696288]
7. STEPwise Approach to NCD risk factor surveillance. World Health Organization. URL: https://www.who.int/teams/noncommunicable-diseases/surveillance/systems-tools/steps [accessed 2021-05-30]
8. Tobacco, alcohol and other drugs 2021. Statistics Norway. URL: https://www.ssb.no/en/helse/helseforhold-og-levevaner/statistikk/royk-alkohol-og-andre-rusmidler [accessed 2021-05-30]
9. Norkost 4 is a dietary survey where a representative sample of the adult Norwegian population aged 18-80 years is invited to participate. Norkost 4. 2022 May 01. URL: https://www.med.uio.no/imb/english/research/projects/norkost/index.html [accessed 2021-10-05]
10. Statistikk om fysisk aktivitetsnivå og stillesitting [Statistics on physical activity level and sitting still]. Helsedirektoratet. URL: https://www.helsedirektoratet.no/tema/fysisk-aktivitet/statistikk-om-fysisk-aktivitetsniva-og-stillesitting [accessed 2021-09-20]
11. Krokstad S, Langhammer A, Hveem K, Holmen TL, Midthjell K, Stene TR, et al. Cohort profile: the HUNT study, Norway. Int J Epidemiol 2013 Aug;42(4):968-977. [doi: 10.1093/ije/dys095] [Medline: 22879362]
12. Jacobsen BK, Eggen AE, Mathiesen EB, Wilsgaard T, Njølstad I. Cohort profile: the Tromso study. Int J Epidemiol 2012 Aug;41(4):961-967 [FREE Full text] [doi: 10.1093/ije/dyr049] [Medline: 21422063]

13. Lov om folkehelsearbeid (folkehelseloven). Lovdata. URL: https://lovdata.no/dokument/NL/lov/2011-06-24-29/KAPITTEL_2#%C2%A75 [accessed 2022-02-01]
14. ICT usage in households 2021. Statistics Norway. URL: https://www.ssb.no/en/statbank/table/12344 [accessed 2021-10-14]
15. Lov om behandling av personopplysninger (personopplysningsloven). Lovdata. URL: https://lovdata.no/dokument/NL/lov/2018-06-15-38 [accessed 2022-02-09]
16. 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]
17. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2020 Jul 1;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
18. Bohn MJ, Babor TF, Kranzler HR. The alcohol use disorders identification test (AUDIT): validation of a screening instrument for use in medical settings. J Stud Alcohol 1995 Jul;56(4):423-432. [Medline: 7674678]
19. Cleland C, Ferguson S, Ellis G, Hunter RF. Validity of the international physical activity questionnaire (IPAQ) for assessing moderate-to-vigorous physical activity and sedentary behaviour of older adults in the United Kingdom. BMC Med Res Methodol 2018 Dec 22;18(1):176 [FREE Full text] [doi: 10.1186/s12874-018-0642-3] [Medline: 30577770]
20. Tobacco control in Norway. The Norwegian Directorate of Health. URL: https://www.helsedirektoratet.no/english/tobacco-control-in-norway [accessed 2022-02-10]
21. Nasjonal alkoholstrategi. Regjeringen. URL: https://www.regjeringen.no/contentassets/7517d7f5a9d34a9ba5954a21a8317803/alkoholstrategien-til-nett.pdf [accessed 2022-01-10]
22. Fysisk aktivitet for barn, unge, voksne, eldre og gravide. The Norwegian Directorate of Health. URL: https://www.helsedirektoratet.no/faglige-rad/fysisk-aktivitet-for-barn-unge-voksne-eldre-og-gravide [accessed 2022-02-10]
23. Norwegian dietary guidelines. The Norwegian Directorate of Health. URL: https://tinyurl.com/35embj77 [accessed 2022-02-10]
24. Muench F, Baumel A. More than a text message: dismantling digital triggers to curate behavior change in patient-centered health interventions. J Med Internet Res 2017 May 26;19(5):e147 [FREE Full text] [doi: 10.2196/jmir.7463] [Medline: 28550001]
25. Pariyo GW, Wosu AC, Gibson DG, Labrique AB, Ali J, Hyder AA. Moving the agenda on noncommunicable diseases: policy implications of mobile phone surveys in low and middle-income countries. J Med Internet Res 2017 May 05;19(5):e115 [FREE Full text] [doi: 10.2196/jmir.7302] [Medline: 28476720]
26. Ali J, Labrique AB, Gionfriddo K, Pariyo G, Gibson DG, Pratt B, et al. Ethics considerations in global mobile phone-based surveys of noncommunicable diseases: a conceptual exploration. J Med Internet Res 2017 May 05;19(5):e110 [FREE Full text] [doi: 10.2196/jmir.7326] [Medline: 28476723]
27. Gram IT, Larbi D, Wangberg SC. Comparing the efficacy of an identical, tailored smoking cessation intervention delivered by mobile text messaging versus email: randomized controlled trial. JMIR Mhealth Uhealth 2019 Sep 27;7(9):e12137 [FREE Full text] [doi: 10.2196/12137] [Medline: 31573935]
28. Hansen AH, Bradway M, Broz J, Claudi T, Henriksen Ø, Wangberg SC, et al. Inequalities in the use of ehealth between socioeconomic groups among patients with type 1 and type 2 diabetes: cross-sectional study. J Med Internet Res 2019;21(5):e13615 [FREE Full text] [doi: 10.2196/13615] [Medline: 31144669]
29. Wangberg SC, Nilsen O, Antypas K, Gram IT. Effect of tailoring in an internet-based intervention for smoking cessation: randomized controlled trial. J Med Internet Res 2011 Dec 15;13(4):e121 [FREE Full text] [doi: 10.2196/jmir.1605] [Medline: 22169631]
30. Westmaas JL, Bontemps-Jones J, Hendricks PS, Kim J, Abroms LC. Randomised controlled trial of stand-alone tailored emails for smoking cessation. Tob Control 2018 Mar;27(2):136-146. [doi: 10.1136/tobaccocontrol-2016-053056] [Medline: 28522745]
31. Geldsetzer P. Use of rapid online surveys to assess people's perceptions during infectious disease outbreaks: a cross-sectional survey on COVID-19. J Med Internet Res 2020 Apr 02;22(4):e18790 [FREE Full text] [doi: 10.2196/18790] [Medline: 32240094]
32. Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, et al. Comparison of sociodemographic and health-related characteristics of UK biobank participants with those of the general population. Am J Epidemiol 2017 Nov 01;186(9):1026-1034 [FREE Full text] [doi: 10.1093/aje/kwx246] [Medline: 28641372]
33. Johnson TP, Wislar JS. Response rates and nonresponse errors in surveys. JAMA 2012 May 2;307(17):1805-1806. [doi: 10.1001/jama.2012.3532] [Medline: 22550194]
34. Eysenbach G. The law of attrition. J Med Internet Res 2005;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
35. Edwards P, Roberts I, Clarke M, DiGuiseppi C, Pratap S, Wentz R, et al. Increasing response rates to postal questionnaires: systematic review. BMJ 2002 May 18;324(7347):1183 [FREE Full text] [Medline: 12016181]
36. Lund E, Gram IT. Response rate according to title and length of questionnaire. Scand J Soc Med 1998 Jun;26(2):154-160. [doi: 10.1177/14034948980260020401] [Medline: 9658516]

37. Skogen JC, Knapstad M, Smith OR, Tell GS, Lie RT, Nilsen TS, et al. How should recruitment to public health surveys be conducted? https://tidsskriftet.no/2019/10/originalartikkel/hvordan-bor-rekrutteringen-til-folkehelseundersokelsene-gjennomfores. URL: http://tidsskriftet.no/article/19-0054 [accessed 2020-02-10]
38. Abel M, Totland T. Kartlegging av kostholdsvanver og kroppsvekt hos voksne i Norge basert på selvrapportering- Resultater fra Den nasjonale folkehelseundersøkelsen. Norwegian Institute of Public Health. URL: https://www.fhi.no/globalassets/dokumenterfiler/rapporter/2021/rapport-nhus-2020.pdf [accessed 2022-08-10]
39. Sund ER, van Lenthe FJ, Avendano M, Raina P, Krokstad S. Does urbanicity modify the relationship between a polygenic risk score for depression and mental health symptoms? cross-sectional evidence from the observational HUNT study in Norway. J Epidemiol Community Health 2021 May;75(5):420-425 [FREE Full text] [doi: 10.1136/jech-2020-214256] [Medline: 32581065]
40. Wettergreen J. Less Norwegians smoke, more use snus 2020. Statistics Norway. URL: https://www.ssb.no/en/helse/artikler-og-publikasjoner/less-norwegians-smoke-more-use-snus [accessed 2021-10-30]
41. Arain M, Campbell MJ, Cooper CL, Lancaster GA. What is a pilot or feasibility study? a review of current practice and editorial policy. BMC Med Res Methodol 2010;10:67 [FREE Full text] [doi: 10.1186/1471-2288-10-67] [Medline: 20637084]
42. The Research Council of Norway. WARIFA: watching the risk factors: artificial intelligence and the personalized prevention and management of chronic conditions. Innovation Place. URL: https://www.innovationplace.eu/news/warifa-project-artificial-intelligence-and-prevention-of-chronic-conditions [accessed 2021-10-29]

Abbreviations

NOK: Norwegian Kroner

Edited by G Eysenbach; submitted 08.11.21; peer-reviewed by J Westmaas; comments to author 29.11.21; revised version received 09.12.21; accepted 31.12.21; published 25.02.22.

Please cite as:
Gram IT, Skeie G, Oyeyemi SO, Borch KB, Hopstock LA, Løchen ML
A Smartphone-Based Information Communication Technology Solution for Primary Modifiable Risk Factors for Noncommunicable Diseases: Pilot and Feasibility Study in Norway
JMIR Form Res 2022;6(2):e33636
URL: https://formative.jmir.org/2022/2/e33636
doi:10.2196/33636
PMID:35212636

Original Paper

Patient Digital Health Technologies to Support Primary Care Across Clinical Contexts: Survey of Primary Care Providers, Behavioral Health Consultants, and Nurses

Oleg Zaslavsky[1], PhD, MHA; Frances Chu[1], MSN, MLIS; Brenna N Renn[2], PhD

[1]School of Nursing, University of Washington, Seattle, WA, United States

[2]Department of Psychology, University of Nevada, Las Vegas, NV, United States

Corresponding Author:
Oleg Zaslavsky, PhD, MHA
School of Nursing
University of Washington
1959 Pacific Ave
Seattle, WA, 98195
United States
Phone: 1 2068493301
Email: ozasl@uw.edu

Abstract

Background: The acceptance of digital health technologies to support patient care for various clinical conditions among primary care providers and staff has not been explored.

Objective: The purpose of this study was to explore the extent of potential differences between major groups of providers and staff in primary care, including behavioral health consultants (BHCs; eg, psychologists, social workers, and counselors), primary care providers (PCPs; eg, physicians and nurse practitioners), and nurses (registered nurses and licensed practical nurses) in the acceptance of various health technologies (ie, mobile apps, wearables, live video, phone, email, instant chats, text messages, social media, and patient portals) to support patient care across a variety of clinical situations.

Methods: We surveyed 151 providers (51 BHCs, 52 PCPs, and 48 nurses) embedded in primary care clinics across the United States who volunteered to respond to a web-based survey distributed in December 2020 by a large health care market research company. Respondents indicated the technologies they consider appropriate to support patients' health care needs across the following clinical contexts: acute and chronic disease, medication management, health-promoting behaviors, sleep, substance use, and common and serious mental health conditions. We used descriptive statistics to summarize the distribution of demographic characteristics by provider type. We used contingency tables to compile summaries of the proportion of provider types endorsing each technology within and across clinical contexts. This study was exploratory in nature, with the intent to inform future research.

Results: Most of the respondents were from urban and suburban settings (125/151, 82.8%), with 12.6% (n=19) practicing in rural or frontier settings and 4.6% (n=7) practicing in rural-serving clinics. Respondents were dispersed across the United States, including the Northeast (31/151, 20.5%), Midwest (n=32, 21.2%), South (n=49, 32.5%), and West (n=39, 25.8%). The highest acceptance for technologies across clinical contexts was among BHCs (32/51, 63%) and PCPs (30/52, 58%) for live video and among nurses for mobile apps (30/48, 63%). A higher percentage of nurses accepted all other technologies relative to BHCs and PCPs. Similarly, relative to other groups, PCPs indicated lower levels of acceptance. Within clinical contexts, the highest acceptance rates were reported among 80% (41/51) of BHCs and 69% (36/52) of PCPs endorsing live video for common mental health conditions and 75% (36/48) of nurses endorsing mobile apps for health-promoting behaviors. The lowest acceptance across providers was for social media in the context of medication management (9.3% [14/151] endorsement across provider type).

Conclusions: The survey suggests potential differences in the way primary care clinicians and staff envision using technologies to support patient care. Future work must attend to reasons for differences in the acceptance of various technologies across providers and clinical contexts. Such an understanding will help inform appropriate implementation strategies to increase acceptability and gain greater adoption of appropriate technologies across conditions and patient populations.

(JMIR Form Res 2022;6(2):e32664) doi:10.2196/32664

KEYWORDS

survey; primary care; acceptance; nurses; primary care providers; behavioral health consultants; mobile health; technology; health promotion; attitudes

Introduction

Primary care plays a central role in managing acute and chronic physical and behavioral health conditions, including patient self-management of such conditions [1]. Digital health, which refers to the use of telehealth, mobile devices, and other wireless technologies to support health care [2], has great potential to augment or enhance such care by supporting behavior change while minimizing barriers such as distance and time. These technologies are intended to enhance education and awareness; support diagnosis and treatment, including self-management; facilitate remote monitoring; and enable remote communication (eg, telehealth) [3].

However, such technologies are far from having established maturity or wide acceptance in primary care [4]. COVID-19, the disease caused by the novel coronavirus (SARS-CoV-2), prompted drastic changes in primary care delivery, including the sudden and unexpected implementation of new tools to expand and support patient care, such as telehealth [5-7]. Innovative digital solutions such as live video, instant chats, and text messages have become essential to continue delivering primary care in times when in-person visits are restricted. Despite nationwide regulatory and reimbursement policy changes concerning health care technologies during the COVID-19 pandemic (eg, less restrictive policies for remote office visits), the implementation of digital tools may have varied across sites, providers, and clinical situations. Disparities in the adoption of digital tools might be due to differences in health care provider perception of acceptability, appropriateness, and feasibility of technologies for specific clinical scenarios [8].

Many studies have examined the acceptance and feasibility of digital health technologies for managing patients' physical and mental health conditions. Notably, an important condition for implementing such technologies is provider attitudes, including acceptance [9]. Indeed, a recommendation from a trusted health care provider is imperative for patients to adopt technologies like mobile apps; however, health care providers' acceptance of technologies varies [10].

Prior work examined mental health professionals' attitudes and interests in using technology in clinical treatment, but this was restricted to websites and mobile apps [9]. Other studies gathered health care providers' (pharmacists, physicians, and advanced practice registered nurses [APRNs]) opinions regarding the use of mobile apps for patients across various clinical contexts [11,12]. Pharmacists tended to recommend mobile apps for smoking cessation, physical activity, diabetes, weight management, and sexual health [11]. By contrast, physicians and APRNs recommended mobile apps for tracking physical activity, diet, and sleep; however, these providers did not view mobile apps as beneficial for monitoring sleep [12].

As primary care medical and behavioral health providers and staff differ in training, experience with technology, and clinical orientation, it is also reasonable to expect differences by provider type in their attitude toward digital health technologies. As such, the purpose of this project was to explore the extent of potential differences between major groups of providers and staff in primary care, including behavioral health consultants (BHCs; eg, psychologists, social workers, and counselors), primary care providers (PCPs; eg, physicians and nurse practitioners), and nurses (eg, registered nurses and licensed practical nurses) in their acceptance of different technologies to support patient care across a variety of clinical situations. This study was exploratory in nature, with the intent to inform future research.

Methods

Study Design and Sampling

We surveyed 151 providers (51 BHCs, 52 PCPs, and 48 nurses) embedded in primary care clinics across the United States who volunteered to respond to a web-based survey invitation. Survey methods are reported in accordance with the Checklist for Reporting Results of Internet E-Surveys (CHERRIES) [13]. Recruitment was overseen by a large health care market research company. The company emailed the survey invitation to their proprietary panel of health care professionals in December 2020. Invitations were unique to each participant to avoid duplicate responses and to coordinate incentive payment through the market research company. The invitation link directed interested participants to our web-based survey portal for screening and participation. Responses were collected and managed using the Research Electronic Data Capture (REDCap) web tool [14]. Screening questions asked participants to verify that they worked in primary care and to select their role (BHC, PCP, nurse, or other; the latter were excluded). Given the preliminary nature of our data collection, we applied a quota of approximately 50 respondents per provider group. Respondents were required to select an answer for all items to complete the survey; however, we included options such as "prefer to not answer" (for potentially sensitive items such as demographics) and occasionally included "unsure/don't know" for select relevant items.

Ethics Approval

The University of Washington Human Subjects Division determined that the survey was not considered research, as defined by federal and state regulations; therefore, no review by the institutional review board was required and participants did not have to provide informed consent. Eligible participants were presented with a basic description of the study and asked to indicate agreement to participate before advancing. No personal or identifying information was collected. Participants received a monetary incentive for their time.

Measures

The parent survey aimed to examine provider use of technologies to support behavioral health since the onset of the

COVID-19 pandemic. The main outcome in the present study was provider acceptance of digital health technologies in primary care, based on responses to a single question. This item asked the respondents to select technologies they consider appropriate to support patients' health care needs in the following clinical contexts: acute and chronic disease, medication management, health-promoting behaviors (diet and physical activity), sleep, substance use (eg, alcohol, nicotine), common mental health disorders (eg, depression, anxiety), and serious mental health conditions (eg, schizophrenia, bipolar disorder). The exact wording of the question was "What technologies can you envision to support behavioral and lifestyle changes in your patients?" Possible technologies included mobile apps, wearables, live video (clinical visits via interactive video), phone, email, instant chat, text messages, social media, and a patient portal. Respondents were presented with a matrix of possible technologies across the 8 clinical contexts and could mark as many technologies as they deemed appropriate within each context.

Statistical Analysis

We used descriptive statistics to summarize the distribution of demographic characteristics by provider type. Contingency tables compiled summaries in terms of a proportion of providers by type endorsing technologies within and across all contexts.

Results

The respondents included 51 BHCs, 48 nurses, and 52 PCPs. Most were located in urban and suburban settings and were dispersed across regions of the United States. See Table 1 and Multimedia Appendix 1 for the participants' demographic details.

We observed potential differences by provider type in the acceptance of technologies within and across all clinical contexts. The highest acceptance rates for technologies across clinical contexts were among BHCs for live video (32/51, 63%), nurses for mobile apps (30/48, 63%), and PCPs for live video (30/52, 58%). In addition to their support for mobile apps, a greater proportion of nurses accepted all other technologies relative to BHCs and PCPs. More than half of the nurse respondents endorsed synchronous technologies such as phone calls or live video, as well as the patient portal. Relative to other groups, PCPs had lower rates of acceptance. Within clinical contexts, the highest acceptance rates were 80% (41/51) of BHCs and 69% (36/52) of PCPs endorsing the use of live video for common mental health conditions and 75% (36/48) of nurses endorsing mobile apps for health-promoting behaviors. Endorsement of technologies was variable, but generally low for serious mental illness across provider types. The lowest acceptance across providers was for social media in the context of medication management (9.3% [14/151] endorsement across provider type). See Figure 1 and Figure 2 for more detail. Figure 1 illustrates the proportion of respondents endorsing a specific technology. The least (lightest blue) to most opaque (darkest blue) shades represent low to high values (0% to 100%). Figure 2 illustrates the proportion of respondents endorsing a specific technology across all clinical contexts.

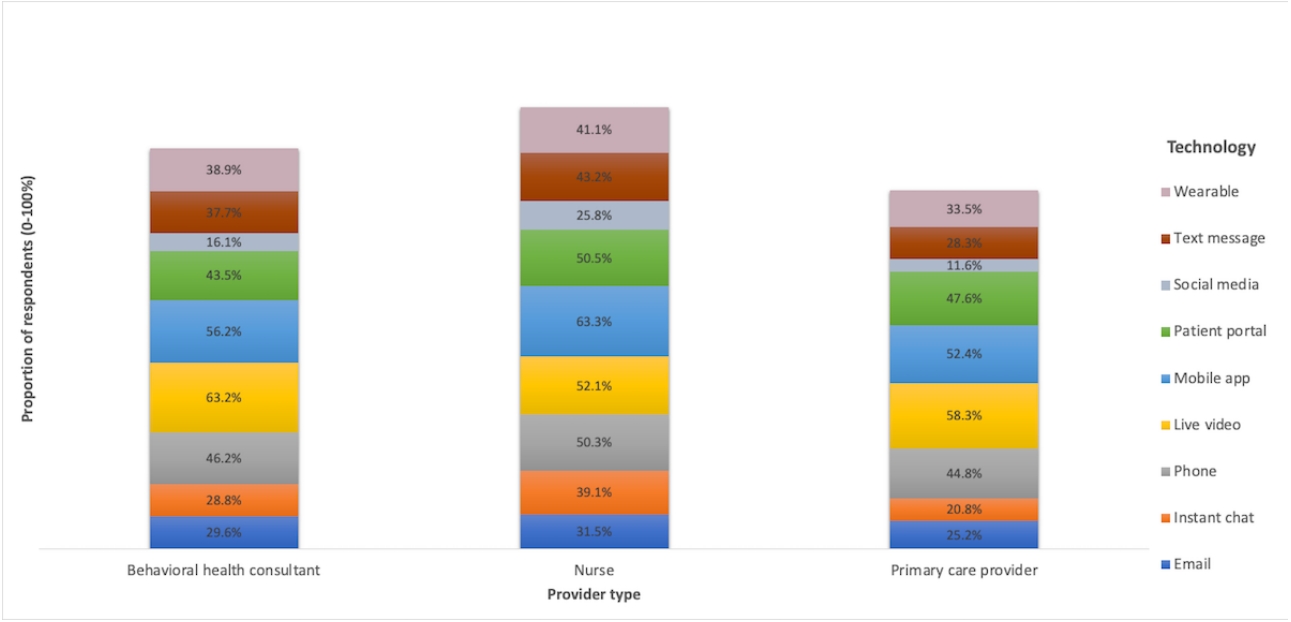
Table 1. Demographic characteristics of behavioral health consultants, nurses, and primary care providers who participated in the study.

Characteristic	Behavioral health consultants (n=51)	Nurses (n=48)	Primary care providers (n=52)	Total sample (N=151)
Race, n (%)				
Black or African American	3 (6)	3 (6)	2 (4)	8 (5.3)
American Indian or Alaska Native	0 (0)	0 (0)	1 (2)	1 (0.7)
Asian	1 (2)	5 (10)	13 (25)	19 (12.6)
Native Hawaiian or Other Pacific Islander	0 (0)	0 (0)	0 (0)	0 (0)
White	45 (88)	33 (69)	30 (58)	108 (71.5)
More than one race	1 (2)	1 (2)	0 (0)	2 (1.3)
Prefer to not answer	1 (2)	6 (13)	6 (12)	13 (8.6)
Ethnicity, n (%)				
Hispanic/Latinx	3 (6)	8 (17)	3 (6)	14 (9.3)
Not Hispanic/Latinx	46 (90)	35 (73)	45 (87)	126 (83.4)
Prefer to not answer	2 (4)	5 (10)	4 (8)	11 (7.3)
Age in years				
Mean (SD)	51.7 (11.9)	47.1 (11.4)	45.1 (10.8)	48.0 (11.7)
Range	30-73	24-67	29-77	24-77
Gender, n (%)				
Female	37 (73)	30 (63)	24 (46)	91 (60.3)
Male	13 (26)	9 (19)	24 (46)	46 (30.5)
No response	1 (2)	9 (19)	4 (8)	14 (9.3)
Clinic setting, n (%)				
Urban	21 (41)	17 (35)	13 (25)	51 (33.8)
Suburban	22 (43)	22 (46)	30 (58)	74 (49)
Rural	4 (8)	4 (8)	9 (17)	17 (11.3)
Rural-serving	3 (6)	4 (8)	0 (0)	7 (4.6)
Frontier	1 (2)	1 (2)	0 (0)	2 (1.3)
Clinic type, n (%)				
Clinic/practice at an academic medical center	5 (10)	10 (21)	10 (19)	25 (16.6)
Clinic/practice affiliated with a university teaching hospital	3 (6)	6 (13)	7 (14)	16 (10.6)
Community health center and/or Federally Qualified Health Center	6 (12)	5 (10)	7 (14)	18 (11.9)
Private health care system	8 (16)	14 (29)	14 (27)	36 (23.8)
Veteran's Affairs medical center or community-based outpatient clinic	2 (4)	2 (4)	0 (0)	4 (2.6)
Other government hospital	2 (4)	1 (2)	0 (0)	3 (1.98)
Private (independent or group) practice	29 (57)	11 (23)	17 (33)	57 (37.7)
Other	1 (2)	3 (6)	1 (2)	5 (3.3)

Figure 1. Acceptance of digital health technologies across various clinical contexts by major primary care health care professional types (n=51 behavioral health consultants, n=52 primary care providers, and n=48 nurses).

Health Professional Category	Technology	Acute disease/ condition management	Chronic disease/ condition management	Common mental health conditions (eg, depression, anxiety, stress)	Health-promoting behaviors (eg, physical activity, health eating)	Medication management	Serious mental illness (eg, schizophrenia, bipolar disorder)	Sleep and insomnia	Substance use (eg, alcohol, nicotine)	All Clinical Contexts
Behavioral health consultant	Email	30.8%	34.6%	40.4%	40.4%	28.9%	7.7%	25.0%	28.9%	**29.6%**
	Instant chat	23.1%	26.9%	40.4%	32.7%	34.6%	15.4%	21.2%	36.5%	**28.8%**
	Live phone	44.2%	38.5%	67.3%	51.9%	46.2%	48.1%	32.7%	40.4%	**46.2%**
	Live video	63.5%	61.5%	80.8%	55.8%	57.7%	63.5%	51.9%	71.2%	**63.2%**
	Mobile app	50.0%	57.7%	67.3%	71.2%	50.0%	25.0%	63.5%	65.4%	**56.2%**
	Patient portal	50.0%	50.0%	42.3%	51.9%	36.5%	36.5%	40.4%	40.4%	**43.5%**
	Social media	7.7%	17.3%	30.8%	28.9%	9.6%	3.9%	9.6%	21.2%	**16.1%**
	Text message	42.3%	42.3%	50.0%	42.3%	34.6%	23.1%	28.9%	38.5%	**37.7%**
	Wearable	40.4%	42.3%	48.1%	59.6%	21.2%	11.5%	53.9%	34.6%	**38.9%**
Nurse	Email	27.1%	35.4%	43.8%	41.7%	25.0%	31.3%	12.5%	35.4%	**31.5%**
	Instant chat	39.6%	39.6%	52.1%	39.6%	35.4%	37.5%	25.0%	43.8%	**39.1%**
	Live phone	58.3%	45.8%	70.8%	41.7%	43.8%	56.3%	27.1%	58.3%	**50.3%**
	Live video	62.5%	52.1%	66.7%	56.3%	37.5%	52.1%	29.2%	60.4%	**52.1%**
	Mobile app	60.4%	64.6%	68.8%	75.0%	58.3%	41.7%	68.8%	68.8%	**63.3%**
	Patient portal	52.1%	56.3%	62.5%	45.8%	54.2%	47.9%	39.6%	45.8%	**50.5%**
	Social media	18.8%	20.8%	39.6%	37.5%	10.4%	22.9%	20.8%	35.4%	**25.8%**
	Text message	45.8%	37.5%	54.2%	43.8%	47.9%	37.5%	27.1%	52.1%	**43.2%**
	Wearable	41.7%	45.8%	45.8%	64.6%	18.8%	25.0%	52.1%	35.4%	**41.1%**
Primary care provider	Email	26.4%	32.1%	24.5%	32.1%	30.2%	11.3%	18.9%	26.4%	**25.2%**
	Instant chat	22.6%	11.3%	37.7%	26.4%	17.0%	17.0%	15.1%	18.9%	**20.8%**
	Live phone	54.7%	43.4%	62.3%	41.5%	47.2%	28.3%	34.0%	47.2%	**44.8%**
	Live video	67.9%	64.2%	69.8%	50.9%	54.7%	54.7%	45.3%	58.5%	**58.3%**
	Mobile app	50.9%	60.4%	54.7%	64.2%	50.9%	26.4%	62.3%	49.1%	**52.4%**
	Patient portal	50.9%	62.3%	47.2%	45.3%	56.6%	35.9%	39.6%	43.4%	**47.6%**
	Social media	7.6%	7.6%	17.0%	22.6%	5.7%	7.6%	11.3%	13.2%	**11.6%**
	Text message	32.1%	30.2%	30.2%	34.0%	30.2%	17.0%	20.8%	32.1%	**28.3%**
	Wearable	28.3%	34.0%	34.0%	62.3%	24.5%	13.2%	43.4%	28.3%	**33.5%**

Figure 2. Proportion of respondents endorsing a specific digital health technology across all clinical contexts (n=51 behavioral health consultants, n=52 primary care providers, and n=48 nurses).



Discussion

This national survey suggests potential differences in the way primary care clinicians, behavioral health consultants, and nursing staff envision using digital health technologies to support patients in primary care. Potential differences were observed across technologies and clinical contexts. Compared to other providers, a higher proportion of BHCs in our sample were receptive to using synchronous interactive technologies such as live video. Some of this acceptability may be a result of the sudden shift to telehealth during the COVID-19 pandemic [15]. Similarly, a high proportion of nurse respondents embraced more diverse ways of connecting and supporting patients through traditional technologies such as phone and email, as well as newer technologies such as mobile apps and instant chats. Relative to BHCs and nurses, PCPs in our sample indicated lower levels of acceptance for digital health technologies across all clinical situations. Across clinical contexts, live video was seen as an acceptable way of connecting with patients, especially for common mental health conditions.

Mobile apps earned high acceptance among nurses, especially in the context of health-promoting behaviors. In general, the acceptance of social media lagged behind other technologies. It is possible that the nature of patient interaction in primary care influenced provider attitudes. Nurses endorsed technologies that support their typical clinical focus on case management, self-management, and promoting health behaviors. By contrast, PCPs and BHCs preferred synchronous video, which aligns with their focus on traditional treatment encounters. We recommend researchers and developers solicit provider needs and preferences when designing digital health technologies to promote the usability and implementation of these tools.

Given that we collected data during the first year of the COVID-19 pandemic, our findings may reflect this pivotal moment for the adoption of digital health tools to provide or enhance care. The nationwide rollout of digital technologies (eg, electronic health records) to support patient care has often faced challenges, and policymakers often struggle to understand how, when, and to what extent technologies could be used. Our preliminary findings highlight potential differences in the acceptance of digital health technologies across providers and clinical situations. Given that acceptance and other attitudinal constructs are considered preconditions for adoption [8], a one-size-fits-all approach to introducing technologies may fail among different providers. Understanding the reasons for such observed differences in acceptance—that is, exploring why the differences exist, perhaps through a qualitative investigation—is an important future direction.

Our study has several limitations. First, the convenience sample may not be representative of all US providers and staff in primary care. Nonresponders may have different opinions about digital health technologies across clinical contexts, while responders may be biased toward using technology in any clinical context. Secondly, the survey was not a validated measure of technology acceptance. Thirdly, our findings are based on a cross-sectional survey that reflects one point in time in the midst of a global pandemic. Longitudinal follow-up is necessary to better ascertain trends in technology acceptance. Fourth, we did not provide context for how technologies would be employed and by whom. Finally, because our sample size is small relative to the number of response items (technologies and clinical contexts), we present descriptive findings and comparisons rather than statistical testing of group differences.

In conclusion, given the potential of technologies to facilitate primary health care delivery, future work must attend to reasons for differences in acceptance of various technologies across providers and clinical contexts. Such an understanding will help inform appropriate implementation strategies to increase acceptability and gain higher adoption rates of appropriate technologies across conditions and patient populations.

Acknowledgments

This project was supported by the National Institute of Mental Health (Grant P50 MH115837) and the National Institute on Aging (Grant K23 AG059912). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest

BR receives unrelated research support from Sanvello Health Inc. The other authors have no competing interests to declare.

Multimedia Appendix 1
Demographic characteristics of health care professionals who participated in the survey, including additional data points.
[DOCX File , 19 KB - formative_v6i2e32664_app1.docx]

References

1. Bodenheimer T, Lorig K, Holman H, Grumbach K. Patient self-management of chronic disease in primary care. JAMA 2002 Nov 20;288(19):2469-2475. [doi: 10.1001/jama.288.19.2469] [Medline: 12435261]
2. World Health Organization Global Observatory for eHealth. mHealth: new horizons for health through mobile technologies: second global survey on eHealth. World Health Organization Institutional Repository for Information Sharing. URL: https://apps.who.int/iris/handle/10665/44607 [accessed 2021-07-13]
3. Pérez Sust P, Solans O, Fajardo JC, Medina Peralta M, Rodenas P, Gabaldà J, et al. Turning the crisis into an opportunity: digital health strategies deployed during the COVID-19 outbreak. JMIR Public Health Surveill 2020 May 04;6(2):e19106 [FREE Full text] [doi: 10.2196/19106] [Medline: 32339998]
4. Jimenez G, Matchar D, Koh CHG, van der Kleij R, Chavannes NH, Car J. The role of health technologies in multicomponent primary care interventions: systematic review. J Med Internet Res 2021 Jan 11;23(1):e20195. [doi: 10.2196/20195]
5. Kopec K, Janney CA, Johnson B, Spykerman K, Ryskamp B, Achtyes ED. Rapid transition to telehealth in a community mental health service provider during the COVID-19 pandemic. Prim Care Companion CNS Disord 2020 Oct 01;22(5) [FREE Full text] [doi: 10.4088/PCC.20br02787] [Medline: 33002350]
6. Sklar M, Reeder K, Carandang K, Ehrhart MG, Aarons GA. An observational study of the impact of COVID-19 and the rapid implementation of telehealth on community mental health center providers. Implement Sci Commun 2021 Mar 11;2(1). [doi: 10.1186/s43058-021-00123-y]

7. 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). [doi: 10.4088/JCP.20m13815] [Medline: 33989463]
8. 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 2010 Oct 19;38(2):65-76. [doi: 10.1007/s10488-010-0319-7]
9. 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]
10. 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. [doi: 10.1016/j.invent.2016.06.004] [Medline: 28090438]
11. Crilly P, Hassanali W, Khanna G, Matharu K, Patel D, Patel D, et al. Community pharmacist perceptions of their role and the use of social media and mobile health applications as tools in public health. Res Social Adm Pharm 2019 Jan;15(1):23-30. [doi: 10.1016/j.sapharm.2018.02.005] [Medline: 29501431]
12. Holtz B, Vasold K, Cotten S, Mackert M, Zhang M. Health care provider perceptions of consumer-grade devices and apps for tracking health: a pilot study. JMIR Mhealth Uhealth 2019 Jan 22;7(1):e9929 [FREE Full text] [doi: 10.2196/mhealth.9929] [Medline: 30668515]
13. 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]
14. 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]
15. Torous J, Jän Myrick K, Rauseo-Ricupero N, Firth J. Digital mental health and COVID-19: using technology today to accelerate the curve on access and quality tomorrow. JMIR Ment Health 2020 Mar 26;7(3):e18848 [FREE Full text] [doi: 10.2196/18848] [Medline: 32213476]

Abbreviations

APRN: advanced practice registered nurse
BHC: behavioral health consultant
CHERRIES: Checklist for Reporting Results of Internet E-Surveys
PCP: Primary care providers
REDCap: Research Electronic Data Capture

Edited by A Mavragani; submitted 05.08.21; peer-reviewed by A Saad, C Gould; comments to author 10.09.21; revised version received 08.10.21; accepted 18.01.22; published 25.02.22.

Please cite as:
Zaslavsky O, Chu F, Renn BN
Patient Digital Health Technologies to Support Primary Care Across Clinical Contexts: Survey of Primary Care Providers, Behavioral Health Consultants, and Nurses
JMIR Form Res 2022;6(2):e32664
URL: https://formative.jmir.org/2022/2/e32664
doi:10.2196/32664
PMID:35212642

Original Paper

Participatory Design of a Mobile App to Safeguard Mental Resilience in the Context of Drug Use in Young Adults: Multi-Method Study

Ofri Ben-Yehuda[1], MA; Efrat Dreazen[1], MA; Danny Koren[1], PhD; Mor Peleg[2], PhD

[1]Department of Psychology, University of Haifa, Haifa, Israel

[2]Department of Information Systems, University of Haifa, Haifa, Israel

Corresponding Author:
Mor Peleg, PhD
Department of Information Systems
University of Haifa
Abba Khoushy Ave 199
Haifa, 3498838
Israel
Phone: 972 048288509
Email: morpeleg@is.haifa.ac.il

Abstract

Background: Existing mental health apps are largely not aimed at generally healthy young people who may be experimenting with addictive substances and mind-altering experiences.

Objective: The aim of this study is to examine the interest and expectations of young people regarding a proposed smartphone app designed to help protect and promote mental health and resilience in the face of risks associated with substance use.

Methods: The study was based on agile system development and had 3 empirical substudies. Our feasibility study (study 1) included an anonymous questionnaire that examined the potential interest of young people in this type of app. It was answered by 339 Israelis aged 18-30 years. The second part of the feasibility study was a pilot study with 1.2% (4/339) of the people who answered the questionnaire and expressed interest in participating in a focus group. They tested and refined the elements planned for the focus groups. Study 2 was a participatory design study involving 7 focus groups of 5 to 7 participants each (young people aged 18-35 years, n=38). Persona development, open discussion, and a Technology Acceptance Model questionnaire were used to elicit user expectations and requirements for the app and to understand the perceived usefulness and usability of the proposed features. Study 3 comprised in-depth interviews with experts in the field of youth mental health and drug use to enlist their professional opinion regarding the value of such an app and recommendations about the features it should include.

Results: The mock-up for the proposed app had five key features: personalized assessment of risk for a drug-associated mental crisis, support for self-monitoring, useful information (eg, warning signs and first-aid guidelines), resilience-building exercises, and a support center. Participants rated highly the usefulness of all 5 main features and 96% (24/25) of the specific features we proposed within those main categories. The participants also suggested additional features as well as a new user persona we had not considered: the parents or family members of the young person. The focus groups rated highly the perceived usability of the app. Most of the experts saw value in all the main features and suggested specific knowledge sources for the app's content. Finally, participants of both the feasibility study and the participatory design study expressed moderate to high interest in using the app for self-help and high interest in using the app to help friends.

Conclusions: The findings provide preliminary encouraging support for the 5 main features suggested by the research team and reinforce recommendations for mobile health apps found in the literature. The findings emphasize the insight that this kind of app should be designed primarily for use by individuals seeking to help others.

(JMIR Form Res 2022;6(2):e34477) doi:10.2196/34477

KEYWORDS

mobile health; mHealth; eHealth; telehealth; mental health; mental resilience; participatory design; mobile phone

Introduction

Background

Adolescence and young adulthood are characterized by major changes in all areas of development: social, emotional, physical, and cognitive [1]. The important changes during this period generate instability and uncertainty and a significant mental health risk [2]. Studies indicate that approximately 75% of the psychological morbidity experienced throughout life erupts during early adulthood [3,4].

Moreover, mental disorders that develop during this age often persist, disrupting the capacity of young people to fulfill their potential [5], limiting access to mental and physical health care [6], and exposing them to poor educational and reduced occupational opportunities [7], stigma, social isolation, discrimination, and violation of human rights [8], as well as higher morbidity and mortality risks (including suicide) than the general population, translating into a striking 10-20 years' reduction in life expectancy [9,10].

The risk factors for developing emotional problems in adolescents and young adults include the use of drugs or other addictive substances, the pursuit of stimulating and spiritual experiences, dangerous behavior and delinquency, unsafe sex, and exposure to trauma [11]. Our main interest here is in the first 2 of these, primarily the use of addictive drugs.

In Israel, as elsewhere, recent years have seen a rise in the number of young people engaging in the use of drugs, mainly cannabis, and in mind-altering techniques, including spiritual and religious practices [12]. Both are particularly common in the context of long-term backpacking—a social phenomenon in which young people travel to exotic (for them) locales, often rich in natural wonders, for periods lasting a few months to a year or more. Long-term backpacking is prevalent among young Israelis, approximately 50,000 of whom set out on such trips every year [13]. Estimates of drug use among Israeli backpackers range from 50% to 90% [13,14]. Each year, approximately 2000 Israeli backpackers are negatively affected by drugs during their trip. Of these, approximately 600 are forced to return to Israel for mental and physical treatment, and dozens require immediate rescue [14].

Alongside their being at increased risk for mental problems, adolescents and young adults still retain a certain degree of physical and psychological plasticity, which facilitates early preventive interventions [15]. Yet, despite the high potential of early identification and intervention to improve young people's physical and mental health, many of those at high risk fail to take advantage of the various professional services on offer [16]. This failure may have many causes, ranging from a lack of information about the risks to psychological inhibitions stemming from social stigma regarding distress at this age [17]. In accordance with this latter explanation, several studies have observed that young adults are more likely to ask for help for their friends than for themselves [18,19]. Likewise, there is evidence that young people are more willing to seek out information when it is offered over the internet [20]. This pattern might reflect both the convenience of accessing information on the web [19] and the fact that seeking help on the web allows the seeker to remain anonymous [21]. This, in turn, may diminish the stigma associated with a sense of being isolated and different, strengthening young people's willingness to access forms of support that may improve their sense of meaning and self-worth [22,23].

For these and other reasons, over the past 2 decades, there has been a substantial rise in the number of mobile health (mHealth) apps focusing on mental health [24-31]. A 2017 survey reported that >10,000 mental health and wellness apps are available for download [32]. Therefore, we first searched the web for an app with a purpose similar to ours—to promote mental health and prevent drug-induced psychosis among young adults. Although we did not find one that meets these requirements, we found many apps that are each related to some of the different mental health and wellness issues. Specifically, we searched for studies related to apps designed to prevent or treat early psychosis, support an individual in mental distress (experiencing anxiety, depression, or posttraumatic stress disorder symptoms), build mental resilience, and minimize the harms of substance use. We focused on apps targeted at young people and that were demonstrated as effective. We found that the main features offered by such apps include information and psychoeducation, tools for developing mental resilience, risk or symptoms assessment, and support for self-monitoring and independent acquisition of therapeutic techniques (Table 1). Some apps are even intended to replace traditional face-to-face therapies, such as those based on cognitive behavioral therapy (CBT) [33]. We also found these features, excluding useful information related to drugs, in an app that is intended to develop resilience in youth without mental disorder diagnoses (Table 1). Apps that focus on harm reduction of drug use (Table 1) mostly offer educational information and contact details for support services but do not offer personalized risk assessment or resilience-building methods. However, the effectiveness of these 2 apps was not evaluated in research. We found supporting studies on apps designed to manage substance abuse disorder, but they do not meet our goals.

Table 1. Features of sample mental health apps and apps that offer drug education and support.

App name	Target end users	Risk assessment	Self-monitoring	Useful information	Mental resilience development	Support center	Reference
Joy Pop	Youth	✓	✓		✓	✓	[25]
RobinZ	Young people at risk for psychosis	✓	✓	✓	✓	✓	[26]
ThinkApp	Youth after first episode of psychosis	✓		✓	✓	✓	[27]
Actissist	Early psychosis relapse	✓	✓	✓	✓	✓	[28]
PTSD[a] coach	PTSD	✓		✓	✓		[29]
SAM[b]	Anxiety disorders		✓		✓	✓	[30]
Optimism	Mood disorders		✓		✓		[31]
MindZone	Drug users			✓		✓	[34]
TripSit	Drug users			✓		✓	[35]

[a]PTSD: posttraumatic stress disorder.

[b]SAM: Self-help for Anxiety Management.

Another shortcoming of existing apps designed to promote mental health is that, despite the extensive research conducted in this area, most of these apps (approximately 98%) have not been accompanied by evaluation studies [24,36]. However, meta-analysis studies found some support for effectiveness in health promotion and management of various emotional problems of young people [37,38].

After examining the existing literature, a central question that remained was whether young people would be disposed to use an app that could protect their mental health (ie, lower the risk of mental breakdowns associated with drug use and engagement in mind-altering spiritual experiences) and whether they would use it for a relatively long period of time (eg, for the entire duration of their backpacking trips, which usually span several months).

Proposed App

In this research, we used agile software development methods [39], including participatory design methods [40], to derive "empathetic solutions that are more desirable to target populations" [41] and to evaluate designs for them, which also draw from behavioral theories and existing apps. The solution that we derived, based on the aforementioned principles, is a smartphone app designed to help young people cope with challenges posed by young adulthood, with an emphasis on substance use, by delivering personalized information, guidance, and support. The proposed SafeGuard app is envisioned as offering five high-level features, similar to those reported in the literature and summarized in Table 1: risk assessment (personalized assessment of both generalized and situational risk for a drug-induced mental crisis), support for self-monitoring (eg, a status assessment and daily recommendations), personalized information related to drug use and treatment of worrying symptoms, exercises and coping strategies (including CBT) to increase mental resilience, and a support center offering quick access to support and help. The main goals of this study are to address the aforementioned open questions by (1) examining whether young people find the idea of the app beneficial as well as the potential interest of young people in using such an app and (2) assessing the expectations of both young people themselves and health care professionals regarding the contents of the app based on the 5 high-level feature categories. Our research objectives are as follows:

- Objective 1a: To obtain young people's views about the proposed SafeGuard app, specifically whether they view it as likely to be beneficial. This was operationalized as (1) perceiving recreational drug use as widespread, (2) perceiving existing information about the risks of drug use as insufficient, and (3) perceiving their friends as likely to use the app. This third condition is based on a large body of literature showing that young adults are more likely to ask for help for their friends than for themselves [18,19].
- Objective 1b: To assess whether young people express interest in using the proposed app, whether for self-help or for the benefit of their friends. This was reflected by reported intentions to download and use the app, and to recommend it to their friends.
- Objective 2: To elicit requirements for the app from young people.
- Objective 3: To obtain experts' opinions about the perceived usefulness of the proposed app and enlist their requirements as well as their opinions about the proposed features gathered according to the first stages of the research: risk assessment, self-monitoring, useful information, resilience development, and a support center.
- Objective 4: To acquire specific knowledge sources that could potentially be used to implement these features.

Objective 1a served as a feasibility assessment and was tested with a large cohort of young adult participants (n=339; study 1). Objective 1b and objective 2 were addressed through qualitative research with participants (n=38) from the same population as study 1, using a participatory design methodology with mock-ups of the proposed app (study 2). Objective 3 and Objective 4 were addressed by means of interviews with health care professionals (N=10) as part of the app development (study 3).

Methods

Overview

The study is based on agile system development—a set of practices in which requirements are gathered and the software is developed, tested, and improved in an iterative process [39]. We first conducted our feasibility study (study 1). We then relied on our knowledge of the literature and the results of study 1 to create initial mock-ups for the app. Next, we used participatory design methods to allow participants of the focus groups from the population of intended users to develop personas and requirements for the app and to evaluate the initial mock-ups through an anonymous questionnaire (study 2). The mock-ups were also evaluated by experts (study 3). The combined methods form triangulation (cross-referencing of data from at least three different independent sources of information, such as observation and interviews) [42,43] and thus help us overcome the limitations of qualitative research.

Ethics Approval

Ethics approval was obtained from the ethics committee of the University of Haifa (#160/19) on May 9, 2019.

Study 1: Feasibility

Participants

Our intention was to recruit 500 respondents to have a representative sample of the 50,000 young Israelis who set out on backpacking treks each year [14]. Participants were recruited using the snowball method through announcements posted in internet discussion groups and forums, representing a convenience sample. The announcements included a short explanation of the research objectives and the option to choose to participate in a short survey. In addition, the announcements indicated that participation in the study would enable participants to enter a lottery for a large trekkers' backpack. Participants comprised 339 Israelis aged 18-30 years who were not currently serving in the military. The participants were accepted into the study between May 19, 2019, and August 14, 2019. All participants signed an electronic informed consent form.

Procedure and Measures

The young people who agreed to participate in the feasibility study were asked to complete a short anonymous questionnaire. Basic sociodemographic background information (age, gender, religious background, occupational status, etc) was collected to ensure the representativeness of the study sample. The main measures comprised three 5-point Likert scale questions. The first two assessed participants' perceptions regarding the extent of drug use among young people and the availability of information on the topic—specifically, "My impression is that use of recreational drugs is a common phenomenon" and "My impression is that young people do not have enough accessible information about the risks associated with drug use" (1=Do not agree at all, 5=Agree completely). The third introduced the main features of the app as described previously and asked, "To what extent do you think your friends will use the proposed app?" (1=Not at all, 5=Daily). At the end of the questionnaire, participants interested in joining a focus group were invited to notify the study organizers (see the *Study 2* section). Of the 339 participants, 4 (1.2%) took part in a 2-hour pilot in which we tested and refined the elements planned for the focus groups.

Data Analysis

Descriptive statistics means and SDs were calculated using SPSS software (version 25.0; IBM Corp). Potential associations between intention to use the app (operationalized as an agreement that friends would use the app) and demographic characteristics were tested using Spearman correlations and analysis of variance (ANOVA).

Study 2: Participatory Design

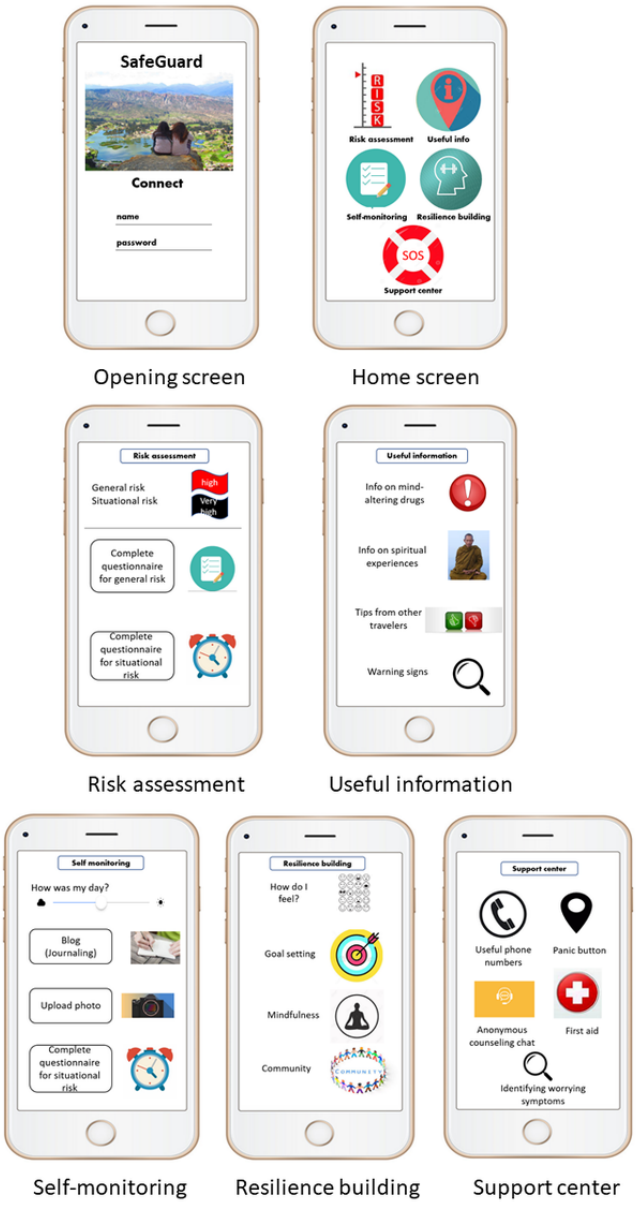
Overview

As mentioned previously, agile software development [39] is a methodology in which requirements are gathered and the software is developed, tested, and further improved in an iterative process. The team that develops the software includes developers and potential users, an approach that is also known as participatory design [40]. We applied the first 2 phases of the Integrate, Design, Assess, and Share (IDEAS) framework of agile strategies, which draws on evidence-based theories to develop more effective digital interventions to change health behavior [41,44]. In the *Integrate* phase, we integrated insights from users through qualitative research, including focus groups and questionnaires, to understand the target population and their needs. The focus of our qualitative study was on understanding the drug-related behavior of young people and their willingness to make use of, and benefit from, an app that could help them develop mental resilience and lower their risk of mental breakdowns, which is increased by substance abuse. To ground the proposed app in behavioral theories, we drew for this study on psychological theories of resilience and related theory-based behavioral strategies, including CBT [45-48]. In the *design* phase of IDEAS, our multidisciplinary coauthors' team carried out the ideation phase by brainstorming creative strategies for translating these theories into requirements for specific app features for strengthening executive functions (planning, awareness, and flexibility) and fostering interpersonal communication. The features included web-based groups (chat style) of the app users, setting personal goals, self-reporting and monitoring, and mindfulness techniques. We also drew requirements from our review of the literature and of related apps, arriving at 5 high-level feature categories of the app, described in the *Introduction* section. On the basis of the extension of IDEAS in the study by Peleg et al [44], we translated the theory-based feature into user interface (UI) mock-ups (Figure 1) that partition the features into a hierarchy

of app screens. The mock-ups were evaluated in studies 2 and 3.

Study 2 used a participatory design methodology. Participatory design involves potential users in eliciting requirements (the first stage of agile development) and evaluating their implementation [49]. Participatory design has often been used to develop mHealth apps [50-53]. Different participatory design techniques exist [49]. We used focus groups, persona development, user feedback, and a questionnaire, as detailed in the next sections.

Figure 1. User interface mock-ups of the app's screens.



Participants

Participants in this study were young Israelis (n=38) aged 18-30 years who were not currently serving in the military. Participants were recruited in two main ways: (1) through direct invitations to participants in study 1 and (2) through advertisements posted in 15 web-based discussion groups and forums. All participants provided informed consent. They were divided into 7 focus groups of 5 to 7 participants each.

Procedure and Measures

Each focus group met for a similar 2-hour session held in either Tel Aviv or Haifa (Israel's main central city and northern city, respectively). The sessions had 3 components, based on standard participatory design practices [49]. First, participants suggested hypothetical personas, characterizing prototypical users and formulating *stories* about situations where they would use the proposed app and their expectations from it. Next, the research team presented the aforementioned UI mock-ups and solicited feedback regarding the proposed features (and any features deemed by participants to be missing) through open discussion.

Discussions regarding the personas and mock-ups were audio recorded and later transcribed for qualitative analysis.

Finally, participants responded anonymously to a written questionnaire (adapted from the Technology Acceptance Model questionnaire [54]) assessing the perceived usefulness of specific features of the app and intentions to use the fully functioning app in the future. In part 1 of the questionnaire (25 items), participants were asked to rate the importance of each feature, whereas in part 2 (11 items), they were asked to rate the usefulness of the app and their intentions to use it themselves or to help friends (see details in the *Study 2*: *Participatory Design* section under *Results*). The results were stored in the research software without linking the details of the participants in the focus groups and their email addresses to the questionnaire data. In addition, the focus group protocol was securely maintained by the research team and was not published or transmitted in any way.

Data Analysis

Objective 1b: Intentions to download and use the app for self-help or for the benefit of their friends and to recommend it to friends were examined through descriptive statistics, including means, SDs, and distribution of the ratings. In addition, we tested whether and to what extent participants intended to use the app themselves or to help friends by subjecting the relevant questionnaire responses to a 2-tailed dependent means *t* test (see the *Results* section). Finally, associations between demographic and drug use characteristics (age, gender, and different types of drug exposure) and intentions to use the app were tested using ANOVA followed by a multivariate linear regression model.

Study 3: Expert Interviews

Interviews are an efficient and focused method of gathering experts' opinions [55] and are used in the IDEAS [41] framework that we followed "to gain a deeper understanding of the selected target population and their needs."

Participants and Procedure

We contacted 12 professionals in fields relevant to the study; of the 12, 7 (58%) were identified through a web-based search for leading professionals in the field of risks to youths' mental health and resiliency in general and due to drug use in particular. Of these 7, 2 (29%) were contacted in earlier phases of the research. Of the 12 professionals, 5 (42%) were recommended by other interviewees. The experts were emailed a brief description of the study, and those who responded were then sent more detailed information. Of the 12 experts, 10 (83%) agreed to be interviewed, representing the following fields: substance abuse prevention, early identification and intervention with young adults at risk for mental health crises due to drug use, search and rescue specialists, and resilience development experts. Of the 10 experts, 8 (80%) were interviewed individually and 2 (20%), who work with *Safe Shore*, a psychedelic harm reduction, education, and peer support project, were interviewed together. All interviews were audio recorded and transcribed for later analysis. All the interviewed experts provided informed consent.

Data Analysis

Of the 10 experts, 1 (10%) agreed to be interviewed only on the topic of resilience development. For the other experts, we used semistructured interviews in which we presented the mock-ups and solicited feedback about the different features while collecting recommendations for specific content. We conducted directed content analyses [56] to assess perceived usefulness, impediments, and general impressions of the app. We used this method [56] also to reveal commonalities and differences among the opinions and suggestions of the experts regarding specific app features. The literature sources suggested by the experts were analyzed to extract content for the app's features.

Results

Study 1: Feasibility

Sociodemographic Characteristics of the Sample

The demographic characteristics of the feasibility study participants are shown in Table 2. Except for the overrepresentation of women, the sample was representative of the population of nonultraorthodox youth in Israel in terms of basic socioeconomic characteristics, including age, occupational status, religion, and religious orientation, based on population totals from the Organization for Economic Cooperation and Development [57] and the Israel Central Bureau of Statistics [58,59]. According to these sources, 50.9% of the adult population in Israel hold an undergraduate degree [57], 75% are Jewish [58], and 45% have a secular orientation [59].

Table 2. Sociodemographic characteristics of the sample in study 1 (N=339).

Characteristics	Values
Age (years), mean (SD; range)	24.3 (2.7; 18-30)
Gender, n (%)	
Male	130 (38.3)
Female	209 (61.7)
Occupational status, n (%)	
Student	170 (50.1)
Full-time job	83 (24.5)
Before or during backpacking trip	50 (14.7)
Premilitary	6 (1.8)
Other	30 (8.8)
Religion, n (%)	
Jewish	288 (85)
Atheist	25 (7.4)
Christian	10 (2.9)
Druze	6 (1.8)
Muslim	3 (0.9)
Other	7 (2.1)
Religious orientation, n (%)	
Secular	228 (67.3)
Traditional	64 (18.9)
Orthodox	25 (7.4)
Other	22 (6.5)

Questionnaire Results

Table 3 presents the distribution and means of responses to 2 of the 3 main questions of the feasibility study. Most participants agreed strongly or completely that the use of recreational drugs is a common phenomenon (Q1) and felt that there is already enough educational information available regarding risks associated with drug use (Q2). As for Q3, the modal response regarding the frequency with which their friends would use an app of this kind was *sometimes* (141/339, 41.6%). Importantly, only 7.9% (27/339) of the participants responded that their friends would not use an app of this kind at all. (97/339, 28.6% indicated that they would seldom use the app, 67/339, 19.8% indicated that they would use the app often, and only 7/339, 2% said they would use it daily. A 1-way ANOVA revealed no significant differences based on gender, occupational status, or religious orientation regarding perceptions about the prevalence of drug use among young people, the accessibility of information about its risks, and friends' intention to use the proposed app.

Table 4 presents the intercorrelations among the 3 survey questions, as well as between each question and age. As can be seen, perceptions about the accessibility of information were correlated with perceptions of friends' intentions to use the proposed app ($P<.001$): the more they thought that there was insufficient information about the risks associated with drug use, the more frequently they thought that their friends would use the app. There was no significant correlation between age and responses to any of the 3 questions.

Table 3. Distribution of replies to the feasibility study questions Q1 and Q2 (N=339).

	Value, mean (SD)	1: Do not agree at all, n (%)	2: Somewhat agree, n (%)	3: Agree, n (%)	4: Strongly agree, n (%)	5: Completely agree, n (%)
Q1: My impression is that the use of recreational drugs is a common phenomenon	4.1 (0.9)	4 (1.2)	14 (4.1)	56 (16.5)	145 (42.8)	120 (35.4)
Q2: My impression is that young people do not have enough accessible information about the risks associated with drug use	2.8 (1.0)	33 (9.7)	105 (31)	120 (35.4)	59 (17.4)	22 (6.5)

Table 4. Matrix of Spearman correlations and 2-tailed *P* values of age and answers to the three study measures.

Variable name	Age	Q1[a]	Q2[b]	Q3[c]
Age				
r	1.00	—[d]	—	—
P value	<.001	—	—	—
Q1				
r	0.051	1.00	—	—
P value	.35	<.001	—	—
Q2				
r	0.001	0.033	1.00	—
P value	.99	.54	<.001	—
Q3				
r	–0.094	0.033	0.255	1.00
P value	.09	.54	<.001	<.001

[a]Q1: My impression is that the use of recreational drugs is a common phenomenon.

[b]Q2: My impression is that young people do not have enough accessible information about the risks associated with drug use.

[c]Q3: To what extent do you think your friends will use the proposed app?

[d]Not applicable.

Study 2: Participatory Design

Overview

Figure 1 shows the mock-up screens that were developed by our team and shown to participants of the focus groups (study 2) and the expert interviewees (study 3). As can be seen, calculated risk is visualized through red and black flags (high risk and very high risk, respectively), following the analogy of warning flags used at Israeli beaches to indicate safety conditions for bathing. As a nonnegligible level of risk for a mental breakdown after drug use always exists, the mock-up does not include a white flag (used on beaches for the lowest level of risk). The actual questionnaires that will be used to assess personal levels of risk [60] were not shown in the mock-up.

Sociodemographic Characteristics of the Sample

The sample for this study (including the pilot group and focus group participants) was demographically similar to the sample for study 1. The mean age of the pilot group participants was 25.8 (SD 4.4) years, and 55% (21/38) were women.

Development of Hypothetical Personas

Hypothetical personas were developed to help formulate the potential user population for the app and the requirements of each user. The participants envisioned 4 hypothetical personas, developed based on their own experiences. Multimedia Appendix 1 presents an evidence trace table containing quotes from the participants describing the personas and the contexts in which they might use the app.

The first persona represented young people using the app for self-help purposes while on a trip abroad. Young people who match this hypothetical persona are in a phase of life marked by self-exploration and working out their own identity as they enter adulthood. They are interested in new and exciting experiences, including recreational drug use. At the same time, they are presumed to be interested in guidance and direction that might give them some perspective before they embark on new experiences.

The second persona represented young people who had been exposed to mental health challenges or mental illness in their family and who want to act responsibly and keep their eyes open while traveling alone or trying new and risky experiences, including drugs or spiritual practices.

The third persona represented parents and family members who have concerns about the well-being of a child or sibling. The proposed app might be able to help these family members approach the young person in a supportive and nonthreatening way, while also helping them cope with their own concerns and distress.

The fourth persona represented a friend in distress. Many participants believed that young people tend to ignore, or be in denial about, their own problems. Concerned individuals might be able to use the app to help friends observed to be in a troubled mental state.

Requirements and Desired Features of the Envisioned App

Table 5 presents the focus groups' ratings for part 1 of the Technology Acceptance Model–based questionnaire in which participants were asked to rate the importance of various features of the app on a scale from 1 (not important at all) to 5 (necessary). As seen in Table 5, potential users (participants of the focus groups) rated as highly important 96% (24/25) of the features we introduced at the prototyping stage. The feature that they did not rate highly was the ability to upload images as part of the self-monitoring component.

Table 5. Means and SDs for responses on the importance of app features (part 1 of the Technology Acceptance Model–based questionnaire)[a].

Feature and details	Value, mean (SD)
Risk assessment	
The app should include a risk-assessment feature	4.13 (1.04)
Self-monitoring	
The app should include a self-monitoring feature	4.13 (0.85)
Image upload[b]	2.24 (1.22)
Icon selection to represent one's current mood	3.74 (1.03)
Personal blog and daily summary	3.38 (1.42)
Status assessment and daily recommendations	4.29 (0.84)
Useful information	
The app should include a useful information feature	4.47 (0.71)
Health and mental risks related to drug use	4.34 (0.79)
Health and mental risks related to spiritual experiences	4.16 (0.89)
Effects and side effects of different types of drugs and combinations	4.66 (0.65)
Red flags for developing mental distress that require immediate attention	4.70 (0.45)
Tips for avoiding or reducing risk	4.47 (0.84)
Location-dependent warnings about sources and unsafe places to consume drugs	4.37 (1.1)
Resilience development	
The app should include a resilience-development feature	4.24 (0.70)
Strengthening self-awareness	3.81 (1.02)
Strengthening executive functions	3.76 (1.16)
Mindfulness and relaxation techniques	3.66 (1.07)
Interpersonal communication	3.92 (1.29)
Support center	
The app should include a support center feature	4.74 (0.43)
Distress button that sends a message to the insurance company	4.39 (0.80)
Distress button that sends a message to people you have listed	4.58 (0.54)
List of useful phone and location assistance centers	4.70 (0.72)
Guidelines for identifying signs of distress	4.51 (0.67)
First-aid guidelines to help a person presenting signs of distress	4.42 (0.95)
Anonymity	
It is important for me to stay anonymous while using the app	4.05 (1.08)

[a]Participants answered all questions on a 5-point Likert scale where 1=not important at all, 2=not important, 3=neutral, 4=very important, and 5=necessary.
[b]The only app feature not rated highly by the participants.

The open discussion portion of the focus group sessions allowed participants to propose new features beyond those on our list. Altogether, 13 new features were suggested, of which 2 (15%) were raised by at least 10 different participants. The first was providing tips for parents and friends on ways to communicate concerns regarding risky drug use through supportive dialogue (12/38, 32%), and the second was a feature that would enable sharing one's own experience or reading about others' experiences of experimenting with drugs (10/38, 26%). In addition, many users suggested alternative ways to present various features proposed by the team. For example, participants proposed various ways for users to report on their current mood beyond the use of icons (eg, *How Do I Feel?*). The participants also suggested incorporating reminders to engage in various activities on the app, such as resilience activities or filling out a situational risk–assessment questionnaire.

Table 6 presents the means, SDs, and frequency distributions for responses to part 2 of the questionnaire. The first 2 sections of part 2 deal with the perceived usability and usefulness of the app. As can be seen, participants overwhelmingly agreed that the app would be easy to use, and most agreed or strongly agreed with 6 statements reflecting the perceived usefulness of the app.

The final section of part 2 contains 3 items reflecting intentions to use the app. Regarding objective 1b, responses to the first 3 intention measures suggest that young people would indeed be willing to use the app, either for self-help or to help their friends. Specifically, 53% (20/38) of the participants agreed or strongly agreed that they would download the app for personal use (mean 3.58, SD 0.98), whereas 76% (29/38) agreed or strongly agreed that they would download the app to help others (mean 4.21, SD 0.80) and 92% (35/38) agreed or strongly agreed that they would recommend the app to a friend (mean 4.29, SD 0.60).

Regarding the frequency of use, 50% (19/38) of the participants responded that they would use the app *sometimes* (13/19, 68%) or *often* (6/19, 32%; mean 2.55, SD 0.87), although none of the users intended to use it daily and 4/38, 11% said they would not use it at all. Interestingly, almost as many participants said that they would *seldom* (15/38, 39%) use the app for their own personal needs as responded either *sometimes* (13/38, 34%) or *often* (6/38, 16%). This finding is in keeping with the fact that the proportion of respondents saying that they would use the app to help others was significantly larger than the proportion saying that they would use it themselves (t_{32}=–3.125;*P*=.004). Both these findings support our operationalization of the feasibility study (Objective 1a) based on the idea that young people are more likely to seek help for their friends than for themselves [18,19].

Table 6. Means, SDs, and frequency distributions of responses on perceived usability, perceived usefulness, and intention to use the app (part 2 of the Technology Acceptance Model–based questionnaire; N=38).

	Value, mean (SD)	1: Completely disagree, n (%)	2: Disagree, n (%)	3: Neutral, n (%)	4: Agree, n (%)	5: Strongly agree, n (%)
Perceived usability						
Do you find the app easy to use?	4.29 (0.60)	0 (0)	0 (0)	3 (8)	21 (55)	14 (37)
Perceived usefulness						
Young people might benefit from this app	4.26 (0.50)	0 (0)	0 (0)	1 (3)	26 (68)	11 (29)
This app will raise awareness of the risks of drug use, both physical and mental	4.23 (0.67)	0 (0)	1 (3)	2 (5)	22 (58)	13 (34)
The app will provide tools to assess levels of personal risk from drug use	4.10 (0.79)	0 (0)	1 (3)	7 (18)	17 (45)	13 (34)
The app will help identify red flags that require immediate attention	4.50 (0.60)	0 (0)	0 (0)	2 (5)	15 (40)	21 (55)
The app will prevent drug use or encourage safe consumption	3.60 (0.85)	0 (0)	3 (8)	15 (40)	14 (37)	6 (16)
The app will *harm* young people by encouraging them to use drugs (low score means high benefit)	1.89 (0.86)	14 (37)	16 (42)	6 (16)	2 (5)	0 (0)
Intention to use the app						
I would download the app and use it for my personal needs	3.58 (0.98)	0 (0)	6 (16)	12 (32)	12 (32)	8 (21)
I would download the app and use it to help others	4.21 (0.80)	0 (0)	2 (6)	2 (6)	16 (49)	13 (39)
I would recommend the app to a friend	4.29 (0.60)	0 (0)	0 (0)	3 (8)	21 (55)	14 (37)

Study 3: Expert Interviews

We interviewed 10 professionals: 2 (20%) in the area of resilience development, 2 (20%) in the area of substance abuse prevention, 2 (20%) in the area of early intervention to prevent drug-induced mental crises, 1 (10%) rehabilitation and treatment expert, and 3 (30%) search and rescue professionals. Table 7 presents a high-level summary of the experts' opinions regarding whether the SafeGuard app should include the 5 high-level features (objective 3), along with ideas about how to implement them, including knowledge sources (objective 4), which are presented as footnotes below the table. Multimedia Appendix 2 [46-50] presents specific suggestions for improving the presentation of the app's features and sources of information other than those listed here.

Table 7. High-level summary of opinions of experts regarding the proposed app (N=10).

Feature	Experts, n (%)
The idea of the SafeGuard app	
In favor	10 (100)
Against	0 (0)
Neutral	0 (0)
Risk assessment[a]	
In favor	4 (40)
Against	2 (20)
Neutral	4 (40)
Self-monitoring[b]	
In favor	4 (40)
Against	0 (0)
Neutral	6 (60)
Useful information[c]	
In favor	8 (80)
Against	0 (0)
Neutral	2 (2)
Resilience building[d]	
In favor	8 (80)
Against	0 (0)
Neutral	2 (20)
Support center[e]	
In favor	9 (90)
Against	0 (0)
Neutral	1 (10)

[a]Risk-assessment questionnaire [60].

[b]Ad hoc guiding questions.

[c]Web-based library of the Israel Anti-Drug Authority [61].

[d]Specific methods mentioned by the experts include guided imagery, breathing exercises, movement exercises, mindfulness exercises, focusing on others when distressed, self-compassion development, and criticism reduction.

[e]Descriptions of symptoms that should raise red flags [62,63]; warnings from the Israel National Security Council [64].

Discussion

Overview

Our findings suggest that young people would indeed use an app designed to deliver personalized information on the effects of drug use, calculate one's personal risk for a mental breakdown, and recommend ways to deal with challenges and pressures in ways that increase mental resilience. The findings also support the 5 higher-order requirements and most of the specific features we envisioned for the app, as well as introducing 2 new features that might be incorporated. All the findings point to the app as being particularly useful in the context of long-term backpacking trips—a common rite of passage in which young adults spend time exploring their identity and trying new experiences in distant and unfamiliar locales, away from family and other supportive adults (eg, teachers).

In the sections that follow, we discuss the findings in relation to previous works in the literature, note the study's strengths and limitations, and offer directions for future research.

Principal Findings

Feasibility

Our participants expressed moderate to high interest in using the app themselves and high interest in using the app to help their friends. These findings are consistent with findings in the literature, which indicate that young people are often more comfortable seeking help for their peers than for themselves [17,18,65]. Although the possibility of accessing support through an app rather than through public services may remove some

of the many barriers to seeking help among young people, a certain reluctance to accept that they themselves might need such services still remains. However, the relatively high interest reported by our participants overall is encouraging.

Interestingly, the participants proposed a new user persona that we did not think of ourselves but which we find valuable: the parents or family members of a young person who might be at risk. Rickwood et al [17] argue that interventions should be targeted at the individuals to whom young people turn for information and help. Indeed, interventions by people in a young person's environment—parents, siblings, friends, and others—may help prevent mental breakdowns, stop mental disorders from worsening, and reduce distress [66,67]. The individuals who interact with a young person at risk are usually aware that they are well positioned to help, but they do not always have the tools, knowledge, and ability to do so [68]. Our proposed app can provide the supportive individuals with the requisite tools and knowledge to intervene at different stages of the young person's experience (eg, before, during, or after experimentation with recreational drugs).

Finally, we asked the participants to share with us the frequency with which they expected to use the app. This was important not only in terms of assessing interest in the app, but also in terms of our ability to improve the app later on by means of user data. For example, data on the effects of resilience-building exercises from people using the app on a daily basis and willing to share certain nonidentifying personal details (eg, changes in mood) could feed a machine learning algorithm, which could allow the app to predict which types of resilience-building activities are most helpful for different types of users. Similarly, we hoped to be able to discover new risk factors by automatically analyzing structured questionnaire replies and free-text reports. In fact, none of the participants expressed an intention to use the app daily, and only 16% (6/38) of the participants said that they would use the app often. Although we were disappointed in this finding, it was valuable for us to know the limits of the app's envisioned prediction features.

Design of the App

As described previously, we organized the app's features into five categories: risk assessment, self-monitoring, useful information, resilience development, and a support center. The findings of the study provided encouraging support for these categories. Through the discussion of the mock-up prototype and the written questionnaire in study 2, the focus group participants volunteered functional requirements and additional features that were organized around these categories. In addition, the participants contributed the nonfunctional requirement of anonymity, consistent with extensive literature [20,69-71]. The participants objected to the photo-upload feature in our mock-up, an objection consistent with the desire for anonymity. In this vein, some participants saw value in sharing written testimonials that could help others while not exposing their identity, in line with the study by Schueller et al [72]. However, we do not know whether all participants understood that to be perceived as credible, app contents (such as tips or advice) should disclose the authority of the author, representing the level of expertise of the person or persons writing the information, as well as information regarding the objectivity of the information (ie, how impartial and unbiased the source is) [73]. It is likely that not all participants were conscious of the conflict between privacy and credibility when sharing data with others and the fact that increasing credibility (by disclosing a name or even a nickname) could compromise privacy.

The participants also mentioned an ability to communicate with a counselor for support or help (Figure 1: "Support Center: Anonymous counseling chat") [26,74] as a desired feature. In fact, our proposed app offers several ways to receive instant support from real people (through the support center) as well as useful information. Our focus group participants also concurred with previous findings [33,75] that the amount of detailed information presented in mental health apps must be balanced with the need for a simple design, which enhances the user experience and promotes engagement.

Among the main features in our proposed app is CBT-based resilience-building activities, along with mindfulness exercises, goal setting, journaling experiences and feelings, and activities for promoting social connections, which is encouraged in the literature [24,76]. Several participants in the focus groups suggested that we incorporate reminders to engage with the app's situational risk–assessment and resilience-building features, a requirement that is congruent with the literature [77] and that we readily adopt.

Concerning objectives objective 3 and objective 4, all the experts we interviewed appreciated the potential usefulness of the proposed SafeGuard app. All the experts agreed on the importance of resilience development and the support center. Most of the experts agreed that self-monitoring, providing useful information, and risk assessment are useful features. The experts were able to provide useful knowledge sources for risk assessment, useful information, resilience building, and the support center. For self-monitoring, they suggested ad hoc guiding questions.

Strengths, Limitations, and Future Directions

Strengths

An important strength of this study is its use of both quantitative (questionnaires) and qualitative data to identify the requirements for the proposed app. This multi-method approach strengthens the reliability and validity of the findings. Furthermore, engaging end users throughout the development and evaluation processes enables a deep form of participation and a better understanding of patients' diverse health needs, as well as supports focusing on, and responding to, what matters to patients [78]. This has been advocated and reflected in regulations such as England's Patient and Public Participation Policy [78]. A related strength of our study relates to the raw nature of the design and the contents of the various features that were presented to the participants. The raw design was intentional because we wanted to gather potential users' opinions before investing time in a more precise design and to minimize potential bias due to too much detail. Consequently, it is impossible to rule out that a certain portion of the variance in the answers stems from different interpretations of the same features.

Limitations

However, our methods included some significant limitations that call for caution in interpreting the results. First, unlike common participatory design processes, the 5 key high-level features of the app were identified without any participant involvement. However, they were based on insights drawn from existing behavioral theories and on a review of similar apps in the literature and in the market. Hence, these features are not arbitrary.

Second, anticipated end users were included in the study; however, their views do not represent the views of all anticipated end users. Although we collected data from young people around the country, we derived our findings based on a limited sample size that overrepresents young secular Jewish female students (it should be noted that overrepresentation of female students is a common finding in web-based mental health research) [79]. Finally, another limitation of this study relates to its exclusive reliance on reporting intentions to use the app, rather than actual behavioral data. This limitation will be addressed in future studies (see the next section).

Future Directions

Future research should address fundamental questions about behavior change regarding drug experiences to broaden understanding regarding the motivations, self-efficacy, and triggers [80] that drive healthy young people to become more engaged with mental health apps and especially to use resilience-building activities to decrease the risk of a potential crisis after drug use.

Future research should also examine the relevance of the proposed app in light of the COVID-19 pandemic. The pandemic has changed many aspects of daily life, affecting people's behavior, lifestyle, and sense of security [81-85]. Young people facing early adulthood during the pandemic have had to cope with increased feelings of uncertainty and anxiety. Studies around the world show that the pandemic has led to increased consumption of drugs, especially cannabis [82-85]. Our proposed app, which offers ways for young people to self-assess their own risk and develop resilience, has the potential to be helpful during the present collective crisis.

Once a beta version of the app is developed, usability and marketing testing studies will be conducted using a small group of real users who will be using the app for several months. The usability testing study will collect the beta testers' feedback regarding their overall impression of the app; the degree to which it meets their needs; and whether the app's UI, design, and features are all deemed necessary. The marketing testing study will collect the beta testers' ideas and feedback on ways to attract attention and distribute the app. Among the key issues we intend to test in this study are (1) the type of messages to use to advertise the app (*mental health safeguard*, *resilience builder*, etc), (2) where to advertise it in the real world (eg, travelers' clinics, backpackers' equipment stores, Vipassana retreats, and trance music festivals) and on the internet (eg, specialized websites and communities of travelers), and (3) which social influencers to use. Once the final version is launched, we will test the degree to which it achieves its goals using a large-scale effectiveness study with a randomized controlled design.

Conclusions

Although they are exploratory in nature, these findings provide preliminary compelling support for the feasibility of the proposed SafeGuard mHealth app. In addition, they provide important insights and information regarding the features that young people, who are the target audience for the app, and experts in the field view as important. Some of these insights (eg, the one related to allowing an option to use the app for a friend) may be useful for the development of other mHealth apps.

Acknowledgments

The study was supported by a National Alliance for Research on Schizophrenia and Depression Independent Investigator Award (2010) to DK. The authors thank Chen Stanger for fruitful discussions. The authors also thank the experts who were interviewed: Dr Haim Mell, Dr Haggit Bonny Noach, Oren Gerstenhaber (MSW), Ziv Demeter (BA), Hilik Magnus, Omri Frish (BA), Prof Mooli Lahad, Prof Danny Hamiel, Nir Tadmor (MA), and Igal Tartakovsky.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Evidence trace table: contexts of use and characteristics of potential users.
[DOCX File , 20 KB - formative_v6i2e34477_app1.docx]

Multimedia Appendix 2
Expert evaluations and feedback.
[DOCX File , 24 KB - formative_v6i2e34477_app2.docx]

References

1. Suvisaari J, Aalto-Setälä T, Tuulio-Henriksson A, Härkänen T, Saarni S, Perälä J, et al. Mental disorders in young adulthood. Psychol Med 2009;39(2):287-299. [doi: 10.1017/S0033291708003632] [Medline: 18507875]
2. Arnett JJ, Žukauskienė R, Sugimura K. The new life stage of emerging adulthood at ages 18-29 years: implications for mental health. Lancet Psychiatry 2014;1(7):569-576. [doi: 10.1016/S2215-0366(14)00080-7] [Medline: 26361316]
3. 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;62(6):593-602. [doi: 10.1001/archpsyc.62.6.593] [Medline: 15939837]
4. Fusar-Poli P. Integrated mental health services for the developmental period (0 to 25 years): a critical review of the evidence. Front Psychiatry 2019;10:355 [FREE Full text] [doi: 10.3389/fpsyt.2019.00355] [Medline: 31231250]
5. McGorry P. Building the momentum and blueprint for reform in youth mental health. Lancet Psychiatry 2019;6(6):459-461. [doi: 10.1016/S2215-0366(19)30050-1] [Medline: 31003907]
6. Wang PS, Angermeyer M, Borges G, Bruffaerts R, Tat Chiu W, DE Girolamo G, et al. Delay and failure in treatment seeking after first onset of mental disorders in the World Health Organization's World Mental Health Survey Initiative. World Psychiatry 2007;6(3):177-185 [FREE Full text] [Medline: 18188443]
7. Mojtabai R, Stuart EA, Hwang I, Eaton WW, Sampson N, Kessler RC. Long-term effects of mental disorders on educational attainment in the National Comorbidity Survey ten-year follow-up. Soc Psychiatry Psychiatr Epidemiol 2015;50(10):1577-1591 [FREE Full text] [doi: 10.1007/s00127-015-1083-5] [Medline: 26082040]
8. Stuart H. Fighting the stigma caused by mental disorders: past perspectives, present activities, and future directions. World Psychiatry 2008;7(3):185-188 [FREE Full text] [doi: 10.1002/j.2051-5545.2008.tb00194.x] [Medline: 18836546]
9. Wasserman D, Cheng Q, Jiang GX. Global suicide rates among young people aged 15-19. World Psychiatry 2005;4(2):114-120 [FREE Full text] [Medline: 16633527]
10. Chesney E, Goodwin GM, Fazel S. Risks of all-cause and suicide mortality in mental disorders: a meta-review. World Psychiatry 2014;13(2):153-160 [FREE Full text] [doi: 10.1002/wps.20128] [Medline: 24890068]
11. Call KT, Riedel AA, Hein K, McLoyd V, Petersen A, Kipke M. Adolescent health and well-being in the twenty-first century: a global perspective. J Res Adolesc 2002;12(1):69-98. [doi: 10.1111/1532-7795.00025]
12. Harel-Fisch Y, Ezrachi Y. Psychoactive drug use among the adult population in Israel National Epidemiological Survey. Jerusalem: Israel Anti-Drug Authority; 2017.
13. Bonny-Noach H, Sagiv-Alayoff M. Rescuing Israeli travellers: effects of substance abuse, mental health, geographic region of rescue, gender and age of rescuees. J Travel Med 2017;24(5):tax045. [doi: 10.1093/jtm/tax045] [Medline: 28931135]
14. Zadok D. Drug use among backpackers: data and treatment frameworks. Knesset Research and Information Center. 2005. URL: https://fs.knesset.gov.il/globaldocs/MMM/ccb2d14b-9832-e811-80de-00155d0a0235/2_ccb2d14b-9832-e811-80de-00155d0a0235_11_9607.pdf [accessed 2021-07-18]
15. Arnett JJ. Emerging adulthood. A theory of development from the late teens through the twenties. Am Psychol 2000;55(5):469-480. [Medline: 10842426]
16. Dubow EF, Lovko KR, Kausch DF. Demographic differences in adolescents' health concerns and perceptions of helping agents. J Clin Child Psychol 1990;19(1):44-54. [doi: 10.1207/s15374424jccp1901_6]
17. Rickwood D, Deane FP, Wilson CJ, Ciarrochi J. Young people’s help-seeking for mental health problems. Aus J Adv Ment Health 2014;4(3):218-251. [doi: 10.5172/jamh.4.3.218]
18. Raviv A, Sills R, Raviv A, Wilansky P. Adolescents' help-seeking behaviour: the difference between self- and other-referral. J Adolesc 2000;23(6):721-740. [doi: 10.1006/jado.2000.0355] [Medline: 11161335]
19. Gilat I. Mental assistance in an online environment. Lexi-Kaye 2018:8-10 [FREE Full text]
20. Psychological therapy apps: digital tools for mental counseling. Tamir Institute of Psychotherapy. 2019. URL: https://www.tipulpsychology.co.il/psychiatric-counseling/smartphones-in-mental-health.html [accessed 2021-07-18]
21. Suler J. The online disinhibition effect. Cyberpsychol Behav 2004;7(3):321-326. [doi: 10.1089/1094931041291295] [Medline: 15257832]
22. Wittenberg N. Applying for online psychic assistance: auxiliary relationships and a therapeutic alliance in an online consultation process. Hebrew Psychology. 2014. URL: https://www.hebpsy.net/articles.asp?id=3223 [accessed 2021-07-18]
23. Manny-Ican I, Roze D, Marmor A, Berger-Tikochinsky T. Youth online - risk and opportunity for treatment. Szold Institute. 2016. URL: https://www.btl.gov.il/Funds/kolotkorim/Documents/skiratsafrotcyberbulling.pdf [accessed 2021-07-18]
24. Lau N, O'Daffer A, Colt S, Yi-Frazier JP, Palermo TM, McCauley E, et al. Android and iPhone mobile apps for psychosocial wellness and stress management: systematic search in app stores and literature review. JMIR Mhealth Uhealth 2020;8(5):e17798 [FREE Full text] [doi: 10.2196/17798] [Medline: 32357125]
25. MacIsaac A, Mushquash AR, Mohammed S, Grassia E, Smith S, Wekerle C. Adverse childhood experiences and building resilience with the JoyPop app: evaluation study. JMIR Mhealth Uhealth 2021;9(1):e25087 [FREE Full text] [doi: 10.2196/25087] [Medline: 33393908]
26. Traber-Walker N, Gerstenberg M, Metzler S, Joris MR, Karr M, Studer N, et al. Evaluation of the combined treatment approach "Robin" (standardized manual and smartphone app) for adolescents at clinical high risk for psychosis. Front Psychiatry 2019;10:384 [FREE Full text] [doi: 10.3389/fpsyt.2019.00384] [Medline: 31244692]

27. Barbeito S, Sánchez-Gutiérrez T, Mayoral M, Moreno M, Ríos-Aguilar S, Arango C, et al. Mobile app-based intervention for adolescents with first-episode psychosis: study protocol for a pilot randomized controlled trial. Front Psychiatry 2019;10:27 [FREE Full text] [doi: 10.3389/fpsyt.2019.00027] [Medline: 30804818]
28. Bucci S, Barrowclough C, Ainsworth J, Machin M, Morris R, Berry K, et al. Actissist: proof-of-concept trial of a theory-driven digital intervention for psychosis. Schizophr Bull 2018;44(5):1070-1080 [FREE Full text] [doi: 10.1093/schbul/sby032] [Medline: 29566206]
29. Owen JE, Jaworski BK, Kuhn E, Makin-Byrd KN, Ramsey KM, Hoffman JE. mHealth in the wild: using novel data to examine the reach, use, and impact of PTSD coach. JMIR Ment Health 2015;2(1):e7 [FREE Full text] [doi: 10.2196/mental.3935] [Medline: 26543913]
30. Matthews P, Topham P, Caleb-Solly P. Interaction and engagement with an anxiety management app: analysis using large-scale behavioral data. JMIR Ment Health 2018;5(4):e58 [FREE Full text] [doi: 10.2196/mental.9235] [Medline: 30287415]
31. Optimism. One Mind PsyberGuide. URL: https://onemindpsyberguide.org/apps/optimism/ [accessed 2021-07-26]
32. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. Evid Based Ment Health 2018;21(3):116-119. [doi: 10.1136/eb-2018-102891] [Medline: 29871870]
33. Bakker D, Kazantzis N, Rickwood D, Rickard N. Mental health smartphone apps: review and evidence-based recommendations for future developments. JMIR Ment Health 2016;3(1):e7 [FREE Full text] [doi: 10.2196/mental.4984] [Medline: 26932350]
34. MindZone home page. Project MindZone. URL: https://mindzone.info/ [accessed 2021-07-26]
35. TripSit wiki page. TripSit. URL: https://wiki.tripsit.me/wiki/Main_Page [accessed 2021-07-26]
36. Anthes E. Mental health: there's an app for that. Nature 2016;532(7597):20-23. [doi: 10.1038/532020a] [Medline: 27078548]
37. Fedele DA, Cushing CC, Fritz A, Amaro CM, Ortega A. Mobile health interventions for improving health outcomes in youth: a meta-analysis. JAMA Pediatr 2017;171(5):461-469 [FREE Full text] [doi: 10.1001/jamapediatrics.2017.0042] [Medline: 28319239]
38. Hollis C, Falconer CJ, Martin JL, Whittington C, Stockton S, Glazebrook C, et al. Annual research review: digital health interventions for children and young people with mental health problems - a systematic and meta-review. J Child Psychol Psychiatry 2017;58(4):474-503. [doi: 10.1111/jcpp.12663] [Medline: 27943285]
39. Lucassen G, Dalpiaz F, van der Werf JM, Brinkkemper S. Improving agile requirements: the Quality User Story framework and tool. Requirements Eng 2016;21(3):383-403. [doi: 10.1007/s00766-016-0250-x]
40. LeRouge C, Dickhut K, Lisetti C, Sangameswaran S, Malasanos T. Engaging adolescents in a computer-based weight management program: avatars and virtual coaches could help. J Am Med Inform Assoc 2016;23(1):19-28 [FREE Full text] [doi: 10.1093/jamia/ocv078] [Medline: 26159466]
41. Mummah SA, Robinson TN, King AC, Gardner CD, Sutton S. IDEAS (integrate, design, assess, and share): a framework and toolkit of strategies for the development of more effective digital interventions to change health behavior. J Med Internet Res 2016;18(12):e317. [doi: 10.2196/jmir.5927] [Medline: 27986647]
42. Mathison S. Why triangulate? Educ Res 1988;17(2):13-17. [doi: 10.3102/0013189x017002013]
43. Golafshani N. Understanding reliability and validity in qualitative research. Qual Rep 2003;8(4):597-606. [doi: 10.46743/2160-3715/2003.1870]
44. Peleg M, Michalowski W, Wilk S, Parimbelli E, Bonaccio S, O'Sullivan D, et al. Ideating mobile health behavioral support for compliance to therapy for patients with chronic disease: a case study of atrial fibrillation management. J Med Syst 2018;42(11):234. [doi: 10.1007/s10916-018-1077-4] [Medline: 30317403]
45. Building your resilience. American Psychological Association. 2012. URL: https://www.apa.org/topics/resilience [accessed 2021-07-18]
46. Fletcher D, Sarkar M. Psychological resilience: a review and critique of definitions, concepts, and theory. Eur Psychol 2013;18(1):12-23. [doi: 10.1027/1016-9040/a000124]
47. Lahad S, Shacham Y, Niv S. Coping and community resources in children facing disaster. In: Shalev AY, Yehuda R, McFarlane AC, editors. International handbook of human response to trauma. Amsterdam: Kluwer Academic Publishers; 2000:389-395.
48. Reivich KJ, Seligman ME, McBride S. Master resilience training in the U.S. Army. Am Psychol 2011;66(1):25-34. [doi: 10.1037/a0021897] [Medline: 21219045]
49. Inal Y, Wake JD, Guribye F, Nordgreen T. Usability evaluations of mobile mental health technologies: systematic review. J Med Internet Res 2020;22(1):e15337 [FREE Full text] [doi: 10.2196/15337] [Medline: 31904579]
50. Jessen S, Mirkovic J, Nes LS. MyStrengths, a strengths-focused mobile health tool: participatory design and development. JMIR Form Res 2020;4(7):e18049 [FREE Full text] [doi: 10.2196/18049] [Medline: 32706651]
51. Ospina-Pinillos L, Davenport TA, Navarro-Mancilla AA, Cheng VW, Cardozo Alarcón AC, Rangel AM, et al. Involving end users in adapting a Spanish version of a Web-based mental health clinic for young people in Colombia: exploratory study using participatory design methodologies. JMIR Ment Health 2020;7(2):e15914 [FREE Full text] [doi: 10.2196/15914] [Medline: 32027313]

52. LaMonica HM, Davenport TA, Roberts AE, Hickie IB. Understanding technology preferences and requirements for health information technologies designed to improve and maintain the mental health and well-being of older adults: participatory design study. JMIR Aging 2021;4(1):e21461 [FREE Full text] [doi: 10.2196/21461] [Medline: 33404509]
53. Müssener U, Thomas K, Linderoth C, Löf M, Åsberg K, Henriksson P, et al. Development of an intervention targeting multiple health behaviors among high school students: participatory design study using heuristic evaluation and usability testing. JMIR Mhealth Uhealth 2020;8(10):e17999 [FREE Full text] [doi: 10.2196/17999] [Medline: 33118942]
54. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989;13(3):319-340. [doi: 10.2307/249008]
55. Bogner A, Littig B, Menz W. Introduction: expert interviews — an introduction to a new methodological debate. In: Bogner A, Littig B, Menz W, editors. Interviewing experts. London: Palgrave Macmillan; 2009:1-13.
56. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
57. Adult education level. Organisation for Economic Co-operation and Development. 2020. URL: https://data.oecd.org/eduatt/adult-education-level.htm#indicator-chart [accessed 2021-07-28]
58. The population of Israel on the eve of 2021. Israel Central Bureau of Statistics. URL: https://www.cbs.gov.il/en/mediarelease/Pages/2020/Population-of-Israel-on-the-Eve-of-2021.aspx [accessed 2021-07-28]
59. Religion and self-definition of the extent of religiosity: selected data from the Society in Israel Report No. 10. Israel Central Bureau of Statistics. 2018. URL: https://www.cbs.gov.il/en/mediarelease/Pages/2018/Religion-And-Self-Definition-Of-Extent-Of-Religiosity-Selected-Data-From-The-Society-In-Israel-Report-No-10.aspx [accessed 2021-07-28]
60. Loewy RL, Pearson R, Vinogradov S, Bearden CE, Cannon TD. Psychosis risk screening with the Prodromal Questionnaire--brief version (PQ-B). Schizophr Res 2011;129(1):42-46 [FREE Full text] [doi: 10.1016/j.schres.2011.03.029] [Medline: 21511440]
61. Archive of the Israel anti-drug authority. Israel Anti-drug Authority. URL: https://tinyurl.com/43swatxw [accessed 2021-07-18]
62. Bonny-Noach H. Backpackers and drug abuse: a documentary, research, treatment and prevention perspective. Jerusalem: Bialik Publishers; 2008.
63. Frish O. Identifying a victim of cannabis. Mako Health. 2020. URL: https://www.mako.co.il/health-wellness/mental-health/Article-ea102ac58d98671026.htm [accessed 2021-07-26]
64. Warnings for travelers. Israel National Security Council. URL: https://www.nsc.gov.il/he/Travel-Warnings/Pages/default.aspx [accessed 2021-07-26]
65. Fox S. The social life of health information. California Healthcare Foundation. 2011. URL: https://www.pewresearch.org/internet/wp-content/uploads/sites/9/media/Files/Reports/2011/PIP_Social_Life_of_Health_Info.pdf [accessed 2021-07-18]
66. Yap MB, Wright A, Jorm AF. First aid actions taken by young people for mental health problems in a close friend or family member: findings from an Australian national survey of youth. Psychiatry Res 2011;188(1):123-128. [doi: 10.1016/j.psychres.2011.01.014] [Medline: 21324529]
67. Keitner GI, Ryan CE, Miller IW, Kohn R, Bishop DS, Epstein NB. Role of the family in recovery and major depression. Am J Psychiatry 1995;152(7):1002-1008. [doi: 10.1176/ajp.152.7.1002] [Medline: 7793434]
68. Jorm AF. Mental health literacy: empowering the community to take action for better mental health. Am Psychol 2012;67(3):231-243. [doi: 10.1037/a0025957] [Medline: 22040221]
69. Bradley KL, Robinson LM, Brannen CL. Adolescent help-seeking for psychological distress, depression, and anxiety using an Internet program. Int J Ment Health Promot 2012;14(1):23-34. [doi: 10.1080/14623730.2012.665337]
70. Farrer L, Gulliver A, Chan JK, Bennett K, Griffiths KM. A virtual mental health clinic for university students: a qualitative study of end-user service needs and priorities. JMIR Ment Health 2015;2(1):e2 [FREE Full text] [doi: 10.2196/mental.3890] [Medline: 26543908]
71. Pretorius N, Rowlands L, Ringwood S, Schmidt U. Young people's perceptions of and reasons for accessing a Web-based cognitive behavioural intervention for bulimia nervosa. Eur Eat Disord Rev 2010;18(3):197-206. [doi: 10.1002/erv.985] [Medline: 20443203]
72. Schueller SM, Neary M, O'Loughlin K, Adkins EC. Discovery of and interest in health apps among those with mental health needs: survey and focus group study. J Med Internet Res 2018;20(6):e10141 [FREE Full text] [doi: 10.2196/10141] [Medline: 29891468]
73. Sbaffi L, Rowley J. Trust and credibility in Web-based health information: a review and agenda for future research. J Med Internet Res 2017;19(6):e218 [FREE Full text] [doi: 10.2196/jmir.7579] [Medline: 28630033]
74. Holtz BE, McCarroll AM, Mitchell KM. Perceptions and attitudes toward a mobile phone app for mental health for college students: qualitative focus group study. JMIR Form Res 2020;4(8):e18347 [FREE Full text] [doi: 10.2196/18347] [Medline: 32667892]
75. Garrido S, Cheers D, Boydell K, Nguyen QV, Schubert E, Dunne L, et al. Young people's response to six smartphone apps for anxiety and depression: focus group study. JMIR Ment Health 2019;6(10):e14385 [FREE Full text] [doi: 10.2196/14385] [Medline: 31579023]

76. Tudor-Sfetea C, Rabee R, Najim M, Amin N, Chadha M, Jain M, et al. Evaluation of two mobile health apps in the context of smoking cessation: qualitative study of Cognitive Behavioral Therapy (CBT) versus non-CBT-based digital solutions. JMIR Mhealth Uhealth 2018;6(4):e98 [FREE Full text] [doi: 10.2196/mhealth.9405] [Medline: 29669708]
77. Chib A, Lin SH. Theoretical advancements in mHealth: a systematic review of mobile apps. J Health Commun 2018;23(10-11):909-955. [doi: 10.1080/10810730.2018.1544676] [Medline: 30449261]
78. NHS England patient and public participation policy. Version 2. National Health Service England. 2017. URL: https://www.england.nhs.uk/wp-content/uploads/2017/04/ppp-policy.pdf [accessed 2021-12-21]
79. Wong HW, Lo B, Shi J, Hollenberg E, Abi-Jaoude A, Johnson A, et al. Postsecondary student engagement with a mental health app and online platform (Thought Spot): qualitative study of user experience. JMIR Ment Health 2021;8(4):e23447 [FREE Full text] [doi: 10.2196/23447] [Medline: 33797395]
80. Fogg BJ. Tiny habits: the small changes that change everything. Boston: Houghton Mifflin Harcourt; 2019.
81. Panchal N, Kamal R, Cox C, Garfield R. The implications of COVID-19 for mental health and substance use. Kaiser Family Foundation. 2021. URL: https://www.kff.org/coronavirus-covid-19/issue-brief/the-implications-of-covid-19-for-mental-health-and-substance-use/ [accessed 2021-07-18]
82. Czeisler M, Lane RI, Petrosky E, Wiley JF, Christensen A, Njai R, et al. Mental health, substance use, and suicidal ideation during the COVID-19 pandemic - United States, June 24-30, 2020. MMWR Morb Mortal Wkly Rep 2020;69(32):1049-1057 [FREE Full text] [doi: 10.15585/mmwr.mm6932a1] [Medline: 32790653]
83. Dumas TM, Ellis W, Litt DM. What does adolescent substance use look like during the COVID-19 pandemic? Examining changes in frequency, social contexts, and pandemic-related predictors. J Adolesc Health 2020;67(3):354-361 [FREE Full text] [doi: 10.1016/j.jadohealth.2020.06.018] [Medline: 32693983]
84. Grass at noon: How the corona affects cannabis consumption in Israel Hebrew. Cannabis Magazine. 2020. URL: https://knb.is/2ycI54k [accessed 2022-02-06]
85. Caroppo E, Mazza M, Sannella A, Marano G, Avallone C, Claro AE, et al. Will nothing be the same again?: changes in lifestyle during COVID-19 pandemic and consequences on mental health. Int J Environ Res Public Health 2021;18(16):8433 [FREE Full text] [doi: 10.3390/ijerph18168433] [Medline: 34444180]

Abbreviations

ANOVA: analysis of variance
CBT: cognitive behavioral therapy
IDEAS: Integrate, Design, Assess, and Share
mHealth: mobile health
UI: user interface

Edited by G Eysenbach; submitted 26.10.21; peer-reviewed by M Brown, H Mehdizadeh; comments to author 18.12.21; revised version received 25.12.21; accepted 29.12.21; published 25.02.22.

Please cite as:
Ben-Yehuda O, Dreazen E, Koren D, Peleg M
Participatory Design of a Mobile App to Safeguard Mental Resilience in the Context of Drug Use in Young Adults: Multi-Method Study
JMIR Form Res 2022;6(2):e34477
URL: https://formative.jmir.org/2022/2/e34477
doi:10.2196/34477
PMID:35212631

Original Paper

Identifying Health-Related Discussions of Cannabis Use on Twitter by Using a Medical Dictionary: Content Analysis of Tweets

Jon-Patrick Allem[1], MA, PhD; Anuja Majmundar[2], MBA, PhD; Allison Dormanesh[1], MS; Scott I Donaldson[1], MS, PhD

[1]Department of Population and Public Health Sciences, Keck School of Medicine, University of Southern California, Los Angeles, CA, United States
[2]Department of Surveillance and Health Equity Science, American Cancer Society, Kennesaw, GA, United States

Corresponding Author:
Jon-Patrick Allem, MA, PhD
Department of Population and Public Health Sciences
Keck School of Medicine
University of Southern California
2001 N Soto Street
SSB 312D
Los Angeles, CA, 90032
United States
Phone: 1 323 442 7921
Email: allem@usc.edu

Abstract

Background: The cannabis product and regulatory landscape is changing in the United States. Against the backdrop of these changes, there have been increasing reports on health-related motives for cannabis use and adverse events from its use. The use of social media data in monitoring cannabis-related health conversations may be useful to state- and federal-level regulatory agencies as they grapple with identifying cannabis safety signals in a comprehensive and scalable fashion.

Objective: This study attempted to determine the extent to which a medical dictionary—the Unified Medical Language System Consumer Health Vocabulary—could identify cannabis-related motivations for use and health consequences of cannabis use based on Twitter posts in 2020.

Methods: Twitter posts containing cannabis-related terms were obtained from January 1 to August 31, 2020. Each post from the sample (N=353,353) was classified into at least 1 of 17 a priori categories of common health-related topics by using a rule-based classifier. Each category was defined by the terms in the medical dictionary. A subsample of posts (n=1092) was then manually annotated to help validate the rule-based classifier and determine if each post pertained to health-related motivations for cannabis use, perceived adverse health effects from its use, or neither.

Results: The validation process indicated that the medical dictionary could identify health-related conversations in 31.2% (341/1092) of posts. Specifically, 20.4% (223/1092) of posts were accurately identified as posts related to a health-related motivation for cannabis use, while 10.8% (118/1092) of posts were accurately identified as posts related to a health-related consequence from cannabis use. The health-related conversations about cannabis use included those about issues with the respiratory system, stress to the immune system, and gastrointestinal issues, among others.

Conclusions: The mining of social media data may prove helpful in improving the surveillance of cannabis products and their adverse health effects. However, future research needs to develop and validate a dictionary and codebook that capture cannabis use–specific health conversations on Twitter.

(JMIR Form Res 2022;6(2):e35027) doi:10.2196/35027

KEYWORDS

cannabis; marijuana; Twitter; social media; adverse event; cannabis safety; dictionary; rule-based classifier; medical; health-related; conversation; codebook

Introduction

The cannabis product and regulatory landscape is changing in the United States. A total of 34 states have legalized medical cannabis, and 10 states have legalized cannabis for adult recreational use (ie, for people aged 21 years or older) [1]. Against the backdrop of these changes, there have been increasing reports on health-related motives for cannabis use [2,3] and adverse events from its use [4]. Examples of motivations for cannabis use include treatment for clinical health conditions (eg, glaucoma, nausea, AIDS-associated anorexia, epilepsy, multiple sclerosis, and chronic pain) [5,6]—a use supported by the US Food and Drug Administration (FDA). Additionally, studies have shown that motivations for cannabis use have been based on the perceived benefits of its use, including its use as a sleep aid [2] and an aid for coping with stress or anxiety [3]. The low perception of harm from cannabis use when compared to that from other psychoactive drugs has also been documented as a motivation for its use [7]. However, cannabis use has been associated with adverse events, such as impaired short-term memory, impaired motor coordination, paranoia, and psychosis [6]; increased levels of depression and anxiety over time; symptoms of chronic bronchitis; addiction; and altered brain development [3,5,6]. Although the literature on the motivations for and effects of cannabis use is developing, medical experts recommend establishing a centralized federal agency for reporting, researching, and regulating cannabis products as a timely public health surveillance strategy [4]. The surveillance of the adverse health effects of cannabis is also a key priority of the US FDA [8]. The FDA's MedWatch program conducts the surveillance of serious adverse effects from cannabis use, but doubts have been raised over how effective this surveillance system is in identifying reports of cannabis safety signals [9].

The surveillance of health-related behaviors includes the use of digital data sources [10]. Publicly accessible data from individuals who post to social media platforms, such as Twitter, have been used to capture and describe the context of cannabis use [11,12]. However, health-related conversations surrounding its use have been understudied, and there has been a lack of cannabis-related studies that use social media data. The mining of social media data permits the collection and analysis of qualitative information, is noninvasive (ie, no demand effect), minimizes recall error, and allows for data to be captured in real time. Twitter has been a growing tool in health research, and it has been used for various purposes, including content analysis, surveillance, recruitment, intervention, and network analysis [13]. Twitter in particular reflects the views, attitudes, and behaviors of millions of people and is used by 22% of US adults (24% of men, 21% of women, 21% of White Americans, 24% of African Americans, and 25% of Hispanic Americans), with 42% of individuals using the platform daily [14].

This study attempted to determine the extent to which a medical dictionary—the Unified Medical Language System Consumer Health Vocabulary (CHV) [15]—could accurately identify cannabis-related motivations for use and health consequences of cannabis use based on Twitter posts in 2020. The findings may be useful to state- and federal-level regulatory agencies as they grapple with identifying cannabis safety signals in a comprehensive and scalable way.

Methods

Study Design

Twitter posts containing the cannabis-related terms *blunt*, *bong*, *budder*, *cannabis*, *cbd*, *ganja*, *hash*, *hemp*, *indica*, *kush*, *marijuana*, *marihuana*, *reefer*, *sativa*, *thc*, and *weed* were obtained from January 1 to August 31, 2020. These terms were informed by prior research that focused on comprehensively collecting cannabis-related posts on Twitter [11]. To treat each observation as independent, retweets were removed, leaving a total of 16,703,751 unique posts that contained these terms during this time. We used the following two dictionaries: (1) the Unified Medical Language System CHV [15], which comprises 13,479 medical terms (symptoms and diseases) that are used by consumers and health care professionals to describe health conditions, and (2) a list of 177 colloquial terms that were generated collaboratively by 2 trained coders and were related to the CHV terms when pertinent (eg, the colloquial expression of *inebriation* is *drunk*). The CHV has been used in prior research for the surveillance of health discussions about e-cigarette use or vaping on Twitter [16]. CHV terms are available at no cost to applicants who have a license, which is assigned upon the completion of a web-based application process. A sample of 609,227 cannabis-related posts referenced at least 1 of these terms.

We then identified and removed posts from social bots (ie, automated Twitter accounts) to reliably describe the public's health-related motivations for cannabis use or the perceived health effects of its use [17]. In order to distinguish nonbots from social bots, we relied upon Botometer (Observatory on Social Media) [18,19]. This program analyzes the features of a Twitter account and provides a score based on how likely the account is to be a social bot. The Botometer threshold was set to ≥4 on an English rating scale of 1 to 5. All Twitter accounts were screened after data were collected (ie, not in real time). During this process, 127,140 accounts responsible for the tweets in our data were deleted from Twitter. As a result, these accounts could not be processed through Botometer, and their posts were removed from our data. Of the 261,134 available accounts, 15,245 were marked as bots and removed. The final analytic sample contained 353,353 posts from 245,889 unique nonbot accounts.

Each post from the final sample was classified into at least 1 of 17 a priori health-related categories [16] by using a rule-based classifier. Each category was defined by the terms in the two dictionaries. The 17 health-related categories included 14 categories from prior research [16] and 3 additional categories that were unique to this study, accounting for the potential psychoactive effects of cannabis use (the "Cognitive" category), topical cannabis products (the "Dermatological" category), and the intersection of cannabis and food additives (the "Poisoning" category). A post could belong to multiple categories. The 17 categories, example keywords, and prevalence of keywords from each category can be found in Table 1.

A stratified random sample of posts (n=1092) was extracted from the corpus (n=353,353) based on the original classifications of the posts by using the rule-based classifier. A coding procedure (Multimedia Appendix 1 contains the complete codebook) was used to determine if each post pertained to a health-related motivation for cannabis use, a perceived adverse health effect of cannabis use, or neither. Two trained coders double coded each post independently, with κ values ranging from 0.790 to 0.856. Discrepancies were resolved by the two coders and the first author. This analysis served as a validation procedure for the rule-based classifier.

Table 1. Health categories, example keywords, and the frequency of occurrence on Twitter (N=353,353).

Health categories	Example keywords	Frequency, n (%)
Cancer	*Cancer*, *tumor*, and *malignant*	13,834 (3.92)
Cardiovascular	*Stroke*, *heart attack*, and *blood pressure*	1810 (0.52)
Cognitive	*Unconscious* and *attention*	8807 (2.49)
Death	*Die*, *kill*, and *lost life*	31,590 (8.95)
Dermatological	*Itchy*, *acne*, and *blister*	1557 (0.44)
Gastrointestinal	*Belly*, *belch*, *vomit*, and *puke*	10,434 (2.95)
Immune System	*Flu*, *common cold*, and *allergy*	12,229 (3.46)
Injury	*Injury*, *rupture*, *wound*, and *bruise*	19,490 (5.52)
Mental health	*PTSD*, *ADHD*, and *jittery*	100,155 (28.34)
Neurological	*Coma*, *dizzy*, and *lightheaded*	56,347 (15.95)
Other	*Anemia*, *jaundice*, and *mumps*	44,111 (12.48)
Pain	*Painful*, *achy*, and *cramping*	38,335 (10.85)
Poisoning	*Toxic*, *poisonous*, and *noxious*	8345 (2.36)
Pregnancy or in utero	*Pregnant*, *preggers*, and *miscarriage*	4760 (1.35)
Respiratory	*Cough*, *wheeze*, and *black lung*	16,616 (4.70)
Stress	*Stressed* and *cortisol*	13,372 (3.78)
Weight	*Fat*, *obese*, *weight*, and *stoutness*	5888 (1.67)

Ethical Considerations

All analyses relied on public, anonymized data; adhered to the terms and conditions, terms of use, and privacy policies of Twitter; and were performed under institutional review board approval from the authors' university. To protect privacy, no tweets were reported verbatim in this report.

Results

The validation process indicated that the medical dictionary could identify health-related conversations in 31.2% (341/1092) of posts (Table 2). Specifically, 20.4% (223/1092) of posts were identified as posts related to a health-related motivation for cannabis use, while 10.8% (118/1092) of posts were identified as posts related to a health-related consequence from cannabis use. The health-related conversations about cannabis use included those about issues with the respiratory system, stress to the immune system, and gastrointestinal issues, among others.

Table 2. The validation results for the rule-based classifier.[a]

Category	Motivations, n (%)	Consequence, n (%)	Neither, n (%)	Total[b], n
Medical term				
Cancer	15 (42.9)	4 (11.4)	16 (45.7)	35
Cardiovascular	1 (20)	1 (20)	3 (60)	5
Cognitive	5 (18.5)	3 (11.2)	19 (70.3)	27
Death	4 (4)	7 (8)	79 (88)	90
Dermatological	0 (0)	0 (0)	4 (100)	4
Gastrointestinal	6 (21)	1 (3)	22 (76)	29
Immune system	2 (6)	2 (6)	31 (88)	35
Injury	1 (2)	5 (9)	49 (89)	55
Mental health	89 (31.8)	19 (6.7)	172 (61.4)	280
Neurological	18 (11.3)	40 (25)	102 (63.7)	160
Other	33 (26.6)	7 (5.6)	84 (67.8)	124
Pain	28 (25.7)	3 (2.8)	78 (71.5)	109
Poison	2 (8.3)	6 (25)	16 (66.7)	24
Pregnant	2 (14.3)	2 (14.3)	10 (71.4)	14
Respiratory	0 (0)	17 (36.2)	30 (63.8)	47
Stress	17 (44.7)	1 (2.6)	20 (52.7)	38
Weight	0 (0)	0 (0)	16 (100)	16
Total[c]	223 (20.4)	118 (10.8)	751 (68.8)	1092[d]

[a]The values in the *Motivations*, *Consequence*, and *Neither* columns show the number and percentage of posts related to health-related motivations for cannabis use, health-related consequences from cannabis use, or neither, respectively, for each medical term.

[b]The *Total* column refers to the total number of tweets coded per medical term.

[c]The values in the *Total* row show the number and percentage of posts related to health-related motivations for cannabis use, health-related consequences from cannabis use, or neither, respectively, for all medical terms.

[d]The total number of tweets in the subgroup.

Discussion

Principal Findings

This study determined the extent to which a commonly used medical dictionary of health effects could accurately identify cannabis-related motivations for use and health consequences of cannabis use based on Twitter posts in 2020. This is the first study to date to use a high-quality medical dictionary of consumer-oriented health terms to capture the public's expressions of health concepts and thereby identify health conversations about cannabis use. The findings suggest that a medical dictionary alone is limited in its ability to identify health-related conversations in a cannabis context. The posts discussed the respiratory system, stress to the immune system, and gastrointestinal problems. The posts also discussed mental health, pain, injuries, and poisonings, among other potential health effects.

Previous research has identified motivations for cannabis use, including using cannabis to treat chronic conditions (eg, glaucoma, nausea, AIDS-associated anorexia, epilepsy, multiple sclerosis, and chronic pain) [2,5,6], using it as a sleep aid [2], and using it to help improve mental health (eg, stress, anxiety, and depression) [3]. Previous research has also identified adverse reactions associated with cannabis consumption based on search engine queries and found that such queries revealed many of the known adverse effects of cannabis use, such as coughing and psychotic symptoms, as well as plausible reactions that could be attributed to cannabis use, such as pyrexia [20]. A prior content analysis of 5000 tweets about "dabbing" (the use of a high-potency cannabis-related product) from a 30-day period in 2015 showed that the most common physiologic effects from this form of cannabis use were the loss of consciousness and respiratory effects, such as coughing [21]. Our study compliments prior research by using a professionally used term dictionary. It also indicates that the public made varied health-related references in their conversations about cannabis on Twitter. However, if the mining of social media data is to be proven helpful in the surveillance of cannabis products and their adverse health effects, the use of a standardized medical term dictionary alone will not suffice in the identification of cannabis safety signals. Future research will need to develop a codebook and term dictionary that incorporate a priori categories and data-driven inductive approaches that capture nuanced cannabis and health-related conversations on Twitter.

Limitations

This study focused on posts to Twitter, and the findings may not extend to other social media platforms. Additionally, the posts in this study were collected from an 8-month period in 2020; thus, the findings may not extend to other time periods. The data collection process relied on Twitter's Streaming application programming interface, which prevented the collection of posts from private accounts. As such, the findings may not generalize to all Twitter users or to the US population. The people responsible for each post in this study were not examined, and as a result, we could not describe the demographics of the Twitter users in this study. Further, Twitter posts can contain misspellings, and our lexicon-based exact matching approach likely missed these expressions. The CHV has also not been updated since 2011, which may in part explain its limited ability to identify health-related conversations in a cannabis context. Finally, this study could not determine modes of cannabis use or whether cannabis use was coupled with other substances or medications, which may impact perceived health effects.

Conclusions

Medical experts and regulatory agencies have called for the improved surveillance of cannabis products and the adverse health effects from cannabis use. Until the limitations with syndromic surveillance and hospital data systems for cannabis (eg, accessibility of data and timeliness) are resolved, the mining of social media data may clarify the public's experiences with cannabis use. The development of a validated dictionary and codebook that capture cannabis-specific health conversations may be key to advancing future efforts in the surveillance of Twitter data. A robust, national-level surveillance system for cannabis-related health effects may benefit from using real-time social media surveillance data on health effects and should consider using data from other sources (eg, emergency room visits and survey data).

Acknowledgments

This project was supported by funds provided by The Regents of the University of California, Research Grants Program Office, Tobacco-Related Diseases Research Program (grant 28KT-0003). The opinions, findings, and conclusions herein are those of the authors and do not necessarily represent those of The Regents of the University of California or any of its programs.

Authors' Contributions

JPA has full access to all of the data in this study and takes responsibility for the integrity of the data and the accuracy of the data analysis. JPA and AM contributed to the concept and design of this study. JPA, AM, and SID were responsible for the acquisition, analysis, and interpretation of the data. JPA drafted the manuscript. JPA, AM, SID, and AD critically revised the manuscript and approved the final version of the manuscript. SID conducted the statistical analysis. JPA obtained funding for this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Codebook for monitoring health-related discussions about cannabis use on Twitter.
[DOCX File , 42 KB - formative_v6i2e35027_app1.docx]

References

1. Hartman M. Cannabis overview. National Conference of State Legislators. 2021 Jul 06. URL: https://www.ncsl.org/research/civil-and-criminal-justice/marijuana-overview.aspx [accessed 2020-10-28]
2. Haug NA, Padula CB, Sottile JE, Vandrey R, Heinz AJ, Bonn-Miller MO. Cannabis use patterns and motives: A comparison of younger, middle-aged, and older medical cannabis dispensary patients. Addict Behav 2017 Sep;72:14-20 [FREE Full text] [doi: 10.1016/j.addbeh.2017.03.006] [Medline: 28340421]
3. Glodosky NC, Cuttler C. Motives matter: Cannabis use motives moderate the associations between stress and negative affect. Addict Behav 2020 Mar;102:106188. [doi: 10.1016/j.addbeh.2019.106188] [Medline: 31706141]
4. Mudan A, DeRoos F, Perrone J. Medical marijuana miscalculation. N Engl J Med 2019 Sep 12;381(11):1086-1087. [doi: 10.1056/NEJMc1907013] [Medline: 31509680]
5. Hill KP. Medical use of cannabis in 2019. JAMA 2019 Sep 10;322(10):974-975. [doi: 10.1001/jama.2019.11868] [Medline: 31397842]
6. Volkow ND, Baler RD, Compton WM, Weiss SRB. Adverse health effects of marijuana use. N Engl J Med 2014 Jun 05;370(23):2219-2227 [FREE Full text] [doi: 10.1056/NEJMra1402309] [Medline: 24897085]
7. Morgan CJA, Noronha LA, Muetzelfeldt M, Feilding A, Curran HV. Harms and benefits associated with psychoactive drugs: findings of an international survey of active drug users. J Psychopharmacol 2013 Jun;27(6):497-506 [FREE Full text] [doi: 10.1177/0269881113477744] [Medline: 23438502]
8. FDA regulation of cannabis and cannabis-derived products, including cannabidiol (CBD). U.S. Food and Drug Administration. URL: https://tinyurl.com/sjnp5zjp [accessed 2021-11-16]

9. Hines MC, Harinstein LM, Kortepeter CM. Reporting adverse events for cannabis to the FDA. N Engl J Med 2020 Jan 02;382(1):98. [doi: 10.1056/NEJMc1913460] [Medline: 31875515]
10. Althouse BM, Scarpino SV, Meyers LA, Ayers JW, Bargsten M, Baumbach J, et al. Enhancing disease surveillance with novel data streams: challenges and opportunities. EPJ Data Sci 2015;4(1):17 [FREE Full text] [doi: 10.1140/epjds/s13688-015-0054-0] [Medline: 27990325]
11. Allem JP, Escobedo P, Dharmapuri L. Cannabis surveillance with Twitter data: Emerging topics and social bots. Am J Public Health 2020 Mar;110(3):357-362. [doi: 10.2105/AJPH.2019.305461] [Medline: 31855475]
12. Lamy FR, Daniulaityte R, Sheth A, Nahhas RW, Martins SS, Boyer EW, et al. "Those edibles hit hard": Exploration of Twitter data on cannabis edibles in the U.S. Drug Alcohol Depend 2016 Jul 01;164:64-70 [FREE Full text] [doi: 10.1016/j.drugalcdep.2016.04.029] [Medline: 27185160]
13. Sinnenberg L, Buttenheim AM, Padrez K, Mancheno C, Ungar L, Merchant RM. Twitter as a tool for health research: A systematic review. Am J Public Health 2017 Jan;107(1):e1-e8. [doi: 10.2105/AJPH.2016.303512] [Medline: 27854532]
14. Perrin A, Anderson M. Share of U.S. adults using social media, including Facebook, is mostly unchanged since 2018. Pew Research Center. 2019 Apr 10. URL: https://tinyurl.com/3un9k66e [accessed 2019-12-04]
15. Zeng QT, Tse T. Exploring and developing consumer health vocabularies. J Am Med Inform Assoc 2006;13(1):24-29 [FREE Full text] [doi: 10.1197/jamia.M1761] [Medline: 16221948]
16. Majmundar A, Allem JP, Cruz TB, Unger JB, Pentz MA. Monitoring health effects of vaping discussed on Twitter in 2018 and 2019. J Gen Intern Med. Epub ahead of print 2021 Apr 09. [doi: 10.1007/s11606-021-06705-9] [Medline: 33835311]
17. Allem JP, Ferrara E. The importance of debiasing social media data to better understand e-cigarette-related attitudes and behaviors. J Med Internet Res 2016 Aug 09;18(8):e219 [FREE Full text] [doi: 10.2196/jmir.6185] [Medline: 27507563]
18. Ferrara E, Varol O, Davis C, Menczer F, Flammini A. The rise of social bots. Commun ACM 2016 Jul;59(7):96-104 [FREE Full text] [doi: 10.1145/2818717]
19. Botometer by OSoMe. Observatory on Social Media. URL: https://botometer.osome.iu.edu/ [accessed 2022-02-16]
20. Yom-Tov E, Lev-Ran S. Adverse reactions associated with cannabis consumption as evident from search engine queries. JMIR Public Health Surveill 2017 Oct 26;3(4):e77 [FREE Full text] [doi: 10.2196/publichealth.8391] [Medline: 29074469]
21. Cavazos-Rehg PA, Sowles SJ, Krauss MJ, Agbonavbare V, Grucza R, Bierut L. A content analysis of tweets about high-potency marijuana. Drug Alcohol Depend 2016 Sep 01;166:100-108 [FREE Full text] [doi: 10.1016/j.drugalcdep.2016.06.034] [Medline: 27402550]

Abbreviations

CHV: Consumer Health Vocabulary
FDA: Food and Drug Administration

Edited by A Mavragani; submitted 17.11.21; peer-reviewed by R Sun, M Navarro; comments to author 29.12.21; revised version received 05.01.22; accepted 21.01.22; published 25.02.22.

Please cite as:
Allem JP, Majmundar A, Dormanesh A, Donaldson SI
Identifying Health-Related Discussions of Cannabis Use on Twitter by Using a Medical Dictionary: Content Analysis of Tweets
JMIR Form Res 2022;6(2):e35027
URL: https://formative.jmir.org/2022/2/e35027
doi:10.2196/35027
PMID:35212637

Original Paper

Text Messaging Intervention for Mental Wellness in American Indian and Alaska Native Teens and Young Adults (BRAVE Study): Analysis of User Engagement Patterns

Julia Wrobel[1], PhD; Joshva Silvasstar[1], MSIS; Roger Peterson[2], BA; Kanku Sumbundu[1], MPH; Allyson Kelley[3], DrPH; David Stephens[2], BSN; Stephanie Craig Rushing[2], PhD; Sheana Bull[1], PhD

[1]Colorado School of Public Health, University of Colorado, Aurora, CO, United States

[2]Northwest Portland Area Indian Health Board, Portland, OR, United States

[3]Allyson Kelley and Associates PLLC, Sisters, OR, United States

Corresponding Author:
Julia Wrobel, PhD
Colorado School of Public Health
University of Colorado
13001 East 17th Place
Mail Stop B119
Aurora, CO, 80045
United States
Phone: 1 307 724 4585
Email: JULIA.WROBEL@cuanschutz.edu

Abstract

Background: Many American Indian and Alaska Native (AI/AN or Native) communities express concern about high rates of suicide and poor mental health. Technology-based health interventions that nurture resilience, coping skills, connectedness, and help-seeking skills may be an effective strategy for promoting health and wellbeing among AI/AN youth. The Northwest Portland Area Indian Health Board designed the BRAVE intervention for AI/AN youth. BRAVE is delivered via SMS text messaging and includes role model videos, mental wellness strategies, links to culturally relevant resources, and social support from family and friends.

Objective: The aim of this study is to explore system data from the BRAVE intervention to determine patterns of user engagement and differences in psychosocial outcomes based on the number of clicks on BRAVE content.

Methods: The BRAVE study included 1030 AI/AN teens and young adults nationwide (15 to 24 years old). The message series in the BRAVE and STEM study arms included 3 to 5 SMS text messages per week, featuring 1 role model video and 1 image per week. Messages were sent out via Mobile Commons (Upland Software Inc), a mobile messaging provider that supports text, picture, and video SMS.

Results: Of the 509 participants in the original BRAVE analysis, 270 had sufficient data to analyze user engagement, with at least 1 trackable click on a study SMS text message. Of the 270, 184 (68.1%) were female, 50 (18.5%) were male, and 36 (13.3%) selected another gender category. The average participant was 20.6 years old, with a minimum and maximum age of 15 and 26 years. Most participants had relatively low engagement measured by the number of clicks (median 2; mean 3.4), although others clicked message content as many as 49 times. Users engaged most frequently with the YouTube-based content (viewing 1 of 7 role model videos), with 64.8% (175/270) of total clicks coming from the role model videos, and earlier episodes receiving the highest number of clicks. Most baseline psychosocial measures were not significantly associated with the number of links clicked. However, help-seeking behavior was highly significant (P<.001), with a rate ratio of 0.82 (0.73, 0.92), indicating that each 1-unit increase in help-seeking score at baseline was associated with an 18% decrease in the expected number of study content clicks.

Conclusions: This is the first study to set initial standards for assessing user engagement in an mHealth intervention. Our work underscores the feasibility of exploring the impact of engagement on intended outcomes, allowing for more precise exploration of the dose-response relationship that may be realized through these low-touch interventions that offer promising potential for reaching high numbers of program participants.

Trial Registration: ClinicalTrials.gov NCT04979481; https://clinicaltrials.gov/ct2/show/NCT04979481

(JMIR Form Res 2022;6(2):e32138) doi:10.2196/32138

KEYWORDS

American Indian; Alaska Native; adolescent; mental health; help-seeking skills, text messaging; mHealth, behavioral intervention; user engagement; feasibility; engagement; low-touch; intervention

Introduction

Background

In the United States, many youths face a mental health crisis, with 1 of every 5 considering suicide each year and 1 million attempting suicide [1], now the third leading cause of death for this age group [2]. The teen years are frequently cited as the time when suicidal ideation begins, which underscores an urgent need to support mental health for teens [3,4]. Suicide prevention and mental health promotion remain a critical challenge for American Indian and Alaska Native (AI/AN or Native) communities, in particular [5]. Among AI/AN youth in the 9th to 12th grade, the past-year prevalence of suicidal thoughts, planning, and attempts was nearly 15% in 2017 [6]. Suicide was the second leading cause of death for Native youth aged 10 to 24 years, a rate that is 2.5 times higher than the national average [7].

Multilevel interventions are critically needed to build protective factors against suicide and violence [8]. According to the recent report, Culture Forward, key factors that protect against suicide include hope, self-efficacy, connectedness to family, community belonging, identity and participation in tribal culture, family living a traditional lifestyle, self-determination, spirituality, connectedness to community and lands, and talking to family and friends about problems [9]. Technology-based health interventions that nurture self-esteem, help-seeking skills, and connectedness to their self, peers, family, community, and the natural environment may be an effective strategy for promoting health and wellbeing for AI/AN youth.

Mobile and Digital Technologies

There is growing evidence of the efficacy of interventions that use mobile and digital technologies to support healthy behaviors, including mental health [10], although efficacy varies across text message programs. Researchers in this domain have limited information and agreement on optimal strategies for designing engaging SMS text message content, measuring engagement, if the timing and dose of messages moderates or mediates outcomes, or if there is a specific threshold for engagement to realize intervention effects.

We know from prior research in mobile and digital health that attention to the design and content of text messages [11] is important to increase engagement with messaging, and that health communication theory can be employed to frame text messages in a way that increases their resonance for diverse demographics (eg, various genders, races, ethnicities) [12]. A meta-analysis of health promotion interventions relying on text messages supported this notion, demonstrating that targeted messages (ie, generating messages that resonate for a specific demographic such as men, younger adults, African American individuals, or Latino individuals) combined with tailored messages (ie, making content specific to individuals based on information they provide), produce significantly greater effects compared to more generic messages [13].

Researchers have used various approaches to measure engagement with mobile and digital health interventions, including in-depth interviews with users, ecological momentary assessment (EMA), and reviews of system use data [14]. However, currently, there is a lack of agreement on the best strategy to measure participants' engagement with health-related text messages. EMA, which deploys brief surveys to occasionally poll users about engagement, requires users to complete the surveys to document their engagement, so although useful, this may be challenging to implement. Another strategy is to rely on backend user data to document engagement, a more passive and potentially less cumbersome approach.

The consideration of dose and the appropriate level of engagement with SMS text messaging is important given assumptions that a certain threshold of engagement is necessary to generate intervention effects [15-17], and evidence that participants demonstrate waning interest in text messages over time, sometimes precipitously [18-22].

In this paper, we focus on a review of system data for a study conducted to assess the efficacy of an SMS text messaging intervention to promote mental wellness in AI/AN teens and young adults. Prior papers have focused on the formative design of the intervention, recruitment methods, and the efficacy of the BRAVE intervention [23,24]. Here, we analyze an array of patterns in user engagement using passively collected backend user data. Our objectives are to understand which content was most popular, analyze the timing of messaging and engagement, and test the association between high engagement and efficacy outcomes.

Methods

Research Partners

The BRAVE intervention was designed by the THRIVE and We R Native adolescent health teams at the Northwest Portland Area Indian Health Board (NPAIHB). The NPAIHB is a regional, tribal nonprofit organization that represents 43 federally recognized tribes in Washington, Oregon, and Idaho. The Northwest Tribal Epidemiology Center is housed under NPAIHB and provides support through research, surveillance, and public health capacity building in partnership with the Northwest Tribes. NPAIHB partnered with the mHealth Impact Lab at the Colorado School of Public Health to develop, implement, and evaluate the BRAVE intervention. NPAIHB recruited study participants and delivered SMS text messages, and the mHealth Impact Lab led the design of data collection tools, data collection, and analysis. The partnership was supported by the Technology & Adolescent Mental Wellness (TAM) program, run by the Social Media and Adolescent Health

Research Team and housed within the Department of Pediatrics at the University of Wisconsin-Madison.

Study Overview, Population, and Recruitment

All data collection methods were approved by the Portland Area Indian Health Service Institutional Review Board in Portland, Oregon (1384639). All instruments and data collection methods were reviewed and approved by the institutional review board before data collection took place.

Youth who enrolled in the study received either 8 weeks of BRAVE text messages or 8 weeks of STEM text messages, then crossed over to the other arm and received the next set of messages. The message series in both study arms included 3 to 5 text messages per week, featuring 1 role model video and 1 image per week.

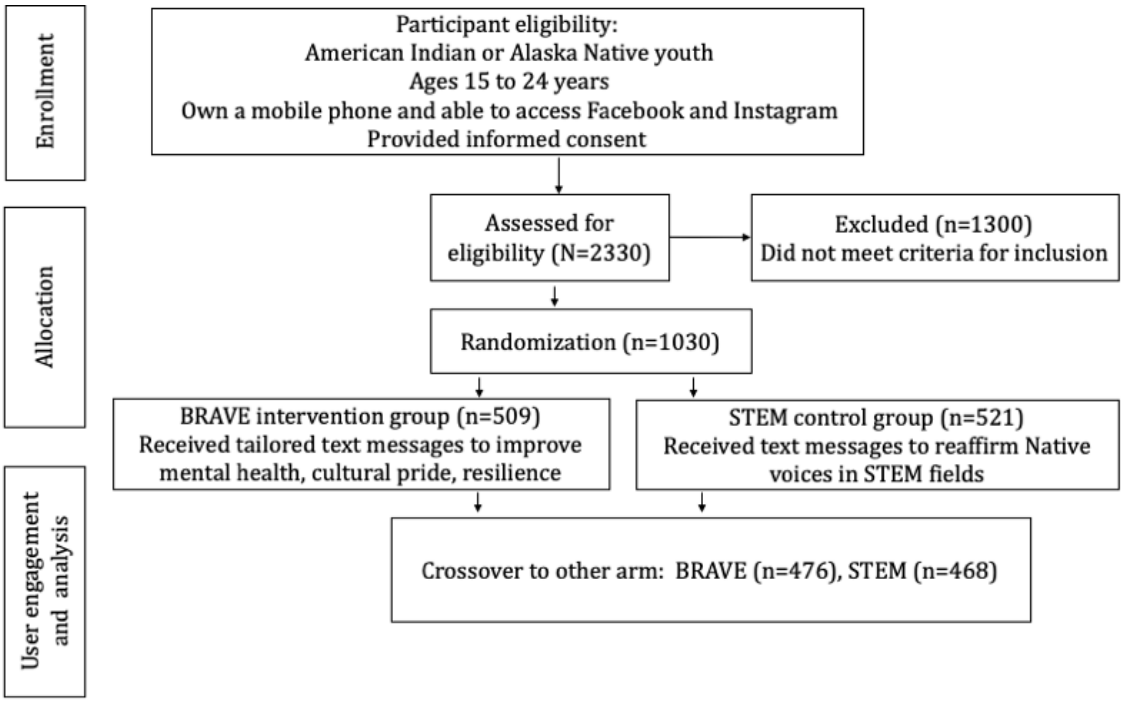
The BRAVE intervention was designed to amplify and reinforce healthy social norms and cultural values, teach suicide warning signs, prepare youth to initiate difficult conversations with peers and trusted adults, encourage youth to access mental health resources (ie, tribal clinics, chat lines), destigmatize mental health services, and connect youth to trusted adults. The message series included links to 7 role model videos (1 to 3 minutes long each) that featured relatable characters experiencing and addressing violent behavior, alcohol misuse, and suicidality (through the eyes of a perpetrator, an intimate partner violence survivor, and a peer bystander), intended to demonstrate important coping strategies and help-seeking skills.

Participant Recruitment and Eligibility

Participant recruitment occurred via We R Native's social media channels (Facebook and Instagram, including text messaging on these platforms) and listservs associated with tribes, tribal health organizations, Indian education, and human service organizations that serve AI/AN teens and young adults. Youth were asked to text the keyword BRAVE to a short code, which triggered a series of eligibility and consent text messages, including a link to a web-based consent form for more information about this study.

The study included self-identified AI/AN youth ages 15 to 24 years. All participants were required to have a cell phone with text message capabilities. To enroll, participants were required to complete the presurvey. Those who met the eligibility requirements and completed the presurvey were randomly assigned to a study arm (n=1030) and invited to complete a baseline survey and 3 follow-up surveys. Participants received a US $10 amazon gift code for each survey they completed, for a compensation of up to US $40 per person in appreciation for their time. Figure 1 summarizes participant recruitment, randomization, and user engagement.

Figure 1. Flowchart summarizing participant recruitment and randomization.



Delivery of BRAVE Messages

For consistency across the 2 study arms, the regular intervention text messages were scheduled to go out in the evenings (between 4 PM and 6 PM EDT). The messages were sent out via Mobile Commons (Upland Software Inc), a mobile messaging provider that supports text, picture, and video SMS [25]. Mobile Commons allows users to track both message delivery and message engagement (ie, link clicks).

To analyze message delivery, we tracked when each message was delivered, its content, and the participant's phone number. Mobile Commons also tracks message engagement by generating a record each time a study participant clicks a link from an SMS text message. For each record we stored the exact time the message was clicked, the message content, and the participant's phone number and IP address.

Data Sets

The Mobile Commons engagement data required substantial data cleaning, which is described below. We produced 2 data sets for analysis. The first data set examined the relationship between the timing of message delivery and timing of message or link clicks for all messages, where link click and message click are used interchangeably and define an event where a subject clicked the link within a message. The second data set merges the engagement data with demographic (gender, age) and efficacy data from the BRAVE study. This enabled us to analyze the number of clicks by gender, message content, and psychosocial health outcomes. All participants were required to complete the baseline survey at the time of enrollment. Enrollees were asked to complete the same survey after the first set of messages. At the end of the intervention period (2 months), participants crossed over to receive the second set of messages and were asked to complete the survey a third time. The study team discontinued communication and asked participants to complete the final survey 90 days later.

Psychosocial Survey Measures

The psychosocial survey measures were influenced by the Healing of the Canoe Survey as part of the Tribal Health: Reaching out InVolves Everyone project [26] and taken from validated survey tools, including the Youth Risk Behavior Surveillance Survey [27], the Youth Coping Responses Inventory [28], the Child and Youth Resilience Measure [29], the Bandura Self-Efficacy Beliefs of Adolescents Scale [30], the Counseling and Help Seeking Questionnaire [31], and the Rosenberg Self-Esteem Scale [32]. Outcomes of interest included health, help-seeking behavior, identification with cultural heritage (cultural identity), self-efficacy, self-esteem, negative coping behavior (alcohol and drug misuse), positive coping behavior, and resilience. Each measure is an aggregate score calculated from multiple survey questions, as defined in [23,24].

Analysis

All statistical analyses were completed using or R, version 4.0 (R Foundation for Statistical Computing) [33]. All statistical tests were considered significant when $P<0.05$. Descriptive statistics were calculated to summarize engagement levels, assess which message content was most popular, and compare engagement across genders. To increase the power of our analyses, we collapsed gender into three categories (male, female, and other).

The analysis of the relationship between user engagement and efficacy measures at baseline including health, resilience, coping skills, self-efficacy, self-esteem, cultural pride and identity, and help-seeking behavior was performed using Poisson regression. A separate Poisson regression model was used for each efficacy measure, with user engagement measured by the number of clicks and composite score of the efficacy measure as the model covariate of interest. We report the exponentiated coefficients and 95% confidence intervals from the Poisson regression models, which are interpreted as rate ratios.

Results

Participant Characteristics and Engagement

Of the 509 participants randomized to the BRAVE arm of the study after enrollment, 270 had 1 trackable click on a study-related message that could be matched to baseline BRAVE survey data using their cell phone number. The gender breakdown of the 270 participants is as follows: 184 were female (68.1%), 50 were male (18.5%), and 36 selected another gender category (13.3%). The average participant was 20.6 years old, with a minimum and maximum age of 15 and 26 years, respectively. Table 1 shows summary statistics for the number of links clicked in the data set, broken down by gender. The data are highly skewed, where most of the 270 users had relatively low engagement as measured by the number of clicks (median 2; mean 3.4), although some users clicked message links as many as 49 times.

Table 1. Maximum, mean, and median number of links clicked for users with at least 1 trackable link on a study-related message.

Gender category	Maximum number of clicks	Mean number of clicks	Median number of clicks
Female (n=184)	36	3.2	2
Male (n=50)	27	3.3	2
Other (n=36)	49	4.4	2
All users (n=270)	49	3.4	2

Users engaged most frequently with the YouTube-based content featuring role model videos, with 64.7% (565/873) of the total clicks derived from the BRAVE video episodes, and earlier episodes receiving the highest number of views. Of the 270 users with at least 1 trackable link, 128 (47.4%) opened the first YouTube video (Episode 1: Alex), 112 (41.5%) opened the second YouTube video (Episode 2: Chris), and 96 (35.6%) clicked a link on the We R Native webpage article about resilience. A breakdown of engagement by type of message is provided in Table 2.

Table 2. Count and frequency of clicks and number of unique users who accessed each type of message.[a]

Message content	Total clicks (n=873), n (%)	Number of unique users (n=270), n (%)	Mean clicks per user (maximum)
We R Native YouTube video: Episode 1: Alex	165 (18.9)	128 (47.4)	1.31 (7)
We R Native YouTube video: Episode 2: Chris	144 (16.5)	112 (41.5)	1.33 (6)
We R Native article: How Does a Person Become Resilient?	128 (14.7)	96 (35.6)	1.36 (5)
We R Native YouTube video: Episode 4: Alex	76 (8.7)	49 (18.1)	1.58 (11)
We R Native YouTube video: Episode 3: Benny	71 (8.1)	59 (21.9)	1.22 (4)
We R Native YouTube video: Episode 5: Chris	66 (7.6)	41 (15.2)	1.67 (9)
Resource article on domestic or dating violence/abuse	62 (7.1)	45 (16.7)	1.43 (8)
We R Native YouTube video: Episode 6: Benny	44 (5)	36 (13.3)	1.3 (4)
We R Native article: Creating Safe Spaces	41 (4.7)	28 (10.4)	1.66 (7)
Resource on healthy relationships and dating	34 (3.9)	27 (10)	1.29 (4)
StrongHearts Native Helpline	24 (2.7)	15 (5.6)	1.69 (6)
Tradition Not Addiction community Facebook page	19 (2.2)	16 (5.9)	1.47 (6)

[a]Numbers are provided for the total (unique users).

User Engagement Times

A time-dependent analysis of the engagement data shows that participants tend to interact with content at specific times of the day and days of the week, and are most likely to click on a message soon after it was sent. Figure 2 shows a histogram of the distribution of clicks by hour of the day (top panel) and by day of the week (bottom panel). Messages were most often clicked in the evening hours. Wednesday is the day of the week when links were most often clicked. An analysis of when users tended to interact with the content compared to when the content is sent is provided Table 3. Table 3 summarizes (in percentiles) the time elapsed (in minutes) from when a message was sent out by the BRAVE team to when a user clicked on it. For example, 50% of messages were clicked on less than 1 minute after they were sent, and 75% of messages were clicked on less than 24 hours after they were sent.

Figure 2. Total clicks by hour and weekday.

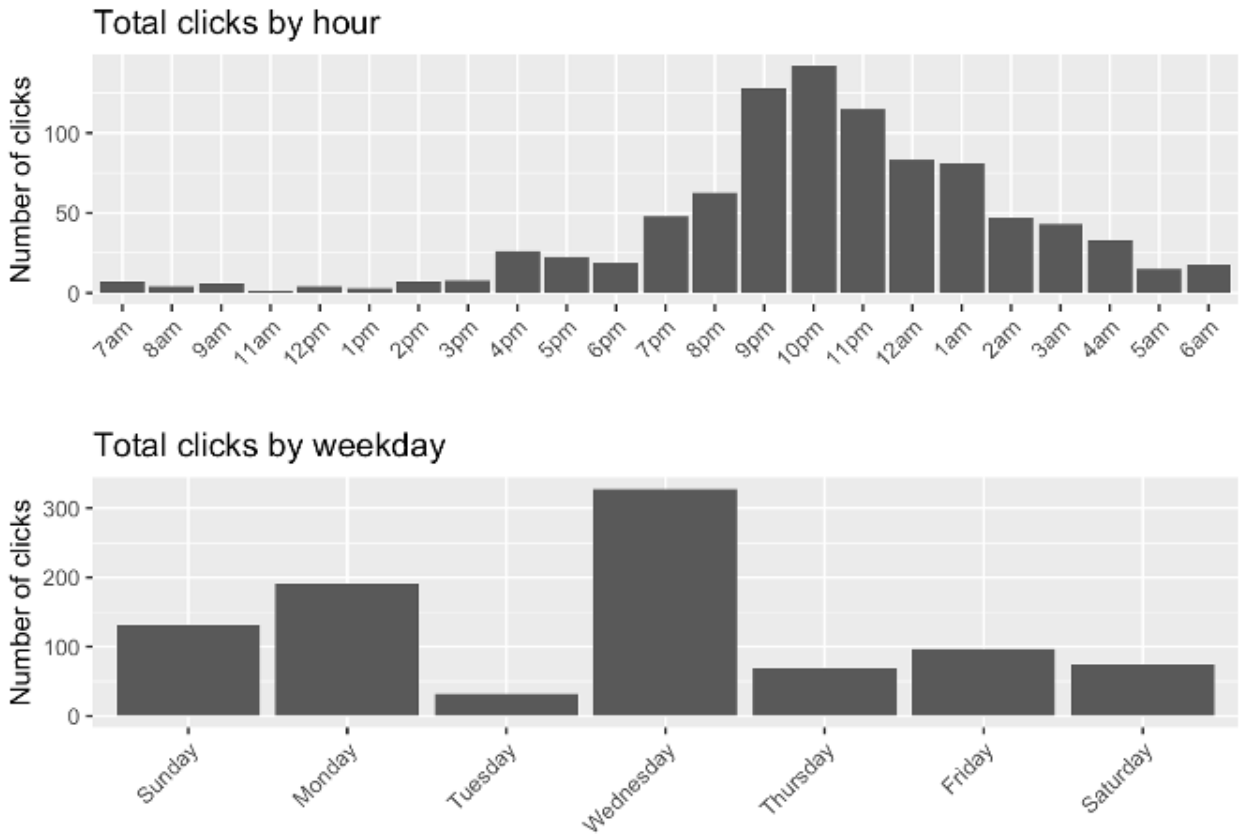
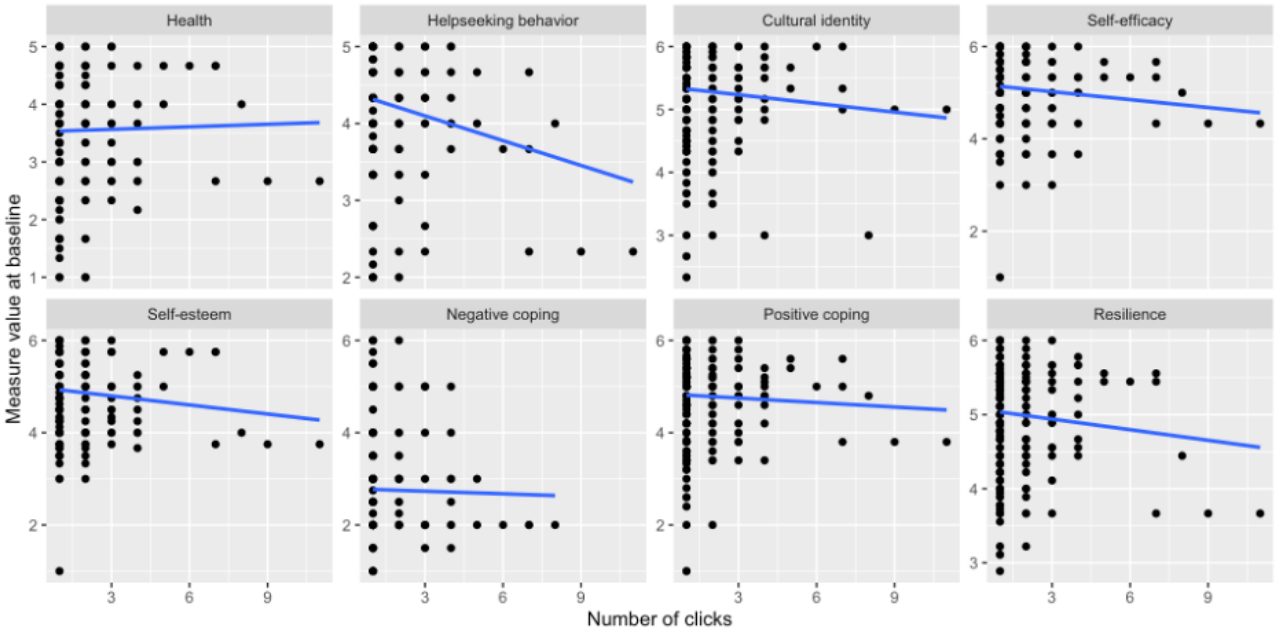


Table 3. Percentile of messages clicked by the time elapsed.[a]

Percentile of messages	Time elapsed
50	<1 minute
75	<24 hours
90	<100 hours
99	<37 days

[a]Time elapsed was calculated based on when the message was sent and when the message was clicked.

Figure 3. Scatter plots of the composite score at baseline for each efficacy measure against the number of clicks.



Finally, we analyzed the correlation between the number of links clicked and composite scores for the psychosocial outcome measures described in the BRAVE efficacy paper [23,24]. The measure scores are on a continuous scale of 1 to 5, which is a composite score across survey questions that were used and reported in BRAVE efficacy paper [23,24]. Plots of composite scores at baseline for each efficacy measure against the number of clicks are shown in Figure 2. The blue line indicates a linear least squares fit through the data. In these plots, although the strength of the association does not appear to be large, the lines show the direction of association for each measure. For the health efficacy measure, a positive association was observed with the number of clicks, and for all other measures the association appears to be flat or negative.

We analyzed each of the efficacy measures separately using a Poisson regression model with number of clicks as the outcome. Exponentiated coefficients from these models can be interpreted as rate ratios and are provided in Table 4, along with 95% confidence intervals and *P* values. Most efficacy measure scores at baseline were not significantly associated with the number of clicks. However, help-seeking behavior was highly significant ($P<.001$), with a rate ratio of 0.82 (0.73, 0.92), indicating that each 1-unit increase in help-seeking score at baseline is associated with an 18% decrease in the expected number of clicks throughout the study. Alternatively, one could say that someone with a help-seeking score of 4 at baseline has, on average, 0.82 times the number of clicks of someone with a help-seeking score of 5 at baseline. Higher help-seeking scores at baseline indicate those who are more likely to seek help for themselves or others; this relationship indicates that having a self-described higher amount of help-seeking behavior at baseline may actually make someone less likely to click frequently. This is a surprising finding.

Table 4. Poisson regression results.[a]

Measure	Rate ratio (95% CI)	*P* value
Health	1.02 (0.93, 1.12)	.74
Resilience	0.89 (0.78, 1.02)	.10
Negative coping	0.99 (0.91, 1.08)	.81
Positive coping	0.95 (0.86, 1.06)	.38
Self-efficacy	0.89 (0.79, 1.01)	.07
Self-esteem	0.90 (0.81, 1.01)	.07
Cultural identity	0.91 (0.8, 1.03)	.14
Help-seeking behavior	0.82 (0.73, 0.92)	<.001

[a]Each row in the table represents results from a Poisson regression model with the number of clicks as the outcome and the indicated BRAVE measure as the covariate of interest.

Discussion

Principal Findings

This paper offers an approach for passively measuring user engagement with a technology-based intervention to support mental health among AI/AN youth. The contributions of the paper are twofold; first, the methods contribute specific strategies for measuring engagement in technology-based health promotion interventions, an important contribution that offers a partial solution to address the limited agreement on how to measure and document engagement with technology-based solutions to promote health. Second, we offer data that help elucidate the breadth and impact of engagement for a nationwide mHealth intervention that successfully recruited participants from a hard-to-reach group, Native youth.

We offer specific strategies to document engagement with an SMS text messaging campaign, including documenting the total number of messages reviewed, time of day, day of the week, and immediacy of engagement (ie, how soon after messages are distributed are they read?). The findings of this paper illustrate important engagement outcomes that can be useful to establish realistic expectations for the level of intensity and population reach for a health-related SMS text messaging campaign. Few SMS text messaging interventions track engagement specifically and then conduct analyses that explore the impact of greater engagement on outcomes. We believe doing so is becoming easier to accomplish and can offer greater programmatic insight into SMS content, intensity, and series length that can become standard measures in future program implementation.

Factors that contribute to differences in user engagement should be explored in future studies. Contextual factors that may explain user engagement include social and environmental priorities within Native populations. Individual and community-level factors that likely influence engagement include phone ownership, network access, cost of devices, network infrastructure, and location [34]. One SMS intervention to support healthy lifestyle interventions among AI families reported different engagement levels based on urban and rural status, where urban participants liked and commented on posts more than rural participants [35] Phone number changes may also account for lower engagement. The Healthy Children, Strong Families 2 study reported that nearly one-third of AI participants changed their phone number during the intervention [35].

Although there was only modest engagement with the clickable SMS messages, it is of particular interest to note the findings that demonstrate a positive association between greater engagement with messages and self-reported help-seeking behavior, suggesting that engagement with 3 or more messages has a positive impact on overall help-seeking behavior for participants. This is among a very limited number of studies we are aware of that is able to elucidate a particular number of engagements that is needed to generate benefit, whether for in-person or virtual interventions [21]. The data showing that greater self-reported help-seeking behavior is associated with less engagement are challenging to interpret. It may be that people with a greater sense of efficacy seeking help at baseline did not consider the BRAVE messages personally relevant as they already had tools they employed to support their mental health and wellbeing.

As the use of SMS text messaging for health promotion becomes more commonplace, having standards for measuring engagement should become more accepted and is expected to help in the evaluation of these interventions. In addition, we believe that more dose-finding studies are needed to ascertain the optimal message delivery for mHealth text messages interventions like BRAVE [15,16,36]. BRAVE was designed to engage users with bidirectional messaging. Previous studies report positive outcomes from 1-way SMS text messaging interventions, including increased clinical appointments, childhood vaccinations, and malaria control [37,38]. SMS interventions are particularly useful for low- to middle-income countries and marginalized groups in the United States because users do not pay for incoming messages. Cost and user considerations should be carefully weighed when designing SMS interventions for these populations.

We assert that our methods have helped to set initial standards for what can be measured, while offering some caution for the limitations of these measures given current technological capabilities for SMS text messaging campaigns. Importantly, our work underscores the feasibility of exploring the impact of

engagement on intended outcomes, allowing for more precise exploration of the dose-response relationship that may be realized through these low-touch initiatives with potential for reaching high numbers of users.

Limitations

There are several limitations of this study. Our sample size was small for some of the engagement measures and some of the data were not easily captured. For example, some phone numbers from clicks were not linked to participants who had completed follow-up surveys, so for these participants, we could not explore a relationship between engagement and demographic or efficacy data. By contrast, for some participants, we had efficacy and demographic data, but no engagement data; thus, we could not determine if the lack of engagement (0 clicks) was truly due to no engagement or because we were not able to recover engagement data for those users. Finally, for participants who interacted with the content by going to the site of the material directly (eg, accessing a video on YouTube) rather than clicking on the link in the text message, we were not able to track content accessed in this way. Overall, the engagement data is likely an underestimation. This suggests having multiple approaches to tracking data that go beyond the passive methods we described here could be an improvement in study design. There are also systems that integrate the backend databases and engagement logs, so that all engagement is tracked through the same system that delivered messages, such as Twilio (Twilio Inc) [39] or Messenger (Meta Platforms Inc) [40], two commercial platforms for scaled text message delivery campaigns.

Conclusions

The BRAVE intervention was designed and promoted by the We R Native team, whose extensive credibility and reach with AI/AN youth was potentially a factor in having users read messages quickly (eg, half of users reading messages within the first minute of distribution) and in having users access content (primarily YouTube videos developed by the team). Despite this credibility, there was still a very modest level of engagement overall, with users only clicking a median of 2 message links from the campaign. This suggests that it may be challenging, even for a well-known and well-respected organization that has strong ties to the communities it serves, to use text messages as a stand-alone intervention strategy.

Although there have been positive outcomes from stand-alone SMS text messaging campaigns, such as those targeting smoking cessation and healthy pregnancy [41], the majority of positive outcomes from SMS text messaging interventions for health promotion have been linked to specific clinical or organizational interventions that facilitate amplification of clinic- or community-based initiatives with supplemental texting to enhance or extend that which occurs in face-to-face settings [13,42-44]. A useful next step would be to consider linking the BRAVE intervention to school-, clinic-, or community-based mental health programs for AI/AN youth as a way to enhance the intervention. Doing so could help build connections to local health services and destigmatize help-seeking, an important goal of the BRAVE SMS series.

Acknowledgments

The project was also supported by the Technology and Adolescent Mental Wellness (TAM) program, run by the Social Media and Adolescent Health Research Team, housed within the Department of Pediatrics at the University of Wisconsin-Madison. TAM is a funding and dissemination mechanism made up of researchers, industry representatives, health care providers, and youth themselves, all of whom are asking and generating novel responses to the following foundational question: How can technology support adolescent mental wellness?

The intervention was developed, in part, with staff funded by grant number H79SM082106 from the Substance Abuse and Mental Health Service Administration (SAMHSA). The views, opinions and content of this publication are those of the authors and contributors, and do not necessarily reflect the views, opinions, or policies of the Centers for Medicare and Medicaid Services, SAMHSA, or the Department of Health and Human Services, and should not be construed as such.

Conflicts of Interest

Authors JW, JS, KS, and SB are employees of the University of Colorado co-led the BRAVE intervention. RP, DS, and SCR are employees of the NPAIHB, which developed the BRAVE intervention.

References

1. Grunbaum J, Kann L, Kinchen S, Williams B, Ross J, Lowry R, et al. Youth risk behavior surveillance--United States, 2001. J Sch Health 2002 Oct;72(8):313-328. [doi: 10.1111/j.1746-1561.2002.tb07917.x] [Medline: 12389372]
2. Anderson RN. Deaths: leading causes for 2000. Natl Vital Stat Rep 2002 Sep 16;50(16):1-85 [FREE Full text] [Medline: 12355905]
3. Kessler RC, Borges G, Walters EE. Prevalence of and risk factors for lifetime suicide attempts in the National Comorbidity Survey. Arch Gen Psychiatry 1999 Jul;56(7):617-626. [doi: 10.1001/archpsyc.56.7.617] [Medline: 10401507]
4. Womble A. Teen life, interrupted. Crisis Text Line. URL: https://www.crisistextline.org/blog/2021/03/23/teen-life-interrupted/ [accessed 2021-06-01]

5. 2018 state of Native youth report: generation Indigenous. Center for Native American Youth. 2018. URL: https://www.cnay.org/wp-content/uploads/2019/09/CNAY_AR2018_FINAL_Web_110618-2.pdf [accessed 2022-01-06]
6. Subica AM, Wu L. Substance use and suicide in Pacific Islander, American Indian, and multiracial youth. Am J Prev Med 2018 Jun;54(6):795-805 [FREE Full text] [doi: 10.1016/j.amepre.2018.02.003] [Medline: 29656915]
7. Wexler L, Chandler M, Gone JP, Cwik M, Kirmayer LJ, LaFromboise T, et al. Advancing suicide prevention research With rural American Indian and Alaska Native populations. Am J Public Health 2015 May;105(5):891-899. [doi: 10.2105/ajph.2014.302517]
8. Bartgis J, Albright G. Online role-play simulations with emotionally responsive avatars for the early detection of Native youth psychological distress, including depression and suicidal ideation. Am Indian Alsk Native Ment Health Res 2016;23(2):1-27. [doi: 10.5820/aian.2302.2016.1] [Medline: 27115130]
9. Center for American Indian Health. Culture forward: a strengths and culture based tool to protect our Native youth from suicide. Center for American Indian Health. URL: https://caih.jhu.edu/programs/cultureforward [accessed 2021-06-01]
10. Kannisto KA, Koivunen MH, Välimäki MA. Use of mobile phone text message reminders in health care services: a narrative literature review. J Med Internet Res 2014 Oct 17;16(10):e222 [FREE Full text] [doi: 10.2196/jmir.3442] [Medline: 25326646]
11. Devine S, Bull S, Dreisbach S, Shlay J. Enhancing a teen pregnancy prevention program with text messaging: engaging minority youth to develop TOP ® Plus Text. J Adolesc Health 2014 Mar;54(3 Suppl):S78-S83 [FREE Full text] [doi: 10.1016/j.jadohealth.2013.12.005] [Medline: 24560081]
12. Psihogios AM, Li Y, Butler E, Hamilton J, Daniel LC, Barakat LP, et al. Text message responsivity in a 2-way short message service pilot intervention With adolescent and young adult survivors of cancer. JMIR Mhealth Uhealth 2019 Apr 18;7(4):e12547 [FREE Full text] [doi: 10.2196/12547] [Medline: 30998225]
13. Head KJ, Noar SM, Iannarino NT, Grant HN. Efficacy of text messaging-based interventions for health promotion: a meta-analysis. Soc Sci Med 2013 Nov;97:41-48. [doi: 10.1016/j.socscimed.2013.08.003] [Medline: 24161087]
14. Short CE, DeSmet A, Woods C, Williams SL, Maher C, Middelweerd A, et al. Measuring engagement in eHealth and mHealth behavior change interventions: viewpoint of methodologies. J Med Internet Res 2018 Nov 16;20(11):e292 [FREE Full text] [doi: 10.2196/jmir.9397] [Medline: 30446482]
15. McVay MA, Bennett GG, Steinberg D, Voils CI. Dose-response research in digital health interventions: concepts, considerations, and challenges. Health Psychol 2019 Dec;38(12):1168-1174 [FREE Full text] [doi: 10.1037/hea0000805] [Medline: 31580127]
16. Evans W, Nielsen PE, Szekely DR, Bihm JW, Murray EA, Snider J, et al. Dose-response effects of the text4baby mobile health program: randomized controlled trial. JMIR Mhealth Uhealth 2015;3(1):e12 [FREE Full text] [doi: 10.2196/mhealth.3909] [Medline: 25630361]
17. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
18. Vandelanotte C, Müller AM, Short CE, Hingle M, Nathan N, Williams SL, et al. Past, present, and future of eHealth and mHealth research to improve physical activity and dietary behaviors. J Nutr Educ Behav 2016 Mar;48(3):219-228.e1. [doi: 10.1016/j.jneb.2015.12.006] [Medline: 26965100]
19. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting From an international workshop. J Med Internet Res 2017 Jun 29;19(6):e232 [FREE Full text] [doi: 10.2196/jmir.7126] [Medline: 28663162]
20. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012 Nov 14;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
21. Devine S, Leeds C, Shlay JC, Leytem A, Beum R, Bull S. Methods to assess youth engagement in a text messaging supplement to an effective teen pregnancy program. J Biomed Inform 2015 Aug;56:379-386 [FREE Full text] [doi: 10.1016/j.jbi.2015.07.003] [Medline: 26173038]
22. Bull S, Schmiege S, Hammes A, Devine S, Gutierrez Y, Shlay J. Measurement and impacts of engagement with a text message program on pregnancy prevention among youth. In: Moreno M, Hoopes A, editors. Technology and Adolescent Health: In Schools and Beyond. Amsterdam, the Netherlands: Elsevier; 2020.
23. Craig Rushing S, Kelley A, Bull S, Stephens D, Wrobel J, Silvasstar J, et al. Efficacy of an mHealth intervention (BRAVE) to promote mental wellness for American Indian and Alaska Native teenagers and young adults: randomized controlled trial. JMIR Ment Health 2021 Sep 15;8(9):e26158 [FREE Full text] [doi: 10.2196/26158] [Medline: 34524092]
24. Stephens D, Peterson R, Singer M, Johnson J, Rushing SC, Kelley A. Recruiting and engaging American Indian and Alaska Native teens and young adults in a SMS help-seeking intervention: lessons learned from the BRAVE study. Int J Environ Res Public Health 2020 Dec 16;17(24):9437 [FREE Full text] [doi: 10.3390/ijerph17249437] [Medline: 33339260]
25. SMS marketing software and API that drives action. Upland. URL: https://uplandsoftware.com/mobile-messaging/
26. Donovan D, Thomas L, Sigo R, Price L, Lonczak H, Lawrence N, et al. Healing of the canoe: preliminary results of a culturally tailored intervention to prevent substance abuse and promote tribal identity for Native youth in two Pacific

Northwest tribes. Am Indian Alsk Native Ment Health Res 2015;22(1):42-76 [FREE Full text] [doi: 10.5820/aian.2201.2015.42] [Medline: 25768390]
27. Grunbaum J, Kann L, Kinchen S, Williams B, Ross J, Lowry R, et al. Youth risk behavior surveillance--United States, 2001. MMWR Surveill Summ 2002 Jun 28;51(4):1-62 [FREE Full text] [Medline: 12102329]
28. Hernandez B, Vigna J, Kelley ML. The Youth Coping Responses Inventory: development and initial validation. J Clin Psychol 2010 Sep;66(9):1008-1025. [doi: 10.1002/jclp.20697] [Medline: 20694963]
29. Ungar M, Liebenberg L. Assessing resilience across cultures using mixed methods: construction of the Child and Youth Resilience Measure. J Mix Methods Res 2011 Apr;5(2):126-149. [doi: 10.1177/1558689811400607]
30. Bandura A. Guide for constructing self-efficacy scales. In: Pajares F, Urban T, editors. Self-Efficacy Beliefs of Adolescents. Charlotte, NC: Information Age Publishing; 2006:307-337.
31. Dauphinais P, LaFromboise T, Rowe W. Perceived problems and sources of help for American Indian Students. CES 1980;20(1):37-44. [doi: 10.1002/j.1556-6978.1980.tb02030.x]
32. Rosenberg M. Society and the Adolescent Self-Image. Princeton, NJ: Princeton University Press; 1965:1-326.
33. R-Project. URL: https://www.R-project.org/
34. Hobson GR, Caffery LJ, Neuhaus M, Langbecker DH. Mobile health for First Nations populations: systematic review. JMIR Mhealth Uhealth 2019 Oct 07;7(10):e14877 [FREE Full text] [doi: 10.2196/14877] [Medline: 31593537]
35. Tomayko EJ, Webber EJ, Cronin KA, Prince RJ, Adams AK. Use of text messaging and Facebook groups to support the Healthy Children, Strong Families 2 healthy lifestyle intervention for American Indian families. Curr Dev Nutr 2021 Jun;5(Suppl 4):32-39 [FREE Full text] [doi: 10.1093/cdn/nzaa110] [Medline: 34222765]
36. Donkin L, Hickie IB, Christensen H, Naismith SL, Neal B, Cockayne NL, et al. Rethinking the dose-response relationship between usage and outcome in an online intervention for depression: randomized controlled trial. J Med Internet Res 2013 Oct 17;15(10):e231 [FREE Full text] [doi: 10.2196/jmir.2771] [Medline: 24135213]
37. Mohammed A, Acheampong PR, Otupiri E, Osei FA, Larson-Reindorf R, Owusu-Dabo E. Mobile phone short message service (SMS) as a malaria control tool: a quasi-experimental study. BMC Public Health 2019 Aug 29;19(1):1193 [FREE Full text] [doi: 10.1186/s12889-019-7336-6] [Medline: 31464623]
38. Linde D, Korsholm M, Katanga J, Rasch V, Lundh A, Andersen MS. One-way SMS and healthcare outcomes in Africa: systematic review of randomised trials with meta-analysis. PLoS One 2019;14(6):e0217485 [FREE Full text] [doi: 10.1371/journal.pone.0217485] [Medline: 31170176]
39. Twilio. Twilio. 2021. URL: https://www.twilio.com/
40. Messenger. URL: https://www.messenger.com/
41. Whittaker R, Matoff-Stepp S, Meehan J, Kendrick J, Jordan E, Stange P, et al. Text4baby: development and implementation of a national text messaging health information service. Am J Public Health 2012 Dec;102(12):2207-2213. [doi: 10.2105/AJPH.2012.300736] [Medline: 23078509]
42. Gurol-Urganci I, de Jongh T, Vodopivec-Jamsek V, Atun R, Car J. Mobile phone messaging reminders for attendance at healthcare appointments. Cochrane Database Syst Rev 2013 Dec 05(12):CD007458 [FREE Full text] [doi: 10.1002/14651858.CD007458.pub3] [Medline: 24310741]
43. de Jongh T, Gurol-Urganci I, Vodopivec-Jamsek V, Car J, Atun R. Mobile phone messaging for facilitating self-management of long-term illnesses. Cochrane Database Syst Rev 2012 Dec 12;12:CD007459 [FREE Full text] [doi: 10.1002/14651858.CD007459.pub2] [Medline: 23235644]
44. Vodopivec-Jamsek V, de Jongh T, Gurol-Urganci I, Atun R, Car J. Mobile phone messaging for preventive health care. Cochrane Database Syst Rev 2012 Dec 12;12:CD007457 [FREE Full text] [doi: 10.1002/14651858.CD007457.pub2] [Medline: 23235643]

Abbreviations

AI/AN: American Indian and Alaska Native
NPAIHB: Northwest Portland Area Indian Health Board
SAMHSA: Substance Abuse and Mental Health Service Administration
TAM: Technology & Adolescent Mental Wellness

Edited by G Eysenbach; submitted 15.07.21; peer-reviewed by K Halttu, C Markham; comments to author 17.08.21; revised version received 01.10.21; accepted 22.11.21; published 25.02.22.

Please cite as:
Wrobel J, Silvasstar J, Peterson R, Sumbundu K, Kelley A, Stephens D, Craig Rushing S, Bull S
Text Messaging Intervention for Mental Wellness in American Indian and Alaska Native Teens and Young Adults (BRAVE Study): Analysis of User Engagement Patterns
JMIR Form Res 2022;6(2):e32138
URL: https://formative.jmir.org/2022/2/e32138
doi:10.2196/32138
PMID:35212633

Original Paper

The Beneficial Effect of the First COVID-19 Lockdown on Undergraduate Students of Education: Prospective Cohort Study

Gili Joseph[1*], PhD; Hadas Schori[1*], PhD

Department of Physical Activity and Movement, Faculty of Science, Seminar Hakibbutzim College of Education, Tel Aviv, Israel

[*]all authors contributed equally

Corresponding Author:
Gili Joseph, PhD
Department of Physical Activity and Movement
Faculty of Science
Seminar Hakibbutzim College of Education
149 Namir Drive
Tel Aviv, 6250769
Israel
Phone: 972 544577095
Email: gilijosephphd@gmail.com

Abstract

Background: The COVID-19 pandemic has been spreading consistently since the beginning of 2020. On February 27, 2020, the first patient with coronavirus was diagnosed in Israel. On March 14, 2020, the Israeli government declared a general lockdown that lasted about a month, which altered the lives of the entire population.

Objective: The objective of this paper is to evaluate the change in the well-being, physical activity, and sleep quality of undergraduate students of education at 2 time points: before (November 2019) and during (April 2020) the first COVID-19 lockdown.

Methods: In total, 533 undergraduate students of education submitted an online questionnaire before the lockdown and at its end. The questionnaire comprised 4 parts: a (1) sociodemographic and (2) weekly exercise questionnaire taken from the International Physical Activity Questionnaire–Short Form; (3) sleep quality, rated using the Mini Sleep Questionnaire; and (4) well-being, rated using the short version of the Mental Health Inventory. This was a pre-post prospective cohort questionnaire study.

Results: It was predicted that there would be a decrease in the aforementioned parameters. Contrary to all expectations, an increase was observed in all 3. Results showed that during the lockdown, there was an increase in the level of exercise students engaged in. Overall, 102 (61.4%) of 166 students engaged in a greater amount of physical activity during the COVID-19 lockdown compared to 150 (40.9%) of 367 students who engaged in a greater amount of physical activity before COVID-19. Levels of sleep quality (mean 5.34 [SD 0.92] vs mean 5.12 [SD 0.46], *P*=.02) and well-being (mean 3.79 [SD 0.62] vs mean 3.67 [SD 0.59], *P*=.02) were also higher during the COVID-19 lockdown.

Conclusions: These findings indicate that undergraduate students seem to have taken advantage of the change in lifestyle due to the lockdown, directing the free time toward improving health by engaging in more physical activity, thus improving sleep quality and well-being.

(JMIR Form Res 2022;6(2):e27286) doi:10.2196/27286

KEYWORDS

sleep quality; exercise; well-being; undergraduate students; COVID-19 lockdown; COVID-19

Introduction

The year 2020 challenged the entire world. Coronavirus (COVID-19) was declared a global epidemic by the World Health Organization [1]. This led many countries to take various precautions to prevent the spread of the virus. The severity of the general closure varied between countries, each implementing different measures ranging from increased enforcement of limitations (limited unessential social interaction, limited mobility, limited work-related activities) up to a general lockdown [2]. On March 14, 2020, the Israeli government declared a general lockdown, which included restricting mobility, reducing working capacity by 10%-20%, and closing

the doors of the entire educational system for all ages [3], lasting approximately 2 months.

The lockdown, as well as many other limitations, had a serious impact on our lifestyle. Studies have shown an increase in depression and self-reported stress and a reduction in well-being during the first COVID-19 lockdown [4-7]. In addition, a reduction in the amount of physical activity performed was found among different populations worldwide; this has had a negative effect on the well-being of the population [8-10]. Furthermore, a decrease in sleep quality was found [11,12].

Sleep is 1 of the essential components of health [13]. Poor sleep quality may adversely affect the immune system, learning abilities, blood pressure, psychological status, and more [14-16]. In general, any changes in our lifestyle may affect our sleep quality [17,18]. In addition, there was a major, sudden transformation in our lifestyle, as imposed upon the population during the first COVID-19 lockdown. The lockdown not only affected sleep quality but also caused sleeplessness [19-22].

Exercise is 1 of the most recommended ways to improve an individual's health [23]. Physical activity can improve sleep quality [24], and regular exercise can be 1 of the methods of treating people with sleep disorders [25,26]. Moreover, studies have shown small-to-moderate beneficial effects of regular exercise on total sleep time, sleep efficiency, sleep onset latency, and sleep quality [25,27]. Acute and long-term exercise has been documented in a number of studies as 1 of the factors that can increase slow-wave sleep and total sleep time, as well as decreasing the period between the initial attempt to fall asleep and the onset of sleep [28-30]. In addition, aerobic physical activity was shown to improve sleep quality for people with sleep disorders [31].

Physical activity is also considered 1 of the factors that can help people improve their well-being; it is often recommended as a way to deal with stress [32]. People with a variety of mental disorders have shown improvement in their mental condition when implementing physical activity in their daily routine [33-35].

Regarding undergraduate students, there are some findings that present a decrease in physical activity as well as well-being during the first COVID-19 lockdown, due to the transition to remote teaching and a major change in their lifestyle [36-38]. However, research conducted on students studying health and science revealed that during the COVID-19 lockdown, they spent more time engaging in physical activity but also increased the amount of time they were sedentary [39].

This study aims to examine the effect of the first COVID-19 lockdown on undergraduate Israeli students of education. These students are characterized by a nonsedentary lifestyle due to their mandatory practical teaching. During the first COVID-19 lockdown, all practical lessons were transferred to remote teaching and therefore students were forced to increase their sitting time. Here, we aim to analyze their physical activity, well-being, and sleep quality before and during the COVID-19 lockdown.

Methods

Subjects

Subjects comprised over 600 students from a college of education in the center of Israel. The questionnaire was distributed to first-, second-, and third-year students in the science and education faculties. The students who replied were included in the study (N=367), while the students who did not reply were excluded (N=233). All 367 (100%) students replied to the pre-COVID-19 questionnaire, and 166 (45.2%) students replied to the questionnaire during the COVID-19 lockdown (postquestionnaire). Altogether, 533 questionnaires were submitted. Ordinarily, during the 4 years of the students' studies, they attend the college on campus at least 3 days a week, and they are obligated to teach at least once a week in a proscribed school during all 4 of their college study years. In their fourth year, they attend the college only once a week, and during the rest of the week, they teach in schools, under the supervision and guidance of experienced and expert teachers from those schools. This study received ethical approval from the ethics committee of the Seminar Hakibutzim College, Tel Aviv, Israel.

Research Tools

The questionnaire was submitted by the students using Google Forms. The questionnaire included 4 sections: sociodemographic questions, questions pertaining to the level and intensity of their weekly exercise, a well-being questionnaire, and a questionnaire regarding their quality of sleep. The Hebrew version of the exercise questionnaire was taken from the International Physical Activity Questionnaire–Short Form (IPAQ-SF) [40]. The students were asked to note how many times and the number of hours they exercise per week and the number of times and hours a week they engage in physical activity such that it causes sweat production and strenuous breathing. The questions were based on a Likert scale of 1-7, where 1 represents never and 7 represents always. The 4 exercise questions were grouped into 1 variable, and then a dichotomous variable was rebuilt and the population was divided into 2 groups by the median of 3.5. Students with a score of 0-3.49 in exercise were categorized as engaging in a lesser amount of exercise. Students with a score of 3.5-7 in exercise were categorized as engaging in a greater amount of exercise. The well-being questionnaire used was the short version of the Mental Health Inventory (MHI) developed by Veit and Ware [41] and validated by Florian and Drori [42] and included 10 questions: Cronbach α=.96. Sleep quality was measured using the Mini Sleep Questionnaire (MSQ), which included 10 questions that relate to quality of sleep using a Likert scale of 1-7 [43].

Research Process

During the second and third weeks of the first semester (November 2019), the researchers entered the classrooms and explained the objectives of the study. Immediately following that, the questionnaire was distributed, and the students submitted it using Google Forms. The second questionnaire was distributed via email because at that time (at the end of the first COVID-19 lockdown, April 2020), the students studied only online and did not study on campus. Since the students answered the questionnaire of their own volition, and the

questionnaire was anonymous, not all of them chose to participate; thus, only 166 (45.2%) of 367 students submitted the second questionnaire. In the second questionnaire, the students were asked to answer the questions with reference to the COVID-19 lockdown period.

Statistical Analysis

An independent *t* test analysis was performed to measure the difference between the amount of exercise, sleep quality, and well-being among undergraduate students of education before and during the COVID-19 lockdown, and the difference between well-being and the quality of sleep by the amount of exercise engaged in by the students. The analysis of the answers to the questionnaires before and during the COVID-19 lockdown were not matched; thus, an independent *t* test analysis was performed. A chi-square analysis was performed to measure the difference in the amount of exercise before and during the COVID-19 lockdown. A 2-way ANOVA was conducted in order to find the main effects and interactions between the period when the questionnaire was submitted (before or during the COVID-19 lockdown), the year of study/field of study/gender, and the 3 dependent variables: amount of exercise, sleep quality, and well-being. Results were statistically significant at the .05 significance level.

Results

Participant Characteristics

The students who participated in this study are undergraduate students of education from the faculties of science and education. The average age of the students was 25.3 (SD 4.5) years. Most of the students were female (291/367, 79.3%) and single (304/367, 82.8%). In addition, 163 (44.4%) of the students were in their first year of studies and 204 (55.6%) in their second and third years of studies. Furthermore, 130 (35.4%) of the students studied in the Department of Physical Education, while the rest studied in other education disciplines, such as science, elementary school, early childhood education, and special education. All 367 (100%) students submitted the first questionnaire, while 166 (45.2%) students submitted the second one (Table 1). There was no statistical difference between the students answering the prequestionnaire and those who answered the postquestionnaire concerning their characteristics.

The 3 main research questions asked were whether there was any difference between the 2 periods (1) in the amount of exercise engaged in, (2) in their well-being, and (3) in their quality of sleep. An independent *t* test analysis was conducted, and the results showed that the students engaged in more exercise during the lockdown compared to the beginning of the same school year (mean 4.27 [SD 1.4] vs mean 3.4 [SD 1.7], $P<.001$). They also had better quality of sleep (mean 5.34 [SD 0.92] vs mean 5.12 [SD 0.46], $P=.02$) and better well-being (mean 3.79 [SD 0.62] vs mean 3.67 [SD 0.59], $P=.02$); see Figure 1.

Figure 1. The difference between exercise, sleep quality and well-being before and during the COVID-19 lockdown. Independent *t* test analysis was engaged in to measure the difference between the amount of exercise, sleep quality and well-being among undergraduate students of Education. N=367 - pre-COVID-19, N=166 - during COVID-19 lockdown, (***$P<.001$, * $P<.05$).

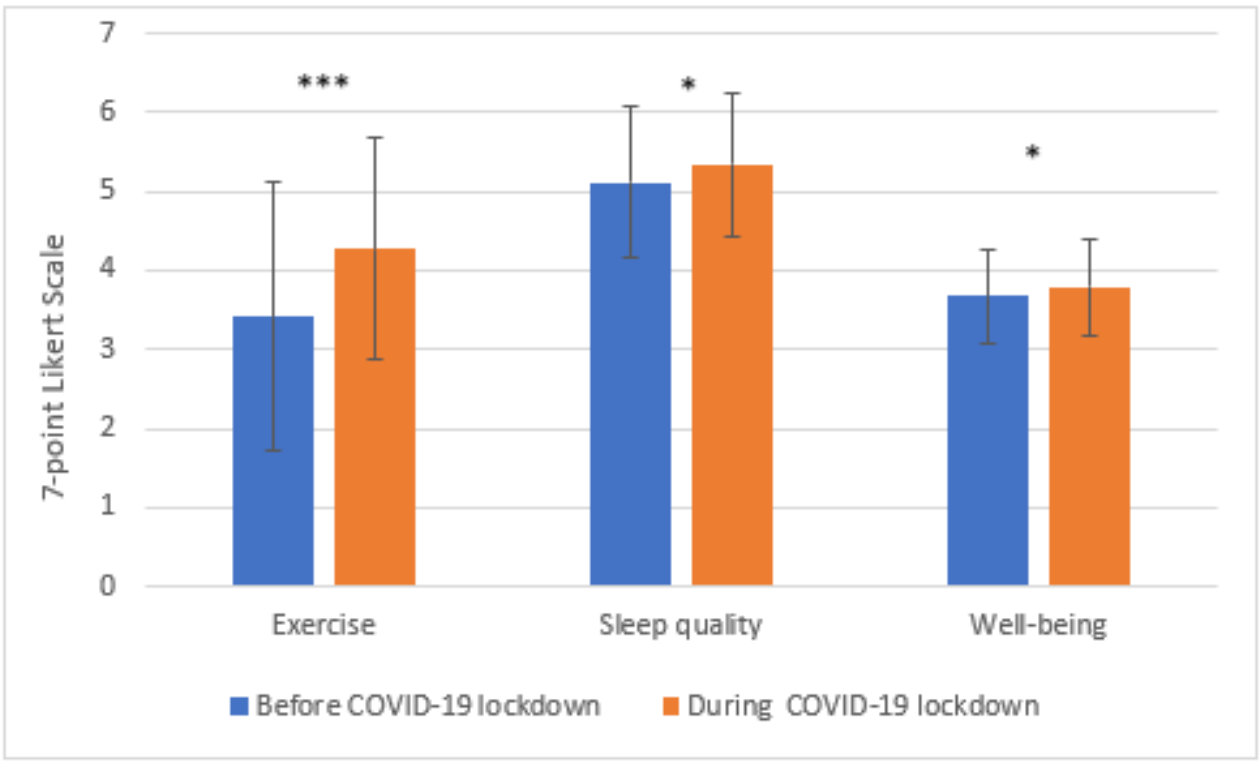


Table 1. Descriptive characteristics of the participants.

Characteristics	Prequestionnaire (N=367)	Postquestionnaire (N=166)
Age (years), mean (SD)	25.3 (4.5)	25.6 (1.5)
Gender, n (%)		
Female	291 (79.3)	122 (73.5)
Male	76 (20.7)	44 (26.5)
Marital status, n (%)		
Single	304 (82.8)	130 (78.3)
Other status	63 (17.2)	36 (21.7)
Field of study, n (%)		
Physical education	130 (35.4)	101 (60.8)
Other study fields	237 (64.6)	65 (39.2)
Study year, n (%)		
First year	163 (44.4)	64 (38.6)
Second and third years	204 (55.6)	102 (61.4)

When analyzing the amount of exercise the students engaged in before and during the COVID-19 lockdown, 150 (40.9%) of the 367 students engaged in a greater amount of exercise before the COVID-19 lockdown. However, during the COVID-19 lockdown, the number of students engaging in a greater amount of exercise increased to 102 (61.4%) of 166 students (Table 2). The chi-square analysis showed a significant difference (χ^2_1=19.41, *P*<.001). The ϕ correlation showed low (0.2) but significant (*P*<.001) strength. An independent *t* test analysis was performed to compare the quality of sleep among the students who engaged in a greater amount of exercise compared to those who engaged in a lesser amount of exercise during the COVID-19 lockdown. Those who engaged in a greater amount of exercise had better sleep quality (mean 5.57 [SD 0.84] vs mean 4.96 [SD 0.93], *P*<.001). No difference was found in their well-being (Table 3).

A 2-way ANOVA was conducted to compare the main effects of the interaction of the time when the questionnaire was submitted (before or during the COVID-19 lockdown) and the year of study/field of study/gender on the amount of exercise, sleep quality, and well-being (Tables 4-6).

As shown in Table 4, the effect of the 2 independent variables (time when the questionnaire was submitted and the year of study) on the amount of exercise, sleep quality, and well-being was studied. Results showed effects that were statistically significant at the 0.05 significance level only for the time when the questionnaire was submitted. Regarding the amount of exercise, the main effect of the time when the questionnaire was submitted yielded an $F_{(1, 527)}$ ratio of 33.0 (*P*<.001), indicating a significant difference between the amount of exercise the students engaged in before the COVID-19 lockdown (mean 3.4 [SD 1.7]) and during the COVID-19 lockdown (mean 4.28 [SD 1.46]). No effect was found of the year of study, and no interaction was found between the time when the questionnaire was submitted and the year of study. When examining the quality of sleep, the main effect of the time when the questionnaire was submitted yielded an $F_{(1, 527)}$ ratio of 6.17 (*P*=.01), indicating a significant difference between the students' quality of sleep before the COVID-19 lockdown (mean 5.12 [SD 0.96]) and during the COVID-19 lockdown (mean 5.35 [SD 0.91]). No effect was found of the year of study, and no interaction was found between the time when the questionnaire was submitted and the year of study. The same trend was found regarding well-being. The main effect of the time when the questionnaire was submitted yielded an $F_{(1, 527)}$ ratio of 5.79 (*P*=.02), indicating a significant difference between the students' well-being before the COVID-19 lockdown (mean 3.67 [SD 0.59]) and during the COVID-19 lockdown (mean 3.8 [SD 0.61]). No effect was found of the year of study, and no interaction was found between the time when the questionnaire was submitted and the year of study.

Table 2. The number and percentage of students engaging in a lesser and greater amount of exercise before and during the COVID-19 lockdown.

Period	Students engaging in a lesser amount of exercise, n (%)	Students engaging in a greater amount of exercise, n (%)	*P* value
Before COVID-19 lockdown (N=367)	217 (59.1)	150 (40.9)	<.001
During COVID-19 lockdown (N=166)	64 (38.6)	102 (61.4)	<.001

Table 3. The difference in well-being and sleep quality in reference to the amount of exercise engaged in by the students during the COVID-19 lockdown (N=166).

Amount of exercise	n (%)	Mean (SD)	*P* value
Well-being			
Lesser amount of exercise	64 (38.5)	3.7 (0.73)	.07
Greater amount of exercise	102 (61.5)	3.9 (0.53)	.07
Sleep quality			
Lesser amount of exercise	64 (38.5)	4.96 (0.93)	<.001
Greater amount of exercise	102 (61.5)	5.6 (0.84)	<.001

Table 4. Results of 2-way ANOVA to compare the main effects of the independent variables "time when the questionnaire was submitted" and "year of study" and their interaction on the amount of exercise, sleep quality, and well-being.

Variables	F_1 ratio	*P* value
Dependent variable: amount of exercise		
Time when the questionnaire was submitted	33.0	<.001
Year of study	0.29	.59
Interaction	0.79	.37
Dependent variable: sleep quality		
Time when the questionnaire was submitted	6.17	.01
Year of study	1.54	.22
Interaction	0.028	.87
Dependent variable: well-being		
Time when the questionnaire was submitted	5.79	.02
Year of study	0.43	.51
Interaction	0.02	.89

Table 5. Results of 2-way ANOVA to compare the main effects of the independent variables "time when the questionnaire was submitted" and "field of study"[a] and their interaction on the amount of exercise, sleep quality, and well-being.

Variables	F_1 ratio	*P* value
Dependent variable: amount of exercise		
Time when the questionnaire was submitted	9.798	.002
Field of study	166.29	<.001
Interaction	8.34	.004
Dependent variable: sleep quality		
Time when the questionnaire was submitted	2.403	.12
Field of study	11.363	<.001
Interaction	0.435	=.51
Dependent variable: well-being		
Time when the questionnaire was submitted	4.81	.03
Field of study	0.010	.92
Interaction	0.003	.96

[a]Field of study is a comparison between physical education students compared to students form all the other fields together.

Table 6. Results of 2-way ANOVA to compare the main effects of the independent variables "time when the questionnaire was submitted" and "gender" and their interaction on the amount of exercise, sleep quality, and well-being.

Variables	F_1 ratio	*P* value
Dependent variable: amount of exercise		
Time when the questionnaire was submitted	8.3	.004
Gender	46.3	<.001
Interaction	7.38	.01
Dependent variable: sleep quality		
Time when the questionnaire was submitted	1.18	.04
Gender	4.09	.27
Interaction	2.018	.16
Dependent variable: well-being		
Time when the questionnaire was submitted	5.29	.02
Gender	2.94	.09
Interaction	0.155	.69

Table 5 presents the results of a 2-way ANOVA when the 2 independent variables this time were the time when the questionnaire was submitted and the field of study. The 2-way ANOVA showed effects on the amount of exercise that were statistically significant (*P*<.05). The main effect of the time when the questionnaire was submitted yielded an $F_{(1, 529)}$ ratio of 9.798 (*P*=.002), indicating a significant difference between the amount of exercise the students engaged in before the COVID-19 lockdown (mean 3.4 [SD 1.7]) and during the COVID-19 lockdown (mean 4.28 [SD 1.46]). The main effect of the field of study yielded an $F_{(1, 529)}$ ratio of 166.29 (*P*<.001), indicating a significant difference between the amount of exercise the physical education students engaged in (mean 4.77 [SD 1.3]) and the students from all the other fields of study (mean 2.85 [SD 1.46]). The interaction effect between the time when the questionnaire was submitted and the field of study was significant ($F_{(1, 529)}$=8.34, *P*=.004), showing an ordinal interaction between the variables, meaning a greater amount of exercise was performed during the COVID-19 lockdown and among the students of physical education (Figure 2). Regarding sleep quality, the main effect of the field of study showed effects that were statistically significant at the 0.05 significance level and yielded an $F_{(1, 529)}$ ratio of =11.363 (*P*=.001), indicating a significant difference between the sleep quality of students of physical education (mean 5.37 [SD 0.87]) compared to students from all the other disciplines (mean 5.05 [SD 0.99]). Regarding sleep quality, no interaction was found between the time when the questionnaire was submitted and the field of study ($F_{(1,529)}$=0.435, *P*=.51).

A 2-way ANOVA was conducted for the influence of another pair of independent variables: time when the questionnaire was submitted and gender (Table 6). Regarding the amount of exercise, the main effect of the time when the questionnaire was submitted yielded an $F_{(1, 529)}$ ratio of 8.3 (*P*=.004), indicating a significant difference between the amount of exercise the students engaged in before the COVID-19 lockdown (mean 3.4 [SD 1.7]) and during the COVID-19 lockdown (mean 4.28 [SD 1.46]). The main effect of gender yielded an $F_{(1, 529)}$ ratio of 46.3 (*P*<.001), indicating a significant difference between the amount of exercise the male students engaged in (mean 4.77 [SD 1.49]) compared to female students (mean 3.4 [SD 1.49]). The interaction effect between the time when the questionnaire was submitted and gender was significant ($F_{(1, 529)}$=7.38, *P*=.01), showing an ordinal interaction between the variables, meaning a greater amount of exercise was performed during the COVID-19 lockdown among male students, although it is interesting to observe that female students showed a more significant change in the amount of exercise during the COVID-19 lockdown (Figure 3 and Table 6).

Figure 2. Interaction effect between the time of submitting the questionnaire (before or during the COVID-19 lockdown) and the field of study on the amount of exercise (2-way ANOVA). *P*<.05.

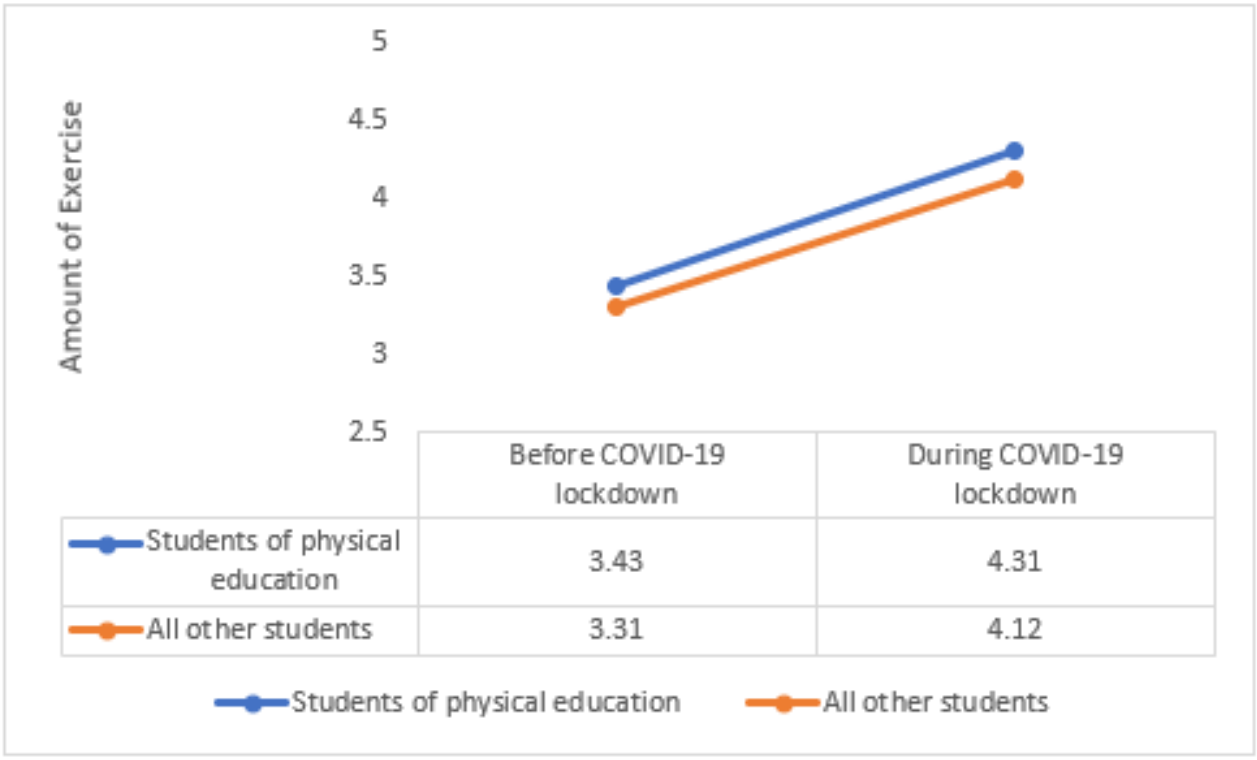
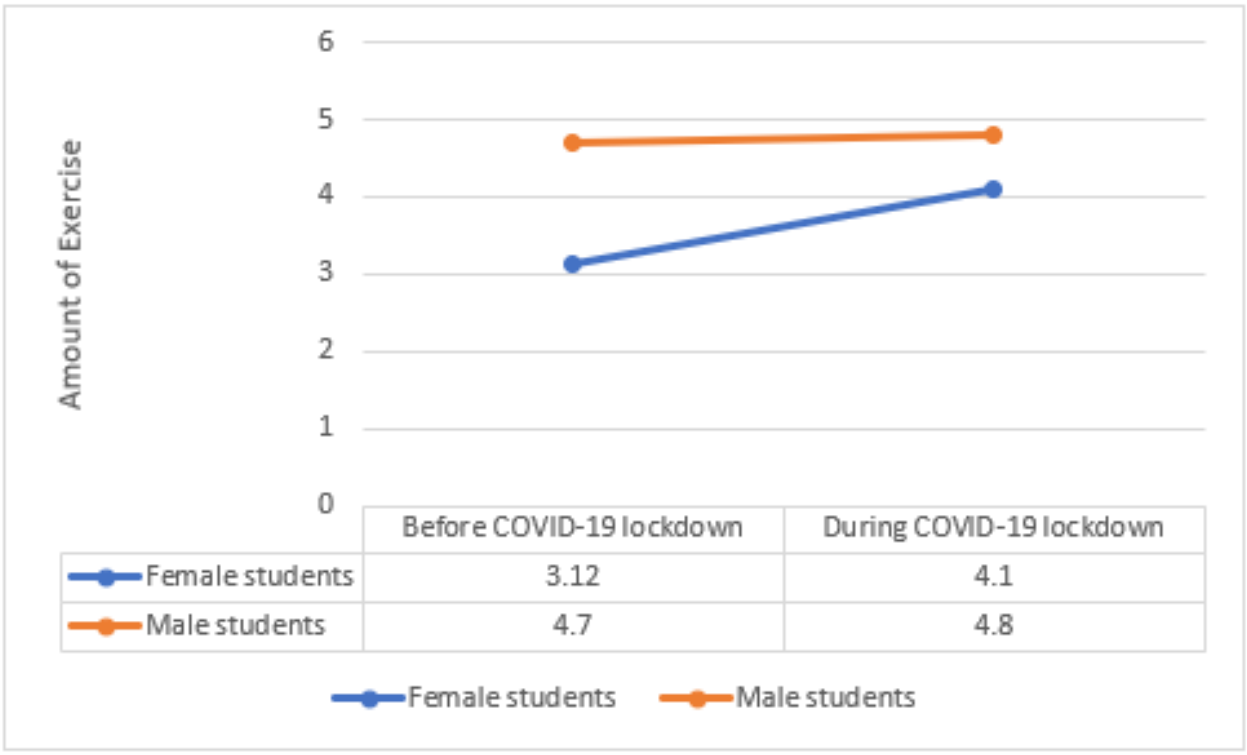


Figure 3. Interaction effect between the time of submitting the questionnaire (before or during the COVID-19 lockdown) and gender on the amount of exercise (2-way ANOVA). Before COVID-19 lockdown: female students, n=367; male students, n=166. During COVID-19 lockdown: female students, n=424; male students, n=109. *P*<.05.



Discussion

Principal Findings

Undergraduate students of education showed improvement in their well-being, sleep quality, and exercise during the first COVID-19 lockdown. These results were contrary to all expectations, since the restriction enforced due to the COVID-19 pandemic was shown to have a serious impact on the population, resulting in an enormous change in lifestyle and in daily routine [44,45]. Moreover, some studies have shown that during the first COVID-19 lockdown or during quarantine, there was a decrease in well-being, sleep quality, and physical activity among most of the general population [5,7,22,46-51]. Our results can be explained by other studies that have shown that the more time people have, the more they invest in their hobbies and in exercising [52]. A study conducted on European Union citizens revealed that 73.1% exercised in their free time [53], and a study in Spain showed that more than 50% enjoyed new positive experiences during the lockdown; specifically, students of health and science reported an increase in both the number of days in which they engaged in physical activity and in the total number of minutes of physical activity they engaged in per week [39].

However, there are studies showing that during the COVID-19 lockdown, people experienced more time alone, increased the use of electronic devices, and stayed indoors more, which all led to a reduction in sleep quality and a decrease in their well-being [4,11,19-22,50,54,55]. This study does not contradict these findings but complements them with additional aspects of life that have not been tested yet, such as the possible benefit of physical activity on well-being and sleep quality in a time when lifestyle changes may cause sudden distress. The students who participated in this study are studying toward earning a teaching certificate in an Israeli education college in Tel Aviv. Typically, they commute between 1 and 3 hours every day to get to college (the college has no dorms) or to the schools where they teach. After school hours, many of them need to go out to work to support themselves. The first COVID-19 lockdown forced the students to stay at home; discontinue working, teaching in the schools, and studying on campus; and switch to distance learning or distance teaching. Although this situation increased many sitting hours, it also led to more free hours because of less travel time and less or no work (88% reported working less than normal times; see Multimedia Appendix 1). The statistical correlation found between exercising more and sleeping better might be explained by the free time the students gained because of less or no work and no commuting, while not having additional personal or family responsibilities. Doing more exercise and sleeping better may have improved their well-being. These results are in line with previous works showing that free-time activities play an important role in subjective well-being [56] and that high levels of free-time exercise are associated with benefits for psychological well-being [57].

Studies have demonstrated that physical activity can improve well-being regardless of gender, socioeconomic status, and health status [8,58-61]. In addition, physical activity (especially high levels) can improve sleep quality [24,25,62,63]. Furthermore, studies have shown that the amount and extent of physical activity are positively associated with the researched population's well-being. During the COVID-19 lockdown period, the student population was at high risk for developing mental health problems, indicating a decrease in their well-being [48,49]. Here, it was shown for the first time that there is not only a correlation between an increase in the intensity and the amount of the students' exercise (102/166 [61.4%] of the students engaged in a greater amount of physical activity) but also a correlation between the students who engaged in a greater amount of exercise and their well-being and sleep quality.

Transitional times in life, such as going to college and being a freshman, include leaving home, gaining more independence, and adhering to less structured schedules, and other stressful conditions can affect sleep patterns and well-being [64-66]. Thus, there was expected to be a difference between first-year and second- and third-year students with regard to maintaining a healthy lifestyle. Here, no interaction was found between the time the questionnaire was submitted and the year of study regarding the amount of exercise, sleep quality, and well-being, emphasizing that all students, regardless of what year of study they are in, showed better sleep quality and well-being and engaged in more exercise during the lockdown compared to the previous period. The study by Romero-Blanco et al [39] showed that first- and second-year students, but not final-year students, who increased the amount of physical activity during COVID-19 accounted for the difference in the exercise time by the fact that third-year students are required to take on a large number of written assignments, hence increasing their sitting time. In contrast, in this study, the undergraduate students of education had a similar curriculum during the first 3 years of studies; thus, no difference was found between the years of study with respect to physical activity.

Finally, an ordinal interaction was found between the time when the questionnaire was submitted and gender regarding the amount of exercise, that is, even though all students exercised more during the COVID-19 lockdown, male students engaged in a greater amount of exercise before and during the lockdown compared to female students, although female students had a more significant change during the COVID-19 lockdown. This is in line with other studies showing that male students, in general, engage in more sports and more physical activity because of the positive sensation they experience from exercise, stimulation, and enjoyment compared to female students [67]. In addition, male students mention enjoyment, challenges, social recognition, affiliation, competition, and strength as motivating factors for exercise, whereas female students mention preventing poor health and maintaining good health, weight management, and a pleasing appearance [68].

Limitations

This study had several limitations. First, the study was conducted on a small and specific population most of whom are single and without children; therefore, our findings cannot be generalized to all undergraduate student populations, particularly those who are married and with children. Second, there was a difference between the students' responses to the first and second questionnaires because the pre- COVID-19 questionnaire was

given to the students by their teacher during 1 of their classes and they were given time to answer and submit during class. During the COVID-19 lockdown, all student-teacher communication was only online. Hence, the second questionnaire was sent via email, and not all students who answered the pre-COVID-19 lockdown questionnaire answered the second one. Although the number of respondents decreased, our findings were statistically significant. Third, it should be noted that there might be other factors that can affect well-being, sleep quality, and the amount of exercise that were not examined, such as health status, economic status, personality type, social status, and even living area (city vs rural locality). Finally, this effect was explored during the first lockdown only and lacks information regarding whether there could be a long-term, during-COVID-19 effect or whether the effect would similar during the second COVID-19 lockdown. This should be studied in continued research.

Conclusion

In conclusion, although this study was conducted on a specific population of undergraduate students of education, and thus it is not possible to draw conclusions about the entire population, we showed a statistical correlation between the tested variables, indicating a beneficial effect on sleep quality, well-being, and amount of exercise in this population during the COVID-19 lockdown. These results can encourage people to understand that there is another positive aspect to crisis situations, such as lockdowns, and one can use such situations to improve the quality of life. The results can also inspire policymakers to propose programs focusing on promoting physical activity, and emphasizing sleep quality and well-being, to maintain a healthy lifestyle during lockdowns.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Working during the COVID-19 lockdown relative to normal days.
[DOCX File , 13 KB - formative_v6i2e27286_app1.docx]

References

1. World Health Organization. Statement on the second meeting of the International Health Regulations (2005) Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV). 2020 Presented at: Second Meeting of the International Health Regulations (2005) Emergency Committee; 30 January 2020; Geneva, Switzerland URL: https://tinyurl.com/55f2kvmu [doi: 10.2807/1560-7917.es.2020.25.5.200131e]
2. Gibney E. Coronavirus lockdowns have changed the way Earth moves. Nature 2020 Apr 31;580(7802):176-177. [doi: 10.1038/d41586-020-00965-x] [Medline: 32235922]
3. Israel Coronavirus, 677 Cases and 3,286 Deaths. Worldometer. URL: https://www.worldometers.info/coronavirus/country/israel/ [accessed 2022-02-15]
4. Rajkumar RP. COVID-19 and mental health: a review of the existing literature. Asian J Psychiatr 2020 Aug;52:102066 [FREE Full text] [doi: 10.1016/j.ajp.2020.102066] [Medline: 32302935]
5. Wynn R. E-health in Norway before and during the initial phase of the Covid-19 pandemic. Stud Health Technol Inform 2020 Jun 26;272:9-12. [doi: 10.3233/SHTI200480] [Medline: 32604587]
6. Roy S. Insomnia Medication Prescriptions Increase with Spread of Coronavirus. Sleep Review. URL: https://www.sleepreviewmag.com/sleep-treatments/pharmaceuticals/prescription-drugs/insomnia-medication-prescriptions-coronavirus/ [accessed 2022-02-15]
7. Mazza C, Ricci E, Biondi S, Colasanti M, Ferracuti S, Napoli C, et al. A nationwide survey of psychological distress among Italian people during the COVID-19 pandemic: immediate psychological responses and associated factors. Int J Environ Res Public Health 2020 May 02;17(9):3165 [FREE Full text] [doi: 10.3390/ijerph17093165] [Medline: 32370116]
8. Maugeri G, Castrogiovanni P, Battaglia G, Pippi R, D'Agata V, Palma A, et al. The impact of physical activity on psychological health during Covid-19 pandemic in Italy. Heliyon 2020 Jun;6(6):e04315 [FREE Full text] [doi: 10.1016/j.heliyon.2020.e04315] [Medline: 32613133]
9. Yamada M, Kimura Y, Ishiyama D, Otobe Y, Suzuki M, Koyama S, et al. Effect of the COVID-19 epidemic on physical activity in community-dwelling older adults in Japan: a cross-sectional online survey. J Nutr Health Aging 2020 Jun 23;24(9):948-950. [doi: 10.1007/s12603-020-1501-6]
10. Dinis J, Bragança M. Quality of sleep and depression in college students: a systematic review. Sleep Sci 2018;11(4):290-301 [FREE Full text] [doi: 10.5935/1984-0063.20180045] [Medline: 30746048]
11. Stanton R, To QG, Khalesi S, Williams SL, Alley SJ, Thwaite TL, et al. Depression, anxiety and stress during COVID-19: associations with changes in physical activity, sleep, tobacco and alcohol use in Australian adults. Int J Environ Res Public Health 2020 Jun 07;17(11):4065 [FREE Full text] [doi: 10.3390/ijerph17114065] [Medline: 32517294]

12. Duncan GE, Avery AR, Seto E, Tsang S. Perceived change in physical activity levels and mental health during COVID-19: findings among adult twin pairs. PLoS One 2020 Aug 13;15(8):e0237695 [FREE Full text] [doi: 10.1371/journal.pone.0237695] [Medline: 32790745]
13. Dabell J. The importance of sleep. SecEd 2019 Nov 02;2019(16):14-15. [doi: 10.12968/sece.2019.16.14]
14. Curcio G, Ferrara M, De Gennaro L. Sleep loss, learning capacity and academic performance. Sleep Med Rev 2006 Oct;10(5):323-337. [doi: 10.1016/j.smrv.2005.11.001] [Medline: 16564189]
15. Lau JTF, Yang X, Tsui HY, Pang E, Wing YK. Positive mental health-related impacts of the SARS epidemic on the general public in Hong Kong and their associations with other negative impacts. J Infect 2006 Aug;53(2):114-124 [FREE Full text] [doi: 10.1016/j.jinf.2005.10.019] [Medline: 16343636]
16. Javaheri S, Storfer-Isser A, Rosen CL, Redline S. Sleep quality and elevated blood pressure in adolescents. Circulation 2008 Sep 02;118(10):1034-1040. [doi: 10.1161/circulationaha.108.766410]
17. Litwiller B, Snyder LA, Taylor WD, Steele LM. The relationship between sleep and work: a meta-analysis. J Appl Psychol 2017 Apr;102(4):682-699. [doi: 10.1037/apl0000169] [Medline: 27893255]
18. Pilcher JJ, Walters AS. How sleep deprivation affects psychological variables related to college students' cognitive performance. J Am Coll Health 1997 Nov;46(3):121-126. [doi: 10.1080/07448489709595597] [Medline: 9394089]
19. Abdulah DM, Musa DH. Insomnia and stress of physicians during COVID-19 outbreak. Sleep Med X 2020 Dec;2:100017 [FREE Full text] [doi: 10.1016/j.sleepx.2020.100017] [Medline: 33860222]
20. Zhang C, Yang L, Liu S, Ma S, Wang Y, Cai Z, et al. Survey of insomnia and related social psychological factors among medical staff involved in the 2019 novel coronavirus disease outbreak. Front Psychiatry 2020 Apr 14;11:306 [FREE Full text] [doi: 10.3389/fpsyt.2020.00306] [Medline: 32346373]
21. Orgilés M, Morales A, Delveccio E, Mazzeschi C, Espada JP. Immediate psychological effects of COVID-19 quarantine in youth from Italy and Spain. SSRN J 2020:24pp. [doi: 10.2139/ssrn.3588552]
22. Radek K, Burkes B, Levitt T, Baxter E, DiNardi A. Global sleep health in a COVID-19 virus-infected world. Int Med 2020;2(2):99. [doi: 10.5455/im.102662]
23. Warburton DER, 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]
24. Ancoli-Israel S. "Sleep is not tangible" or what the Hebrew tradition has to say about sleep. Psychosom Med 2001;63(5):778-787. [doi: 10.1097/00006842-200109000-00011] [Medline: 11573026]
25. Sherrill DL, Kotchou K, Quan SF. Association of physical activity and human sleep disorders. Arch Intern Med 1998 Sep 28;158(17):1894-1898. [doi: 10.1001/archinte.158.17.1894] [Medline: 9759685]
26. Morgenthaler T, Kramer M, Alessi C, Friedman L, Boehlecke B, Brown T, American Academy of Sleep Medicine. Practice parameters for the psychological and behavioral treatment of insomnia: an update. An American academy of sleep medicine report. Sleep 2006;29(11):1415-1419. [doi: 10.1093/sleep/29.11.1415]
27. Kredlow MA, Capozzoli MC, Hearon BA, Calkins AW, Otto MW. The effects of physical activity on sleep: a meta-analytic review. J Behav Med 2015 Jun;38(3):427-449. [doi: 10.1007/s10865-015-9617-6] [Medline: 25596964]
28. Kubitz KA, Landers DM, Petruzzello SJ, Han M. The effects of acute and chronic exercise on sleep. A meta-analytic review. Sports Med 1996 Apr;21(4):277-291. [doi: 10.2165/00007256-199621040-00004] [Medline: 8726346]
29. O'Connor PJ, Youngstedt SD. Influence of exercise on human sleep. Exerc Sport Sci Rev 1995;23:105-134. [doi: 10.1249/00003677-199500230-00006]
30. Youngstedt S, O'Connor PJ, Dishman R. The effects of acute exercise on sleep: a quantitative synthesis. Sleep 1997 Mar;20(3):203-214. [doi: 10.1093/sleep/20.3.203] [Medline: 9178916]
31. Reid KJ, Baron KG, Lu B, Naylor E, Wolfe L, Zee PC. Aerobic exercise improves self-reported sleep and quality of life in older adults with insomnia. Sleep Med 2010 Oct;11(9):934-940 [FREE Full text] [doi: 10.1016/j.sleep.2010.04.014] [Medline: 20813580]
32. Nguyen-Michel ST, Unger JB, Hamilton J, Spruijt-Metz D. Associations between physical activity and perceived stress/hassles in college students. Stress Health 2006 Aug;22(3):179-188. [doi: 10.1002/smi.1094]
33. Cooney G, Dwan K, Greig C. Exercise for depression. Cochrane Database Syst Rev 2013:12 Published online. [doi: 10.1002/14651858.cd004366.pub6]
34. Asmundson GJG, Fetzner MG, Deboer LB, Powers MB, Otto MW, Smits JAJ. Let's get physical: a contemporary review of the anxiolytic effects of exercise for anxiety and its disorders. Depress Anxiety 2013 Apr 08;30(4):362-373. [doi: 10.1002/da.22043] [Medline: 23300122]
35. Gorczynski P, Faulkner G. Exercise therapy for schizophrenia. Cochrane Database Syst Rev 2010:12 Published online. [doi: 10.1002/14651858.cd004412.pub2]
36. Luciano F, Cenacchi V, Vegro V, Pavei G. COVID-19 lockdown: physical activity, sedentary behaviour and sleep in Italian medicine students. Eur J Sport Sci 2021 Oct 06;21(10):1459-1468. [doi: 10.1080/17461391.2020.1842910] [Medline: 33108970]
37. Romero-Blanco C, Rodríguez-Almagro J, Onieva-Zafra MD, Parra-Fernández ML, Prado-Laguna MDC, Hernández-Martínez A. Sleep pattern changes in nursing students during the COVID-19 lockdown. Int J Environ Res Public Health 2020 Jul 20;17(14):5222 [FREE Full text] [doi: 10.3390/ijerph17145222] [Medline: 32698343]

38. Marelli S, Castelnuovo A, Somma A, Castronovo V, Mombelli S, Bottoni D, et al. Impact of COVID-19 lockdown on sleep quality in university students and administration staff. J Neurol 2021 Jan 11;268(1):8-15 [FREE Full text] [doi: 10.1007/s00415-020-10056-6] [Medline: 32654065]
39. Romero-Blanco C, Rodríguez-Almagro J, Onieva-Zafra MD, Parra-Fernández ML, Prado-Laguna MDC, Hernández-Martínez A. Physical activity and sedentary lifestyle in university students: changes during confinement due to the COVID-19 pandemic. Int J Environ Res Public Health 2020 Sep 09;17(18):6567 [FREE Full text] [doi: 10.3390/ijerph17186567] [Medline: 32916972]
40. Booth M. Assessment of physical activity: an international perspective. Res Q Exerc Sport 2000 Jun;71 Suppl 2:114-120. [doi: 10.1080/02701367.2000.11082794] [Medline: 25680021]
41. Veit CT, Ware JE. The structure of psychological distress and well-being in general populations. J Consult Clin Psychol 1983;51(5):730-742. [doi: 10.1037/0022-006X.51.5.730]
42. Florian V, Drori Y. Mental Health Inventory (MHI): psychometric characteristics and normative data regarding the Israeli population. Psychologia Isr J Psychol 1990;2(1):26-35 [FREE Full text]
43. Zomer J, Peled R, Rubin A, Lavie P. Mini Sleep Questionnaire (MSQ) for screening large populations for EDS complaints in sleep. In: Koella W, Ruther E, Schulz H, editors. Sleep. New York, NY: Gustav Fischer Verlag; 1985:1984-1470.
44. Sinha M, Pande B, Sinha R. Impact of COVID-19 lockdown on sleep-wake schedule and associated lifestyle related behavior: a national survey. J Public Health Res 2020 Jul 28;9(3):1826 [FREE Full text] [doi: 10.4081/jphr.2020.1826] [Medline: 32874967]
45. Đogaš Z, Lušić Kalcina L, Pavlinac Dodig I, Demirović S, Madirazza K, Valić M, et al. The effect of COVID-19 lockdown on lifestyle and mood in Croatian general population: a cross-sectional study. Croat Med J 2020 Aug;61(4):309-318. [doi: 10.3325/cmj.2020.61.309]
46. Robinson E, Boyland E, Chisholm A, Harrold J, Maloney NG, Marty L, et al. Obesity, eating behavior and physical activity during COVID-19 lockdown: a study of UK adults. Appetite 2021 Jan 01;156:104853 [FREE Full text] [doi: 10.1016/j.appet.2020.104853] [Medline: 33038479]
47. López-Bueno R, Calatayud J, Andersen LL, Balsalobre-Fernández C, Casaña J, Casajús JA, et al. Immediate impact of the COVID-19 confinement on physical activity levels in Spanish adults. Sustainability 2020 Jul 16;12(14):5708. [doi: 10.3390/su12145708]
48. Sundarasen S, Chinna K, Kamaludin K, Nurunnabi M, Baloch GM, Khoshaim HB, et al. Psychological impact of COVID-19 and lockdown among university students in Malaysia: implications and policy recommendations. Int J Environ Res Public Health 2020 Aug 27;17(17):6206 [FREE Full text] [doi: 10.3390/ijerph17176206] [Medline: 32867024]
49. 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]
50. Altena E, Baglioni C, Espie CA, Ellis J, Gavriloff D, Holzinger B, et al. Dealing with sleep problems during home confinement due to the COVID-19 outbreak: practical recommendations from a task force of the European CBT-I Academy. J Sleep Res 2020 Aug 04;29(4):e13052 [FREE Full text] [doi: 10.1111/jsr.13052] [Medline: 32246787]
51. Di Stefano V, Battaglia G, Giustino V, Gagliardo A, D'Aleo M, Giannini O, et al. Significant reduction of physical activity in patients with neuromuscular disease during COVID-19 pandemic: the long-term consequences of quarantine. J Neurol 2021 Jan 13;268(1):20-26 [FREE Full text] [doi: 10.1007/s00415-020-10064-6] [Medline: 32661716]
52. Sandín B, Valiente RM, García-Escalera J, Chorot P. Impacto psicológico de la pandemia de COVID-19: efectos negativos y positivos en población española asociados al periodo de confinamiento nacional. RPPC 2020 Sep 01;25(1):1. [doi: 10.5944/rppc.27569]
53. Martínez-González MA, Varo JJ, Santos JL, De Irala J, Gibney M, Kearney J, et al. Prevalence of physical activity during leisure time in the European Union. Med Sci Sports Exerc 2001 Jul;33(7):1142-1146. [doi: 10.1097/00005768-200107000-00011] [Medline: 11445761]
54. Shigemura J, Ursano RJ, Morganstein JC, Kurosawa M, Benedek DM. Public responses to the novel 2019 coronavirus (2019-nCoV) in Japan: mental health consequences and target populations. Psychiatry Clin Neurosci 2020 Apr;74(4):281-282 [FREE Full text] [doi: 10.1111/pcn.12988] [Medline: 32034840]
55. Majumdar P, Biswas A, Sahu S. COVID-19 pandemic and lockdown: cause of sleep disruption, depression, somatic pain, and increased screen exposure of office workers and students of India. Chronobiol Int 2020 Aug 13;37(8):1191-1200. [doi: 10.1080/07420528.2020.1786107] [Medline: 32660352]
56. Brajša-Žganec A, Merkaš M, Šverko I. Quality of life and leisure activities: how do leisure activities contribute to subjective well-being? Soc Indic Res 2010 Sep 30;102(1):81-91. [doi: 10.1007/s11205-010-9724-2]
57. Molina-García J, Castillo I, Queralt A. Leisure-time physical activity and psychological well-being in university students. Psychol Rep 2011 Oct 01;109(2):453-460. [doi: 10.2466/06.10.13.pr0.109.5.453-460]
58. McMahon EM, Corcoran P, O'Regan G, Keeley H, Cannon M, Carli V, et al. Physical activity in European adolescents and associations with anxiety, depression and well-being. Eur Child Adolesc Psychiatry 2017 Jan 9;26(1):111-122. [doi: 10.1007/s00787-016-0875-9] [Medline: 27277894]

59. Haapasalo V, de Vries H, Vandelanotte C, Rosenkranz RR, Duncan MJ. Cross-sectional associations between multiple lifestyle behaviours and excellent well-being in Australian adults. Prev Med 2018 Nov;116:119-125. [doi: 10.1016/j.ypmed.2018.09.003] [Medline: 30218725]
60. Steptoe A, Butler N. Sports participation and emotional wellbeing in adolescents. Lancet 1996 Jun;347(9018):1789-1792. [doi: 10.1016/s0140-6736(96)91616-5]
61. Ströhle A. Physical activity, exercise, depression and anxiety disorders. J Neural Transm (Vienna) 2009 Jun;116(6):777-784. [doi: 10.1007/s00702-008-0092-x] [Medline: 18726137]
62. Wunsch K, Kasten N, Fuchs R. The effect of physical activity on sleep quality, well-being, and affect in academic stress periods. NSS 2017 Apr;9:117-126. [doi: 10.2147/nss.s132078]
63. Sullivan Bisson AN, Robinson SA, Lachman ME. Walk to a better night of sleep: testing the relationship between physical activity and sleep. Sleep Health 2019 Oct;5(5):487-494 [FREE Full text] [doi: 10.1016/j.sleh.2019.06.003] [Medline: 31358470]
64. Lund HG, Reider BD, Whiting AB, Prichard JR. Sleep patterns and predictors of disturbed sleep in a large population of college students. J Adolesc Health 2010 Feb;46(2):124-132. [doi: 10.1016/j.jadohealth.2009.06.016] [Medline: 20113918]
65. Taylor DJ, Bramoweth AD. Patterns and consequences of inadequate sleep in college students: substance use and motor vehicle accidents. J Adolesc Health 2010 Jun;46(6):610-612. [doi: 10.1016/j.jadohealth.2009.12.010] [Medline: 20472221]
66. Taylor DJ, Bramoweth AD, Grieser EA, Tatum JI, Roane BM. Epidemiology of insomnia in college students: relationship with mental health, quality of life, and substance use difficulties. Behav Ther 2013 Sep;44(3):339-348. [doi: 10.1016/j.beth.2012.12.001] [Medline: 23768662]
67. Lauderdale ME, Yli-Piipari S, Irwin CC, Layne TE. Gender differences regarding motivation for physical activity among college students: a self-determination approach. TPE 2015;72:153-172. [doi: 10.18666/tpe-2015-v72-i5-4682]
68. Gao Z, Xiang P. College students' motivation toward weight training: an application of expectancy-value model. J Teach Phys Educ 2008;27(3):399-415. [doi: 10.1123/jtpe.27.3.399]

Edited by T Leung; submitted 20.01.21; peer-reviewed by B Wang, F Gomez, R Fang; comments to author 08.06.21; revised version received 20.07.21; accepted 13.01.22; published 23.02.22.

Please cite as:
Joseph G, Schori H
The Beneficial Effect of the First COVID-19 Lockdown on Undergraduate Students of Education: Prospective Cohort Study
JMIR Form Res 2022;6(2):e27286
URL: https://formative.jmir.org/2022/2/e27286
doi:10.2196/27286
PMID:35072635

Original Paper

A New Performance Metric to Estimate the Risk of Exposure to Infection in a Health Care Setting: Descriptive Study

Kimia Hadian[1,2], BSc; Geoff Fernie[1,2,3], PhD; Atena Roshan Fekr[1,2], PhD

[1]The KITE Research Institute, Toronto Rehabilitation Institute, University Health Network, Toronto, ON, Canada

[2]Institute of Biomedical Engineering, University of Toronto, Toronto, ON, Canada

[3]Department of Surgery, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Kimia Hadian, BSc
The KITE Research Institute
Toronto Rehabilitation Institute
University Health Network
550 University Ave
Toronto, ON, M5G 2A2
Canada
Phone: 1 4164558109
Email: kimia.hadian@mail.utoronto.ca

Abstract

Background: Despite several measures to monitor and improve hand hygiene (HH) in health care settings, health care-acquired infections (HAIs) remain prevalent. The measures used to calculate HH performance are not able to fully benefit from the high-resolution data collected using electronic monitoring systems.

Objective: This study proposes a novel parameter for quantifying the HAI exposure risk of individual patients by considering temporal and spatial features of health care workers' HH adherence.

Methods: Patient exposure risk is calculated as a function of the number of consecutive missed HH opportunities, the number of unique rooms visited by the health care professional, and the time duration that the health care professional spends inside and outside the patient's room without performing HH. The patient exposure risk is compared to the entrance compliance rate (ECR) defined as the ratio of the number of HH actions performed at a room entrance to the total number of entrances into the room. The compliance rate is conventionally used to measure HH performance. The ECR and the patient exposure risk are analyzed using the data collected from an inpatient nursing unit for 12 weeks.

Results: The analysis of data collected from 59 nurses and more than 25,600 records at a musculoskeletal rehabilitation unit at the Toronto Rehabilitation Institute, KITE, showed that there is no strong linear relation between the ECR and patient exposure risk (r=0.7, P<.001). Since the ECR is calculated based on the number of missed HH actions upon room entrance, this parameter is already included in the patient exposure risk. Therefore, there might be scenarios that these 2 parameters are correlated; however, in several cases, the ECR contrasted with the reported patient exposure risk. Generally, the patients in rooms with a significantly high ECR can be potentially exposed to a considerable risk of infection. By contrast, small ECRs do not necessarily result in a high patient exposure risk. The results clearly explained the important role of the factors incorporated in patient exposure risk for quantifying the risk of infection for the patients.

Conclusions: Patient exposure risk might provide a more reliable estimation of the risk of developing HAIs compared to ECR by considering both the temporal and spatial aspects of HH records.

(JMIR Form Res 2022;6(2):e32384) doi:10.2196/32384

KEYWORDS

hand hygiene; health care-acquired; infection control; compliance; electronic monitoring; exposure; risk; hygiene; monitoring; surveillance; performance; metric; method; estimate; predict; development

Introduction

Health care-acquired infections (HAIs) occur during the process of care in hospitals or health care centers and were not present at the time of patient admission. HAIs cause more than 99,000 deaths and account for an additional cost of US $28-$45 billion annually in the United States [1,2]. In Canada, 48,653 cases of HAI were reported between 2017 and 2018 [3]. A study conducted in 2017 indicated that the prevalence of HAIs in Canada was 7.9% [4]. The hands of health care workers (HCWs) play an important role in spreading the pathogens responsible for HAIs in health care settings [5]. Studies have shown that hand hygiene (HH) is one of the most effective ways to reduce infection transmission in health care settings, and both alcohol-based hand sanitizers and soaps are effective in disrupting the chain of infection transmission (including viruses such as SARS-CoV-2) [5-9]. Currently, the growing number of COVID-19 cases highlights the importance of HH in infection prevention [10].

Several organizations including the World Health Organization and the Ministry of Health and Long-Term Care (MOHLTC) in Ontario, Canada have provided guidelines and recommendations for HH practices in health care environments [11,12]. According to MOHLTC's standard precautions, HCWs should clean their hands in four moments or opportunities: (1) before initial patient or patient environment contact; (2) before aseptic procedures; (3) after body fluid exposure risk; and (4) after patient or patient environment contact [12]. Most types of HAIs can be avoided by complying to the standard HH protocols [13]. HCWs' adherence to recommended HH procedures varies from 5% to 89% with an overall average of as low as 38.7% [13]. Identifying the poor performers as well as the population at greater risk of infection are the two key elements to understanding the underlying reasons for low compliance and directing efforts to prevent the spread of infection.

Detection of HH actions by a trained observer is currently the gold standard in HH monitoring; however, this method is not only expensive but also suffers from inadequate staffing, delayed data feedback, and an overestimate of performance between 200% and 300% [14,15]. With the recent advances in electronic monitoring systems (EMS), it is possible to track individual HCW's HH actions automatically [16-21]. These systems provide additional information such as the length of time spent in each room, consecutive missed HH opportunities, and the number and time of entrances into and exits from each zone. This information allows us to augment existing measurements of HH performance by extending the focus from single moments to continuous time-dependent and space-dependent features in HH records. The compliance rate, which focuses on the adherence of HCWs to HH protocols, is commonly used to estimate each individual's HH performance [5]. However, this parameter masks the temporal and spatial aspects of HH involved in infection transmission. In this study, a novel HH performance metric is proposed to quantify the individual patient's exposure risk to infection. This parameter, called patient exposure risk, changes the focus from improving HCWs' compliance rate to reducing the patient's risk of infection. Different examples are discussed to show the likely ability of the new proposed metric compared to the traditional compliance rate in estimating the infection risk for patients.

Methods

Overview

In this paper, a new performance metric, patient exposure risk, is introduced to quantify the risk of exposure to an infection for each individual patient at a health care environment. This parameter, which is estimated for each patient's room during 24 hours, is compared with the entrance compliance rate (ECR), a localized version of the conventional compliance rate. ECR is defined as the ratio of the number of times HH actions were performed by the HCWs upon entering a specific room to the total number of times HH action was required. Therefore, the higher the ECR, the better the performance of HH for that specific room. In other words, an ECR of 100% shows that all HCWs who entered and exited the patient's room washed their hands.

Conventional compliance rate considers HH opportunities as isolated binary events. However, the risk of infection transmission to the patient is a multifactorial concept depending on the events prior to entering the patient's room as well as the events happening inside the room. The risk of developing an infection in a patient depends on the number of microorganisms on the HCW's hand, the type of microorganism, the duration of exposure, the activity performed by the HCW, and host susceptibility [22]. The new proposed patient exposure risk estimates the risk of exposure to an infection by continuously monitoring five major factors that are responsible in the chain of infection transmission but were not included in the current compliance measure. These factors include the number of consecutive missed HH opportunities, the number of unique rooms visited by the HCWs with missed HH opportunities, the length of time spent inside each patient room without performing HH, the length of time spent outside of the patient room without performing HH, and the risk factor associated with the type of pathogens present in each specific room.

Analyzing the data collected from 59 nurses at a musculoskeletal rehabilitation unit at the Toronto Rehabilitation Institute, KITE, highlighted the important role of the factors incorporated in patient exposure risk estimation. These data were collected from 10 single rooms and 10 double rooms in three 30-day trials from 2016 to 2017 using the Buddy Badge (Hygienic Echo Inc) EMS [23]. The Buddy Badge EMS provides information such as the number of consecutive missed HH opportunities, the HCWs' location, and the time each HCW spent inside or outside of each patient's room [24]. After cleaning the data, 85 days were considered in our analysis. Python 3.6 (Python Software Foundation) and the Spearman rank-order correlation were used for the analysis. The Spearman correlation measures the monotonic relationship between 2 variables (ie, if 1 variable increases or decreases, the other should also increase or decrease to produce a high correlation coefficient). The values in each signal will be assigned a rank variable with respect to each other. The Spearman correlation coefficient is defined as the Pearson correlation coefficient between the ranked values and is computed as:

where r_s denotes the Spearman correlation coefficient, $cov(rs_1,rs_2)$ is the covariance of the rank variables, and σ_{rs_1} and σ_{rs_2} are the standard deviation of the rank variables [25].

Preliminary Analysis of the Contributing Factors to the Patient Exposure Risk

The overall risk of infection lies in the cumulative high frequency of hands touching surfaces and patients instead of individual HH indications [22]. The authors in [23] showed that more than 80% of missed HH opportunities occur as a part of a chain with a median length of 4 and an IQR of 8, instead of a single isolated event. In addition, the results of our analysis showed that 28% (1954/6887) of the nurses entering a specific room without washing hands had a chain of missed HH opportunities (Figure 1).

It was also observed that 40% (2774/6887) of missed HH chains are made up of consecutive visits to the same room (eg, when the nurses exit the room to pick up clean supplies from the clean utility room or dispose soiled linen in the soiled utility room). Table 1 indicates the number of unique rooms visited by the nurses without handwashing before entering a specific room.

Figure 1. Number of entrances with and without a chain of missed hand hygiene opportunities prior to the entrance over 85 days. HH: hand hygiene.

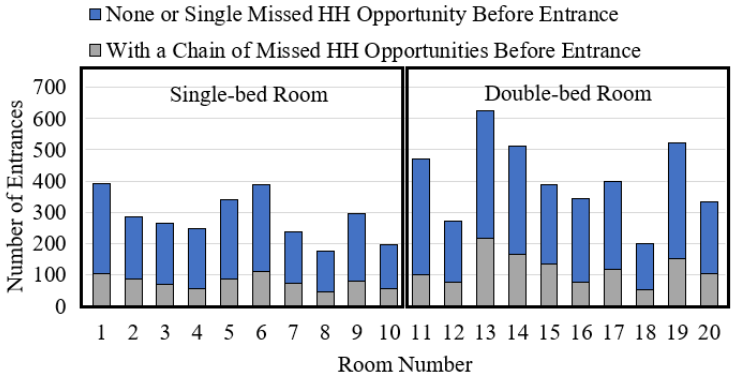


Table 1. Histogram of the number of unique rooms visited in the chain of missed hand hygiene opportunities over 85 days.

Number of unique rooms visited	Total number of unique rooms visited (%)
1	65.82
2	21.48
3	5.71
4	2.51
5	1.47
6	0.94
7	0.68
8	0.42
9	0.29
10	0.26
11	0.20
12	0.07
13	0.12
14	0.03

This analysis shows that in more than 34% (2354/6887) of the entrances in which HCWs did not perform HH, they had failed to perform handwashing actions in 2 or more previous visits to other rooms as well. In other words, in about 34% of entrances, the patient may have been exposed to the risk of infection from 2 or more types of infectious agents acquired in previously visited rooms.

Another important factor responsible in the transmission risk is the length of time spent in a patient room without a handwashing action. Bacterial hand contamination increases linearly over time [26]. Thus, the more time the nurse spends in each room, the more likely the patient will become infected if HH action is missed. Figure 2 depicts the average daily time duration spent in each room without HH action for the entire data set. For instance, considering room #1, the nurses spent totally about 908 minutes (~15 hours) inside this room without washing their hands upon entry, during 85 days. Therefore, on average, the nurses spent about 11 minutes during 24 hours in room #1 without washing their hands. This means that, on

average, the patient in this room was exposed to risk of infection transmission from the HCWs' hands for 11 minutes in a single day.

Clack et al [27] reported that, on average, 14.2 hand-to-surface exposures per minute happened in an intensive care unit, where 46% were outside the patient zone. This indicates that the time duration spent outside the patient room, including visiting the previous patients or staying in the hallways, can play an important role in infection transmission since this will give the HCWs enough time to contact the environment. Figure 3 shows the time duration spent outside the patients' rooms before entering a specific room without HH action.

It was observed that more than 35% (2449/6887) of the nurses spent less than 5 minutes outside the patients' rooms with no handwashing action. In addition, about 75% (5091/6887) of the nurses spent 0-25 minutes outside the rooms without performing HH. It is likely that the conventional performance metric, compliance rate, is unable to accurately estimate the patient's exposure risk to an infection without considering all these factors.

Figure 2. The average duration and unwashed duration spent in each room during 24 hours over 85 days; HH: hand hygiene; min: minutes.

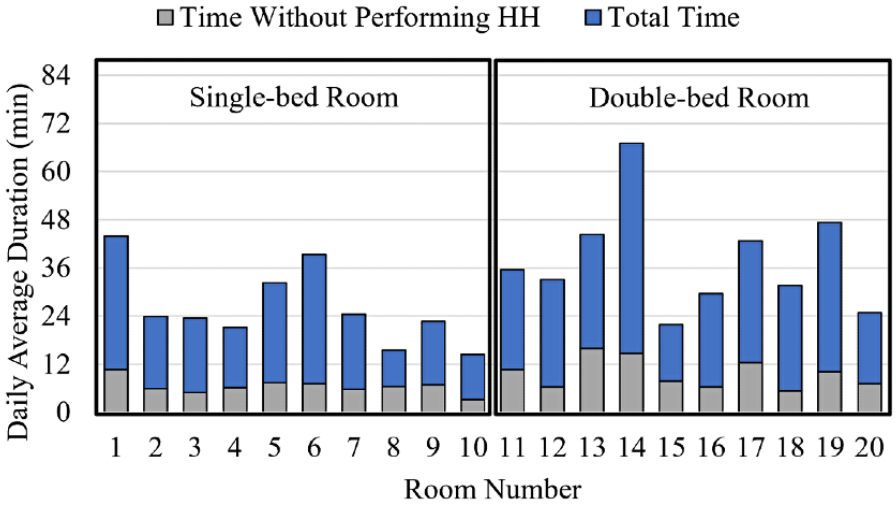
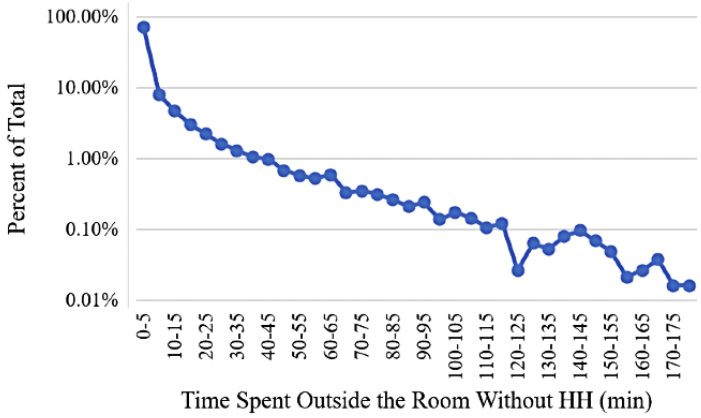


Figure 3. Distribution of time between previous hand hygiene and entrance to a room. HH: hand hygiene; min: minutes.

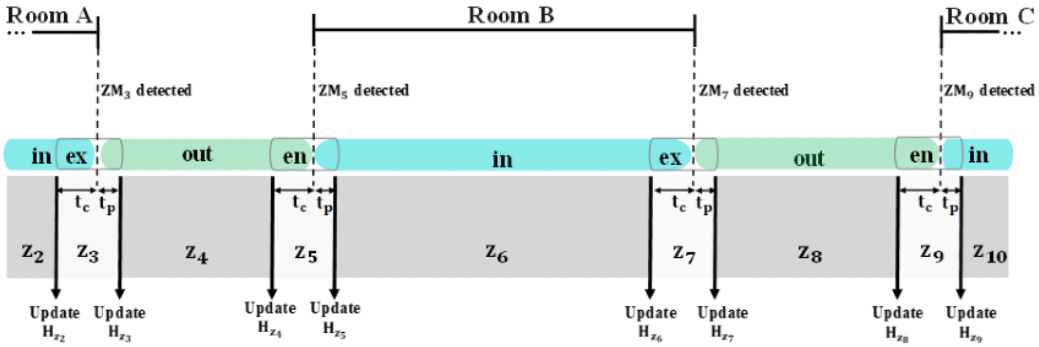


Formal Definition

The patient exposure risk is defined considering 4 zones shown as z_i (Figure 4), including outside the room ("out"), inside the room ("in"), room entrance ("en"), and room exit ("ex"). The "en" or "ex" zones (z_i with $i=2k+1$, where $k \in W$) are specified by 2 time constants: t_c and t_p.

An HH opportunity is compliant if HCWs wash their hands within t_c seconds before and t_p seconds after entering or exiting the room detected by zone markers (ZM_i) (Figure 4). t_c is defined for the scenarios where the first and the fourth moments of HH are combined. For instance, if the HCW performs HH upon exiting a room (moment 4) and immediately enters another room (moment 1) without touching other surfaces, moment 1 and moment 4 are combined, and one handwashing action is sufficient in this scenario. Therefore, the missed HH opportunities that occur within t_c seconds after the last HH action are not counted. HCWs can wash their hands after t_p seconds of their entrance if they forget to do so. This parameter can also be used to issue a reminder for missed HH opportunities.

Figure 4. Zones and timestamps used in calculating the patient exposure risk. in: inside; ex: exit; en: entrance; z: zone; H: History; ZM: zone marker.



The patient exposure risk, denoted by PER, for room x is calculated as follows:

where t denotes time and w is a binary representation of handwashing action. $w=1$ and $w=0$ show that handwashing is performed and missed, respectively. $t_j^k - t_j^{en}$ is the time duration spent in room x by j^{th} entrance without HH. Two scenarios ($k \in \{w, ex\}$) are modeled in Equation 2. The first scenario is when the caregiver washes their hands inside the room after t_j^w seconds ($t_j^w - t_j^{en}$), and the second scenario is when they do not wash their hands until they exit the room. In this case, the total time that they spent in this room is considered as ($t_j^{ex} - t_j^{en}$). The parameter R_n (Equation 3) represents the risk associated with each HCW entering room x and is calculated as follows:

H_{n,z_i} denotes the history of n^{th} staff entering i^{th} zone. This variable represents the risk of infection that staff 'n' accumulates over time before entering z_i. l_{n,z_i} is a set of rooms visited by n^{th} staff without performing HH before entering z_i. Therefore, $|\cup l_{n,z_i}|$ calculates the number of unique rooms visited by this staff without any handwashing action before entering z_i. For the "en" or "ex" zones, where $i=2k+1$, the H_{n,z_i} is calculated as Equation 4, and for the "in" or "out" zones, where $i=2k$, the H_{n,z_i} is calculated as Equation 5.

α_{z_i} and ZM_i are the risk factor and zone marker for zone i. We assume that each zone might have different risk factors based on the infectious agents that might be present in each individual zone (α_{z_i}). For example, the risk factor for the isolation room will be different from that of regular patient rooms, or the risk factor for the hallway will be different from that of patient rooms. In the "en" or "ex" zones, the history of HCW will be increased if they do not perform any handwashing action ($w=0$) upon entering or exiting the room (moment 1 and moment 4). In the "in" or "out" zones, where $i=2k$, we incorporated the time duration that the HCWs did not wash their hands. There are 2 cases that are modeled in Equation 5. The first case is that the caregiver washes their hands inside the zone after $t_{n,w}$ seconds ($t_{ZM_{i+1}} - t_{n,w}$), and the second case is when they do not wash their hands until they exit the zone. In this case, the total time that they spent in this zone ($t_{ZM_{i+1}} - t_{ZM_{i-1}}$) is calculated and added to their current history ($H_{n,z_{i-1}}$). The history (H_{n,z_i}) resets whenever handwashing action is performed ($w=1$)

Results

In this section, we compared the ECR and the proposed patient exposure risk with risk factors set to $\alpha_{z_i}=1$. Since the ECR is calculated based on the number of HH actions performed upon room entrance, this parameter is already included in our patient exposure risk. Therefore, we expect to see some correlations between patient exposure risk and ($100 - ECR$). However, the critical points are where these 2 parameters are not in agreement; this indicates that other factors may play an important role in the prediction model, which were ignored in the ECR. As an example, Figure 5 provides a comparison between patient exposure risk and ($100 - ECR$) for room 1104, in October 2017. Although the correlation coefficient between these 2 variables was r=0.83 (P<.001), there are 7 points where the 2 parameters provided conflicting results. For example, considering October 10 and 11, the ECR values on these 2 days showed that the HH performance was significantly higher on October 11 versus October 10, whereas the analysis of patient exposure risk indicated that the patient was likely exposed to a substantially greater risk of infection on October 11 compared with October 10.

On October 10, 3 out of 7 entrances to Room 1104 had a chain of missed HH opportunities. As summarized in Table 2, these 3 HCWs visited a total of 5 rooms, missed 6 HH opportunities, and spent about 16.45 minutes outside the room without washing their hands before entering Room 1104. However, in total, they all spent only 1 minute in the patient's room. Thus, the overall patient exposure risk for Room 1104 was as low as 38. On this day, 57% (4 out of 7) entrances into this room were compliant with HH.

Figure 5. Entrance compliance rate (ECR) and patient exposure risk for room 1104 in October 2017.

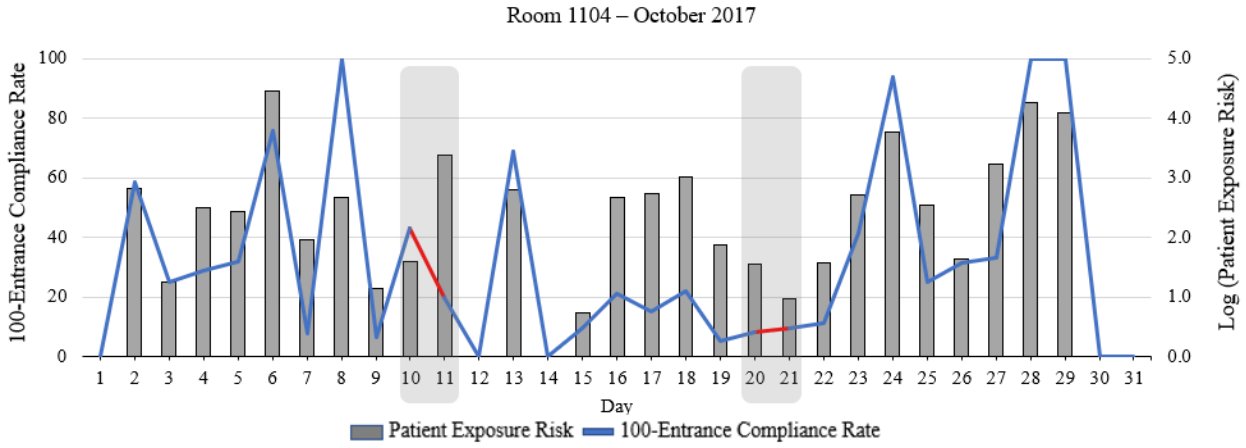


Table 2. An example to compare the patient exposure risk and entrance compliance rate (ECR).

Days	#ent[a]	#ent with H[b]>1	$\|\cup l_{n,z_i}\|$	#missed HH[c]	Time outside the room (minutes)	Time inside the room (minutes)	Patient exposure risk	100-ECR (%)
Oct 10	7	3	{1, 3, 1}	(1, 4, 1)	(5.68, 9.21, 1.55)	(0.03, 0.93, 0.05)	37.59	42.86
Oct 11	25	5	{1, 2, 2, 5, 2}	(2, 3, 3, 19, 5)	(20.2, 37.38, 53.05, 71.6, 13.92)	(3.6, 5.65, 1.16, 3.57, 2.95)	2388.24	20

[a]ent: entrance.

[b]H: History.

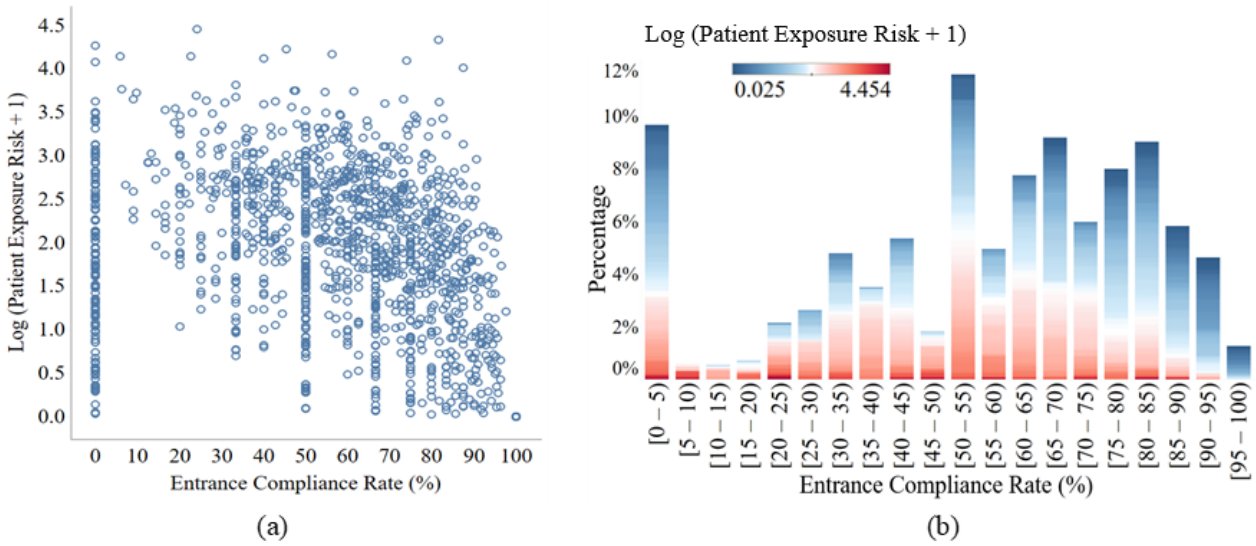
[c]HH: hand hygiene.

On October 11, there were 25 entrances into Room 1104, where 5 nurses did not wash their hands upon entering the room (ECR=80%). As summarized in Table 2, these 5 nurses visited 12 rooms, missed 32 HH opportunities, and spent almost 3.5 hours without washing their hands outside the room. In total, they spent about 17 minutes inside the patient room. This example obviously showed that the patient was in greater risk of becoming infected on October 11 compared with October 10 because these staff were more likely to carry infection. However, the ECR was approximately 25% higher on October 11, meaning that the performance of HH was better on this day. This example clearly explained that the traditional compliance rate only focuses on the adherence of HCWs to HH protocols and is not able to quantify and assess the potential patient exposure to infection. Similarly, October 20 and 21 showed another example in which ECR values were almost the same for both days, but the patient exposure risk was significantly different.

Figure 6 (image A) depicts the reported daily ECR and its corresponding patient exposure risk for each room in the data set. There is no strong linear relation between these 2 parameters (r=0.7, P<.001).

The more the 2 parameters are uncorrelated, the better the role of factors incorporated in the patient exposure risk and excluded in the ECR are highlighted. Figure 6 (image B) shows the distribution of the daily ECR for all the days and rooms in the data set. The distribution of the patient exposure risk in each bar is illustrated using colors. Even in the columns that represent an acceptable range for ECR (eg, 90-95%) there exists a noticeable proportion of high patient exposure risk values (red areas). Moreover, the columns with a significantly low ECR (eg, 0-5%) can have values with low patient exposure risk (blue areas). This is due to the fact that patient exposure risk considers more than just missed HH opportunities upon entrance. For example, if all the nurses who enter a room, except 1 nurse, perform HH upon entrance, the ECR can demonstrate a high HH performance. However, if this nurse has a high "History"—has missed several HH opportunities, visited various rooms, or did not wash his hands for a long time—and spends a considerable amount of time inside the patient's room, the patient may be exposed to a high risk of infection. On the other hand, if only one nurse enters a patient's room and does not perform HH but spends only a few seconds inside that room, the ECR will be as low as 0 whereas the patient exposure risk is negligible.

Figure 6. (a) Scatter plot of the log of the patient exposure risk and the entrance compliance rate (ECR); (b) distribution of the daily ECR for 20 rooms within 85 days (bars) and their corresponding patient exposure risk. The color change in this graph is proportional to log10 (patient exposure risk + 1).



Discussion

Principal Findings

The examples provided in this paper showed that the ECR cannot provide accurate information about the risk of contracting HAIs. On the other hand, patient exposure risk is designed to estimate the risk of infection for each individual patient. The patient exposure risk keeps track of consecutive missed HH opportunities and the risk accumulated by spending time inside and outside the patients' rooms. It considers the source of infectious agents by identifying high-risk areas (eg, type of room and the medical condition of the patient inside the room). Areas such as soiled utility rooms, which are intended for decontamination, storage, and disposal of used equipment and waste, can be a potential source of infection. Moreover, the risk of transmission varies between different types of contagions present in the patient's room [22,28]. Obviously, the ECR does not recognize the high risk caused by these areas whereas the patient exposure risk assigns a risk factor to each zone depending on its location and condition.

Several studies suggested that the risk of infection transmission is cumulative, yet health care settings rely on monitoring HH as individual binary events. In this paper, patient exposure risk is introduced to bridge the gap between the capabilities of new HH monitoring systems and infection prevention and control.

Limitations and Future Work

While introducing this concept, we encountered different limitations such as a lack of prognostic data to validate the relationship between the patient exposure risk and infection rates. However, we plan to address these limitations in our future studies where we will include data from various hospital settings. We will present some examples of these limitations in the following section:

1. In this study, the patient exposure risk was estimated only for single-bed rooms since the current EMSs are not able to identify the exact location of the caregiver inside the room to detect the patient receiving care in multi-bed rooms. Incorporating localization inside the patient room is a possible future research area to overcome this limitation.
2. The proposed performance metric monitors the potential transmission of infectious agents through the HCWs' hands by tracing their history of actions before entering each patient room. This is estimated using the number of consecutive missed HH opportunities, the number of unique rooms visited, and the time the HCWs spent without washing their hands before entering the patient's room. In addition to these 3 parameters, the type of activities performed by the caregivers plays an important role in infection transmission. For instance, changing a patient's dressing will potentially put the patient at greater risk of infection as opposed to assisting the patient with a wheelchair. The current EMSs are not able to identify the activities performed by the HCWs, and our data set does not contain this information. Therefore, the patient exposure risk concept will be extended in the future to include the types of activities performed by each staff, using artificial intelligence. This will be a new era of intelligent activity and behavior detection of frontline staff, which can help us identify the 4 moments of HH and the risk of exposure to infection accurately.
3. The analyzed dataset did not provide information regarding the patients in each room, and the patient exposure risk was not reported for each patient and was instead calculated for each patient's room. However, this might lead to inaccuracies since a room can host different patients in a single day. Including this additional information, the patient exposure risk can be calculated for each patient without any changes in the formulation.
4. HH policies vary depending on the HCWs' group assignment. For example, the staff responsible for pick-up and delivery of meal trays can go room to room without performing HH if there are no contacts with the patient or patient environment. In this paper, the patient exposure risk

was only calculated for the nursing group who will need to perform HH in all the 4 moments of MOHLTC precautions. Future studies could explore the calculation of the patient exposure risk for other HCW groups according to their specific HH protocol. In addition, similar to the patient exposure risk that measures the risk of infection for the patients, a new concept will be introduced in the future to measure the risk of infection for HCWs who are at a high risk of infection.

5. Although in the patient exposure risk definition we considered the time duration that the HCWs spent in each patient's rooms without handwashing action, it is also critical to include cases where the HH actions are performed but the contact time is over 15 minutes. For example, a close contact of more than 15 minutes is thought to increase the risk of infection transmission in the case of COVID-19 from staff to patient and vice versa [29]. Therefore, we will extend the patient exposure risk concept to a more general model that will incorporate such considerations.
6. Depending on the patient's susceptibility, the number of microorganisms required to cause infection varies [22]. In other words, 2 patients can be exposed to the same number of pathogens but, depending on their immune status, the risk of developing an HAI might be different for them. This is another potential for future work to explore.
7. Although the patient exposure risk is introduced and tested in a rehab setting in this study, it can be easily extended to other environments and hospital settings. The integrated risk factors in the patient exposure risk formula can be customized depending on the hospital setting, the type of room, and the contagion in the room. By incorporating activity recognition into the system and accounting for patient's susceptibility, we will be able to calculate the patient exposure risk tailored to each unit or hospital in the future. Currently, we are collecting data to test the patient exposure risk concept in different hospital environments.
8. Finally, in the future, we will investigate how to provide a standard rating system on patient exposure risk to enable the public to better understand the risk of infection exposure.

Several studies have demonstrated that performance feedback is effective in improving HH adherence [30-32]. Maintaining a positive culture is critical for sustaining improvements [33]. Criticizing staff by concentrating on the deviation of their compliance rate from a desired rate may result in a negative disciplinary tone. Infection prevention among patients is considered to be the most important reason among HCWs to perform HH [34]. We believe that by introducing patient exposure risk and reinforcing the need to decrease the risk of infection for both patients and HCWs [35], we will see further improvements in HH performance.

Conclusion

Measuring the risk of HAIs for patients is an essential step for devising effective interventions for infection control. Controlling infection in health care settings requires continuous monitoring of HCWs' handwashing behavior. HH behavior should be studied as a series of linked events in the chain of infection transmission. While conventional performance measures consider HH opportunities as binary events, the patient exposure risk enables us to evaluate the risk of missed HH opportunities based on time and location. As supported by different examples, the patient exposure risk helps predict the likelihood of patients becoming infected with HAIs. Future work will focus on providing a better estimation of the risk of contracting HAIs for patients by including additional factors such as activities performed by the staff in the room, estimated using sensors and AI. The same concept will be extended to estimate the HCWs' exposure to infection risk.

Acknowledgments

We thank Pamela Holliday and Steven Pong for collecting the data used in this study. Grants for this study were provided by the Canadian Institutes of Health Research Foundation (FDN-148450).

Conflicts of Interest

GF receives support as the Creaghan Family Chair in Prevention and Healthcare Technologies. GF is also a founder of a company that has commercialized an electronic HH system to bring the results of research to market. It is possible that the concept described in this publication may be incorporated in future products after development and validation is completed. ARF and KH declare no conflict of interest.

References

1. Klevens RM, Edwards JR, Richards CL, Horan TC, Gaynes RP, Pollock DA, et al. Estimating health care-associated infections and deaths in U.S. hospitals, 2002. Public Health Rep 2007;122(2):160-166 [FREE Full text] [Medline: 17357358]
2. Stone PW. Economic burden of healthcare-associated infections: an American perspective. Expert Rev Pharmacoecon Outcomes Res 2009 Oct;9(5):417-422 [FREE Full text] [doi: 10.1586/erp.09.53] [Medline: 19817525]
3. Measuring Patient Harm in Canadian Hospitals. With What can be done to improve patient safety? Canadian Institute for Health Information. 2016. URL: https://secure.cihi.ca/free_products/cihi_cpsi_hospital_harm_en.pdf [accessed 2021-12-08]
4. Mitchell R, Taylor G, Rudnick W, Alexandre S, Bush K, Forrester L, Canadian Nosocomial Infection Surveillance Program. Trends in health care-associated infections in acute care hospitals in Canada: an analysis of repeated point-prevalence surveys. CMAJ 2019 Sep 09;191(36):E981-E988 [FREE Full text] [doi: 10.1503/cmaj.190361] [Medline: 31501180]

5. Boyce JM, Pittet D, Healthcare Infection Control Practices Advisory Committee, HICPAC/SHEA/APIC/IDSA Hand Hygiene Task Force. Guideline for Hand Hygiene in Health-Care Settings. Recommendations of the Healthcare Infection Control Practices Advisory Committee and the HICPAC/SHEA/APIC/IDSA Hand Hygiene Task Force. Society for Healthcare Epidemiology of America/Association for Professionals in Infection Control/Infectious Diseases Society of America. MMWR Recomm Rep 2002 Oct 25;51(RR-16):1-45, quiz CE1 [FREE Full text] [Medline: 12418624]
6. Allegranzi B, Pittet D. Role of hand hygiene in healthcare-associated infection prevention. J Hosp Infect 2009 Dec;73(4):305-315. [doi: 10.1016/j.jhin.2009.04.019] [Medline: 19720430]
7. Golin AP, Choi D, Ghahary A. Hand sanitizers: A review of ingredients, mechanisms of action, modes of delivery, and efficacy against coronaviruses. Am J Infect Control 2020 Sep;48(9):1062-1067 [FREE Full text] [doi: 10.1016/j.ajic.2020.06.182] [Medline: 32565272]
8. Leslie RA, Zhou SS, Macinga DR. Inactivation of SARS-CoV-2 by commercially available alcohol-based hand sanitizers. Am J Infect Control 2021 Mar;49(3):401-402 [FREE Full text] [doi: 10.1016/j.ajic.2020.08.020] [Medline: 32818578]
9. Kratzel A, Todt D, V'kovski P, Steiner S, Gultom M, Thao TTN, et al. Inactivation of Severe Acute Respiratory Syndrome Coronavirus 2 by WHO-Recommended Hand Rub Formulations and Alcohols. Emerg Infect Dis 2020 Jul;26(7):1592-1595 [FREE Full text] [doi: 10.3201/eid2607.200915] [Medline: 32284092]
10. Infection prevention and control during health care when novel coronavirus (nCoV) infection is suspected. World Health Organization. URL: https://www.who.int/publications/i/item/10665-331495 [accessed 2021-12-08]
11. Sax H, Allegranzi B, Uçkay I, Larson E, Boyce J, Pittet D. 'My five moments for hand hygiene': a user-centred design approach to understand, train, monitor and report hand hygiene. J Hosp Infect 2007 Sep;67(1):9-21. [doi: 10.1016/j.jhin.2007.06.004] [Medline: 17719685]
12. Your 4 Moments for Hand Hygiene. Public Health Ontario. 2008. URL: https://www.publichealthontario.ca/-/media/documents/9/2008/4-moments-poster-series.pdf?la=en [accessed 2021-07-23]
13. WHO Guidelines on Hand Hygiene in Health Care First Global Patient Safety Challenge Clean Care Is Safer Care. Geneva: World Health Organization; 2009.
14. Chen LF, Carriker C, Staheli R, Isaacs P, Elliott B, Miller BA, et al. Observing and improving hand hygiene compliance: implementation and refinement of an electronic-assisted direct-observer hand hygiene audit program. Infect Control Hosp Epidemiol 2013 Feb;34(2):207-210 [FREE Full text] [doi: 10.1086/669084] [Medline: 23295569]
15. McLaws M, Kwok YLA. Hand hygiene compliance rates: Fact or fiction? Am J Infect Control 2018 Aug;46(8):876-880. [doi: 10.1016/j.ajic.2018.03.030] [Medline: 29778435]
16. Benudis A, Stone S, Sait AS, Mahoney I, Price LL, Moreno-Koehler A, et al. Pitfalls and Unexpected Benefits of an Electronic Hand Hygiene Monitoring System. Am J Infect Control 2019 Sep;47(9):1102-1106. [doi: 10.1016/j.ajic.2019.03.011] [Medline: 31005345]
17. Boyce JM, Cooper T, Yin J, Li F, Arbogast JW. Challenges encountered and lessons learned during a trial of an electronic hand hygiene monitoring system. Am J Infect Control 2019 Dec;47(12):1443-1448. [doi: 10.1016/j.ajic.2019.05.019] [Medline: 31324492]
18. Iversen A, Kavalaris CP, Hansen R, Hansen MB, Alexander R, Kostadinov K, et al. Clinical experiences with a new system for automated hand hygiene monitoring: A prospective observational study. Am J Infect Control 2020 May;48(5):527-533. [doi: 10.1016/j.ajic.2019.09.003] [Medline: 31635879]
19. Doll ME, Masroor N, Cooper K, Trimmer T, Pryor R, Auricchio J, et al. A comparison of the accuracy of two electronic hand hygiene monitoring systems. Infect Control Hosp Epidemiol 2019 Oct;40(10):1194-1197. [doi: 10.1017/ice.2019.209] [Medline: 31407648]
20. Karimpour N, Karaduman B, Ural A, Challengerl M, Dagdeviren O. IoT based hand hygiene compliance monitoring. 2019 Nov 21 Presented at: Int Symp Networks, Comput Commun ISNCC. IEEE; 2019; Istanbul, Turkey p. 1-6.
21. Wang C, Sarsenbayeva Z, Chen X, Dingler T, Goncalves J, Kostakos V. Accurate Measurement of Handwash Quality Using Sensor Armbands: Instrument Validation Study. JMIR Mhealth Uhealth 2020 Mar 26;8(3):e17001 [FREE Full text] [doi: 10.2196/17001] [Medline: 32213469]
22. Damani N. Manual of Infection Prevention and Control. New York, USA: Oxford University Press; 2019.
23. Pong S, Holliday P, Fernie G. Secondary measures of hand hygiene performance in health care available with continuous electronic monitoring of individuals. Am J Infect Control 2019 Jan;47(1):38-44 [FREE Full text] [doi: 10.1016/j.ajic.2018.07.004] [Medline: 30195406]
24. Introducing the Buddy Badge System. Hygienic Echo Inc. URL: https://hygienicecho.com/ [accessed 2021-07-23]
25. Zar J. Spearman Rank Correlation. Encyclopedia of Biostatistics. Chichester, UK: John Wiley & Sons, Ltd; 2005.
26. Pittet D, Dharan S, Touveneau S, Sauvan V, Perneger TV. Bacterial contamination of the hands of hospital staff during routine patient care. Arch Intern Med 1999 Apr 26;159(8):821-826. [doi: 10.1001/archinte.159.8.821] [Medline: 10219927]
27. Clack L, Scotoni M, Wolfensberger A, Sax H. "First-person view" of pathogen transmission and hand hygiene - use of a new head-mounted video capture and coding tool. Antimicrob Resist Infect Control 2017;6:108 [FREE Full text] [doi: 10.1186/s13756-017-0267-z] [Medline: 29093812]
28. Grassly NC, Fraser C. Mathematical models of infectious disease transmission. Nat Rev Microbiol 2008 Jun;6(6):477-487 [FREE Full text] [doi: 10.1038/nrmicro1845] [Medline: 18533288]

29. Public health management of cases and contacts associated with COVID-19. Government of Canada. URL: https://tinyurl.com/4p8pawu6 [accessed 2021-07-23]
30. Ofek Shlomai N, Rao S, Patole S. Efficacy of interventions to improve hand hygiene compliance in neonatal units: a systematic review and meta-analysis. Eur J Clin Microbiol Infect Dis 2015 May;34(5):887-897. [doi: 10.1007/s10096-015-2313-1] [Medline: 25652605]
31. Naikoba S, Hayward A. The effectiveness of interventions aimed at increasing handwashing in healthcare workers - a systematic review. J Hosp Infect 2001 Mar;47(3):173-180. [doi: 10.1053/jhin.2000.0882] [Medline: 11247676]
32. Gould DJ, Moralejo D, Drey N, Chudleigh JH, Taljaard M. Interventions to improve hand hygiene compliance in patient care. Cochrane Database Syst Rev 2017 Sep 01;9:CD005186 [FREE Full text] [doi: 10.1002/14651858.CD005186.pub4] [Medline: 28862335]
33. Cumbler E, Castillo L, Satorie L, Ford D, Hagman J, Hodge T, et al. Culture change in infection control: applying psychological principles to improve hand hygiene. J Nurs Care Qual 2013;28(4):304-311. [doi: 10.1097/NCQ.0b013e31829786be] [Medline: 23669615]
34. Larson E, Killien M. Factors influencing handwashing behavior of patient care personnel. Am J Infect Control 1982 Aug;10(3):93-99. [doi: 10.1016/0196-6553(82)90019-0] [Medline: 6922685]
35. McDiarmid M. Advocating for the Health Worker. Ann Glob Health 2019 Jan 28;85(1):16 [FREE Full text] [doi: 10.5334/aogh.2461] [Medline: 30741511]

Abbreviations

ECR: entrance compliance rate
EMS: electronic monitoring system
HAI: health care-acquired infection
HCW: health care worker
HH: hand hygiene
MOHLTC: Ministry of Health and Long-Term Care

Edited by G Eysenbach; submitted 27.07.21; peer-reviewed by R Bajpai, S Nagavally; comments to author 18.08.21; revised version received 12.10.21; accepted 17.11.21; published 02.02.22.

Please cite as:
Hadian K, Fernie G, Roshan Fekr A
A New Performance Metric to Estimate the Risk of Exposure to Infection in a Health Care Setting: Descriptive Study
JMIR Form Res 2022;6(2):e32384
URL: https://formative.jmir.org/2022/2/e32384
doi:10.2196/32384
PMID:35107424

Original Paper

Intensive Longitudinal Data Collection Using Microinteraction Ecological Momentary Assessment: Pilot and Preliminary Results

Aditya Ponnada[1,2], BDes, PhD; Shirlene Wang[3], BA; Daniel Chu[3], PhD; Bridgette Do[3], MPH; Genevieve Dunton[3], MPH, PhD; Stephen Intille[1,2], PhD

[1]Khoury College of Computer Sciences, Northeastern University, Boston, MA, United States

[2]Bouve College of Health Sciences, Northeastern University, Boston, MA, United States

[3]Department of Population and Public Health Sciences, Keck School of Medicine, University of Southern California, Los Angeles, CA, United States

Corresponding Author:
Aditya Ponnada, BDes, PhD
Khoury College of Computer Sciences
Northeastern University
360 Huntington Avenue
Boston, MA, 02130
United States
Phone: 1 6173061610
Email: ponnada.a@northeastern.edu

Abstract

Background: Ecological momentary assessment (EMA) uses mobile technology to enable in situ self-report data collection on behaviors and states. In a typical EMA study, participants are prompted several times a day to answer sets of multiple-choice questions. Although the repeated nature of EMA reduces recall bias, it may induce participation burden. There is a need to explore complementary approaches to collecting in situ self-report data that are less burdensome yet provide comprehensive information on an individual's behaviors and states. A new approach, microinteraction EMA (µEMA), restricts EMA items to *single, cognitively simple questions* answered on a smartwatch with single-tap assessments using a quick, glanceable microinteraction. However, the viability of using µEMA to capture behaviors and states in a large-scale longitudinal study has not yet been demonstrated.

Objective: This paper describes the µEMA protocol currently used in the Temporal Influences on Movement & Exercise (TIME) Study conducted with young adults, the interface of the µEMA app used to gather self-report responses on a smartwatch, qualitative feedback from participants after a pilot study of the µEMA app, changes made to the main TIME Study µEMA protocol and app based on the pilot feedback, and preliminary µEMA results from a subset of active participants in the TIME Study.

Methods: The TIME Study involves data collection on behaviors and states from 246 individuals; measurements include passive sensing from a smartwatch and smartphone and intensive smartphone-based hourly EMA, with 4-day EMA bursts every 2 weeks. Every day, participants also answer a nightly EMA survey. On non–EMA burst days, participants answer µEMA questions on the smartwatch, assessing momentary states such as physical activity, sedentary behavior, and affect. At the end of the study, participants describe their experience with EMA and µEMA in a semistructured interview. A pilot study was used to test and refine the µEMA protocol before the main study.

Results: Changes made to the µEMA study protocol based on pilot feedback included adjusting the *single-question* selection method and smartwatch vibrotactile prompting. We also added sensor-triggered questions for physical activity and sedentary behavior. As of June 2021, a total of 81 participants had completed at least 6 months of data collection in the main study. For 662,397 µEMA questions delivered, the compliance rate was 67.6% (SD 24.4%) and the completion rate was 79% (SD 22.2%).

Conclusions: The TIME Study provides opportunities to explore a novel approach for collecting *temporally dense* intensive longitudinal self-report data in a sustainable manner. Data suggest that µEMA may be valuable for understanding behaviors and states at the individual level, thus possibly supporting future longitudinal interventions that require within-day, temporally dense self-report data as people go about their lives.

(JMIR Form Res 2022;6(2):e32772) doi:10.2196/32772

KEYWORDS

intensive longitudinal data; ecological momentary assessment; experience sampling; microinteractions; smartwatch; health behavior research; mobile phone

Introduction

Background

Mobile technologies create new opportunities to design personalized health interventions in 2 broad ways. First, mobile technologies such as smartphones and smartwatches can be used to improve the assessment of behavioral, psychological, and contextual states, reducing reliance on known-to-be-problematic retrospective recall surveys that may poorly capture within-day variations [1-3]. Second, real-time measures gathered using mobile technologies can be used to design just-in-time adaptive interventions [4,5] that are sensitive to everyday changes in behaviors and states. State or behavior assessment and personalized interventions both require computational models that can represent the interrelationships between different behaviors and states unique to individuals. Such models could be created by using (1) a hypothetico-deductive approach to study relationships between predictors and outcomes of interest based on prior knowledge [6,7], (2) data-driven discovery from a large number of constructs measured in intensive longitudinal data (ILD) studies [8,9], or (3) an appropriate combination of both [10,11]. Creating models of behavior that capture relationships between behavior and context that can happen quickly and many times in a day requires methods for sustainable data gathering on, and modeling of, within-day changes of variables (eg, physical activity [PA], sleep, sedentary behavior [SB], and affect) [12,13].

Ideally, health-related constructs could be measured using passive sensors from easy-to-wear devices, without requiring any end-user effort other than wearing the devices and charging them. However, *we currently need self-reportinput* to measure many subjective experiences (eg, perceived pain, hunger, and fatigue) that play an important role in predicting health behavior changes [14]. A popular practice for obtaining self-report data on mobile devices is using experience sampling methods [15], also known as ecological momentary assessment (EMA) [16], to gather participant perspectives on in situ experiences that sensors cannot measure directly. As smartphones have become ubiquitous, EMA has become affordable to deploy with most adult populations [17].

In the most common EMA protocols used to collect ILD, participants' smartphones prompt with audio or vibration several times a day, each time presenting a set of questions that may take a few seconds or minutes to answer, depending upon the number and complexity of the questions. Typically, questions are multiple choice, but sometimes they involve open-ended responses or sliding scales. The sampling frequency in EMA varies widely based on study goals, as does potential burden. Burden results from the annoyance of each EMA interruption and the time and mental effort required to answer the questions presented at the prompt [18,19]. To minimize response burden, researchers often reduce the number of questions in each question set, the complexity of the questions, the frequency of prompting, or all three, thereby limiting the scope and temporal density of the behaviors and states measured in an ILD study. Although intensive EMA protocols can be sustained for short-term studies (eg, 7-30 days), burden will accumulate in ILD collection studies that last several months or years [20]. This accumulated burden, if not managed, could prevent the conducting of large-scale longitudinal studies or interventions that rely on frequent self-report input.

Although compliance-contingent compensation is a common practice to boost or maintain EMA response rates, it can be expensive for large-scale longitudinal studies involving thousands of participants. Beyond observational studies, future personalized health interventions that may require self-report of participant states and contexts must balance information gathering with burden to ensure that long-term engagement does not wane. For instance, an intervention to reduce stress and increase PA may need to query users regularly about their perceived stress and physical exertion (among other subjective experiences) to adjust the intervention in real time. For such an intervention to be sustainable, the system may have to gather these data frequently at the rate at which stress and activity can change—that is, many times throughout the day—while keeping end-user burden manageable. For such an intervention to be affordable, financially compensating participants to motivate a reasonable response rate is not a viable long-term option. In fact, compliance-contingent compensation may result in poor data quality if the participants are only motivated by money [21]. Thus, new strategies are needed for acquiring temporally dense self-report on multiple constructs with sustainable burden, even for studies or interventions deployed for months or years.

A strategy for acquiring longitudinal self-report data is using microinteraction EMA (µEMA or micro-EMA) [22]. *Microinteractions* are actions lasting for 3-4 seconds (eg, checking the time on a watch or turning on a lamp); microinteractions are so short lived that they can be completed without disrupting ongoing activity [23]. µEMA is a type of EMA where all prompts are for *single questions* with *single-tap answers* (eg, *Nervous right now?* with answers *Yes*, *Sort of*, or *No*) [22]. Unlike EMA, where questions are prompted in a set, in µEMA a device prompts only a single question per interruption and there is a guarantee that the question can be answered with a microinteraction [23]. Because of this microinteraction guarantee, µEMA may enable self-report data collection at a much higher frequency than EMA. This at-a-glance response property is achieved in 2 ways. First, µEMA is deployed on wearable devices such as smartwatches that permit quick access to question content; unlike mobile phones, the questions can be seen with a flip of the wrist without additional time required to access a mobile phone that may be in a bag or out of sight [24,25]. Prior studies have demonstrated that smartwatches are more suitable for glanceable microinteractions (eg, checking notifications) than smartphones [26,27]. Second, µEMA questions are intended to be cognitively simple to answer (eg, *Feeling stressed?—Yes*, *Sort of*, or *No*)

and with a limited answer set that fits on a small smartwatch screen and therefore does not require a scrolling interface. Prior work shows that reducing the number of answers (eg, from a 5-point Likert scale to a 3-point ordinal or nominal scale) improves response time without necessarily changing the perceived item difficulty [28]. The benefit of leveraging microinteractions is also supported by 2 laws in user interface (UI) design. First, Hick's law posits that the more options there are to choose from, the longer the response time on the UI will be [29,30]. Second, Fitt's law shows that the navigation time between 2 targets on a UI is directly proportional to the distance between the targets and inversely proportional to the size of the target [31,32]. However, the downside of restricting all prompted interactions to cognitively simple, glanceable, single-question microinteractions is that information obtained from a single µEMA question is more limited than what can be obtained from unconstrained EMA questions with multiple answer options (eg, *In the past one hour, how stressed did you feel?—Extremely*, *Quite a bit*, *Moderately*, *A little*, or *Not at all*). Presenting such EMA questions on a smartwatch screen would make smartwatch interaction cumbersome, requiring either a font size that would be too difficult for most people to read, especially for those who need reading glasses, or requiring scrolling. Furthermore, using self-report to capture feelings aggregated over longer time windows (eg, in the past hour or over the past day) introduces recall burden [18,33] and cognitive complexity. Therefore, EMA questions nearly always require adjustment to achieve the microinteraction property we seek.

Prior work has used wearable devices such as smartwatches and heads-up displays (eg, Google Glass) to deploy EMA question sets with the goal of making EMA easier. However, typically, such work has directly adopted surveys (with back-to-back questions) from mobile phone–based EMA surveys that require users to engage in prolonged interactions on the devices (eg, sliding or scrolling) [34-37] and thus do not result in microinteractions. Converting an EMA survey with multiple questions with multiple answers per question to a small smartwatch display has indeed required additional scrolling to make the questions and answers readable, thereby likely slowing down—not speeding up—answer selection and resulting in the smartwatch interaction becoming more burdensome than answering the same surveys on the mobile phone.

In 2 prior 4-week between-subject pilot studies, we found that despite experiencing approximately 4-6 times more interruption when using µEMA instead of mobile phone–based EMA, participants using µEMA on a smartwatch had significantly higher compliance, reported lower perceived burden, and answered with faster response times than participants when measuring the same constructs using EMA [22,38]. Prior work also demonstrates that for a single construct measured at high frequency, µEMA can yield good criterion validity when compared against a research-grade passive sensor (eg, in the domain of PA measurement) [39]. Table 1 summarizes key differences between the mobile phone–based EMA and smartwatch-based µEMA in a prior study [22] where both EMA and µEMA were used to measure positive and negative affect (using the Positive and Negative Affect Schedule [40] over a period of 4 weeks [41]).

With µEMA, a user is guaranteed that the worst-case interaction required to answer a prompt will always be a single, glance-and-tap microinteraction; thus, it is nearly as easy to answer a microinteraction prompt as it is to manually dismiss it (by swiping on the screen) or ignore it. If researchers develop simple questions appropriate for microinteractions that provide valuable information about constructs of interest, µEMA may support ILD studies that gather dense, within-day information on behavior. So far, researchers have used µEMA to measure only 1-2 constructs per study, such as hyperarousal [42], stress [43], and subjective comfort [44]; µEMA has also been used with situated displays in home settings [45,46]. µEMA is well suited for constructs that may require frequent self-report, such as chronic pain assessment [47]. µEMA can be assessed at such a high temporal density (ie, 4 times per hour) that it can also measure multiple constructs in a single day while maintaining reasonable temporal density of each construct.

Despite the promise of µEMA, much remains to be explored to determine the viability of the technique. In this paper, we present the µEMA protocol for the Temporal Influences on Movement & Exercise (TIME) Study. The TIME Study has a primary protocol that uses EMA, which is presented elsewhere (Wang, S, unpublished data, January 2022). Here, we present the protocol for a secondary, exploratory study within the TIME Study on the viability of using µEMA in an ILD study. The TIME µEMA substudy is the first large-scale ILD study (duration: 1 year) examining how µEMA might be used to measure multiple health behaviors such as PA, SB, and sleep along with time-varying subjective states (eg, stress, fatigue, and happiness). First, we describe the overall goal of the TIME µEMA substudy. Next, we detail the µEMA app we designed for the study, followed by reporting the qualitative feedback received from our pilot study participants. Subsequently, we present details of changes made to the protocol in response to pilot testing. Finally, we present preliminary compliance results from a subset of participants in the main study (81/246, 32.9%).

Table 1. Differences between mobile phone–based ecological momentary assessment (EMA) and smartwatch-based micro-EMA (µEMA; from prior work [22]).

	Mobile phone–based EMA	Smartwatch-based µEMA
Prompts per day	≤7	≤48
Prompting frequency	Once in 2 hours	Four times an hour
Number of questions per prompt	≤6	Only 1 question
Example of question framing	Over the past hour, how stressed did you feel?	Feeling stressed?
Number of response options per question	≥4 (including *choose all that apply* responses)	≤3 (eg, *Yes*, *Sort of*, and *No*)
Interruption burden	High (must access mobile phone, unlock it, and then start to answer)	Low (smartwatch always accessible with glance)
Response burden	High, because of multiple questions with more answer options	Low, because of only 1 question with limited answer options; aim is cognitive simplicity

TIME µEMA Substudy Objectives

The goal of the TIME Study is to examine daily and within-day microtemporal processes (eg, feeling stressed, increased workload, being with family, and being at home) that may influence PA, SB, and sleep in young adults [48]. EMA is being deployed to capture reflective processes, reactive processes, internal factors, and external factors and assess how they affect health behavior adoption and maintenance (Wang, S, unpublished data, January 2022). Reflective processes are those that are slow and require careful deliberation (eg, intention to engage in healthy behaviors) [49]. In contrast, reactive processes are quick and automatic (eg, being on a regular routine) [49]. Internal factors are physiological and emotional sensations that originate internally (eg, positive and negative affect, pain, and fatigue). External factors are social, situational, and physical settings or events that originate externally to the individual (eg, an increase in workload or meeting a friend). An exploratory aim of the TIME Study is to assess the feasibility of using µEMA to gather similar health behavior data from participants. The fundamental difference between using EMA and µEMA for ILD can be summarized as follows. Traditional EMA is a method that interrupts less often than µEMA but asks for substantially more information with each interruption. In contrast, µEMA is a method that interrupts more often than EMA but asks only 1 simple question with each interruption. The TIME µEMA substudy will permit exploration of what we can learn about health behavior from small amounts of information gathered at high frequency using µEMA; the TIME substudy data will allow investigators to study the following research questions:

1. How sustainable is µEMA for ILD studies compared with mobile phone–based EMA?
   Interrupting participants 4 times more often with a single, simple µEMA question could be perceived as far more burdensome than interrupting less often with a longer, more complex EMA survey. Can participants in a full-year ILD collection study sustain µEMA with high compliance compared with EMA?
2. How do contextual factors influence µEMA compliance?
   Prior EMA literature has shown that contextual factors such as time of day, day of week, and location influence whether the participant is able to complete the EMA survey [50]. Behavior, such as being active, could also influence response rates. The TIME µEMA substudy will permit exploration of the contextual factors that may influence µEMA compliance and comparison of such effects between µEMA and EMA.
3. Can intermittent µEMA questions provide information on an individual's overall behavior and state?
   Sustaining intensive mobile phone–based EMA longitudinally may not be realistic, requiring long temporal gaps between measurements and more retrospective recall. The TIME µEMA substudy will permit exploration of whether using µEMA *between* EMA burst periods could provide information on the diurnal patterns of behaviors and states not captured in the EMA bursts alone.
4. How can we use µEMA data *together with* data acquired from the passive sensors and EMA to design predictive models of health behavior change and maintenance?
   Each instance of µEMA self-report provides information on a single variable at a particular time, but µEMA prompts will be temporally dense, measuring different behaviors and states throughout the day. The TIME µEMA substudy will permit exploration of whether the amount and quality of µEMA data acquired in the TIME µEMA substudy is sufficient to build ideographic predictive models of behavior, either alone or in combination with EMA and passive sensing data.

In the remainder of the paper, we describe the protocol for the TIME µEMA substudy, from which ILD will be derived that will support such work.

µEMA Design Overview

Deploying µEMA requires designing a specific smartwatch-based graphical UI and the question scheduling strategy.

Custom µEMA App Interface

We developed a custom µEMA app (called the TIME app or *app*) that runs on a smartwatch (Android Wear OS version 2.0) with a linked app on the paired smartphone (Android OS version ≥7.0). The Android OS was chosen because it provides flexibility in gathering raw sensor data from both smartphone and smartwatch continuously in the background. The apps work together to present µEMA questions on the smartwatch, collect

and process passive sensor data, and transfer data to a research server each day. Each µEMA prompt presents only 1 question at a time that can be answered in a glance and a tap (a microinteraction). Example questions are *Feeling stressed?* and *Feeling productive today?* with three answer options: *Yes*, *Sort of*, and *No*. The smartwatch prompts the participant with a vibration pattern lasting for 3 seconds. The question screen displays at the start of the vibration (Figure 1). The brightness of the screen is determined automatically by the smartwatch based on ambient brightness or the smartwatch settings selected by the participant.

The question stays visible on the screen for 20 seconds, after which the question disappears and a missed response is recorded. If participants answer a question, they are presented with an acknowledgment screen with a short thank-you message and an *Undo?* button (Figure 1). The thank-you messages on the acknowledgment screen are selected from >250 unique variations designed to reduce repetition and boost engagement—they are all pithy, and some are quirky. The undo option is available for 3 seconds. If participants tap on the undo button, they are taken back to the question screen and they can then change their answer within 20 seconds. If they do not undo within 3 seconds, the app records the final answer and the question disappears. Participants have only 1 chance to undo an answer. If the app is dismissed (by swiping right on the screen) when the original question appears, it is recorded as *never started*, and if the app is dismissed on the undo screen, it is recorded as *completed then dismissed*. Similarly, if the user does not answer the question again after selecting the undo button, then the prompt status is recorded as *partially completed.*

Figure 1. (Left) A microinteraction ecological momentary assessment question screen with 3 answer options, which will display for 20 seconds or until an answer is selected. (Right) A thank-you screen with the Undo? button indicating that the user selected the Sort of answer; the screen displays for 3 seconds with a countdown timer before it disappears.

µEMA Prompt Scheduling

µEMA prompts occur during waking hours, including when the smartwatch is being charged or is off the wrist. Participants self-report their upcoming sleep time and the next day's wake time using EMA on their smartphone (Figure 2); this information is used to compute the waking hours and is sent to the smartwatch in real time. If participants do not respond to the sleep- and wake-time questions on the smartphone or if the smartphone cannot connect to the smartwatch for any reason, the smartwatch uses the previous day's wake and sleep times as the default waking schedule to determine the µEMA prompting schedule. The purpose of the sleep-wake questions is to predetermine the prompting schedule for the day. Once answered, these sleep- and wake-time questions are presented again after 10 hours, in case participants want to change their schedule. Moreover, participants can use the app to manually change the sleep-wake times at any time of the day. µEMA prompting starts 15 minutes after each day's wake time and ends 15 minutes before that day's sleep time. For instance, if the participant plans to sleep at 11 PM and wake up the next day at 6 AM, then µEMA prompting will stop at 10:45 PM and resume at 6:15 AM the next day. Thus, days with longer waking hours will result in more µEMA prompts than days with shorter waking hours. µEMA questions are prompted 4 times an hour at random, with at least eight minutes guaranteed between 2 consecutive prompts using the following formula. The app generates this schedule for the 24-hour period using the following equations and then only prompts during waking hours.

$$P_n = \square \in [0, MaxTimeAvailable_n) + 8 + P_{n-1}$$

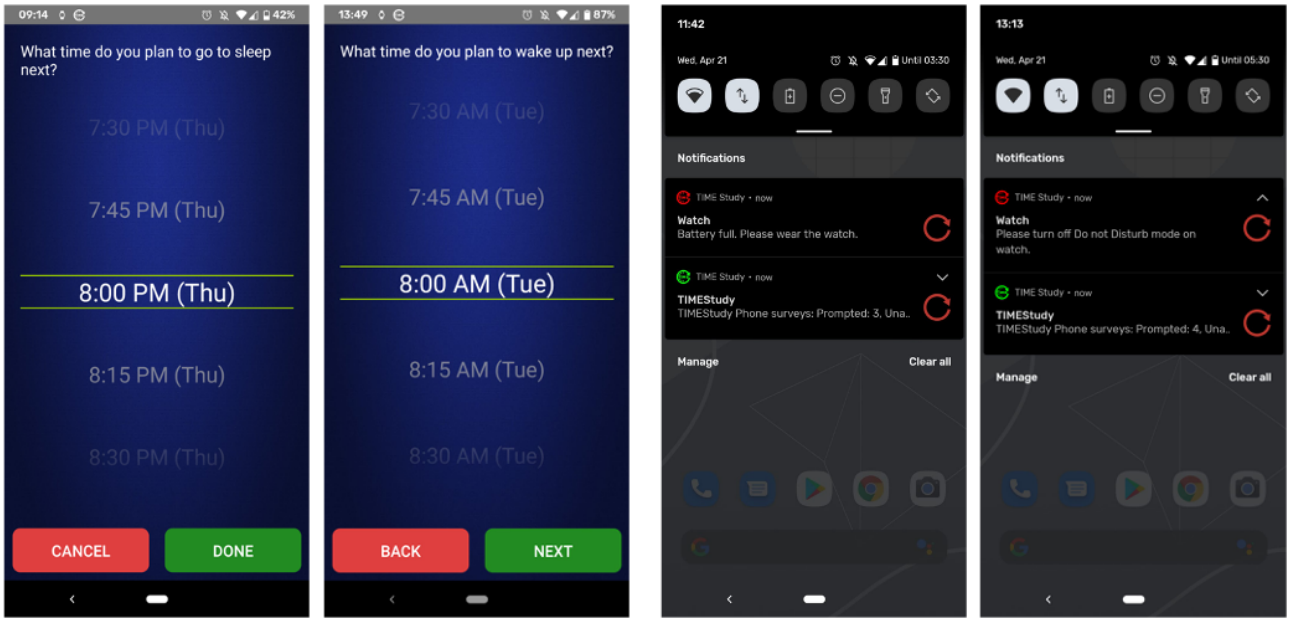
$$MaxTimeAvailable_n = 55 - ((4 - n + 1) \times 8) - P_{n-1}$$

First, for each hour (eg, 8 AM to 9 AM), the app computes the maximum remaining time ($MaxTimeAvailable_n$) in minutes to schedule a prompt (P_n, where $n=1,...4$, $P_0=0$) by subtracting the minimum required time gap (ie, 8-minute gap per prompt) and the time elapsed since the previous prompt within that hour (P_{n-1}). Then, the app selects a random time from the $MaxTimeAvailable_n$ and adds the 8-minute gap and the prompt time of the previous prompt (P_{n-1}) to determine the current prompt time (P_n). The µEMA smartwatch app respects the settings on the smartwatch; therefore, prompting is paused when a participant turns on the do-not-disturb (DND) mode on the smartwatch. However, if the smartwatch is in DND mode for >60 consecutive minutes, participants receive a notification on the mobile phone to disable the DND mode on the smartwatch so that prompting can be resumed without further data loss (Figure 2). If the smartwatch is off the wrist, including when it is being charged, µEMA prompting continues normally.

Participants receive a notification on the mobile phone and smartwatch to (1) connect the smartwatch if it is disconnected from the mobile phone (through Bluetooth) for >60 consecutive minutes, (2) wear the smartwatch if the system detects that the smartwatch is fully charged but on the charger, (3) charge the smartwatch when the smartwatch battery reaches ≤15% capacity (Figure 2), and (4) update the software when the mobile phone

or smartwatch software requires an update. Mobile phone and smartwatch notifications disappear as soon as the problem that was flagged is resolved. In addition, the mobile phone app presents a persistent notification indicating that the TIME app is running in the background collecting data and highlighting the countdown of when the next burst period on the mobile phone will start.

Figure 2. (Left) A mobile phone survey asking for prospective sleep and wake times. These times are used to adjust the prompt scheduling. (Right) Example notifications: (1) a reminder to wear the smartwatch prompted when it is still on the charger and is 100% charged and (2) a reminder to turn off the do-not-disturb mode on the smartwatch, which appears if that mode is active for >60 minutes consecutively.



System Components

The system has three components: mobile phone, smartwatch, and a remote server. The mobile phone prompts users with EMA surveys and end-of-day surveys. In addition, the mobile phone passively collects sensor data such as acceleration, location, and mobile phone use (eg, number of apps used). The details of mobile phone–related functionality are beyond the scope of this paper and are reported elsewhere (Wang, S, unpublished data, January 2022). The smartwatch prompts µEMA questions and passively collects raw accelerometer data (described in the next section). The data from the smartwatch are sent to the mobile phone either (1) once every 4 hours or (2) whenever the smartwatch is on the charger. The mobile phone then encrypts the data and sends it to a remote server daily. When the smartwatch or mobile phone are not transferring data automatically as expected, they each provide an option that permits participants to force the data transfer to the mobile phone; this functionality is used to resolve data transfer problems working remotely with the research staff.

Methods

Pilot Study

With the µEMA system described in the previous sections, we conducted a pilot study with 15 participants who answered µEMA questions for up to 3 consecutive weeks.

Pilot Study Participants

Participants for the pilot study were recruited by means of flyers posted on university campuses in the Los Angeles metropolitan area and through targeted ads on Facebook. Participants were eligible to participate if they (1) were aged 18-24 years, (2) owned an Android smartphone with OS version ≥6.0, (3) were fluent in English, 4) were not planning to change their smartphone in the next 1 month, (4) were willing and able to wear a smartwatch and answer questions on their smartphone for a period of 1 month, and (5) were currently engaged in recommended levels of PA or intending to do so in the next 12 months.

Pilot Study Design

The research staff met with the eligible participants in person (in 2019) and installed the TIME app on their smartphones. The app prompted participants with EMA surveys in 2 measurement burst periods. Each burst period lasted 4 days (called burst days). If participants completed at least eight EMA surveys per day in the first burst period, they were loaned a Fossil Gen 4 smartwatch (Fossil Group, Inc) and instructed on how to install the TIME smartwatch app. For the remaining 3 weeks, participants also collected data on the smartwatch. On the nonburst days, the smartwatch prompted with µEMA questions about a myriad of behaviors, including positive and negative affect, location, and PA. With each prompt, the smartwatch TIME app selected a question at random and presented it on the screen. In addition, questions about momentary PA, SB, and location (ie, *Physically active now?*, *Sedentary right now?*, and *At home right now?* with answer options *Yes*, *Sortof*, and *No*) were also selected randomly throughout the day. Thus, it was possible for a question to repeat consecutively when sampled at random. The goal of the pilot study is to assess the app's performance, improve the study protocol, and fine-tune question wording and the overall user experience. At the end of the pilot study, participants took part in a structured interview

about their experience of answering questions on the smartphone and smartwatch.

Pilot Study µEMA Response Behavior

A total of 15 participants completed 1 month of data collection (3 weeks with µEMA prompting on the smartwatch). With 15,120 µEMA questions delivered, the µEMA response rate was 76.4% (SD 22.3%), with a mean µEMA response time of 4.3 (SD 1.0) seconds.

Pilot Study Qualitative Feedback

At the end of the 1-month pilot study, participants completed a structured interview lasting for approximately 30 minutes to enable us to gather feedback on how to improve the user experience with the TIME app and to adjust the main study protocol accordingly. When asked about positive experiences with µEMA, participants highlighted the ease of access of the smartwatch and the simplicity of answering just 1 question. A participant described the experience as follows:

> *Honestly, it was the easiest to answer surveys on the watch because it's right on your wrist and it vibrates [and] there is a survey, I answer the question and I am done...The watch questions are actually less burdensome than the phone questions. Watch questions are pretty convenient because you can just touch [tap] on the watch and you're done basically.* [F, aged 24 years]

However, several participants felt that the µEMA question did not stay available for long enough. A participant stated as follows:

> *I like how it was really simple, it wasn't like a series of questions. They were just "yes" or "no." What made it difficult was the timing...it seems like it only lasts for five seconds and it disappears. Sometimes I may be doing something like resting and then survey comes up. Then I have to stick my arm out of the blanket to answer but then it goes away before I could.* [F, aged 22 years]

Initially, the prompt only remained active for 15 seconds, which we later updated during the pilot to 20 seconds. Participants generally noticed the smartwatch vibration on their wrists with a delay, despite a relatively aggressive vibration pattern. In fact, pilot participants described the µEMA vibration pattern (lasting for 11 seconds, intensifying progressively) as "too loud." Although the smartwatch vibration creates a quiet buzzing sound when on the arm, if the smartwatch is sitting on a surface (eg, charging), the resulting surface vibration can make a surprisingly loud sound that is difficult to ignore.

Participants reported difficulty answering µEMA questions when both hands were occupied, especially while driving. However, they were instructed to only answer µEMA questions when it was safe to do so, and they were given the specific example of driving as a situation in which they should ignore the prompts. Participants also reported difficulty answering µEMA questions while doing other cognitively demanding activities such as writing. In fact, a participant mentioned missing prompts while being in an examination room, stating as follows:

> *Sometimes if I was taking a quiz or test in class, I didn't want to be caught using my watch, so I didn't touch [ie, answer] it.* [M, aged 19 years]

Participants found the ability to undo their responses useful, especially when there were accidental taps on the smartwatch. A participant stated as follows:

> *Yes [Undo] was useful because sometimes on the watch...maybe because of the clothes I'm wearing...but [if] an option would be chosen so I'd be able to go "undo" and choose the right answer.* [F, aged 24 years]

Participants also found the *Sort of* answer option helpful in instances when they were not sure of their experience at the moment and would have had difficulty entering a more limited *Yes* or *No* response. A participant stated as follows:

> "*Sort of*" *made me think about my answer a little more rather than an outright yes or no, which is how I kind of determine how I'm going to answer yes or no...if you're feeling such and such, [then] I do like the "sort of" option because you're not feeling fully something. And you are not "not" feeling it completely either. So ["Sort of"] is a pretty easy option for people who are feeling like in the middle.* [F, aged 24 years]

In response to such observations with the goal of achieving cognitive simplicity to support microinteraction, we designed all the questions to have a *middle* answer such as *sort of*.

In the pilot study, each question presented was randomly sampled from the pool of questions. This resulted in questions either being overrepresented or repeated too consecutively for some individuals, adding monotony, which some participants noted. A participant stated as follows:

> *I think there were questions based on time or place. Questions like how was your day?...It would ask multiple times a day. I don't know if it was intentional, but it feels like I'm answering the same question over and over again. Same thing for places. It would ask me multiple times "is this your home?" or something like this.* [M, aged 18 years]

Changes to the Main Study Protocol

Overview

Table 2 summarizes the changes made to the µEMA app and protocol based on the pilot participants' feedback. In summary, there were 4 major changes to the main protocol based on the data and participant feedback in the pilot study. These changes are explained in detail in the next sections.

Table 2. Changes made to the microinteraction ecological momentary assessment (μEMA) protocol or technology before the main study in response to the pilot participants' feedback.

	Pilot study	Main study
Physical activity and sedentary behavior questions	Asked at random	Asked as sensor-triggered questions based on wrist-accelerometer activity
Location-based questions	Asked at random	Removed from μEMA and included as part of mobile phone EMA to obtain semantic location labels
μEMA question selection	At random	Using filter-based sampling algorithm
Person-level characteristics and validation questions	Not included in pilot study	Included in main study to assess validity and add variety to questions on μEMA days
Thank-you messages after answering a μEMA question	Messages selected at random from a bank of 10 messages	Messages selected sequentially from a bank of >250 messages
μEMA prompt wait time	15 seconds	Increased to 20 seconds
μEMA prompt vibration	11 seconds of progressively intense vibration pattern	6 seconds (2-second vibration with 1-second pause)

Changed μEMA Question-Selection Algorithm

In the pilot study, each μEMA prompt selected a question at random. As a result, the app had less control over preventing overrepresentation of questions within the day. On the basis of the interviews, we learned that participants found this repetitiveness of the questions monotonous. Thus, for the main study we developed a new question-selection algorithm that (1) guarantees that a given question is never asked consecutively and (2) ensures that a question is not answered more than a predetermined number of times.

Added Sensor-Triggered Questions

In the pilot study, we randomly sampled questions on PA, SB, and the participant's current location. However, these questions also felt repetitive to the pilot participants. Thus, in the main study protocol, these questions were asked when the passive sensors (using accelerometer and GPS) detected a relevant event, gathering information *only* when it was contextually relevant.

Added μEMA Validation and Engagement Questions

We added 2 additional types of questions for μEMA. First, we included *validation questions* to check whether the participants are paying attention to the μEMA prompts. Second, we added questions related to the slowly changing characteristics of the participants; these questions add variety to the question pool while also gathering useful information. Finally, on the μEMA undo screen, we added more variety in the thank-you messages that participants receive after answering a μEMA prompt. As the main study would last for 12 months per participant, we deemed reducing burden and increasing engagement to be important considerations.

Optimized μEMA Interface

We increased the time duration that a μEMA question stays visible on the smartwatch face from 15 seconds to 20 seconds. To address participant concerns about the loud vibration, we changed the pattern from an 11-second-long vibration to a 6-second pattern of 2-second vibrations separated by 1-second pauses.

Main TIME μEMA Substudy

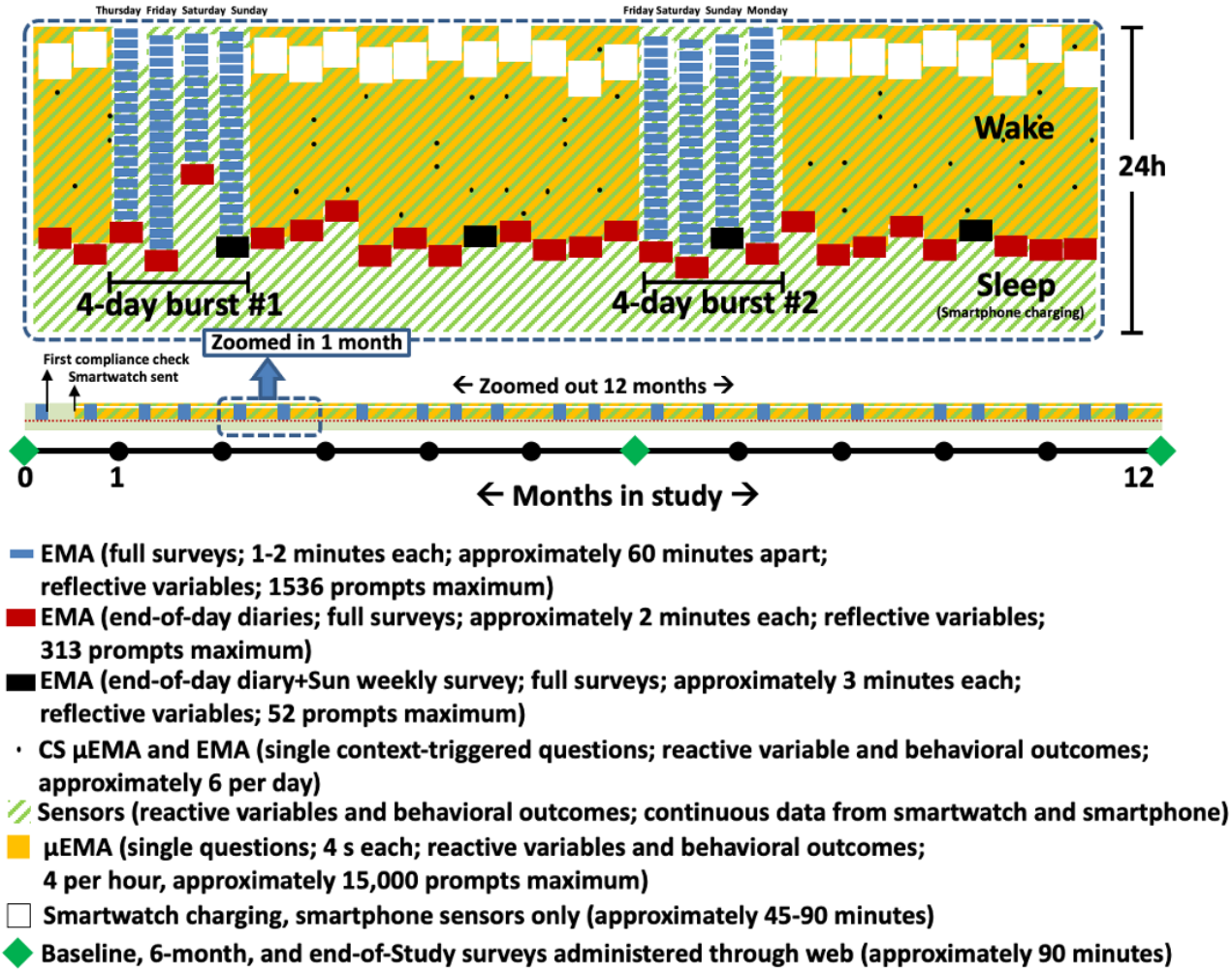
The protocol described here is based on changes made to the μEMA app after the pilot study feedback.

TIME Study Design

The TIME Study uses a nested design with multiple measurement bursts of smartphone-based EMA across a 12-month study period (Figure 3). Each measurement burst occurs over 4 days and always includes Saturday and Sunday. Measurement bursts occur every 2 weeks. The app ensures that there is at least a one-week gap between 2 consecutive burst periods. Thus, for a 12-month period (ie, 52 weeks), we expect up to 26 EMA burst periods (104 EMA burst days). The remaining nonburst days (or μEMA days) are reserved for μEMA prompting on the smartwatch; there can be up to 261 such days, with the number per participant dependent upon the receipt date of the smartwatch that is mailed only after the first successful burst period. The smartphone automatically sets the study schedule and controls the smartwatch so that μEMA questions are only prompted on the appropriate days and at the appropriate times. End-of-day surveys are prompted on all days, irrespective of measurement bursts. No μEMA prompts occur on the EMA burst days, and no other smartphone-based EMA prompts occur on μEMA days except for the end-of-day survey and sensor-triggered location surveys. Details of the smartphone-based EMA protocol are beyond the scope of this paper; they can be found in the main protocol paper (Wang, S, unpublished data, January 2022). Participants in the TIME Study are exposed to more interruptions from μEMA than from EMA, both within a day (4 μEMA prompts vs 1 EMA prompt per hour with possible additional reprompts) and across months (approximately 261 days for μEMA vs approximately 104 days for EMA for a year). Participants are also asked to complete baseline, 6-month, and 12-month surveys on the web using REDCap (Research Electronic Data Capture) [51]; the 6-month and 12-month surveys include questions about the perceived burden [52] of responding to μEMA and EMA questions. After the completion of the study, participants undergo a semistructured interview designed to gather additional information about their overall experiences answering questions on the devices. Participants who voluntarily withdraw from the

study or who are asked to withdraw because of low EMA compliance are also asked to complete an exit survey to enable us to learn more about their experiences and reasons for withdrawal or poor compliance.

Figure 3. Temporal Influences on Movement & Exercise Study protocol with nested ecological momentary assessment (EMA) bursts and microinteraction EMA (µEMA) on nonburst days; CS: context-sensitive.



Measures

Here we describe the self-report measures, question-selection algorithms, and simulation results to validate the filtering algorithm to select questions for µEMA.

Self-report µEMA Measures

Core-Construct Questions

Core-construct questions measure different behaviors, momentary states, and contexts that cannot be objectively captured using sensors and are expected to change throughout the day. The questions are selected to capture reflective processes, reactive processes, and internal and external processes that can predict or explain behavior adoption and maintenance [49]. Example constructs include feelings of pain, stress, procrastination, happiness, and positive and negative affect. These questions are sampled from a question bank (based on a procedure defined later) consisting of 30 questions (Multimedia Appendix 1); they are presented with signal-contingent prompts based on the aforementioned schedule. Each day during waking hours, 90% (55/62) of the scheduled prompts are reserved for these core-construct questions. In the question bank of 30 core-construct questions, 6 (20%) measure reflective processes, 5 (17%) measure reactive processes, 6 (20%) measure external factors, and 13 (43%) measure internal factors.

Engagement and Validation Questions

In addition to the core-construct questions, 8% (5/62) of the scheduled prompts during waking hours are used to present questions about person-level characteristics. These are questions that gather data on attributes such as whether a participant owns a car, cares for a pet, self-describes as a morning person, or has other specific personality characteristics. The questions serve two purposes: gathering information that changes relatively slowly and breaking monotony (thus, perhaps helping with participant engagement) by *surprising* participants with approximately 4 novel questions a day. These questions follow the wording requirements of other µEMA questions (eg, *Own a pet?* with answers *Yes* or *No* or *Bike to work*? with answers *Frequently*, *Sometimes*, and *Rarely*). The system uses a question bank of >280 questions so that they will not repeat for approximately 3½ months (assuming 60 µEMA prompts per day).

The remaining 2% (1/62) of the waking-day prompts are used to present validation questions designed to assess whether participants are paying attention to the µEMA questions and answering them carefully (ie, not randomly tapping on the screen to dismiss a question without looking at it, carelessly answering prompts incorrectly, or accidently pressing buttons on the smartwatch without realizing it). The validation questions are like other µEMA questions—cognitively simple single questions with 2-3 answers. However, these questions always have an unambiguously correct answer to verify response accuracy. Validation questions include simple math problems (*1 + 3 = 5?—Yes* or *No*), trivial trivia questions (*Sun rises in the east?—Yes* or *No*), intentionally silly questions (*Do pigs fly?—Yes* or *No*), and simple attention-assessment questions (*Please press A:—A*, *B*, or *C*). To break the question monotony and enhance engagement, validation questions rarely repeat. Our item bank has >300 validation questions for an estimated 7½ months of question prompting without repetition.

PA and SB Questions

Sensor-triggered questions are prompted to measure PA and SB, which are indirectly, passively measured using sensors but may require self-report for validation. Participants are asked to report whether they are currently engaging in PA or SB based on their activity level as inferred from the smartwatch raw acceleration data (collected continuously at 50 Hz). In real time, we compute the area under the curve (AUC) of the high-passed raw accelerometer signal for 10-second epochs [53]. The AUC is then used to determine the prompting of sensor-triggered questions. If the AUC exceeds an empirically determined threshold in a continuous 10-minute window, a PA question is eligible to be asked. If the AUC falls below a threshold for a continuous 60-minute window, an SB question is eligible to be asked [53].

Sensor-triggered questions are prompted *only* when a specific sensor event of interest occurs that suggests contextual relevance. The questions are *Physically active [Δ] min ago?* and *Sedentary [Δ] min ago?*, with answer options *Yes*, *Sort of*, and *No*, where Δ is the time difference from the current time to the end of the activity plus 25% of the activity bout length (ie, 2½ minutes for a 10-minute PA bout or 15 minutes for a 60-minute sedentary bout). This additional 25% of bout length is used to ensure that the question is asked at approximately the moment in time when the behavior was likely taking place but not at the behavioral transition.

At each scheduled µEMA prompt time, if there is an eligible sensor-triggered question, the system will trigger the question only 50% of the time. This is done to avoid continuous and monotonous sensor-triggered prompting for prolonged sedentary and physical activities. When the sensor-triggered questions are eligible to be prompted, they take priority over the core-construct, validation, and engagement questions. For instance, if the next scheduled prompt has a core-construct question queued but the app has identified a PA event, then 50% of the time, that core-construct question is replaced by the PA confirmation question. If both PA and SB are identified, the question about the most recent event is presented.

Question Selection for Prompting

A way to gather self-report data on a comprehensive set of behaviors using µEMA is to randomly select a question to ask the user—as was done in our pilot study. However, it may be useful to measure internal factors (eg, perceived stress, fatigue, and nervousness) more frequently within a day rather than external factors (eg, being with a friend). The app uses a strategic sampling algorithm to maximize the value of data collected about some constructs by limiting presentation of some questions, ensure that questions differ within an hour, and guarantee that each day includes questions from each of four broad construct categories (ie, internal and external factors and reflective and reactive processes).

Selecting Day-Level Subset for Core-Construct Questions

Each µEMA day, from our bank of 30 µEMA questions of core constructs, the app selects 4 (13%) questions related to internal factors, 2 (7%) to external factors, 2 (7%) to reflective processes, and 2 (7%) to reactive processes. More internal factors than other categories are selected because they are expected to vary throughout the day [54]. Rather than selecting any question at random, this procedure ensures that at least one question measuring each of the different factors is presented each day, with a temporally dense presentation. Each question is expected to be presented ≥4 times in a day, allowing measurement of within-day changes.

Filters Applied for Core-Construct Question Selection

Maximum Allowable Prompting per Question

Each µEMA question is assigned a maximum number of presentations per day (Multimedia Appendix 1). This ensures that some questions are not overrepresented in a day at the expense of others. For instance, questions capturing frequently changing constructs are assigned a higher limit (eg, *Feeling stressed?*) than questions that assess more stable constructs (eg, *Slept well yesterday?*). If the number of answers received for a question reaches the question's limit, that question is no longer presented that day.

Minimum Time Gap for Question Repetition

Once a particular µEMA question has been answered, it will not be presented again until at least 60 minutes have elapsed. For instance, if the question *Feeling stressed?* is answered, it is not asked again for at least one hour. However, if the question is not answered, it could be presented at the next prompt, within the hour.

Filtering Algorithm When Selecting a µEMA Question

The app uses the following algorithm to select µEMA questions (Textbox 1). For each µEMA day (*µEMA_prompting_day*, line 1), the app first generates a day-level subset (*day_level_subset*, line 2) of 10 questions from the pool of 30 µEMA questions (*all_* µEMA *_questions*). Next, at each µEMA prompt, day-level questions that have reached their limit (*remove_maxed_out_questions*, line 4) or have been asked recently (*remove_questions_answered_within_one_hour*, line 5) are filtered. From the remaining list, a question is selected at random (*select_question_at_random*, line 6). If the question

prompted is answered, the question's prompt count is updated (*update_prompt_count*, line 8), as is the last-answered time (*update_answer_time*, line 9). If the question is not answered, then it is available for selection in the next scheduled µEMA prompt.

Textbox 1. Microinteraction ecological momentary assessment (µEMA) question-filtering algorithm at each prompt on nonburst days.

Algorithm used to select µEMA questions

1. For each µEMA_prompting_day:
2. day_level_subset=subset (all_ µEMA _questions)
3. For each µEMA prompt:
4. filtered_questions=remove_maxed_out_questions (day_level_subset)
5. filtered_questions=remove_questions_answered_within_one_hour (filtered_questions)
6. question_to_ask=select_question_at_random (filtered_questions)
7. if question_to_ask is answered:
8. update_prompt_count (question_to_ask)
9. update_answer_time (question_to_ask)

µEMA Question-Selection Simulation

To verify our filtering algorithm (Textbox 1), we simulated question sampling and compared the distributions of questions asked through random sampling (used in the pilot study) with our proposed protocol.

Question-Sampling Algorithm Simulation

We simulated question selection for 365 days for 300 users, where 261 days were reserved for µEMA prompting, resulting in 78,300 total µEMA days. We assumed 8 hours of sleep and 16 hours of wake time per day, resulting in 15½ hours available for µEMA prompting. Targeting 4 prompts an hour results in 62 µEMA prompting opportunities. Of these, 90% (55/62) were available for core constructs, resulting in 55 slots for core-construct questions each day. We assumed that 75% (46/62) of the scheduled prompts would be answered (ie, compliance).

Algorithm Simulation Results

Compared with random sampling (Multimedia Appendix 1), the filtering algorithm increased the frequency of presentation of core-construct questions in a day. For instance, when using random sampling, certain items such as sleep satisfaction are repeated at the expense of other variables of interest (eg, internal factors). Although with random sampling a question is prompted on more days (approximately 75% vs approximately 33%), the questions are repeated only twice (median frequency) on a given day. This presentation is insufficient to study within-day changes in slopes when we need at least three observations. The filtering algorithm ensures that on days when questions are presented, they are repeated 4 times (median frequency) a day. Later, we validate this algorithm against the real-world data from our main study participants (Multimedia Appendix 2).

TIME Study Main Trial Participants and Recruitment

Participants were eligible for the TIME Study main trial if they (1) owned an Android smartphone running Android version ≥6.0 as their only personal mobile phone with no intention to switch to a non-Android smartphone, (2) did not wear a smartwatch already, (3) were aged 18-29 years and living in the United States, (4) were currently engaged in recommended levels of PA (or intended to within the next 12 months) [55], (5) spoke and read English, (6) resided in an area with Wi-Fi connectivity, (7) did not have any physical or cognitive limitations that prevented participation, and (8) were able wear a smartwatch and answer real-time smartphone and smartwatch surveys. All study procedures were approved by the institutional review board at the University of Southern California (USC; HS-18-00605). Participants were recruited using several strategies: (1) sending emails to individuals enrolled in the Happiness & Health Study, a USC longitudinal cohort study of young adults [56]; (2) posting flyers in the greater Los Angeles metropolitan area; (3) purchasing web and social media advertisements; (4) sending emails to addresses on file from other institutional review board–approved USC studies; and (5) contacting participants identified using ResearchMatch [57]. As of June 2021, 246 participants had been enrolled, among whom 81 (32.9%) had completed at least 6 months of data collection from their respective start dates (Table 3). We recruited 246 participants who received a smartwatch after the initial run-in period. This sample size was determined based on the objectives of the main TIME Study protocol (Wang, S, unpublished data, January 2022). The µEMA substudy was added (Figure 3) for exploratory purposes.

Table 3. Participant demographic survey for those who had completed 6 months as of June 2021 (N=81).

Demographics	Values
Age (years), mean (SD)	21.7 (2.4)
Sex, n (%)	
Male	45 (55)
Female	36 (45)
Ethnicity, n (%)	
Non-Hispanic	51 (63)
Hispanic	30 (37)
Race[a], n (%)	
White	41 (54)
Asian or Pacific Islander	35 (46)
Black	7 (9)
American Indian or Alaska Native	5 (7)
Education, n (%)	
High school	12 (15)
Some college	47 (58)
College graduate	22 (27)
Work status[b], n (%)	
Employed	36 (45)
Out of work	16 (20)
Student	53 (66)
Unable to work	4 (5)

[a]Participants could select >1 answer option from American Indian or Alaska Native, Hawaiian or Pacific Islander, Black or African American, White, Asian, Unknown, and Would prefer not to answer.

[b]Participants could select >1 answer option; for example, "Student" and "Employed."

Study Procedures

Because of COVID-19 restrictions on in-person recruitment for human subjects research, all participant recruitment and onboarding was conducted remotely. After screening, research staff individually met with each participant remotely (through Zoom) to obtain informed consent. Staff then guided the participants through the TIME app installation on their personal smartphones. Researchers then observed EMA burst compliance for the first 4-day EMA burst (ie, run-in period) during the next 2 weeks. If the compliance for the run-in period was <8 surveys per day on all days, the participants were withdrawn from the study. Otherwise, they could continue in the study and were mailed a smartwatch (Fossil Gen 4 or Gen 5 model). Once the participants received the smartwatch, a staff member scheduled a second remote orientation session with the participant to guide them through the TIME smartwatch setup. Expectations for smartwatch wear (including wearing it during sleep), charging it, and answering µEMA questions were explained. Participants were instructed that to remain compliant they should wear the smartwatch for 23 hours a day and charge it every day. They were also instructed to wear the smartwatch while sleeping. For µEMA prompting, participants were instructed as follows:

> *"You will be prompted on the smartwatch with single questions of the type Are you walking right now? Yes | Sort of | No. These questions will take only 2-4 seconds to respond to—just like checking time on the smartwatch. There may be up to six such question prompts in an hour, on average, and the frequency of question prompts may vary at different times of your waking day. All the smartwatch questions will be outside of the 4-day smartphone survey burst periods. Each watch prompt will wait for up to 20 seconds for you to respond; if the question is not answered within 20 seconds, the question may disappear from the watch. Unlike phone surveys, watch questions are not re-prompted."*

After the smartwatch orientation, participants could use the smartwatch as they deemed fit (if it did not interfere with TIME app functionality), while answering µEMA questions when they appeared. Participants receive US $20 per month for wearing the smartwatch for 23 hours on at least 24 days of the month. Participants could earn up to an additional US $80 per month for EMA burst compliance (Wang, S, unpublished data, January 2022). In addition, if at the end of the study, participants have >50% compliance with µEMA, they were able to keep the

smartwatch as their personal device. However, no monthly or regular compensation was provided for µEMA compliance. Participants do not receive feedback on their µEMA compliance, and their withdrawal from the study does not depend on µEMA compliance.

Response Behavior Measures

We measure compliance rate, completion rate, undo rate, and validation rate to characterize participant response behavior when answering µEMA questions.

Compliance Rate

Compliance rate is measured as the percentage of µEMA questions answered out of all the scheduled questions, including those not prompted because of the device being turned off or in DND mode.

Compliance rate (%) = #Questions answered/#Questions scheduled × 100

Completion Rate

Completion rate is measured as the percentage of µEMA questions answered out of all the delivered questions, excluding those not prompted because of the device being turned off or in DND mode.

Compliance rate (%) = #Questions answered/#Questions delivered × 100

Undo Rate

Undo rate is measured as the total number of µEMA questions answered when the *Undo?* option was selected after providing an initial answer.

Undo rate (%) = #Undo count/#Questions answered × 100

Validation Rate

Validation rate is measured as the percentage of µEMA validation questions answered correctly out of all the validation questions answered.

Validation rate (%) = #Validation correctly answered/#Validation answered × 100

Results

Compliance, Completion, and Validation Rates

Overall, the compliance rate (n=81 at ≥6 months) for µEMA was 67.7% (SD 13.7%) and the completion rate was 80.23% (SD 13.3%). The µEMA validation rate for the participants (n=81) was 92.9% (SD 23.7%)—indicating that they were paying attention to the µEMA question content and not carelessly tapping to dismiss them.

Response Behavior

Table 4 presents a summary of the response behavior. We also verified the filter-based question-selection algorithm's performance against the real-world data from 2% (2/81) of the study participants who completed 6 months in the study (Multimedia Appendix 2). We found that no core-construct question was asked more than its maximum allowable number of times. In addition, the median number of times a core-construct question was asked was consistent with our simulation results (Multimedia Appendix 1).

Table 4. Main study microinteraction ecological momentary assessment (µEMA) response behavior for participants completing 6 months as of June 2021 (N=81).

	Values
µEMA days, n	13,415
Expected µEMA questions, n	790,388
Delivered µEMA questions, n, %	662,397 (83.81)
Answered µEMA questions, n, %	535,430 (80.83)
Mean daily µEMA compliance rate, % (SD)	67.6 (24.4)
Mean daily µEMA completion rate, % (SD)	78.5 (22.2)
Mean participant µEMA compliance rate, % (SD)	67.4 (13.7)
Mean participant µEMA completion rate, % (SD)	80.2 (13.3)
Mean µEMA response time, seconds (SD)	4.8 (1.4)
Total µEMA undos (% of total µEMA questions answered)	22,202 (4.2)
Mean µEMA question validation rate, % (SD)	92.9 (23.7)

Discussion

Our work with µEMA demonstrates that the technique may enable temporally dense ILD collection with manageable burden to support longitudinal studies or interventions, although some limitations of this work will require further investigation.

Summary and Strengths

Traditional mobile phone–based EMA can lead to interruption and response burden. To prevent this burden, researchers often compromise on the temporal density of prompts (ie, by prompting less frequently) and reduce the number of questions or constructs being measured. Complementary forms of EMA are needed where both the researchers' needs for comprehensive

understanding of health behaviors and users' concern for burden are taken into account. µEMA provides such an opportunity, where the quick microinteractions not only make it less burdensome for the users to answer questions on the smartwatch, but also enable data gathering at a higher temporal density.

The TIME µEMA substudy combines smartphone-based EMA with low-burden µEMA on a smartwatch to gather real-time self-report data from naturalistic settings using the personal smartphones of participants. Participants are encouraged to use their smartphone and the loaned smartwatch normally with only a few limitations (ie, avoiding apps that interfere with TIME app functionality, such as fitness trackers). Participants are incentivized to wear the smartwatch, but they can pause smartwatch survey prompting as needed. Participants are not directly compensated for high µEMA compliance, although they are told that they can keep the smartwatch if they have >50% µEMA compliance at the end of 1 year. EMA compliance, by contrast, is reinforced with an explicit and compliance-contingent reward provided at regular monthly intervals.

Our preliminary results from the pilot and ongoing TIME µEMA substudy show that despite a high interruption rate, µEMA could be sustainable for gathering self-report data longitudinally. Controlled filtering of µEMA questions enables measurement of a comprehensive set of behaviors repeatedly during waking hours. The validation questions make it possible to identify careless responding on both µEMA and smartphone-based EMA.

The TIME µEMA substudy can be used to assess the utility and sustainability of µEMA. Comparing response rates with the qualitative data from the exit interviews, for example, suggests that µEMA can be used to gather temporally dense self-report information. In addition, passively collected data on wrist motion (using accelerometers) and location (using GPS) may help explore contextual factors that affect µEMA compliance. The filtering-based µEMA question-selection algorithm may be used to capture diurnal patterns of different constructs during waking hours. These diurnal patterns could be used to explore (1) different clusters of individuals who have similar diurnal patterns for variables of interest [58] and (2) how similar, or different, diurnal patterns are that were captured using µEMA versus those captured using EMA. With 4 prompts per hour (or 70 prompts for an 18-hour wake period), µEMA may enable denser measurement than EMA. Combined with passive sensing, µEMA may support study of diurnal patterns, which might be compared with EMA at the end of the day.

Overall, the purpose of µEMA is not to replace traditional EMA but to complement it. The TIME Study's µEMA substudy protocol provides an opportunity to explore how small amounts of information gathered at high frequency can help us learn about health behavior change and maintenance, even as constructs related to behavior change deviate by the hour. Beyond observation studies, TIME Study data could enable exploration of the viability of µEMA for longitudinal interventions (eg, just-in-time adaptive interventions), where self-report on many different behaviors may be required. For instance, when optimizing mobile health interventions as part of microrandomized trials [59], µEMA may provide a low-burden self-report interface that can gather information about behaviors that sensors cannot measure yet (eg, fatigue, procrastination, and pain), especially when such measurements are needed in close time intervals (eg, within an hour) and across changing contexts.

Limitations and Opportunities

There will be several limitations of the TIME µEMA substudy that will provide opportunities for future research. Findings about the µEMA methodology may not generalize to other population age groups because the main protocol is designed to study health behavior adoption and maintenance in young adults (age 18-29 years). There has been an increased interest in exploring the acceptance of wearable technologies in older adults [60-62], youth [63], and children [64] and more research is needed to explore the acceptance of µEMA as a data collection method for both longitudinal observational and intervention studies with these demographic groups. As µEMA is deployed on a wearable smartwatch, our data collection excludes participants who may not be allowed to wear, or be comfortable wearing, the smartwatch during work hours; in some professions, answering µEMA or EMA surveys may not be appropriate (eg, professions that require being in intensive care units, using cleanrooms, extended driving, operating heavy machinery, and working in construction).

We had to restrict our data collection to Android users because our research app used advanced sensor capabilities not programmatically available on iOS devices; thus, the TIME µEMA substudy excludes iOS users. Nevertheless, we observed more demographic diversity in the Android users who participated in our screening surveys (of the 746 participants, 393, 52.7%, were women; 372, 49.9%, were White; 216, 28.9%, were Asian; 119, 15.9%, were Black; and 184, 24.6%, were Hispanic; 630, 84.5%, engage in PA) versus the iOS users (of the 395 participants, 296, 75.1%, were women; 177, 44.9%, were White; 99, 25.1%, were Asian; 71, 17.8%, were Black, and 130, 32.9% were Hispanic; 365, 92.4%, engage in PA). We have no reason to believe that the platform itself would affect µEMA response patterns.

Although µEMA may keep burden more manageable than EMA despite intensive prompting, each question provides a limited answer set (*Yes*, *Sort of*, and *No*) that may be less sensitive to construct variance than mobile phone–based EMAs with more response options. Yet, with a median of ≥4 measurements per construct within a day (Multimedia Appendix 1), µEMA provides opportunities to use mixed effects location scale models to study changes in variances and slopes of different constructs [65,66]. This also provides opportunities to explore approaches beyond multilevel models, for instance, using dynamic Bayesian networks [67,68] or temporal networks [69,70] to build individual-specific models.

Unlike EMA, where multiple questions are asked back to back, thus allowing for the simultaneous comparison of different behaviors, µEMA presents only 1 question at a time to ensure that each prompt can be responded to as a microinteraction. Therefore, different construct measurements are spread throughout waking hours at different times. Although this approach does not allow for testing concurrent associations

among constructs, there are alternative ways of exploring temporally lagged associations of different µEMA measures (eg, Granger Causality [71]).

Conclusions

The TIME µEMA substudy provides an opportunity to explore a new method of collecting temporally dense self-report data. Although EMA provides information on temporal dynamics of behavior, the burden of accessing the mobile phone, unlocking it, and then answering multiple complex questions compromises engagement. In this formative research, we explored the feasibility of using µEMA to measure multiple constructs in an ILD study. When deploying µEMA, each prompt is a cognitively simple single question that can be answered with a quick glanceable microinteraction with an always-accessible smartwatch. Because of this simplicity, µEMA permits a higher interruption rate than EMA without a corresponding increase in perceived burden. The TIME Study is the first ILD study to deploy µEMA for an entire year; thus, it could provide insights on methodological properties of µEMA—especially the ability to sustain participant engagement. Data from the study could be used to study what influences µEMA compliance, what can be learned about an individual's behavior using µEMA, and how µEMA might be used along with EMA and passive sensors to richly characterize human behavior, especially how it may change throughout the day in response to rapidly changing contexts.

Acknowledgments

This work is supported by the National Institutes of Health (NIH)/NHLBI U01HL146327. The content is solely the responsibility of the authors and does not represent the official views of the NIH. Krit Goyal, Adeeb M Arif, Binod Thapa Chhetry, Jixin Li, and Qu Tang contributed to the development of the TIME app and associated data processing algorithms. We thank the pilot and main study participants for their time.

Authors' Contributions

AP led the development of the TIME app, the microinteraction ecological momentary assessment (µEMA) components of the study design, µEMA deployment and troubleshooting, and the first draft of this manuscript. SW, DC, and BD designed the EMA components of the TIME study, baseline and follow-up survey designs, participant recruitment, and analysis of participant demographics and contributed to manuscript writing. GD and SI are the co–primary investigators of the research grant and contributed to research study design, technology design, research oversight, and manuscript writing.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Random versus filter-based question sampling.
[DOCX File , 19 KB - formative_v6i2e32772_app1.docx]

Multimedia Appendix 2
Filter-based sampling simulation versus main study participant results.
[DOCX File , 23 KB - formative_v6i2e32772_app2.docx]

References

1. Nusser S, Intille S, Maitra R. Emerging technologiesnext generation intensive longitudinal data collection. In: Models for Intensive Longitudinal Data. Oxford, United Kingdom: Oxford University Press; 2006.
2. Intille S. The precision medicine initiative and pervasive health research. IEEE Pervasive Comput 2016 Jan;15(1):88-91. [doi: 10.1109/mprv.2016.2]
3. Collins FS, Varmus H. A new initiative on precision medicine. N Engl J Med 2015 Feb 26;372(9):793-795 [FREE Full text] [doi: 10.1056/NEJMp1500523] [Medline: 25635347]
4. Intille SS. Ubiquitous computing technology for just-in-time motivation of behavior change. In: Studies in Health Technology and Informatics. Amsterdam, Netherlands: IOS Press; 2004.
5. Nahum-Shani I, Smith S, Spring B, Collins L, Witkiewitz K, Tewari A, et al. Just-in-time adaptive interventions (JITAIs) in mobile health: key components and design principles for ongoing health behavior support. Ann Behav Med 2018 May 18;52(6):446-462 [FREE Full text] [doi: 10.1007/s12160-016-9830-8] [Medline: 27663578]
6. Yang C, Maher J, Ponnada A, Dzubur E, Nordgren R, Intille S, et al. An empirical example of analysis using a two-stage modeling approach: within-subject association of outdoor context and physical activity predicts future daily physical activity levels. Transl Behav Med 2021 Apr 26;11(4):912-920 [FREE Full text] [doi: 10.1093/tbm/ibaa107] [Medline: 33159452]

7. Maher JP, Dzubur E, Nordgren R, Huh J, Chou C, Hedeker D, et al. Do fluctuations in positive affective and physical feeling states predict physical activity and sedentary time? Psychol Sport Exerc 2019 Mar;41:153-161 [FREE Full text] [doi: 10.1016/j.psychsport.2018.01.011] [Medline: 30853854]
8. Marsch LA. Digital health data-driven approaches to understand human behavior. Neuropsychopharmacology 2021 Jan;46(1):191-196 [FREE Full text] [doi: 10.1038/s41386-020-0761-5] [Medline: 32653896]
9. Wongvibulsin S, Martin SS, Saria S, Zeger SL, Murphy SA. An individualized, data-driven digital approach for precision behavior change. Am J Lifestyle Med 2020;14(3):289-293 [FREE Full text] [doi: 10.1177/1559827619843489] [Medline: 32477031]
10. Yang X, Ram N, Gest S, Lydon-Staley D, Conroy D, Pincus A, et al. Socioemotional dynamics of emotion regulation and depressive symptoms: a person-specific network approach. Complexity 2018;2018:5094179 [FREE Full text] [doi: 10.1155/2018/5094179] [Medline: 30613129]
11. Oyibo K, Adaji I, Vassileva J. Social cognitive determinants of exercise behavior in the context of behavior modeling: a mixed method approach. Digit Health 2018;4:2055207618811555 [FREE Full text] [doi: 10.1177/2055207618811555] [Medline: 30479828]
12. Dunton G, Rothman A, Leventhal A, Intille S. How intensive longitudinal data can stimulate advances in health behavior maintenance theories and interventions. Transl Behav Med 2021 Feb 11;11(1):281-286 [FREE Full text] [doi: 10.1093/tbm/ibz165] [Medline: 31731290]
13. Spruijt-Metz D, Hekler E, Saranummi N, Intille S, Korhonen I, Nilsen W, et al. Building new computational models to support health behavior change and maintenance: new opportunities in behavioral research. Transl Behav Med 2015 Sep;5(3):335-346 [FREE Full text] [doi: 10.1007/s13142-015-0324-1] [Medline: 26327939]
14. Baldwin W. Information no one else knows: the value of self-report. In: The Science of Self-report: Implications for Research and Practice. NJ, USA: Lawrence Erlbaum Associates Publishers; 2000.
15. 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]
16. Stone A, Shiffman S. Ecological momentary assessment (EMA) in behavioral medicine. Annals Behav Med 1994;16(3):199-202. [doi: 10.1093/abm/16.3.199]
17. van Berkel N, Ferreira D, Kostakos V. The experience sampling method on mobile devices. ACM Comput Surv 2018 Nov 30;50(6):1-40. [doi: 10.1145/3123988]
18. Fuller-Tyszkiewicz M, Skouteris H, Richardson B, Blore J, Holmes M, Mills J. Does the burden of the experience sampling method undermine data quality in state body image research? Body Image 2013 Sep;10(4):607-613. [doi: 10.1016/j.bodyim.2013.06.003] [Medline: 23856302]
19. Courvoisier DS, Eid M, Lischetzke T. Compliance to a cell phone-based ecological momentary assessment study: the effect of time and personality characteristics. Psychol Assess 2012 Sep;24(3):713-720. [doi: 10.1037/a0026733] [Medline: 22250597]
20. Burke LE, Shiffman S, Music E, Styn MA, Kriska A, Smailagic A, et al. Ecological momentary assessment in behavioral research: addressing technological and human participant challenges. J Med Internet Res 2017 Mar 15;19(3):e77 [FREE Full text] [doi: 10.2196/jmir.7138] [Medline: 28298264]
21. Stone A, Kessler R, Haythornthwaite J. Measuring daily events and experiences: decisions for the researcher. J Pers 1991 Sep;59(3):575-607. [doi: 10.1111/j.1467-6494.1991.tb00260.x] [Medline: 1960643]
22. Intille S, Haynes C, Maniar D, Ponnada A, Manjourides J. µEMA: microinteraction-based Ecological Momentary Assessment (EMA) using a smartwatch. Proc ACM Int Conf Ubiquitous Comput 2016 Sep;2016:1124-1128 [FREE Full text] [doi: 10.1145/2971648.2971717] [Medline: 30238088]
23. Ashbrook D. Enabling Mobile Microinteractions. Atlanta, GA: College of Computing, Georgia Institute of Technology; 2010.
24. Ashbrook D, Clawson J, Lyons K, Starner T, Patel N. Quickdraw: the impact of mobility and on-body placement on device access time. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2008 Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; Apr 5 - 10, 2008; Florence, Italy. [doi: 10.1145/1357054.1357092]
25. Dey A, Wac K, Ferreira D, Tassini K, Hong J, Ramos J. Getting closer: an empirical investigation of the proximity of user to their smartphones. In: Proceedings of the 13th international conference on Ubiquitous computing. 2011 Presented at: Proceedings of the 13th international conference on Ubiquitous computing; Sep 17 - 21, 2011; Beijing, China. [doi: 10.1145/2030112.2030135]
26. Chun J, Dey A, Lee K, Kim S. A qualitative study of smartwatch usage and its usability. Hum Factors Man 2018 Feb 16;28(4):186-199. [doi: 10.1002/hfm.20733]
27. Pizza S, Brown B, McMillan D, Lampinen A. Smartwatch in vivo. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. 2016 Presented at: Proceedings of the CHI Conference on Human Factors in Computing Systems; May 7 - 12, 2016; San Jose, California, USA. [doi: 10.1145/2858036.2858522]

28. Schneid SD, Armour C, Park YS, Yudkowsky R, Bordage G. Reducing the number of options on multiple-choice questions: response time, psychometrics and standard setting. Med Educ 2014 Oct;48(10):1020-1027. [doi: 10.1111/medu.12525] [Medline: 25200022]
29. Hick WE. On the rate of gain of information. Q J Exper Psychol 2018 Jan 01;4(1):11-26. [doi: 10.1080/17470215208416600]
30. Proctor RW, Schneider DW. Hick's law for choice reaction time: a review. Q J Exp Psychol (Hove) 2018 Jun;71(6):1281-1299. [doi: 10.1080/17470218.2017.1322622] [Medline: 28434379]
31. Fitts PM. The information capacity of the human motor system in controlling the amplitude of movement. J Exp Psychol 1954 Jun;47(6):381-391. [Medline: 13174710]
32. Bi X, Li Y, Zhai S. FFitts law: modeling finger touch with fitts' law. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2013 Presented at: CHI '13: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; Apr 27- May 2, 2013; Paris, France. [doi: 10.1145/2470654.2466180]
33. Schembre SM, Liao Y, O'Connor SG, Hingle MD, Shen S, Hamoy KG, et al. Mobile ecological momentary diet assessment methods for behavioral research: systematic review. JMIR Mhealth Uhealth 2018 Nov 20;6(11):e11170 [FREE Full text] [doi: 10.2196/11170] [Medline: 30459148]
34. Hawkley LC, Burleson MH, Berntson GG, Cacioppo JT. Loneliness in everyday life: cardiovascular activity, psychosocial context, and health behaviors. J Pers Soc Psychol 2003 Jul;85(1):105-120. [doi: 10.1037/0022-3514.85.1.105] [Medline: 12872887]
35. Kikuchi H, Yoshiuchi K, Yamamoto Y, Komaki G, Akabayashi A. Does sleep aggravate tension-type headache?: an investigation using computerized ecological momentary assessment and actigraphy. Biopsychosoc Med 2011 Aug 12;5:10 [FREE Full text] [doi: 10.1186/1751-0759-5-10] [Medline: 21835045]
36. Hernandez J, McDuff D, Infante C, Maes P, Quigley K, Picard R. Wearable ESM: differences in the experience sampling method across wearable devices. In: Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services. 2016 Presented at: MobileHCI '16: Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services; Sep 6 - 9, 2016; Florence, Italy. [doi: 10.1145/2935334.2935340]
37. Timmermann J, Heuten W, Boll S. Input methods for the Borg-RPE-scale on smartwatches. In: Proceedings of the 2015 9th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth). 2015 Presented at: 2015 9th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth); May 20-23, 2015; Istanbul, Turkey. [doi: 10.4108/icst.pervasivehealth.2015.259220]
38. Ponnada A, Haynes C, Maniar D, Manjourides J, Intille S. Microinteraction ecological momentary assessment response rates: effect of microinteractions or the smartwatch? Proc ACM Interact Mob Wearable Ubiquitous Technol 2017 Sep;1(3):1-16 [FREE Full text] [doi: 10.1145/3130957] [Medline: 30198012]
39. Ponnada A, Thapa-Chhetry B, Manjourides J, Intille S. Measuring criterion validity of microinteraction ecological momentary assessment (Micro-EMA): exploratory pilot study with physical activity measurement. JMIR Mhealth Uhealth 2021 Mar 10;9(3):e23391 [FREE Full text] [doi: 10.2196/23391] [Medline: 33688843]
40. 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 Jun;54(6):1063-1070. [doi: 10.1037//0022-3514.54.6.1063] [Medline: 3397865]
41. Dunton GF, Liao Y, Intille S, Huh J, Leventhal A. Momentary assessment of contextual influences on affective response during physical activity. Health Psychol 2015 Dec;34(12):1145-1153. [doi: 10.1037/hea0000223] [Medline: 26053885]
42. Larsen J, Eskelund K, Christiansen T. Active self-tracking of subjective experience with a one-button wearable: a case study in military PTSD. In: Proceedings of the CHI EA '17: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems. 2017 Presented at: CHI EA '17: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems; May 6 - 11, 2017; Denver Colorado USA.
43. King ZD, Moskowitz J, Egilmez B, Zhang S, Zhang L, Bass M, et al. micro-Stress EMA: a passive sensing framework for detecting in-the-wild stress in pregnant mothers. Proc ACM Interact Mob Wearable Ubiquitous Technol 2019 Sep;3(3):91 [FREE Full text] [doi: 10.1145/3351249] [Medline: 32432212]
44. Jayathissa P, Quintana M, Abdelrahman M, Miller C. Humans-as-a-sensor for buildings—intensive longitudinal indoor comfort models. Buildings 2020 Oct 01;10(10):174. [doi: 10.3390/buildings10100174]
45. Paruthi G, Raj S, Gupta A, Huang C, Chang Y, Newman M. HEED: situated and distributed interactive devices for self-reporting. In: Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers. 2017 Presented at: UbiComp '17: Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers; Sep 11 - 15, 2017; Maui, Hawaii. [doi: 10.1145/3123024.3123162]
46. Paruthi G, Raj S, Baek S, Wang C, Huang C, Chang Y, et al. Heed: exploring the design of situated self-reporting devices. Proc ACM Interact Mob Wearable Ubiquitous Technol 2018 Sep 18;2(3):1-21. [doi: 10.1145/3264942]
47. Ono M, Schneider S, Junghaenel DU, Stone AA. What affects the completion of ecological momentary assessments in chronic pain research? An individual patient data meta-analysis. J Med Internet Res 2019 Feb 05;21(2):e11398 [FREE Full text] [doi: 10.2196/11398] [Medline: 30720437]

48. Dunton GF. Sustaining health-protective behaviors such as physical activity and healthy eating. JAMA 2018 Aug 21;320(7):639-640 [FREE Full text] [doi: 10.1001/jama.2018.6621] [Medline: 29852046]
49. Rothman AJ, Sheeran P, Wood W. Reflective and automatic processes in the initiation and maintenance of dietary change. Ann Behav Med 2009 Dec;38 Suppl 1:S4-17. [doi: 10.1007/s12160-009-9118-3] [Medline: 19787308]
50. van Berkel N, Goncalves J, Hosio S, Sarsenbayeva Z, Velloso E, Kostakos V. Overcoming compliance bias in self-report studies: a cross-study analysis. Int J Human Comput Stud 2020 Feb;134:1-12. [doi: 10.1016/j.ijhcs.2019.10.003]
51. 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]
52. Suh H, Shahriaree N, Hekler E, Kientz J. Developing and validating the user burden scale: a tool for assessing user burden in computing systems. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. 2016 Presented at: CHI '16: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems; May 7 - 12, 2016; San Jose, CA. [doi: 10.1145/2858036.2858448]
53. Dunton GF, Dzubur E, Intille S. Feasibility and performance test of a real-time sensor-informed context-sensitive ecological momentary assessment to capture physical activity. J Med Internet Res 2016 Jun 01;18(6):e106 [FREE Full text] [doi: 10.2196/jmir.5398] [Medline: 27251313]
54. Nosakhare E, Picard R. Toward assessing and recommending combinations of behaviors for improving health and well-being. ACM Trans Comput Healthcare 2020 Mar 02;1(1):1-29. [doi: 10.1145/3368958]
55. Piercy KL, Troiano RP. Physical activity guidelines for Americans from the US department of health and human services. Cardiovascular Q Outcomes 2018 Nov;11(11):e005263. [doi: 10.1161/circoutcomes.118.005263]
56. Leventhal AM, Strong DR, Kirkpatrick MG, Unger JB, Sussman S, Riggs NR, et al. Association of electronic cigarette use with initiation of combustible tobacco product smoking in early adolescence. JAMA 2015 Aug 18;314(7):700-707 [FREE Full text] [doi: 10.1001/jama.2015.8950] [Medline: 26284721]
57. Harris PA, Scott KW, Lebo L, Hassan N, Lightner C, Pulley J. ResearchMatch: a national registry to recruit volunteers for clinical research. Academic Med 2012;87(1):66-73. [doi: 10.1097/acm.0b013e31823ab7d2]
58. Ram N, Brinberg M, Pincus AL, Conroy DE. The questionable ecological validity of ecological momentary assessment: considerations for design and analysis. Res Hum Dev 2017;14(3):253-270 [FREE Full text] [doi: 10.1080/15427609.2017.1340052] [Medline: 30613195]
59. Klasnja P, Hekler EB, Shiffman S, Boruvka A, Almirall D, Tewari A, et al. Microrandomized trials: an experimental design for developing just-in-time adaptive interventions. Health Psychol 2015 Dec;34S:1220-1228 [FREE Full text] [doi: 10.1037/hea0000305] [Medline: 26651463]
60. Mercer K, Giangregorio L, Schneider E, Chilana P, Li M, Grindrod K. Acceptance of commercially available wearable activity trackers among adults aged over 50 and with chronic illness: a mixed-methods evaluation. JMIR Mhealth Uhealth 2016 Jan 27;4(1):e7 [FREE Full text] [doi: 10.2196/mhealth.4225] [Medline: 26818775]
61. Farivar S, Abouzahra M, Ghasemaghaei M. Wearable device adoption among older adults: a mixed-methods study. Int J Inf Manage 2020 Dec;55:102209 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2020.102209] [Medline: 32834339]
62. Wu A, Munteanu C. Understanding older users' acceptance of wearable interfaces for sensor-based fall risk assessment. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Presented at: Proceedings of the CHI Conference on Human Factors in Computing Systems; Apr 21 - 26, 2018; Montreal, QC, Canada. [doi: 10.1145/3173574.3173693]
63. 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]
64. Schaefer SE, Van Loan M, German JB. A feasibility study of wearable activity monitors for pre-adolescent school-age children. Prev Chronic Dis 2014 May 22;11:E85 [FREE Full text] [doi: 10.5888/pcd11.130262] [Medline: 24854236]
65. Hedeker D, Nordgren R. MIXREGLS: a program for mixed-effects location scale analysis. J Stat Softw 2013 Mar;52(12):1-38 [FREE Full text] [doi: 10.18637/jss.v052.i12] [Medline: 23761062]
66. Dzubur E, Ponnada A, Nordgren R, Yang C, Intille S, Dunton G, et al. MixWILD: a program for examining the effects of variance and slope of time-varying variables in intensive longitudinal data. Behav Res Methods 2020 Aug;52(4):1403-1427 [FREE Full text] [doi: 10.3758/s13428-019-01322-1] [Medline: 31898295]
67. Xu S, Thompson W, Kerr J, Godbole S, Sears DD, Patterson R, et al. Modeling interrelationships between health behaviors in overweight breast cancer survivors: applying Bayesian networks. PLoS One 2018;13(9):e0202923 [FREE Full text] [doi: 10.1371/journal.pone.0202923] [Medline: 30180192]
68. Chih M, Patton T, McTavish FM, Isham AJ, Judkins-Fisher CL, Atwood AK, et al. Predictive modeling of addiction lapses in a mobile health application. J Subst Abuse Treat 2014 Jan;46(1):29-35 [FREE Full text] [doi: 10.1016/j.jsat.2013.08.004] [Medline: 24035143]
69. Wright AG, Woods WC. Personalized models of psychopathology. Annu Rev Clin Psychol 2020 May 07;16:49-74. [doi: 10.1146/annurev-clinpsy-102419-125032] [Medline: 32070120]

70. Beltz AM, Wright AG, Sprague BN, Molenaar PC. Bridging the nomothetic and idiographic approaches to the analysis of clinical data. Assessment 2016 Aug;23(4):447-458 [FREE Full text] [doi: 10.1177/1073191116648209] [Medline: 27165092]
71. Molenaar PC. Granger causality testing with intensive longitudinal data. Prev Sci 2019 Apr;20(3):442-451. [doi: 10.1007/s11121-018-0919-0] [Medline: 29858760]

Abbreviations

AUC: area under the curve
DND: do not disturb
EMA: ecological momentary assessment
ILD: intensive longitudinal data
PA: physical activity
REDCap: Research Electronic Data Capture
SB: sedentary behavior
TIME: Temporal Influences on Movement & Exercise
UI: user interface
µEMA: microinteraction ecological momentary assessment

Edited by A Mavragani; submitted 09.08.21; peer-reviewed by D Fulford, J Lai; comments to author 29.09.21; revised version received 24.11.21; accepted 17.12.21; published 09.02.22.

Please cite as:
Ponnada A, Wang S, Chu D, Do B, Dunton G, Intille S
Intensive Longitudinal Data Collection Using Microinteraction Ecological Momentary Assessment: Pilot and Preliminary Results
JMIR Form Res 2022;6(2):e32772
URL: https://formative.jmir.org/2022/2/e32772
doi:10.2196/32772
PMID:35138253

Original Paper

Automated Pulmonary Embolism Risk Assessment Using the Wells Criteria: Validation Study

Nasen Jonathan Zhang[1], MD; Philippe Rameau[1]; Marsophia Julemis[2], MD; Yan Liu[1], PhD; Jeffrey Solomon[1]; Sundas Khan[1], MD; Thomas McGinn[1], MPH, MD; Safiya Richardson[1], MPH, MD

[1]Northwell Health, Manhasset, NY, United States

[2]Latin American School of Medicine, Havana, Cuba

Corresponding Author:
Nasen Jonathan Zhang, MD
Northwell Health
600 Community Dr
Manhasset, NY, 11020
United States
Phone: 1 (516) 470 3377
Fax: 1 (516) 600 1756
Email: nasenz@gmail.com

Abstract

Background: Computed tomography pulmonary angiography (CTPA) is frequently used in the emergency department (ED) for the diagnosis of pulmonary embolism (PE), while posing risk for contrast-induced nephropathy and radiation-induced malignancy.

Objective: We aimed to create an automated process to calculate the Wells score for pulmonary embolism for patients in the ED, which could potentially reduce unnecessary CTPA testing.

Methods: We designed an automated process using electronic health records data elements, including using a combinatorial keyword search method to query free-text fields, and calculated automated Wells scores for a sample of all adult ED encounters that resulted in a CTPA study for PE at 2 tertiary care hospitals in New York, over a 2-month period. To validate the automated process, the scores were compared to those derived from a 2-clinician chart review.

Results: A total of 202 ED encounters resulted in a completed CTPA to form the retrospective study cohort. Patients classified as "PE likely" by the automated process (126/202, 62%) had a PE prevalence of 15.9%, whereas those classified as "PE unlikely" (76/202, 38%; Wells score >4) had a PE prevalence of 7.9%. With respect to classification of the patient as "PE likely," the automated process achieved an accuracy of 92.1% when compared with the chart review, with sensitivity, specificity, positive predictive value, and negative predictive value of 93%, 90.5%, 94.4%, and 88.2%, respectively.

Conclusions: This was a successful development and validation of an automated process using electronic health records data elements, including free-text fields, to classify risk for PE in ED visits.

(JMIR Form Res 2022;6(2):e32230) doi:10.2196/32230

KEYWORDS

health informatics; pulmonary embolism; electronic health record; quality improvement; clinical decision support systems

Introduction

Computed tomography pulmonary angiography (CTPA) is the gold standard test for diagnosing patients with pulmonary embolism (PE), a potentially deadly condition that often presents with nonspecific signs and symptoms [1,2]. Fast, sensitive, and specific, CTPA use has rapidly proliferated since it supplanted ventilation-perfusion scanning for the diagnosis of PE in the 1990s [3,4]. While CTPA testing is widely available and easy to perform, its utility must be weighed against harm from ionizing radiation and intravenous contrast. Potential harm from CTPA includes a 14% risk of contrast-induced nephropathy and lifetime radiation-induced malignancy risk as high as 2.76% [5,6].

Clinical prediction rules, such as the Wells criteria for pulmonary embolism, or the Wells score, have been developed and widely validated to assist providers with the decision to perform CTPA [7]. By estimating pretest probability and

recommending CTPA only when suspicion is appropriately high, the use of such tools can reduce the number of tests performed without missing diagnoses of PE [8]. The incorporation of prediction rules into electronic health record (EHR) systems as clinical decision support (CDS) has been shown in multiple studies to significantly improve CTPA yield by 30% to 98% [9-11]. However, acceptance of CDS varies among physicians as CDS use is viewed as time-consuming [12]. At our institution, although users of a CDS tool incorporating the Wells score had CTPA yields of 38% higher than nonusers, the tool was dismissed in 65% of the cases [13]. Rather than requiring burdensome review of fragmented clinical data [14-16] and manual input of score components by providers, a CDS tool that presents a Wells score automatically calculated from existing EHR data could improve efficiency and usability and thereby tool acceptance [17,18].

Previous analysis of the Wells score concluded that it is less amenable to automatic calculation due to the inclusion of variables that either require clinical gestalt (PE as or more likely than alternative diagnosis) or are likely to be embedded in unstructured data (clinical signs and symptoms of deep venous thrombosis [DVT] and hemoptysis) [19]. Yet, using narrow definitions of the Wells score components based only on structured data can lead to decreased sensitivity for relevant clinical documentation [17]. The objective of our study was to design an automated process that incorporates information from unstructured data, and to validate its accuracy.

Methods

Study Setting and Design

This retrospective cohort study of emergency department (ED) encounters took place at 2 academic tertiary-care hospitals in New York. The study included all consecutive adult ED visits during the months of May and June of 2019 where a CTPA was performed to evaluate for PE. CTPA studies carried out for other indications, for example aortic dissection, were excluded. Patient demographic and clinical characteristics not part of the Wells Score were not collected. The study was approved by the institutional review board as minimal-risk research using data collected for routine clinical practice and the requirement for informed consent was waived. Data were collected from the enterprise EHR (Sunrise Clinical Manager, Allscripts, Chicago, IL, U.S.) reporting database.

Automatic Score Design

We designed the automated process for the Wells score calculation with usability for an ED-based CDS tool as our goal. The process was therefore limited to only information in each encounter that was recorded prior to the CTPA order. We incorporated all 7 components of the Wells score, which are clinical signs and symptoms of DVT (3 points), PE being as or more likely than other diagnoses (clinical gestalt, 3 points), heart rate greater than 100 beats per minute (1.5 points), immobilization for at least 3 days or surgery in the prior 4 weeks (1.5 points), previous objectively diagnosed PE or DVT (1.5 points), hemoptysis (1 point), and active malignancy (with treatment within 6 months or palliative, 1 point).

Clinical signs and symptoms consistent with DVT or hemoptysis were taken from the chief complaint fields of the ED nurse triage note. This note is completed on patient arrival, before assessment by a provider, and includes both free text and discrete options for documentation of chief complaint. By combining a list of anatomic terms describing parts of the lower extremity, such as “leg” or “thigh,” terms of laterality, and prefix and suffix descriptors, such as “pain in” or “edema” (Textbox 1), we generated a list of 192 search phrases (Table S1 in Multimedia Appendix 1) for the signs and symptoms of DVT component of the Wells score. The list included common abbreviations referring to the lower extremity, such as “LE,” and included indicators of laterality. A similar list of 7 phrases describing hemoptysis (“hemoptysis,” “coughing blood,” “coughing up blood,” “blood-tinged sputum,” “bloody sputum,” “blood in sputum,” and “blood in phlegm”) was created. These lists were supplemented with phrasing encountered during a preliminary review of ED nurse triage notes and ED provider notes from encounters with CTPA in a period prior to our period of study.

Textbox 1. Terms used in combination to generate list of key phrases indicating signs and symptoms of deep venous thrombosis.

List of anatomic terms describing parts of the lower extremity

- Prefix descriptors
  - edema to
  - swollen
  - pain in
- Indicators of laterality
  - bilateral
  - b/l
  - bl
  - right
  - left
  - L
  - R
- Anatomic terms
  - LE
  - lower extremity
  - leg
  - lower leg
  - thigh
  - calf
- Suffix descriptors
  - swelling
  - swollen
  - edema
  - pain
  - discomfort

We assumed 3 points for clinical gestalt for all encounters, assuming high concern for PE by the provider as a CTPA was performed. The remaining 4 components of the score were derived from structured data. For the heart rate criterion, the maximum value prior to the CTPA order was extracted, and 1.5 points were given if greater than 100. For the immobilization criterion, the EHR was queried for any intensive care unit stays and operative notes (specifying use of general anesthesia) within the preceding 30 days, as well as International Classification of Disease (ICD)-10 codes corresponding to quadriplegia. For history of PE/DVT, the problem list in Sunrise was queried for relevant ICD codes (for PE: ICD-9 codes 415.1, 415.11, 415.12, 415.13, 415.19, V12.55; ICD-10 code I26.99; and for DVT: ICD-9 codes 453.4-453.9, V12.51; ICD-10 code I82.409). For active malignancy, the Sunrise problem list was similarly queried for ICD-10 codes corresponding to a malignancy diagnostic group. The query was limited to problem list items documented prior to the index CTPA order. Once all score components became available, each encounter was classified as “PE likely” (Wells score greater than 4) or “PE unlikely” (Wells score less than or equal to 4) based on the two-tier model of risk stratification.

Chart Review

A 2-reviewer manual chart review was conducted to validate the automatically derived Wells scores. A review process and standardized data collection sheet were first designed by the senior investigator and were trialed by 2 clinician-investigators (NZ and PR) in a preliminary review using data from a period prior to our period of study. In this preliminary review, the data collection process was refined and standardized. Subsequently, 2 investigators independently reviewed data from the study period. The review of each chart included the entirety of available data, including vital signs, laboratory values, radiology reports, problem list, and nurse and provider notes. Clinician notes were reviewed if they were linked to the patient encounter even if documentation was completed after CTPA order. It was assumed that the provider would have been aware of all findings

documented in the history and physical exam before CTPA order. Three points were given for clinical gestalt in all cases. During the review, D-dimer ordering and the results of the CTPA (whether positive for PE) were also noted. Interreviewer agreement was measured by comparing risk classifications for encounters based on the Wells score using the Cohen kappa coefficient. Discrepancies were resolved by consensus.

Measures and Data Analysis

The automated Wells score components based on the queries designed as above were then compared to manually derived score components as the gold standard to arrive at sensitivity and specificity data for each component, as well as sensitivity, specificity, positive predictive value, and negative predictive value with respect to risk classification based on the two-tier model. Positive and negative predictive values for the automatic score were also calculated with regards to risk stratification as "PE likely," as this category is recommended to proceed directly to CTPA. For the lower risk category, a D-dimer is recommended to be performed first, and if normal, to stop further PE evaluation. To assess the ability of the automated score to stratify risk, we calculated the CTPA yield for each automated risk category. CTPA yield was calculated as the number of PE diagnoses divided by the number of CTPA exams. Ordering of the CTPA study was considered guideline concordant if the patient either had a Wells score in the "PE likely" category or in the "PE unlikely" category and if a D-dimer was subsequently ordered and was above the upper limit of normal (>230 ng/mL). Otherwise, the CTPA order was considered not guideline concordant. All data analysis was performed in Microsoft Excel.

Results

Prevalence of PE and Score Components

A total of 202 ED encounters resulted in a completed CTPA to form the retrospective study cohort. There was a high interreviewer agreement with a Cohen kappa 0.93. Based on chart review of CTPA results, the overall prevalence of PE was 12.9% (26/202 encounters). Patients classified as "PE unlikely" by the automated Wells score had a PE prevalence of 7.9% (6/76 encounters), whereas those classified as "PE likely" had a PE prevalence of 15.9% (20/126 encounters). This compares to 8.1% (6/74) and 15.6% (20/128) based on chart review. By chart review, the prevalence of positive Wells score components ranged from 3.5% (7/202) for hemoptysis to 44% (87/202) for tachycardia (Table 1).

Table 1. Proportion of Wells score components captured by the automated process.

Wells score components	Present on chart review, n/N (prevalence rate %)	Captured by automated process, n/N (sensitivity rate %)	Erroneously captured, n/N (false positive rate %)
Clinical signs or symptoms of DVT[a] (3 points)	28/202 (14)	15/28 (54)	4/174 (2.3)
Pulse >100 (1.5 points)	87/202 (44)	87/87 (100)	1/115 (0.9)
Immobilization x 3 days or surgery in previous 4 weeks (1.5 points)	22/202 (11)	15/22 (68)	0/180 (0)
History of PE[b] or DVT (1.5 points)	32/202 (16)	29/32 (91)	5/170 (2.9)
Hemoptysis (1 point)	7/202 (3.5)	6/7 (86)	1/195 (0.5)
Malignancy with treatment within 6 months or palliative (1 point)	47/202 (23)	41/47 (87)	8/155 (5.2)
PE as likely or more than alternate diagnoses (3 points, assumed to be true)	202/202 (100)	202/202 (100)	0/0 (0)

[a]DVT: deep venous thrombosis.

[b]PE: pulmonary embolism.

Sensitivity of the Automated Process

Of these components, the sensitivity of the automated process ranged from 54% (15/28) for signs and symptoms of DVT, to 100% (87/87) for pulse greater than 100. Of the 13 instances where signs and symptoms of DVT were missed by the automated process (out of the 28 found by chart review), 1 was due to a description not covered by our search strategy and 12 were due to descriptions not being present in the "Chief Complaint" field of the ED nurse triage note, but only in the "History of Present Illness" or "Physical Exam" field of the ED provider note (Table S2 in Multimedia Appendix 1). Moreover, 7 out of 22 instances of immobility were only described in the "History of Present Illness" fields of ED provider notes and not captured in intensive care unit stays, operative notes, or quadriplegia diagnoses.

Specificity of the Automated Process

False positive rates were low across all automated score components, ranging from 0% (0/180) for immobility to 5.2% (8/155) for active malignancy, corresponding to specificities of 94.8% to 100%. Several false positive findings of diagnoses of PE/DVT or active malignancy were due to erroneous entries in the problem list for the former, and inactive, past diagnoses for the latter. For all individual score components, excluding clinical gestalt, which was assumed positive in all cases, overall accuracy was 96% (1163/1212).

Overall Performance of the Automated Process

With respect to classification of the patient as "PE likely," the automated process achieved an accuracy of 92.1% (186 correct classifications out of 202 encounters) when compared to chart review, with sensitivity of 93% (119/128) and specificity of 90% (67/74) (Table 2 and Table 3). Positive predictive value was 94.4% (119/126), and negative predictive value was 88% (67/76). Out of a total of 202 patient encounters, there were 16 (7.9%) instances where there was discrepancy between automated and manual classifications. Moreover, 9 false negatives included 5 where signs and symptoms of DVT were present but not mentioned in the ED nurse triage note, 2 where history of PE was not documented in the problem list but was described by the ED provider note, and 2 where patients had recent surgeries described in notes but were not captured by operative notes in the EHR. In addition, 7 false positives included 3 where a search phrase was present in the ED nurse triage note but was preceded by words of negation, 3 due to erroneous entries of PE or DVT in the EHR problem list, and 1 due to an erroneous pulse entry.

Table 2. Performance of the automated risk classifications—confusion matrix.

Total 202 encounters	PE[a] likely by chart review (n=128)	PE unlikely by chart review (n=74)
PE likely by automated process	119	7
PE unlikely by automated process	9	67

[a]PE: pulmonary embolism.

Table 3. Automated process performance measures.

Performance measure	Formula	Value
Accuracy	(119 + 67) / (119 + 7 + 9 + 67)	92.1%
Sensitivity	119 / (119 + 9)	93.0%
Specificity	67 / (67 + 7)	90.5%
Positive predictive value	119 / (119 + 7)	94.4%
Negative predictive value	67 / (67 + 9)	88.2%

Guideline Concordance

Based on the automated process, 151 of 202 CTPA orders were guideline concordant, resulting in a concordance rate of 74.8%, compared to 153 of 202 (75.7%) based on chart review. Of the 76 cases classified as "PE unlikely" by the automated Wells score, 28 (37%) had a D-dimer ordered, with 25 (33%) resulting above the upper limit of normal. The 3 cases (1.5%) where D-dimer was within normal range and the 48 cases (24%) where no D-dimer was ordered prior to CTPA were considered nonguideline concordant.

Discussion

Principal Findings

In this study we created an automated process to calculate Wells score with the aim of improving CDS tool usability and evaluating provider guideline concordance. We found that our process was 96% accurate with respect to individual instances of score components, and 92% accurate with respect to risk classification when compared to a manual chart review standard. The process achieved high positive predictive value (94.4%) while preserving negative predictive value (88%). To address 2 important score components that tend to reside outside of structured EHR elements (signs and symptoms of DVT and hemoptysis), we designed an innovative key phrase search method that made use of free-text fields in notes without requiring advanced natural language processing (NLP) techniques. The automated score was able to stratify risk within ED encounters where CPTA was performed, with cases classified as "PE likely" having a CTPA yield that is double the yield of cases classified as "PE unlikely."

Significant prior work has been carried out for the automated calculation of clinical risk prediction scores using EHR data. Much success has been achieved for scores whose variables can be exclusively derived from structured data, such as CURB-65 (confusion, uremia, respiratory rate, BP, age ≥65 years) for pneumonia severity [20], CHADS2-VASc (congestive heart failure, hypertension, age ≥75 years, diabetes mellitus, stroke, vascular disease, age 65-74 years, sex category [female]) for stroke risk [21], PESI (pulmonary embolism severity index) for PE severity [22], SOFA (sequential organ failure assessment) in sepsis [23], and the Padua prediction score for risk of venous thromboembolism [16,24]. These studies in general found very high degrees of accuracy when compared to scores derived manually, with near perfect accuracy with regards to scalar variables such as vital signs or laboratory values.

Certain isolated areas of lower sensitivity and specificity are related to the incompleteness of administrative and EHR databases of diagnoses—a well-recognized issue [24,25]—as well as the temporality of EHR-documented diagnoses, due to the accumulation of resolved, inactive medical problems and incomplete documentation of recent, acute ones. For example, Navar-Boggan et al [20] found lower specificity for active stroke risk factors in the CHA2DS2-VASc score due to detection of historical, resolved diagnoses, and Pavon et al [15] found lower sensitivity for active cancer, infection, myocardial infarction, and stroke diagnoses for the Padua score due to election of only using items from the "Admitting Diagnoses" list. The issue of

temporality also affected the active malignancy component of our automatic score, which resulted in 8 false positives due to the detection of inactive, resolved cancer. Using EHR problem list diagnoses relies on their accuracy, which may introduce error [18], as it did in our score when false positive history of PE or DVT were introduced by erroneous entries, although the false positive rate was very low (2.9%).

More challenging is the extraction of information necessary for variable components more likely to be found in unstructured data, such as signs and symptoms of DVT in our study. Significant work has been carried out to address this issue, with Grouin et al [26] using NLP techniques to extract criteria for the CHA2DS2-VASc score from clinical notes, including negation and speculation handling, achieving an accuracy of 97.6% for score components and 85.7% for scores. Deleger et al [27] similarly used NLP techniques to extract data from ED physician notes and incorporated structured lab values to calculate an automated pediatric appendicitis score. The automated score achieved a sensitivity of 86.9% and positive predictive value of 86.3% when compared to manual chart review. Without the option to implement advanced NLP techniques on EHR data in real time, we devised a less complex and more transparent search strategy using a list of phrases generated from combinations of possible descriptors of signs and symptoms of DVT and hemoptysis. With the additional constraint of being only applied to EHR documentation present prior to ED CTPA ordering, our process achieved sensitivity, specificity, and overall accuracy of 54%, 98%, and 90.6% percent for signs and symptoms of DVT, and 86%, 99%, and 96.5% for hemoptysis when compared to chart review.

While our automated process shows remarkable promise for use in CDS for CTPA ordering for PE, given its high accuracy and positive and negative predictive values, it also reveals the limits of automated information retrieval. False positives and negatives are likely to occur when applied to hundreds of cases, and the automated process is naive regarding information not documented in its targeted fields. Therefore, in addition to further refining our automated process, the practical limitations of its information retrieval may be better remedied with a hybrid approach combining automated variables that are captured with high accuracy and manual input variables, as suggested by Pavon et al [15] in their study of automated Padua scores for venous thromboembolism risk classification. Additionally, inaccuracies in information retrieval need to be checked by human knowledge. After the clinician interviews and examines the patient, he or she is likely to have gathered pertinent information not captured by the EHR, information that can rectify false positives and false negatives presented by automated information retrieval. Ideally, a CDS tool incorporating the automated Wells score should also display the source of each extracted element to create transparency and enable validation by the human user.

Limitations

Our study has several limitations. We did not study ED encounters where PE may have been considered as a diagnosis but CTPA was ultimately not performed; therefore we do not know the performance of our score in such patients. Our study was retrospective in nature, relying on manual chart review to establish a standard against which our automatic score was compared. Although our chart review was thorough, it likely represents an imperfect standard due to likely incomplete documentation of information available to the provider at the time of CTPA ordering. Our study was performed using data from 2 urban tertiary care hospitals within a single New York health care system, using Allscripts Sunrise EHR, and may not be generalizable to other settings.

Conclusions

We successfully designed and validated an automated process to calculate a Wells criteria score for PE to risk-stratify patients prior to CTPA. Our study achieved high accuracy as well as positive and negative predictive values, demonstrating its potential for use in augmenting CDS tools with automated information retrieval. When implemented in a CDS tool, our score could serve as the foundation for a hybrid approach combining automated prepopulation of variables with the option for manual input by the provider. Our method was novel in using a keyword search strategy using a list of phrases formed from combinations of terms used to describe signs and symptoms of DVT, applied to existing EHR notes prior to CTPA ordering.

Acknowledgments

This study was supported by the Agency for Healthcare Research and Quality (grant 1 R18 HS026196) and the National Heart, Lung, and Blood Institute (grant 1K23HL145114).

Authors' Contributions

NJZ, PR, MJ, and SR contributed to the study concept and design. NJZ, PR, and YL contributed to data acquisition. NJZ and SR contributed to analysis and interpretation of the data. NJZ, PR, and SR contributed to the drafting of the manuscript. All authors contributed to the critical revision of the manuscript. NJZ and SR contributed to statistical expertise. TM and SR contributed to the acquisition of funding.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplemental material.
[DOCX File , 17 KB - formative_v6i2e32230_app1.docx]

References

1. Bajaj N, Bozarth AL, Guillot J, Kojokittah J, Appalaneni SR, Cestero C, et al. Clinical features in patients with pulmonary embolism at a community hospital: analysis of 4 years of data. J Thromb Thrombolysis 2014 Apr 17;37(3):287-292. [doi: 10.1007/s11239-013-0942-8] [Medline: 23681675]
2. Perera M, Aggarwal L, Scott IA, Cocks N. Underuse of risk assessment and overuse of computed tomography pulmonary angiography in patients with suspected pulmonary thromboembolism. Intern Med J 2017 Oct 09;47(10):1154-1160. [doi: 10.1111/imj.13524] [Medline: 28635149]
3. Stein PD, Woodard PK, Weg JG, Wakefield TW, Tapson VF, Sostman HD, PIOPED II Investigators. Diagnostic pathways in acute pulmonary embolism: recommendations of the PIOPED II Investigators. Radiology 2007 Jan;242(1):15-21. [doi: 10.1148/radiol.2421060971] [Medline: 17185658]
4. Burge A, Freeman K, Klapper P, Haramati L. Increased diagnosis of pulmonary embolism without a corresponding decline in mortality during the CT era. Clin Radiol 2008 Apr;63(4):381-386. [doi: 10.1016/j.crad.2007.10.004] [Medline: 18325357]
5. Mitchell A, Jones A, Tumlin J, Kline J. Prospective study of the incidence of contrast-induced nephropathy among patients evaluated for pulmonary embolism by contrast-enhanced computed tomography. Acad Emerg Med 2012 Jun;19(6):618-625 [FREE Full text] [doi: 10.1111/j.1553-2712.2012.01374.x] [Medline: 22687176]
6. Niemann T, Zbinden I, Roser H, Bremerich J, Remy-Jardin M, Bongartz G. Computed tomography for pulmonary embolism: assessment of a 1-year cohort and estimated cancer risk associated with diagnostic irradiation. Acta Radiol 2013 Sep;54(7):778-784. [doi: 10.1177/0284185113485069] [Medline: 23761544]
7. Wells PS, Anderson DR, Rodger M, Stiell I, Dreyer JF, Barnes D, et al. Excluding pulmonary embolism at the bedside without diagnostic imaging: management of patients with suspected pulmonary embolism presenting to the emergency department by using a simple clinical model and d-dimer. Ann Intern Med 2001 Jul 17;135(2):98-107. [doi: 10.7326/0003-4819-135-2-200107170-00010] [Medline: 11453709]
8. Wang RC, Bent S, Weber E, Neilson J, Smith-Bindman R, Fahimi J. The Impact of Clinical Decision Rules on Computed Tomography Use and Yield for Pulmonary Embolism: A Systematic Review and Meta-analysis. Ann Emerg Med 2016 Jun;67(6):693-701.e3. [doi: 10.1016/j.annemergmed.2015.11.005] [Medline: 26747217]
9. Mills AM, Ip IK, Langlotz CP, Raja AS, Zafar HM, Khorasani R. Clinical decision support increases diagnostic yield of computed tomography for suspected pulmonary embolism. Am J Emerg Med 2018 Apr;36(4):540-544 [FREE Full text] [doi: 10.1016/j.ajem.2017.09.004] [Medline: 28970024]
10. Raja AS, Ip IK, Prevedello LM, Sodickson AD, Farkas C, Zane RD, et al. Effect of computerized clinical decision support on the use and yield of CT pulmonary angiography in the emergency department. Radiology 2012 Feb;262(2):468-474 [FREE Full text] [doi: 10.1148/radiol.11110951] [Medline: 22187633]
11. Deblois S, Chartrand-Lefebvre C, Toporowicz K, Chen Z, Lepanto L. Interventions to Reduce the Overuse of Imaging for Pulmonary Embolism: A Systematic Review. J Hosp Med 2018 Jan;13(1):52-61. [doi: 10.12788/jhm.2902] [Medline: 29309438]
12. Drescher FS, Chandrika S, Weir ID, Weintraub JT, Berman L, Lee R, et al. Effectiveness and acceptability of a computerized decision support system using modified Wells criteria for evaluation of suspected pulmonary embolism. Ann Emerg Med 2011 Jun;57(6):613-621. [doi: 10.1016/j.annemergmed.2010.09.018] [Medline: 21050624]
13. Richardson S, Cohen S, Khan S, Zhang M, Qiu G, Oppenheim MI, et al. Higher Imaging Yield When Clinical Decision Support Is Used. J Am Coll Radiol 2020 Apr;17(4):496-503 [FREE Full text] [doi: 10.1016/j.jacr.2019.11.021] [Medline: 31899178]
14. Mamykina L, Vawdrey DK, Stetson PD, Zheng K, Hripcsak G. Clinical documentation: composition or synthesis? J Am Med Inform Assoc 2012 Nov 01;19(6):1025-1031 [FREE Full text] [doi: 10.1136/amiajnl-2012-000901] [Medline: 22813762]
15. Pavon J, Sloane R, Pieper C, Colón-Emeric CS, Cohen H, Gallagher D, et al. Automated versus Manual Data Extraction of the Padua Prediction Score for Venous Thromboembolism Risk in Hospitalized Older Adults. Appl Clin Inform 2018 Jul;9(3):743-751 [FREE Full text] [doi: 10.1055/s-0038-1670678] [Medline: 30257260]
16. Bates DW, Kuperman GJ, Wang S, Gandhi T, Kittler A, Volk L, et al. Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality. J Am Med Inform Assoc 2003;10(6):523-530 [FREE Full text] [doi: 10.1197/jamia.M1370] [Medline: 12925543]
17. Koziatek CA, Simon E, Horwitz LI, Makarov DV, Smith SW, Jones S, et al. Automated Pulmonary Embolism Risk Classification and Guideline Adherence for Computed Tomography Pulmonary Angiography Ordering. Acad Emerg Med 2018 Sep 25;25(9):1053-1061 [FREE Full text] [doi: 10.1111/acem.13442] [Medline: 29710413]
18. Aakre C, Dziadzko M, Keegan M, Herasevich V. Automating Clinical Score Calculation within the Electronic Health Record. Appl Clin Inform 2017 Dec 21;08(02):369-380. [doi: 10.4338/aci-2016-09-ra-0149]

19. Jones BE, Jones J, Bewick T, Lim WS, Aronsky D, Brown SM, et al. CURB-65 pneumonia severity assessment adapted for electronic decision support. Chest 2011 Jul;140(1):156-163 [FREE Full text] [doi: 10.1378/chest.10-1296] [Medline: 21163875]
20. Navar-Boggan AM, Rymer JA, Piccini JP, Shatila W, Ring L, Stafford JA, et al. Accuracy and validation of an automated electronic algorithm to identify patients with atrial fibrillation at risk for stroke. Am Heart J 2015 Jan;169(1):39-44.e2. [doi: 10.1016/j.ahj.2014.09.014] [Medline: 25497246]
21. Vinson D, Morley J, Huang J, Liu V, Anderson M, Drenten CE, et al. The accuracy of an electronic Pulmonary Embolism Severity Index auto-populated from the electronic health record. Appl Clin Inform 2017 Dec 19;06(02):318-333. [doi: 10.4338/aci-2014-12-ra-0116]
22. Aakre C, Franco PM, Ferreyra M, Kitson J, Li M, Herasevich V. Prospective validation of a near real-time EHR-integrated automated SOFA score calculator. Int J Med Inform 2017 Jul;103:1-6. [doi: 10.1016/j.ijmedinf.2017.04.001] [Medline: 28550994]
23. Elias P, Khanna R, Dudley A, Davies J, Jacolbia R, McArthur K, et al. Automating Venous Thromboembolism Risk Calculation Using Electronic Health Record Data upon Hospital Admission: The Automated Padua Prediction Score. J Hosp Med 2017 Apr;12(4):231-237. [doi: 10.12788/jhm.2714] [Medline: 28411291]
24. Quan H, Li B, Saunders L, Parsons GA, Nilsson CI, Alibhai A, IMECCHI Investigators. Assessing validity of ICD-9-CM and ICD-10 administrative data in recording clinical conditions in a unique dually coded database. Health Serv Res 2008 Aug;43(4):1424-1441 [FREE Full text] [doi: 10.1111/j.1475-6773.2007.00822.x] [Medline: 18756617]
25. Weiskopf NG, Hripcsak G, Swaminathan S, Weng C. Defining and measuring completeness of electronic health records for secondary use. J Biomed Inform 2013 Oct;46(5):830-836 [FREE Full text] [doi: 10.1016/j.jbi.2013.06.010] [Medline: 23820016]
26. Grouin C, Deléger L, Rosier A, Temal L, Dameron O, Van Hille P, et al. Automatic computation of CHA2DS2-VASc score: information extraction from clinical texts for thromboembolism risk assessment. AMIA Annu Symp Proc 2011;2011:501-510 [FREE Full text] [Medline: 22195104]
27. Deleger L, Brodzinski H, Zhai H, Li Q, Lingren T, Kirkendall ES, et al. Developing and evaluating an automated appendicitis risk stratification algorithm for pediatric patients in the emergency department. J Am Med Inform Assoc 2013 Dec;20(e2):e212-e220 [FREE Full text] [doi: 10.1136/amiajnl-2013-001962] [Medline: 24130231]

Abbreviations

CDS: clinical decision support
CHADS2-VASc: congestive heart failure, hypertension, age ≥75 years, diabetes mellitus, stroke, vascular disease, age 65-74 years, sex category (female)
CTPA: computed tomography pulmonary angiography
CURB-65: confusion, uremia, respiratory rate, BP, age ≥65 years
DVT: deep venous thrombosis
ED: emergency department
EHR: electronic health record
ICD: International Classification of Disease
NLP: natural language processing
PE: pulmonary embolism
PESI: pulmonary embolism severity index
SOFA: sequential organ failure assessment

Edited by G Eysenbach; submitted 22.07.21; peer-reviewed by S Prior, J Walsh; comments to author 26.10.21; revised version received 05.12.21; accepted 27.12.21; published 28.02.22.

Please cite as:
Zhang NJ, Rameau P, Julemis M, Liu Y, Solomon J, Khan S, McGinn T, Richardson S
Automated Pulmonary Embolism Risk Assessment Using the Wells Criteria: Validation Study
JMIR Form Res 2022;6(2):e32230
URL: https://formative.jmir.org/2022/2/e32230
doi:10.2196/32230
PMID:35225812

Original Paper

Mobile Phone Apps for Intimate Partner and Sexual Violence Prevention and Response: Systematic Search on App Stores

Jessica Draughon Moret[1*], RN, PhD; Angela Todd[1], FNP-BC, RN, PhD; Lauren Rose[2], BS; Erin Pollitt[3], SANE-A, RN, BSN, MHA; Jocelyn Anderson[2*], SANE-A, RN, PhD

[1]Betty Irene Moore School of Nursing, University of California, Davis, Sacramento, CA, United States

[2]College of Nursing, Pennsylvania State University, University Park, PA, United States

[3]District of Columbia Forensic Nurse Examiners, Washington, DC, United States

[*]these authors contributed equally

Corresponding Author:
Jessica Draughon Moret, RN, PhD
Betty Irene Moore School of Nursing
University of California, Davis
2570 48th Street
Suite 2600
Sacramento, CA, 95817
United States
Phone: 1 916 734 0511
Fax: 1 916 452 2112
Email: jdmoret@ucdavis.edu

Abstract

Background: Since the 2008 advent of the smartphone, more than 180 billion copies of apps have been downloaded from Apple App Store, with more than 2.6 million apps available for Android and 2.2 million apps available for iOS. Many violence prevention and response apps have been developed as part of this app proliferation.

Objective: This study aims to evaluate the prevalence and quality of freely available mobile phone apps targeting intimate partner violence (IPV) and sexual violence (SV) prevention and response.

Methods: We conducted a systematic search of violence prevention and response mobile phone apps freely available in Apple App Store (iOS; March 2016) and Google Play Store (Android; July 2016). Search terms included violence prevention, sexual assault, domestic violence, intimate partner violence, sexual violence, forensic nursing, wife abuse, and rape. Apps were included for review if they were freely available, were available in English, and had a primary purpose of prevention of or response to SV or IPV regardless of app target end users.

Results: Using the Mobile Application Rating Scale (MARS), we evaluated a total of 132 unique apps. The majority of included apps had a primary purpose of sharing information or resources. Included apps were of low-to-moderate quality, with the overall subjective quality mean for the reviewed apps being 2.65 (95% CI 2.58-2.72). Quality scores for each of the 5 MARS categories ranged from 2.80 (engagement) to 4.75 (functionality). An incidental but important finding of our review was the difficulty in searching for apps and the plethora of nonrelated apps that appear when searching for keywords such as “rape” and “domestic violence” that may be harmful to people seeking help.

Conclusions: Although there are a variety of mobile apps available designed to provide information or other services related to SV and IPV, they range greatly in quality. They are also challenging to find, given the current infrastructure of app store searches, keyword prioritization, and highlighting based on user rating. It is important for providers to be aware of these resources and be knowledgeable about how to review and recommend mobile phone apps to patients, when appropriate.

(JMIR Form Res 2022;6(2):e28959) doi:10.2196/28959

KEYWORDS

rape; intimate partner violence; gender-based violence; smartphone; mobile phone app

Introduction

Both sexual violence (SV) and intimate partner violence (IPV) continue to be major public health problems in the United States and worldwide. Every 68 seconds, a US resident is sexually assaulted [1]. In the 2015 National Intimate Partner and Sexual Violence Survey, 1 in 3 (33%) people reported experiencing lifetime physical violence or SV [2]. When psychological abuse is considered, the numbers are closer to 1 in 2 (50%) people. Rates are often higher in lesbian, gay, bisexual, transgender, and queer (LGBTQ) populations, with 50% of transgender people, 61% of bisexual women, 44% of lesbian women, 37% of bisexual men, and 26% of gay men reporting an experience of IPV in their lifetime [3,4].

Documented direct and indirect health outcomes linked to IPV and SV include physical, mental, and sexual health sequelae. In addition to physical injury, these types of violence are associated with chronic stress, chronic immune system activation, and inflammation [5]; accelerated cellular aging [6]; and cardiovascular disease risk [7]. Depression, acute stress disorder, and posttraumatic stress disorder are common comorbidities [8]. IPV and SV are also associated with substance use, alcohol use, and sexual risk taking, all of which are documented risk factors for HIV and other sexually transmitted infections. In addition, experience of SV and IPV is associated with diminished control over sexual and reproductive health decisions [9], unplanned pregnancies, preterm labor, low-birthweight babies, and maternal morbidity and mortality [10].

Although rates of SV and IPV have remained relatively stable over the past decade, the ways that people access and gather information have changed. Social media and other mobile apps are most people's preferred source of information [11]. Smartphones are ubiquitous among adolescents and young adults, with 98% of Generation Z owning a smartphone [12]—overlapping with those at highest risk for sexual assault ages 12-34 years [13]. As researchers and clinicians working on SV and IPV, we recognized this shift toward internet- and smartphone-available information was imminent in our field. For example, in 2012, early media reports of the UAskDC app garnered attention, showing that these spaces were being utilized and that information in apps could be vetted in partnership with reputable health and advocacy service providers [14-16].

The UAskDC app effectively curated the many disparate resources from each higher education campus Title IX and student affairs office, and health, advocacy, and criminal justice services across the District of Columbia into 1 place. The app provides more accurate and trauma-informed information than a Google search for "rape" or "sexual assault" and "District of Columbia" would, and the platform allows for rapid updates. Resource sharing apps, such as UAskDC, are focused on secondary and tertiary prevention—connecting a survivor to resources for safety and health. These community-specific apps, while designed for potential survivors/patients, are also an invaluable resource for health care providers, friends, and family members who may be trying to direct a patient or loved one to appropriate resources. Because most health care providers are not experts in IPV or SV, receiving an average of only 1-5 hours of training in these topics during their prelicensure training [17], resources used by health systems (eg, handouts, apps, and websites) become heavily-relied-upon sources of information. Therefore, it is of utmost importance that the quality of these apps be known and maintained to achieve their goals.

Although IPV and SV prevention and response apps are widely available, the literature focused on these apps remains limited. Of the studies including IPV- or SV-related apps, most examined apps that were directed at college-aged women offering resources for use during or after SV to support safety and decision making [18-21], while 2 were directed toward education in recognition and prevention of child sexual abuse and trafficking [22,23]. Overall, these studies found potential for IPV- or SV-related apps to educate users about prevention, recognition, harm reduction, safety measures, and resources for victims of IPV or SV [18-23].

Perhaps the most documented IPV app in the literature is MyPlan [24]. MyPlan draws on elements of social cognitive and decision-making theories through self-monitoring, social support, and priority setting [25]. MyPlan further integrates safety-planning strategies and tools used by IPV advocates for decades [26,27]. This app allows survivors to evaluate their relationship and safety while designing a plan tailored to their individual needs and simultaneously receiving resources with embedded links. It allows survivors to return to their plan and review and update information over time to coincide with changes within their abusive relationship. In prospective clinical trials, MyPlan and its precursor, the computer-based decision aid Internet Resource for Information and Safety (IRIS), both found improvements in decisional conflict, use of relationship safety strategies, and ending unsafe relationships [28-30].

Given the proliferation of apps and our prior experience developing and testing a mobile app for IPV and SV response, we are aware of the multiple challenges with app dissemination and maintenance [20]. Therefore, this paper aims to determine the prevalence and quality of freely available mobile smartphone apps that include a primary goal of addressing prevention and response. A secondary aim was to determine priority recommendations for health care providers interested in integrating mobile apps within patient care.

Methods

Study Design

We conducted a search of Apple App Store (March 2016) and Google Play Store (July 2016) using the following search terms: violence prevention, sexual assault, domestic violence, intimate partner violence, sexual violence, forensic nursing, wife abuse, and rape. Complete lists of results were downloaded to Microsoft Excel for review. SV and IPV apps were included in this analysis as they both commonly co-occur (approximately 18% of women and 8% of men report lifetime intimate partner sexual violence [31]) and are commonly addressed by services that are colocated or multipurpose (eg, a community's IPV shelter also provides rape crisis accompaniment services to health care facilities).

Titles were reviewed by a member of the study team to determine whether inclusion criteria were met. When the title was unclear, they continued to the next step, which was review of the app's general information available in the publisher's app store without downloading the app. For apps that clearly met the inclusion criteria or in which it was unclear from the information available in the app store Information section, we continued to the final step: full review via download of the app to a mobile phone or tablet.

Inclusion criteria for our analysis were (1) available in English, (2) free version available, and (3) directed toward 1 of the following audiences: the general public at risk for SV or IPV, people who have experienced IPV or SV, or health or advocacy providers who work with people who experience violence.

App Review

Apps were reviewed for quality using the Mobile Application Rating Scale (MARS) [32]. Since its initial publication in 2015, MARS has been used to evaluate the quality of smartphone apps on a wide range of health-related topics. These include health promotion topics, such as fitness [33], nutrition and weight management [34-36], mental health [37], and mindfulness [38], as well as self-management of medical conditions, such as diabetes [39], sleep disorders [40], pain management [41], heart failure [42], and asthma [43].

For this analysis, modifications to MARS were made to ensure fit for the SV and IPV content area. These included using "sexual or intimate partner violence" to fill in the content areas for the target health behavior in the Perceived Impact of Health Behavior Change section, including instructions for categorizing violence advocacy and service agencies when addressing the item on credibility, and adding features that we knew to be potentially common or relevant to the goals of violence-specific apps (eg, Global Positioning System [GPS], linking to service providers, emergency exit features). Our full data collection instrument is available in Multimedia Appendix 1.

All data were entered into an internet-based survey form, which also collected date and time information as well as which research team member was entering the data. At the search onset, the team selected 4 apps to all independently review and discuss during a team meeting to create shared definitions and consistency within the team. Subsequently, each app was reviewed by 1 team member with consultation to the team, as needed. Apple platform apps were reviewed between April 2016 and September 2016, and Google platform apps were reviewed between October 2017 and February 2018.

Data Analysis

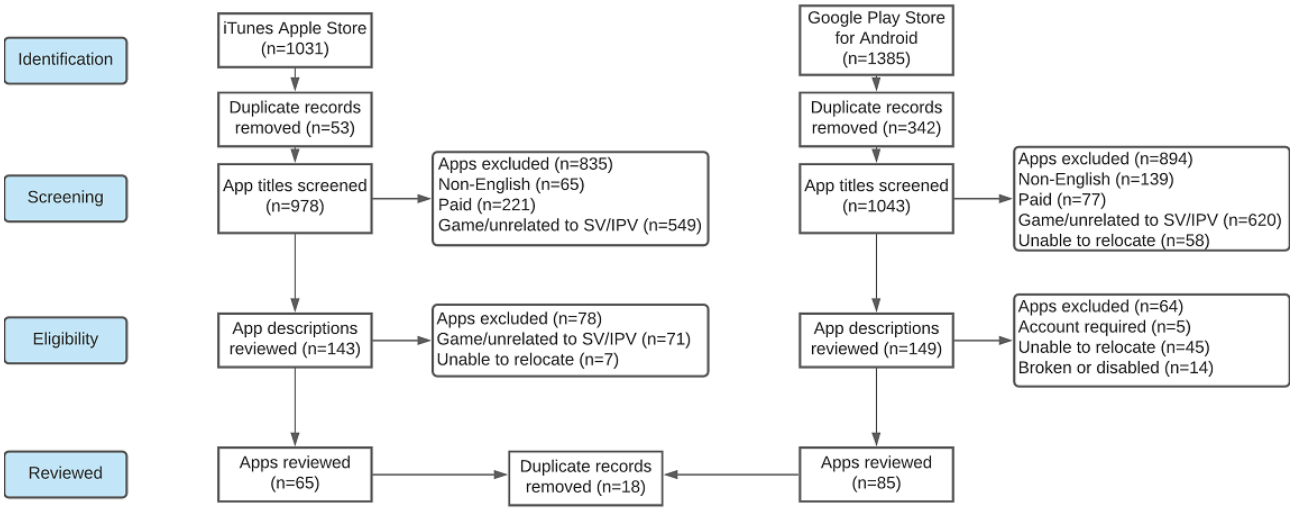
Descriptive analyses were completed to summarize the apps reviewed. Mean scores were calculated for each of the MARS categories (engagement, functionality, aesthetics, information, and subjective quality). Individual item frequencies and proportions were also calculated for nonscale items. Data were analyzed using SPSS Statistics version 25 (IBM Corp.) [44]. App classification data were recoded by consensus of 2 of the team members (authors JCA and LR) to better summarize results and due to the large number of "Other" responses in some categories (initial *n*=38, 28.8% for focus area; *n*=14, 10.6% for theoretical background/strategies) based on the initial use of MARS with minor modifications.

Results

Search Results

In our initial searches, 978 unique apps were identified from Apple App Store and 1043 from Google Play Store. Of the app titles screened, 835 (85.4%) Apple apps and 894 (85.7%) Google apps were excluded (see Figure 1 for the search flow diagram), resulting in 143 (14.6%) Apple apps and 149 (14.3%) Google apps for store description abstract information review. Following this step, 65 (45.5%) Apple and 85 (57%) Google apps met the inclusion criteria and remained for full analysis. Of these 150, 18 (12%) apps appeared in both app stores, and duplicates were removed from the list, resulting in 132 (86%) apps in the final analysis.

Figure 1. Flow diagram of the app search and inclusion process. IPV: intimate partner violence; SV: sexual violence.



Although it was not an aim of this project, an additional important finding came to light during the app store search process. Unlike typical search engines (eg, Google) or research databases (eg, PubMed), Boolean operators ("and," "or") do not work when searching in app stores. Searching in app stores is heavily reliant on developer tagging of titles and keywords, ratings and reviews from other users, and advertising/marketing monies spent to promote apps [45,46]. Because of how these search functions work, searching for the word "rape" in Apple App Store brings up hundreds of results for voice changer and rap music apps (presumably because of the similar spelling of "rape" and "rap"). Searching for sexual assault, synonymous but more formal terminology, in the same Apple App Store brings up only a fraction of the results and far more that are designed for SV providers, advocates, prevention, and response. These differences in search strategies and logic have important implications for providers who may be recommending use of apps to individuals they work with.

Descriptives

Table 1 summarizes the apps' targeted age groups, features, focus areas, behavior change strategies, and organizational affiliations. Over three-quarters (101/132, 76.4%) of the apps focused on resource or information sharing, with at least 10% of apps focusing on each of the following: IPV (37/132, 28%), crisis intervention or mental health (29/132, 22.2%), sexual assault (18/132, 13.6%), relationship conflict/health (17/132, 12.9%), and peer support (15/132, 11.4%). Apps primarily targeted adults (87/132, 65.9%), young adults (36/132, 27.3%), or no specific age group (31/132, 23.5%). In addition, 1 in 3 (45/132, 34.1%) apps was affiliated with a nonprofit or nongovernment agency and 1 in 4 (36/132, 27.3%) with a government agency. Nearly 1 in 5 (26/132, 19.7%) apps had developers or content that did not allow their affiliation to be determined. The most common features or functions observed in the reviewed apps included requiring internet service to operate (21/132, 15.9%), location or GPS services (20/132, 15.2%), and emergency exit/panic features (17/132, 12.9%); however, these were still only present in 13-20 (10%-15%) apps. Additional functions, such as logins, passwords, and reminders, were each present in a minority of apps. Apps overwhelmingly used information and education as a behavior change (104/132, 78.8%), with safety monitoring/tracking (15/132, 11.4%) and goal setting/safety planning (14/132, 10.6%), each following with approximately 1-10 (0.7%-7.6%) apps using these strategies.

Table 1. App overview information (N=132).

Category	Apps, n (%)
App focus areas[a]	
Resource or information sharing	100 (76.4)
IPV[b]	37 (28.0)
Crisis intervention/mental health	29 (22.2)
Sexual assault	18 (13.6)
Relationship conflict/health	17 (12.9)
Peer support	15 (11.4)
General violence risk	9 (6.8)
Behavior change	6 (4.5)
Goal setting	6 (4.5)
Physical health	3 (2.3)
Safety planning	24 (18.2)
Entertainment	1 (0.8)
Education	4 (3.2)
Legal	2 (1.6)
Other	2 (1.6)
Target age groups[a]	
General	31 (23.5)
Adults	87 (65.9)
Young adults	36 (27.3)
Teens	13 (9.8)
Children (<12 years)	2 (1.5)
Affiliations	
Nongovernment organization/nonprofit	45 (34.1)
Government organization	36 (27.3)
Unknown	26 (19.7)
University/educational organization	13 (9.8)
Commercial organization	6 (4.5)
Health care organization	2 (1.2)
App features or functions[a]	
Web required	21 (15.9)
Location services	20 (15.2)
Panic/exit	17 (12.9)
Social media sharing	9 (6.8)
Login	8 (6.1)
Password	7 (5.3)
Reminders	6 (4.5)
Integration with phone (eg, calendar or reminders)	2 (1.5)
App community	1 (0.8)
App theoretical background/intervention strategies[a]	
Information/education	104 (78.8)

Category	Apps, n (%)
Safety monitoring/tracking	15 (11.4)
Goal setting/safety planning	14 (10.6)
Other	12 (9.1)
Location tracking	10 (7.6)
Decision making	8 (6.1)
Feedback	3 (2.3)
Assessment	2 (1.5)

[a]Categories are not mutually exclusive.

[b]IPV: intimate partner violence.

App Quality

MARS classifies app quality into 5 categories: engagement, functionality, aesthetics, information, and overall subjective quality. The scale also includes a sixth domain of questions related to the perceived potential impact of an app on behavior change. App mean quality scores in this study ranged from 2.80 (engagement) to 4.75 (functionality), and the app perceived potential impact mean score was 3.02 (95% CI 2.84-3.20). The overall subjective quality mean for the reviewed apps was 2.65 (95% CI 2.58-2.72). The individual item means ranged from 1.08 (evidence base) to 4.15 (quantity of information), both items being within the information domain. Tables 2 and 3 summarize category and individual item scores for the reviewed apps.

Table 2. Mobile Application Rating Scale (MARS) app quality subscale ratings.

MARS section	Mean (SD)	95% CI
Engagement		
Overall	2.08 (0.676)	1.96-2.20
Entertainment	1.94 (0.935)	1.78-2.11
Interest	2.14 (1.077)	1.95-2.33
Customization	1.56 (0.981)	1.39-1.74
Interactivity	1.54 (0.896)	1.38-1.70
Target group	3.26 (0.734)	3.13-3.39
Functionality		
Overall	3.73 (0.957)	3.57-3.90
Performance	3.45 (1.163)	3.25-3.66
Ease of use	3.95 (0.847)	3.80-4.10
Navigation	3.95 (0.932)	3.79-4.12
Gestural design	3.98 (0.704)	3.85-4.10
Aesthetics		
Overall	3.38 (0.757)	3.25-3.52
Layout	3.83 (0.853)	3.68-3.98
Graphics	3.32 (0.829)	3.17-3.47
Visual appeal	3.08 (0.771)	2.94-3.22
Information		
Overall	2.71 (0.731)	2.58-2.84
Accuracy of description	3.74 (1.023)	3.56-3.92
Goals	2.15 (1.628)	1.86-2.44
Quality of information	4.02 (1.402)	3.76-4.27
Quantity of information	4.15 (1.186)	3.92-4.38
Visual information	1.7 (1.453)	1.44-1.96
Credibility	2.94 (1.148)	2.74-3.15
Evidence base	1.08 (0.302)	1.03-1.13

Table 3. Mobile Application Rating Scale (MARS) app subjective quality ratings and perceived impact scores.

Subjective quality and perceived impact items	Mean (SD)	95% CI
Subjective quality		
Overall score	2.65 (0.398)	2.58-2.72
Would you recommend this app to people who might benefit from it?	3.37 (1.147)	3.16-3.57
How many times do you think you would use this app in the next 12 months if it was relevant to you?	1.62 (0.778)	1.48-1.76
What is your overall star rating of the app?	3.06 (0.905)	2.90-3.22
Perceived impact		
Overall score	3.02 (1.005)	2.84-3.20
Awareness: This app is likely to increase awareness of the importance of sexual assault and IPV[a].	2.97 (1.101)	2.77-3.16
Knowledge: This app is likely to increase knowledge/understanding of sexual assault and IPV.	2.96 (1.165)	2.75-3.17
Attitudes: This app is likely to change attitudes toward improving sexual assault and IPV.	3.34 (1.043)	3.15-3.52
Intention to change: This app is likely to increase intentions/motivation to address sexual assault and IPV.	3.11 (1.061)	2.92-3.30
Help seeking: Use of this app is likely to encourage further help seeking for sexual assault and IPV (if it is required).	2.73 (1.188)	2.52-2.94
Behavior change: Use of this app is likely to increase/decrease sexual assault and IPV (of their sequelae).	3.10 (1.032)	2.91-3.28

[a]IPV: intimate partner violence.

Discussion

Principal Findings

Despite reviewing over 100 freely available, English language mobile apps targeted at SV and IPV prevention and response, the overall quality was average. There were few apps that we ourselves as experienced forensic examiners would use as clinicians or recommend to our patients after a physical or sexual assault. We recognize the limitations of a dated search in a rapidly evolving mobile app space. However, our primary findings related to (1) identifying relevant apps and (2) high-quality evidence-based apps remain salient.

The focus of the apps was largely on education and information sharing. Although individuals are spending more time on mobile phones, if the focus is largely on education and information sharing, we fear that these apps will not be successful in meeting their desired goal. As with any health promotion and prevention content, mobile app content must be regularly reviewed, and updated for accuracy—and the mobile app platform adds additional technology hurdles to overcome regarding maintaining the infrastructure and content in ways that are accessible and engaging for users. Although not part of any of the MARS subscales, the tool does include collecting data on the number of times an app was rated and the current app rating. Of the 132 included apps, 104 (78.8%) had at least 1 rating listed (median user rating across rated apps was 4.20; however, the median number of user ratings across rated apps was 2, with a range of 1 to >21,000). These variations in how apps are marketed, downloaded, and shared among networks highlight 1 key area of their usefulness and 1 challenge in their dissemination in violence prevention and response work [20,47].

Marketing and sharing are key variables in how app-sharing platforms disseminate content to users and are not necessarily a skill set that violence advocates and health care providers have been trained in or possess. Notably, we believe the most concerning finding of this search was incidental. We were disturbed during the search process at the juxtaposition of violence prevention and response apps with zombie-killing games (any search including the term “violence”), dating sims (“intimate”), and the aforementioned voice changer app (when searching “rape”). The potential for retraumatization of our patients if they search these app stores looking for appropriate resources is high.

Recommendations for Health Care Providers

Clinicians caring for patients after sexual or physical violence interested in sharing a mobile app–based resource with their patient population should treat any app similar to any other resource. The resource should be vetted by the health care team before adoption. As things in the mobile app industry change at an ever-increasing pace, any recommended app should routinely be reviewed to ensure it is still up to date and has not gone defunct. Several of the apps we identified in the initial search were not available by the time we returned to review them (see Figure 1).

Clinicians interested in providing mobile app resources to their patients experiencing IPV or SV should consider preidentifying a few select mobile app resources and sharing them directly with interested patients via a QR code or direct link to prevent patients from searching the app stores on their own. This would reduce the potential for retraumatization related to inappropriate or unexpected results of a search of the major app stores. Alternatively, if the patient allows or prefers, providers could search for and download the app directly onto the patient’s phone or device on their behalf.

Clinicians interested in developing and launching a mobile health app for their target population should consider using or adopting an existing tool versus creating a new one. The costs of developing and maintaining an app must be weighed against the other services that could be rendered with those funds. There are costs associated with creation and design, as well as costs

to publish an app on each individual platform and maintenance costs to ensure the app remains relevant [48]. One piece of data that may have supported this in our analysis was the large number of apps that were initially found in our title searches but were not able to be relocated by the time we undertook our full analysis. There are multiple reasons that apps are removed from app stores. Primary reasons are related to apps not being compatible with current hardware or software requirements. As technology moves extraordinarily rapidly, maintaining apps requires diligent attention to these requirements to stay current. Our results were also consistent with a 2016 examination of the turnover of mental health apps, which found that in approximately a 6-month period, there was a 50% turnover (eg, apps were found on the initial search and not found on subsequent searches) in search results on the Android platform, whereas in iOS, approximately 90% of apps remained in the app store throughout the entire 9-month study period [49].

Recommendations for Research

Few of the apps included had any scholarship or evidence associated with their effectiveness. It is difficult to recommend an app for use in a clinical setting when there is no evidence related to whether it achieves its stated goals. Many of the apps were targeted at information sharing; something as simple as a test-retest knowledge assessment would provide at least basic data regarding whether the app is effective in increasing knowledge. There is also a precedent for evaluating mobile apps in their target population as well as with relevant service providers [20,21,29,50-54].

As we discovered during our analysis, MARS may not be the best tool for evaluating violence prevention and response apps. Although we adapted the tool for our use, we would recommend further adjustments in the future. For example, several of the features noted anecdotally may be worth formally evaluating (eg, how and when GPS is integrated into the app, the presence of an "emergency exit" button). It would also be prudent to assess and understand the limits of data confidentiality, as GPS can be used by apps to assist people in help seeking but also by IPV perpetrators to track their victims. We also did not further adapt MARS for items such as whether the apps used a trauma-informed approach [55]. Factors that may make an app useful to a provider or patient who has experienced violence, such as whether it is designed with trauma-informed principles in mind (eg, is the information not only correct but also written using language that is nonjudgmental and easily understandable during a traumatic situation), are not currently captured in MARS and would be beneficial to include in future work on violence and trauma-related mobile apps.

There is also the continued difficulty of many apps placing the onus of violence prevention on the potential victim. Many apps are dependent on a potential victim taking a precautionary behavior: downloading the app, setting up a network, and holding a button on an app down until they are "safe." These types of interventions perpetuate victim blaming, both blaming by others and self-blaming. Blaming is a form of retraumatization, which is in direct conflict with providing trauma-informed care [55].

Limitations

Mobile apps and the mobile space are changing rapidly. The amount of time people spend on mobile apps increased by 35% in 2019 [12]. Unfortunately, health care research has historically moved at a much slower pace, and conducting a systematic search and analysis took an incredible investment of time. Based on our own data, by the time these data exist in the world, many of the included apps will likely no longer be accessible to the public, demonstrating the incredibly fast nature of how mobile apps come and go compared to how research is conducted. Conducting a search of a constantly changing medium required adjustment to traditional methods. We were unable to evenly divide up the apps for review due to device and platform availability at our respective institutions.

Additional limitations of this review included both the limitations inherent to the MARS tool and specifically its usefulness as a tool for evaluating violence apps. Although MARS standardizes language (eg, "This app is likely to increase awareness of the importance of address [insert target health behavior]), this still requires a reviewer to make numerous subjective decisions and assumptions. MARS also does not contain violence-specific content. This presented challenges in completely evaluating the aspects of apps that violence victims, survivors, or providers may find most important.

A final significant limitation is the often overlapping yet distinct needs of the people who interact with violence apps. Providers, friends, family, and survivors may all benefit from rapid collated access to local service information, but survivors may additionally want, need, or benefit from specific guided planning and resources. Providers or friends and family using an app to assist a patient or loved one may instead find the most benefit from tailored educational information and trauma-informed response information [19,56]. Although we broadly included all apps for these audiences in our search and analysis, we did not collect data to determine which apps appeared to specify which target audiences.

Conclusion

In assessing freely available smartphone apps related to SV and IPV prevention or response, we note first the incredible amount of information that one needs to sift through before even getting to relevant apps. Over 2000 titles were assessed, including first-person shooter games and voice changer apps. Once narrowed to the 132 relevant and included apps, we must highlight that despite the number of apps in this space, the lack of quality and evidence base leaves much work to be done. As with any other item in our toolbox as health care providers and advocates, apps are 1 tool and will likely be most useful when implemented in the correct settings and with the appropriate knowledge, training, and skill sets.

Acknowledgments

The authors thank Daniel J. Sheridan, Joe Crowley, and Kristen Bettega for contributions to design and data collection. JCA received funding support from the National Institutes of Alcohol Abuse and Alcoholism (K23AA027288) and the National Institute of Child Health and Human Development (T32HD087162). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Authors' Contributions

JDM was responsible for conceptualization, methodology, formal analysis, investigation, resources, data curation, writing (original draft), writing (review and editing), visualization, supervision, and project administration. AT performed writing (original draft) and writing (review and editing). LR conducted visualization, writing (original draft), and writing (review and editing). EP performed conceptualization, investigation, and writing (review and editing). JCA was responsible for conceptualization, methodology, investigation, formal analysis, writing (original draft), writing (review and editing), and supervision.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Adapted Mobile App Rating Scale (MARS).
[PDF File (Adobe PDF File), 166 KB - formative_v6i2e28959_app1.pdf]

References

1. Truman J, Morgan R. Criminal Victimization, 2015.: Bureau of Justice Statistics, U.S. Department of Justice; 2018. URL: https://bjs.ojp.gov/content/pub/pdf/cv15.pdf [accessed 2021-03-10]
2. Smith SG, Zhang X, Basile KC, Merrick MT, Wang J, Kresnow M, et al. The National Intimate Partner and Sexual Violence Survey (NISVS): 2015 Data Brief - Updated Release. Atlanta, GA: National Center for Injury Prevention and Control, Centers for Disease Control and Prevention; 2018. URL: https://www.cdc.gov/violenceprevention/pdf/2015data-brief508.pdf [accessed 2021-03-10]
3. James SE, Herman JL, Rankin S, Keisling M, Mottet L, Anafi M. The Report of the 2015 U.S. Transgender Survey. Washington, DC: National Center for Transgender Equality; 2016. URL: https://transequality.org/sites/default/files/docs/usts/USTS-Full-Report-Dec17.pdf [accessed 2021-02-15]
4. Walters ML, Chen J, Breiding MJ. The National Intimate Partner and Sexual Violence Survey: 2010 Findings on Victimization by Sexual Orientation. Atlanta, GA: National Center for Injury Prevention and Control, Centers for Disease Control and Prevention; 2013. URL: https://www.cdc.gov/ViolencePrevention/pdf/NISVS_SOfindings.pdf [accessed 2020-09-25]
5. Yim IS, Kofman YB. The psychobiology of stress and intimate partner violence. Psychoneuroendocrinology 2019 Jul;105:9-24 [FREE Full text] [doi: 10.1016/j.psyneuen.2018.08.017] [Medline: 30170928]
6. Humphreys J, Epel ES, Cooper BA, Lin J, Blackburn EH, Lee KA. Telomere shortening in formerly abused and never abused women. Biol Res Nurs 2012 Apr;14(2):115-123 [FREE Full text] [doi: 10.1177/1099800411398479] [Medline: 21385798]
7. Chandan JS, Thomas T, Bradbury-Jones C, Taylor J, Bandyopadhyay S, Nirantharakumar K. Risk of cardiometabolic disease and all-cause mortality in female survivors of domestic abuse. J Am Heart Assoc 2020 Feb 18;9(4):e014580 [FREE Full text] [doi: 10.1161/JAHA.119.014580] [Medline: 32063124]
8. World Health Organization. Global and Regional Estimates of Violence Against Women: Prevalence and Health Effects of Intimate Partner Violence and Non-partner Sexual Violence. Geneva, Switzerland: Word Health Organization; 2013. URL: https://www.who.int/publications/i/item/9789241564625 [accessed 2020-09-01]
9. Paterno MT, Draughon Moret JE, Paskausky A, Campbell JC. Exploring reproductive coercion in relationship contexts among young adult, primarily African American women at three women's health clinics. J Interpers Violence 2021 Feb;36(3-4):NP2248-2271NP. [doi: 10.1177/0886260518756116] [Medline: 29460674]
10. Silverman JG, Decker MR, Reed E, Raj A. Intimate partner violence victimization prior to and during pregnancy among women residing in 26 U.S. states: associations with maternal and neonatal health. Am J Obstet Gynecol 2006 Jul;195(1):140-148. [doi: 10.1016/j.ajog.2005.12.052] [Medline: 16813751]
11. Perrin A, Anderson M. Share of U.S. Adults Using Social Media, Including Facebook Is Mostly Unchanged since 2018.: Pew Research Center; 2019 Apr 10. URL: https://www.pewresearch.org/fact-tank/2019/04/10/share-of-u-s-adults-using-social-media-including-facebook-is-mostly-unchanged-since-2018/ [accessed 2019-09-16]
12. App Annie. State of Mobile 2020.: App Annie URL: https://www.appannie.com/en/go/state-of-mobile-2020/ [accessed 2020-11-24]

13. Greenfield LA. An Analysis of Data on Rape and Sexual Assault: Sex Offenses and Offenders. Washington DC: Bureau of Justice Statistics, US Department of Justice; 1997. URL: https://bjs.ojp.gov/content/pub/pdf/SOO.PDF [accessed 2020-11-24]
14. The Oswegonian. UASK App Proves To Be Useful Resource On College Campuses. 2016 Oct 6. URL: https://www.oswegonian.com/2016/10/06/u-ask-app-proves-to-be-useful-resource-on-college-campuses/?fbclid=IwAR0roarew2wLui9i17PKJWF59yUVHdLun189cuEoHR37VEDAjeuUKl4rs6Q [accessed 2020-11-30]
15. Geffen S. This App is Changing the Way College Students Respond to Sexual Assault Internet.: MTV News; 2015. URL: http://www.mtv.com/news/2058624/ask-app-college-sexual-assault/?fbclid=IwAR25lJLlODhbFAxBP-WhQPXFp7RSIVqN1V9A71LMoDXXs3wA8h5N05LkSUs [accessed 2021-09-27]
16. Florian A. UASK Apps Aim to Help Students Reach Sexual Assault Resources on Campus.: USA Today; 2016 Sep 20. URL: https://www.usatoday.com/story/college/2016/09/20/u-ask-apps-aim-to-help-students-reach-sexual-assault-resources-on-campus/37422127/?fbclid=IwAR3Zmqpsy9QYR-XnzboJnUK9WkTAXCP0w_BMDMAcYQSX-pbJmv52gAVagYs [accessed 2021-10-04]
17. Sande MK, Broderick KB, Moreira ME, Bender B, Hopkins E, Buchanan JA. Sexual assault training in emergency medicine residencies: a survey of program directors. West J Emerg Med 2013 Sep 17;14(5):461-466 [FREE Full text] [doi: 10.5811/westjem.2013.2.12201] [Medline: 24106543]
18. Akash SA, Al-Zihad M, Adhikary T, Razzaque MA, Sharmin A. HearMe: A smart mobile application for mitigating women harassment. 2016 Presented at: 2016 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE) IEEE; December 2016; AISSMS, Pune, Maharashtra, India p. 87-90. [doi: 10.1109/wiecon-ece.2016.8009093]
19. Bloom T, Gielen A, Glass N. Developing an app for college women in abusive same-sex relationships and their friends. J Homosex 2016 Jun;63(6):855-874. [doi: 10.1080/00918369.2015.1112597] [Medline: 26515797]
20. Anderson JC, Pollitt E, Crowley J, Holbrook D, Draughon Moret JE. A mixed-methods evaluation of college student and provider perspectives on a smartphone application for help-seeking after violence. J Am Coll Health 2021;69(6):668-674. [doi: 10.1080/07448481.2019.1705839] [Medline: 31944911]
21. Glass N, Clough A, Case J, Hanson G, Barnes-Hoyt J, Waterbury A, et al. A safety app to respond to dating violence for college women and their friends: the MyPlan study randomized controlled trial protocol. BMC Public Health 2015 Sep 08;15:871 [FREE Full text] [doi: 10.1186/s12889-015-2191-6] [Medline: 26350482]
22. Goldman S, Goyal D. Knowledge regarding child victims of commercial sexual exploitation and the feasibility of using a smartphone application: a pilot study. J Forensic Nurs 2019 Jun;15(2):103-109. [doi: 10.1097/JFN.0000000000000240] [Medline: 30985544]
23. Moon KJ, Park KM, Sung Y. Sexual abuse prevention mobile application (SAP_MobAPP) for primary school children in Korea. J Child Sex Abus 2017 Jul;26(5):573-589. [doi: 10.1080/10538712.2017.1313350] [Medline: 28661824]
24. Johns Hopkins University School Of Nursing. myPlan. URL: https://www.myplanapp.org/ [accessed 2021-03-08]
25. Alhusen J, Bloom T, Clough A, Glass N. Development of the MyPlan Safety Decision app with friends of college women in abusive dating relationships. J Technol Hum Serv 2015 Aug 12;33(3):263-282. [doi: 10.1080/15228835.2015.1037414]
26. Davies J, Lyon E, Monti-Catania D. Safety Planning with Battered Women: Complex Lives/Difficult Choices. Thousand Oaks, CA: Sage; 1998.
27. Campbell J, Glass N. Safety planning, danger, lethality assessment. In: Mitchell C, Anglin D, editors. Intimate Partner Violence: A Health Based Perspective. New York: Oxford University Press; 2009:319-334.
28. Eden KB, Perrin NA, Hanson GC, Messing JT, Bloom TL, Campbell JC, et al. Use of online safety decision aid by abused women: effect on decisional conflict in a randomized controlled trial. Am J Prev Med 2015 Apr;48(4):372-383 [FREE Full text] [doi: 10.1016/j.amepre.2014.09.027] [Medline: 25547929]
29. Glass NE, Clough A, Messing JT, Bloom T, Brown ML, Eden KB, et al. Longitudinal IMPACT of the myPlan app on health and safety among college women experiencing partner violence. J Interpers Violence 2021 Feb 12:886260521991880. [doi: 10.1177/0886260521991880] [Medline: 33576291]
30. Glass NE, Perrin NA, Hanson GC, Bloom TL, Messing JT, Clough AS, et al. The longitudinal impact of an internet safety decision aid for abused women. Am J Prev Med 2017 May;52(5):606-615. [doi: 10.1016/j.amepre.2016.12.014] [Medline: 28108189]
31. Breiding MJ, Smith SG, Basile KC, Walters ML, Chen J, Merrick MT. Prevalence and characteristics of sexual violence, stalking, and intimate partner violence victimization: national intimate partner and sexual violence survey, United States, 2011. MMWR Surveill Summ 2014 Sep 05;63(8):1-18 [FREE Full text] [Medline: 25188037]
32. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar 11;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
33. Martín Payo R, Harris J, Armes J. Prescribing fitness apps for people with cancer: a preliminary assessment of content and quality of commercially available apps. J Cancer Surviv 2019 Jun;13(3):397-405. [doi: 10.1007/s11764-019-00760-2] [Medline: 31030308]

34. Bardus M, van Beurden SB, Smith JR, Abraham C. A review and content analysis of engagement, functionality, aesthetics, information quality, and change techniques in the most popular commercial apps for weight management. Int J Behav Nutr Phys Act 2016 Mar 10;13(1):35 [FREE Full text] [doi: 10.1186/s12966-016-0359-9] [Medline: 26964880]
35. Brown J, Siddiqi A, Froome H, Arcand J. What's cooking? A content and quality analysis of food preparation mobile applications (p16-050-19). Curr Dev Nutr 2019 Jun 13;3(Suppl 1):nzz050.P16-050-19. [doi: 10.1093/cdn/nzz050.p16-050-19]
36. Patel R, Sulzberger L, Li G, Mair J, Morley H, Shing MN, et al. Smartphone apps for weight loss and smoking cessation: quality ranking of 120 apps. N Z Med J 2015 Sep 04;128(1421):73-76. [Medline: 26370762]
37. Salehi F, Kermani ZA, Khademian F, Aslani A. Critical appraisal of mental health applications. Stud Health Technol Inform 2019;261:303-308. [Medline: 31156135]
38. Mani M, Kavanagh DJ, Hides L, Stoyanov SR. Review and evaluation of mindfulness-based iPhone apps. JMIR Mhealth Uhealth 2015 Aug 19;3(3):e82 [FREE Full text] [doi: 10.2196/mhealth.4328] [Medline: 26290327]
39. Chavez S, Fedele D, Guo Y, Bernier A, Smith M, Warnick J, et al. Mobile apps for the management of diabetes. Diabetes Care 2017 Oct;40(10):e145-e146 [FREE Full text] [doi: 10.2337/dc17-0853] [Medline: 28774944]
40. Choi YK, Demiris G, Lin SY, Iribarren SJ, Landis CA, Thompson HJ, et al. Smartphone applications to support sleep self-management: review and evaluation. J Clin Sleep Med 2018 Oct 15;14(10):1783-1790 [FREE Full text] [doi: 10.5664/jcsm.7396] [Medline: 30353814]
41. Salazar A, de Sola H, Failde I, Moral-Munoz JA. Measuring the quality of mobile apps for the management of pain: systematic search and evaluation using the mobile app rating scale. JMIR Mhealth Uhealth 2018 Oct 25;6(10):e10718 [FREE Full text] [doi: 10.2196/10718] [Medline: 30361196]
42. Woods LS, Duff J, Roehrer E, Walker K, Cummings E. Patients' experiences of using a consumer mHealth app for self-management of heart failure: mixed-methods study. JMIR Hum Factors 2019 May 02;6(2):e13009 [FREE Full text] [doi: 10.2196/13009] [Medline: 31045504]
43. Ramsey RR, Caromody JK, Voorhees SE, Warning A, Cushing CC, Guilbert TW, et al. A systematic evaluation of asthma management apps examining behavior change techniques. J Allergy Clin Immunol Pract 2019;7(8):2583-2591 [FREE Full text] [doi: 10.1016/j.jaip.2019.03.041] [Medline: 30954644]
44. IBM Corp. IBM SPSS Statistics for Windows. Armonk, NY: IBM Corp; 2017.
45. Müller RM, Kijl B, Martens JKJ. A comparison of inter-organizational business models of mobile app stores: there is more than open vs. closed. J Theor Appl Electron Commer Res 2011;6(2):13-14. [doi: 10.4067/s0718-18762011000200007]
46. Valvi AC, West DC. Mobile applications (apps) in advertising: A grounded theory of effective uses and practices. In: Kubacki K, editor. Ideas in Marketing: Finding the New and Polishing the Old. Proceedings of the 2013 Academy of Marketing Science (AMS) Annual Conference. Cham: Springer International; 2015:349-352.
47. Klibanoff E. App Links Sex Assault Survivors To Help, But Who Downloads It? . 2015. URL: https://www.npr.org/sections/health-shots/2015/01/24/379410663/what-do-i-do-for-sexual-assault-survivors-an-app-has-all-the-answers [accessed 2021-02-17]
48. Van Velthoven MH, Smith J, Wells G, Brindley D. Digital health app development standards: a systematic review protocol. BMJ Open 2018 Aug 17;8(8):e022969 [FREE Full text] [doi: 10.1136/bmjopen-2018-022969] [Medline: 30121614]
49. Larsen ME, Nicholas J, Christensen H. Quantifying app store dynamics: longitudinal tracking of mental health apps. JMIR Mhealth Uhealth 2016 Aug 09;4(3):e96 [FREE Full text] [doi: 10.2196/mhealth.6020] [Medline: 27507641]
50. Decker MR, Wood SN, Kennedy SR, Hameeduddin Z, Tallam C, Akumu I, et al. Adapting the myPlan safety app to respond to intimate partner violence for women in low and middle income country settings: app tailoring and randomized controlled trial protocol. BMC Public Health 2020 May 29;20(1):808 [FREE Full text] [doi: 10.1186/s12889-020-08901-4] [Medline: 32471469]
51. Bagwell-Gray M, Loerzel E, Dana Sacco G, Messing J, Glass N, Sabri B, et al. From myPlan to ourCircle: adapting a web-based safety planning intervention for Native American women exposed to intimate partner violence. J Ethn Cult Divers Soc Work 2020 Jun 22;30(1-2):163-180. [doi: 10.1080/15313204.2020.1770651]
52. Alvarez C, Debnam K, Clough A, Alexander K, Glass NE. Responding to intimate partner violence: healthcare providers' current practices and views on integrating a safety decision aid into primary care settings. Res Nurs Health 2018 Apr;41(2):145-155. [doi: 10.1002/nur.21853] [Medline: 29441596]
53. Decker MR, Wood SN, Hameeduddin Z, Kennedy SR, Perrin N, Tallam C, et al. Safety decision-making and planning mobile app for intimate partner violence prevention and response: randomised controlled trial in Kenya. BMJ Glob Health 2020 Jul;5(7):e002091 [FREE Full text] [doi: 10.1136/bmjgh-2019-002091] [Medline: 32675229]
54. Sabri B, Njie-Carr VPS, Messing JT, Glass N, Brockie T, Hanson G, et al. The weWomen and ourCircle randomized controlled trial protocol: a web-based intervention for immigrant, refugee and indigenous women with intimate partner violence experiences. Contemp Clin Trials 2019 Jan;76:79-84 [FREE Full text] [doi: 10.1016/j.cct.2018.11.013] [Medline: 30517888]
55. Substance Abuse and Mental Health Services Administration. SAMHSA's Concept of Trauma and Guidance for a Trauma-Informed Approach. HHS Publication No. (SMA) 14-4884. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2014 Jul. URL: https://ncsacw.samhsa.gov/userfiles/files/SAMHSA_Trauma.pdf [accessed 2020-06-24]

56. Griffiths KM, Crisp DA, Barney L, Reid R. Seeking help for depression from family and friends: a qualitative analysis of perceived advantages and disadvantages. BMC Psychiatry 2011 Dec 15;11:196 [FREE Full text] [doi: 10.1186/1471-244X-11-196] [Medline: 22171567]

Abbreviations

GPS: Global Positioning System
IPV: intimate partner violence
IRIS: Internet Resource for Information and Safety
LGBTQ: lesbian, gay, bisexual, transgender, and queer
MARS: Mobile Application Rating Scale
SV: sexual violence

Edited by G Eysenbach; submitted 29.03.21; peer-reviewed by L Schwab-Reese, M Testa; comments to author 17.05.21; revised version received 06.10.21; accepted 30.11.21; published 08.02.22.

Please cite as:
Draughon Moret J, Todd A, Rose L, Pollitt E, Anderson J
Mobile Phone Apps for Intimate Partner and Sexual Violence Prevention and Response: Systematic Search on App Stores
JMIR Form Res 2022;6(2):e28959
URL: https://formative.jmir.org/2022/2/e28959
doi:10.2196/28959
PMID:35133285

Original Paper

An Association of Influenza Epidemics in Children With Mobile App Data: Population-Based Observational Study in Osaka, Japan

Yusuke Katayama[1], MD, PhD; Kosuke Kiyohara[2], DrPH; Tomoya Hirose[1], MD, PhD; Kenichiro Ishida[3], MD; Jotaro Tachino[1], MD, PhD; Shunichiro Nakao[1], MD, MSc, PhD; Tomohiro Noda[4], MD; Masahiro Ojima[3], MD; Takeyuki Kiguchi[5], MD, PhD; Tasuku Matsuyama[6], MD, PhD; Tetsuhisa Kitamura[7], MD, DrPH

[1]Department of Traumatology and Acute Critical Medicine, Osaka University Graduate School of Medicine, Suita, Japan

[2]Department of Food Science, Faculty of Home Economics, Otsuma Women's University, Tokyo, Japan

[3]Department of Acute Medicine and Critical Care Center, Osaka National Hospital, National Hospital Organization, Osaka, Japan

[4]Department of Traumatology and Critical Care Medicine, Osaka City University Graduate School of Medicine, Osaka, Japan

[5]Division of Trauma and Surgical Critical Care, Osaka General Medical Center, Osaka, Japan

[6]Department of Emergency Medicine, Kyoto Prefectural University of Medicine, Kyoto, Japan

[7]Division of Environmental Medicine and Population Sciences, Department of Social and Environmental Medicine, Osaka University Graduate School of Medicine, Suita, Japan

Corresponding Author:
Yusuke Katayama, MD, PhD
Department of Traumatology and Acute Critical Medicine
Osaka University Graduate School of Medicine
2-15 Yamada-oka
Suita
Japan
Phone: 81 6 6879 5707
Email: orion13@hp-emerg.med.osaka-u.ac.jp

Abstract

Background: Early surveillance to prevent the spread of influenza is a major public health concern. If there is an association of influenza epidemics with mobile app data, it may be possible to forecast influenza earlier and more easily.

Objective: We aimed to assess the relationship between seasonal influenza and the frequency of mobile app use among children in Osaka Prefecture, Japan.

Methods: This was a retrospective observational study that was performed over a three-year period from January 2017 to December 2019. Using a linear regression model, we calculated the R^2 value of the regression model to evaluate the relationship between the number of "fever" events selected in the mobile app and the number of influenza patients ≤14 years of age. We conducted three-fold cross-validation using data from two years as the training data set and the data of the remaining year as the test data set to evaluate the validity of the regression model. And we calculated Spearman correlation coefficients between the calculated number of influenza patients estimated using the regression model and the number of influenza patients, limited to the period from December to April when influenza is prevalent in Japan.

Results: We included 29,392 mobile app users. The R^2 value for the linear regression model was 0.944, and the adjusted R^2 value was 0.915. The mean Spearman correlation coefficient for the three regression models was 0.804. During the influenza season (December–April), the Spearman correlation coefficient between the number of influenza patients and the calculated number estimated using the linear regression model was 0.946 (P<.001).

Conclusions: In this study, the number of times that mobile apps were used was positively associated with the number of influenza patients. In particular, there was a good association of the number of influenza patients with the number of "fever" events selected in the mobile app during the influenza epidemic season.

(JMIR Form Res 2022;6(2):e31131) doi:10.2196/31131

KEYWORDS

syndromic surveillance; mobile app; influenza; epidemic; children

Introduction

Influenza can still spread in recent years and can cause death, especially in the elderly and infants. In addition, the spread of influenza not only harms people's health but also has a significant social and economic impact due to absenteeism and missed work. Therefore, early surveillance to prevent the spread of influenza is a major public health concern. However, traditional infectious disease reports require 10 days to 2 weeks. Various early surveillance models using absenteeism records [1-3], pharmacy drug sales [4,5], and visits to emergency departments [6,7] have been studied. In recent years, there have also been early prediction models for influenza using Internet search engine data [3,8-12]. In Japan, based on the Infectious Disease Control Law, patients with influenza are reported by medical institutions to public health centers, and the public health departments of each prefecture announce the results. However, such traditional surveillance of infectious diseases is associated with high costs and considerable time from data collection to publication. Thus, we have previously revealed the relationship between telephone triage service data and the number of infectious disease patients in Osaka Prefecture, Japan [12,13].

In 2015, we developed a self-triage mobile app for the residents of Osaka Prefecture that determines the urgency of symptoms in children with sudden illness or injury and guides them to ambulances and appropriate medical institutions based on the urgency of their symptoms. The profile of users of this self-triage app has been revealed in detail [14]. This mobile app has been available for free download from Google Play and the App Store in Japan since 2016, and by the end of 2019, it had been downloaded 23,732 times and used a total of 63,230 times. If the relationship between the frequency of mobile app use and the number of influenza patients can be clarified and a prediction model can be structured, it may be possible to forecast influenza earlier and more easily. While traditional infectious disease reports actually require 10 days to 2 weeks, syndromic surveillance with mobile app data allows aggregation and forecasting to be done programmatically, enabling low-cost, real-time forecasting. The purpose of this study was to assess the relationship between seasonal influenza and the frequency of mobile app use among children in Osaka Prefecture, Japan.

Methods

Study Design, Population, and Setting

This was a retrospective observational study during a three-year period from January 2017 to December 2019. Osaka Prefecture has the largest urban area in western Japan, with an area of 1905.14 km^2, a population of 8.8 million, and 1.09 million children of ≤14 years of age [15]. In this study, we included cases in which the child's parents or guardians used the mobile app in Osaka Prefecture. Informed consent was obtained from the mobile app users at the time of mobile app use. This study was approved by the ethics committee of the Osaka University Graduate School of Medicine (approval no. 20313). This manuscript was written based on the STROBE statement (strengthening the reporting of observational studies in epidemiology) [16].

Outpatient Surveillance of Influenza-Like Illness in Japan

The Infectious Disease Surveillance Program in Japan, initiated in 1981, formed the basis for influenza surveillance for outpatients [17,18]. This program was revised and updated to its present form following the revision of the Infectious Disease Control Law in 1999 [17-20]. The system is currently called the National Epidemiological Surveillance for Infectious Diseases, which includes a mandatory reporting system for nationally notifiable diseases and sentinel surveillance systems for various types of infectious diseases [21].

Influenza falls under the sentinel surveillance arm of the program. Weekly numbers of influenza patients have been reported to local health centers from 5000 medical institutions nationwide. Sentinel sites were designated according to their geographic distribution, type of medical institution (clinic or hospital), and population density. These sentinels use the following criteria for reporting influenza-like illness (ILI): (1) sudden onset of illness, (2) fever >38°C, (3) symptoms of upper respiratory inflammation, and (4) systemic symptoms such as general fatigue. A case is considered to meet the reporting criteria if the patient meets all symptoms from (1) to (4) or at least one of the symptoms in combination with a positive rapid diagnostic test [19]. Sentinel sites report the age group and sex of patients on a weekly basis. The report does not include personal information (eg, names or addresses). This information is transferred from local health centers to the prefectural government, where it is aggregated into a prefectural report. The report is then forwarded to the National Institute of Infectious Diseases in Tokyo, which is affiliated with the Ministry of Health, Labor and Welfare. Within Osaka Prefecture, 300 medical institutions report influenza patients to 10 local health centers [22]. In this study, the main endpoint was the weekly number of influenza patients in Osaka Prefecture. These data were acquired from the website of the Information Center of Infectious Diseases in Osaka Prefecture [22].

Mobile App for Emergency Pediatric Patients

The details of the mobile self-triage app that was used in this study have been described previously [14]. First, the age and sex of the children were selected in this mobile app. Next, the user selects either "sickness" or "injury, poisoning, foreign substances, and others." When either of these is selected, the list of chief complaints shown is displayed in the mobile app, and the user selects the relevant chief complaint. For example, if "fever" is selected, relevant signs and symptoms with high urgency, such as "fever of ≥41°C," are displayed in the app. If none of these are selected, relevant signs and symptoms with moderate urgency, such as "decreased urine volume," are displayed in the app. If none of them apply, the related signs and symptoms corresponding to "low urgency" are further displayed, and the urgency is determined based on the selected signs and symptoms. The app provides emergency medical services, such as the ability to call an ambulance or telephone triage center and information on available hospitals and clinics. Based on the judgment of urgency. If there is another chief

complaint, such as “convulsion” when “fever” is selected, the app will move to the urgency assessment for the other complaint. Only hospitals and clinics in Osaka Prefecture that have agreed to register their information in the app will be displayed as available hospitals and clinics. In addition, the GPS feature of the user’s cellphone also provides a list of hospitals and clinics in order of proximity to the location where the app is being used. The Android version of this app was released in January 2016, while the iOS version was released in April 2016. The mobile app can be downloaded free from Google Play and the App Store in Japan.

Endpoint

The endpoint of this study was the number of influenza patients≤14 years of age per week in the Osaka Prefecture. The number of influenza patients per week was obtained from data published on the website of the Osaka Institute of Public Health [22].

Statistical Analysis

Using a linear regression model, we calculated the R^2 value of the regression model to evaluate the relationship between the number of “fever” events selected in the mobile app and the number of influenza patients ≤14 years of age. We also calculated the Spearman correlation coefficient between the number of influenza patients and the calculated number of influenza patients ≤14 years of age, the adjusted R^2 value, and the *P* value. The age groups were classified as follows: infants and toddlers (0-4 years), kids (5-9 years), and teenagers (10-14 years). The seasons were categorized as winter (January-March), spring (April-June), summer (July-September), and autumn (October-December). Then, we evaluated the interaction between influenza season (December-April) and the number of “fever” events selected in the mobile app. In addition, to assess the regression model in a smaller area, we divided Osaka Prefecture into eight regions (Toyono, Mishima, North-Kawachi, Middle-Kawachi, South-Kawachi, Sakai, Senshu, and Osaka City) (Multimedia Appendix 1). The eight regions were classified based on the medical care plan of the Osaka Prefectural Government [23]. Finally, we calculated Spearman correlation coefficients between the calculated number of influenza patients estimated using the regression model and the number of influenza patients, limited to the period from December to April, when influenza is prevalent in Japan. *P* values of <.05 were considered to indicate statistical significance. All statistical analyses were performed using SPSS version 23.0J (IBM Corp).

Results

From 2017 to 2019, the mobile app was used a total of 59,375 times; 29,392 (49.5%) of these uses occurred in Osaka Prefecture. On the other hand, the number of influenza patients ≤14 years of age in the same period was 188,590 (Multimedia Appendix 2). Table 1 shows the characteristics of the subjects in this study. The median age was 1 year (IQR 0-3 years), and the most frequent age group was infants with 17,401 (59.2%) uses, followed by toddlers with 8999 (30.6%) uses. A total of 15,387 (52.4%) subjects were male, 13,788 (46.9%) were female, and 217 (0.7%) were of unknown sex. The mobile app was used 6453 (22.0%) times in 2017, 10,724 (36.5%) times in 2018, and 12,215 (41.6%) times in 2019. By season, the mobile app was used 6142 (20.9%) times in winter, 7714 (26.2%) times in spring, 7673 (26.1%) times in summer, and 7863 (26.8%) times in autumn. The region in which the app was most frequently used was Osaka City (13,570 times, 46.2%), and the region in which the app was used the least was the Mishima area (1156 times, 3.9%). The chief complaint most frequently selected on the mobile app was “fever” (14,777 times, 39.1%), followed by “cough” (2592 times, 6.9%) and “head and neck injury” (2523 times, 6.7%).

Figure 1 shows the weekly number of influenza patients ≤14 years of age and the weekly number of times that “fever” was selected in the mobile app. Figure 2 shows the weekly number of influenza patients ≤14 years of age and the calculated number of influenza patients ≤14 years of age estimated using the linear regression model. The red line shows the weekly number of influenza patients; the yellow line shows the number of times that “fever” was selected per week in the mobile app, and the blue line shows the calculated number of influenza patients ≤14 years of age estimated using the linear regression model. The regression coefficient of the weekly number of times that “fever” was selected in the mobile app was 2.977 (95% CI 0.680-5.314). The R^2 for the linear regression model was 0.944, and the adjusted R^2 value was 0.915. Spearman correlation coefficient between the number of influenza patients and the calculated number estimated using the regression model was 0.946. There was a significant interaction between influenza season and the number of times “fever” was selected in the mobile app (regression coefficient 5.873, 95% CI 2.521 to 9.226).

Table 2 shows the result of subgroup analysis by age group. The regression coefficient of the weekly number of times that “fever” was selected in the mobile app was 0.955 (95% CI –0.045 to 1.954) in infants and toddlers, 12.684 (95% CI 6.880 to 18.487) in kids, and 4.609 (95% CI –0.730 to 9.949) in teenagers, respectively.

Table 1. Demographic and clinical characteristics (N=29,392).

Characteristics	Values
Age[a] (years), n (%)	
Infants and toddlers (0-4)	25,279 (86.0)
Kids (5-9)	3032 (10.3)
Teenagers (10-14)	1081 (3.7)
Sex, n (%)	
Male	15,387 (52.4)
Female	13,788 (46.9)
Unknown	217 (0.7)
Time of telephone consultation and triage, n (%)	
Daytime (9:00 to 17:59)	11,818 (40.2)
Nighttime (18:00 to 8:59)	17,574 (59.8)
Year, n (%)	
2017	6453 (22.0)
2018	10,724 (36.5)
2019	12,215 (41.6)
Season, n (%)	
Winter (January to March)	6142 (20.9)
Spring (April to June)	7714 (26.2)
Summer (July to September)	7673 (26.1)
Autumn (October to December)	7863 (26.8)
Area, n (%)	
Osaka City Area	13,570 (46.2)
Middle-Kawachi Area	3071 (10.4)
Toyono Area	2893 (9.8)
Sakai Area	2546 (8.7)
Senshu Area	2315 (7.9)
North-Kawachi Area	2097 (7.1)
South-Kawachi Area	1744 (5.9)
Mishima Area	1156 (3.9)
Selected chief complaint in the mobile app, n (%)	
Fever	14,777 (39.1)
Cough	2592 (6.9)
Head and Neck Injury	2523 (6.7)
Nausea/Vomiting	2222 (5.9)
Convulsion	1635 (4.3)
Nasal Discharge	1571 (4.2)
Rash	1431 (3.8)
Diarrhea	1321 (3.5)
Face and extremities injury	995 (2.6)
Dyspnea	886 (2.3)
Other	7882 (20.8)

[a]The median age was 1 year, with an IQR of 0 to 3 years.

Figure 1. The weekly number of influenza patients and that of telephone triage counts due to "fever".

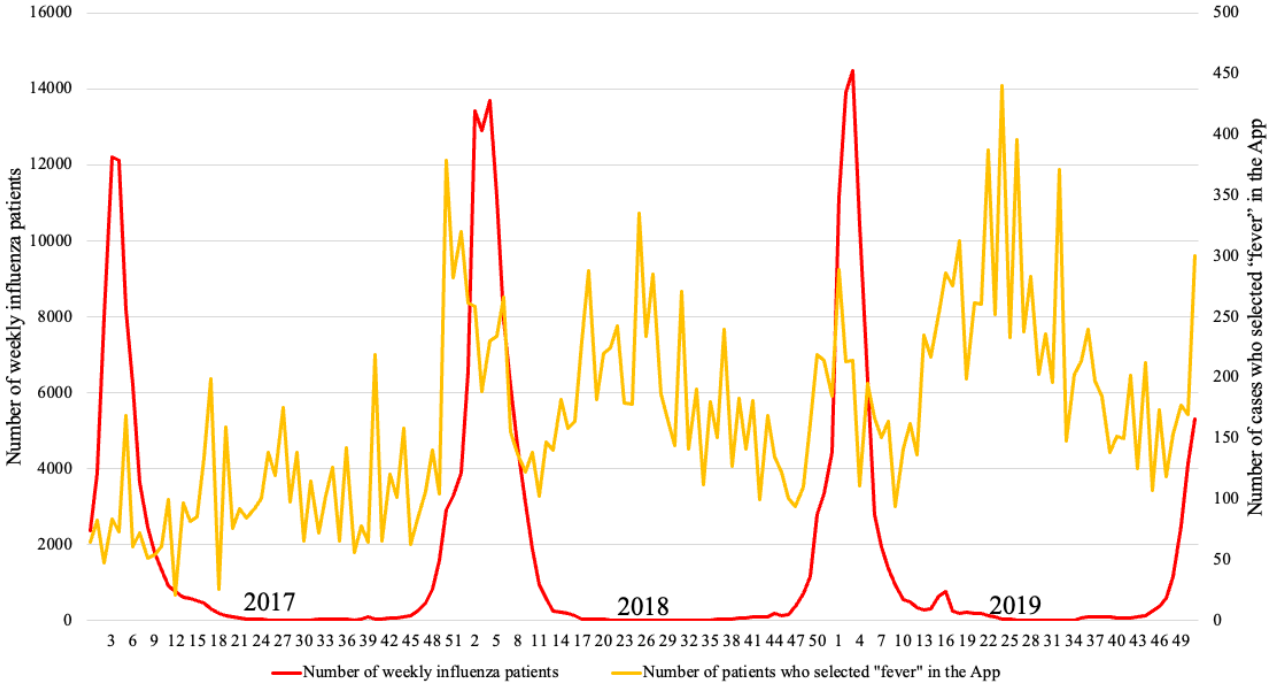


Figure 2. The weekly number of influenza patients and the calculated number of influenza patients from the linear regression model.

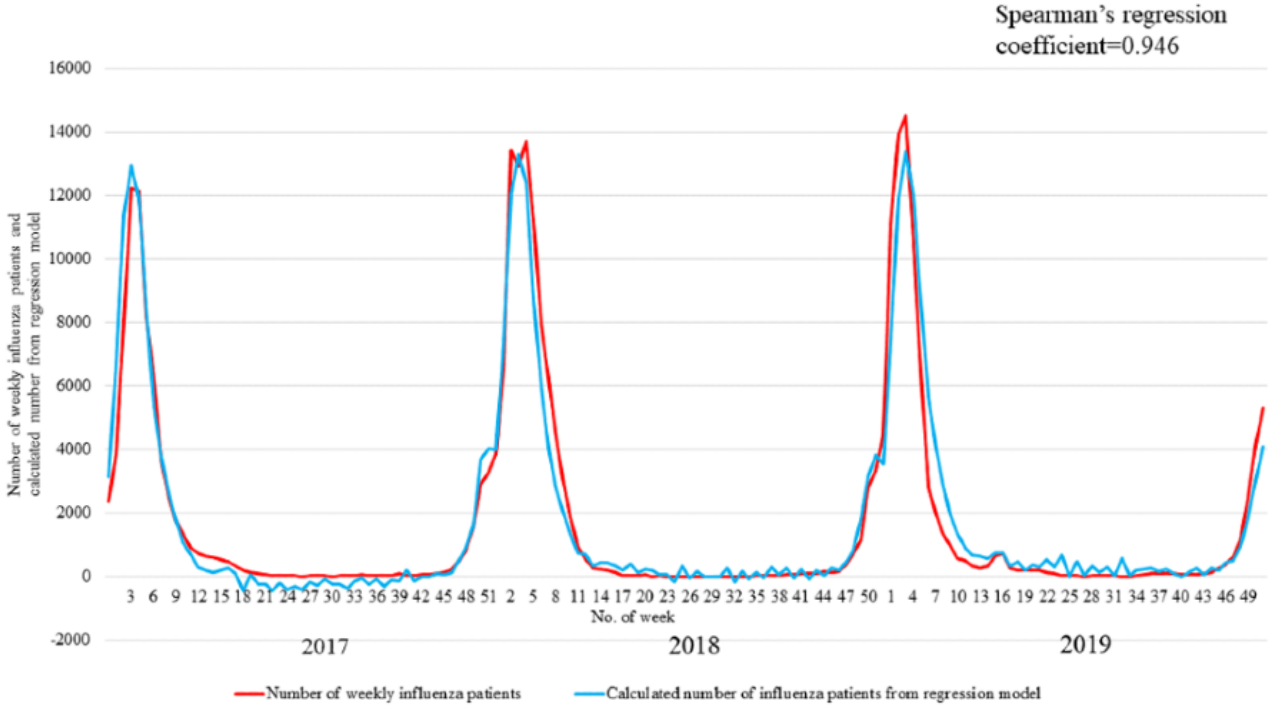


Table 2. The subgroup analysis by each age group.

Regression model	Regression coefficient of the number of times "fever" is selected	95% CI
Infants and toddlers (0-4 years old)	0.955	–0.045 to 1.954
Kids (5-9 years old)	12.684	6.880 to 18.487
Teenagers (10-14 years old)	4.609	–0.730 to 9.949

Figures 3 and 4 show the number of influenza patients ≤14 years of age and the calculated number of patients estimated using the linear regression model for each region. The best correlation was observed in the Toyono area (Spearman correlation coefficient: 0.918; *P*<.001), and the worst correlation was observed in the Senshu area (Spearman correlation coefficient: 0.768; *P*<.001).

Figure 5 shows the relationship between the number of influenza patients ≤14 years of age and the calculated number estimated using the linear regression model during the influenza season (December–April). From December to April, the Spearman correlation coefficient between the number of influenza patients and the calculated number estimated using the linear regression model was 0.946 (*P*<.001).

Figure 3. The weekly number of influenza patients and the calculated number of influenza patients from the linear regression model for each region.

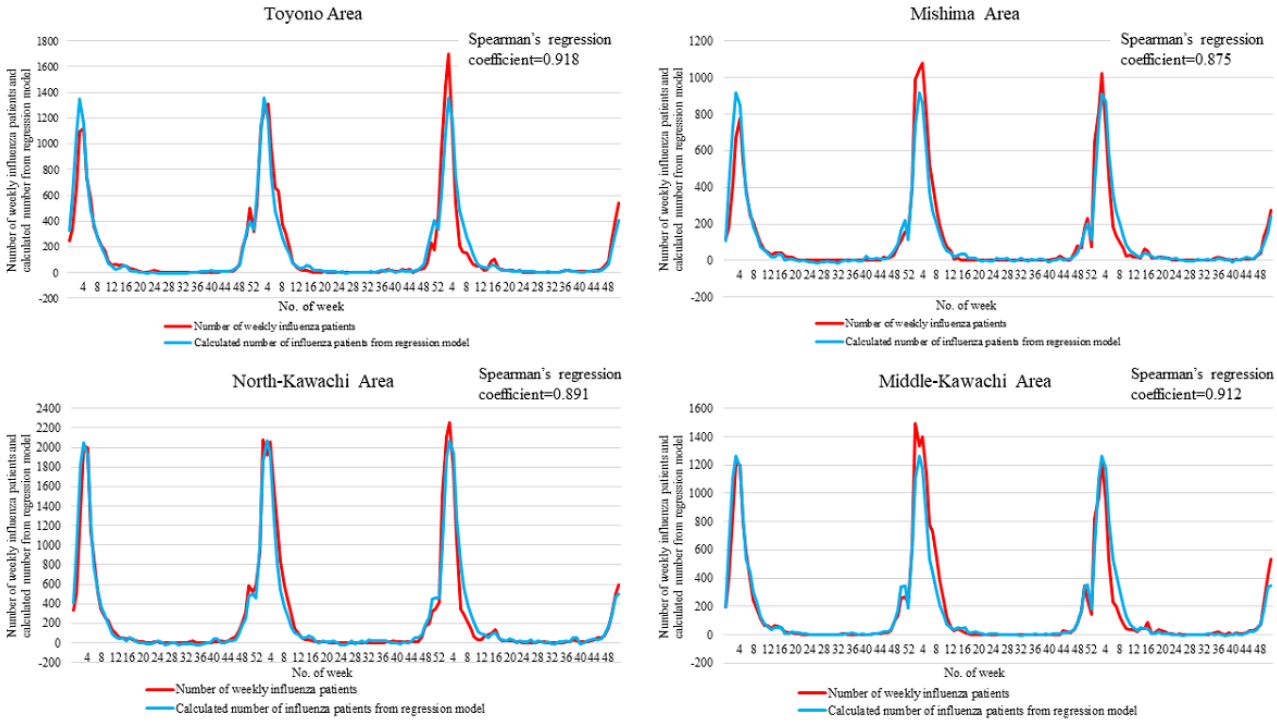


Figure 4. The weekly number of influenza patients and the calculated number of influenza patients from the linear regression model for each region.

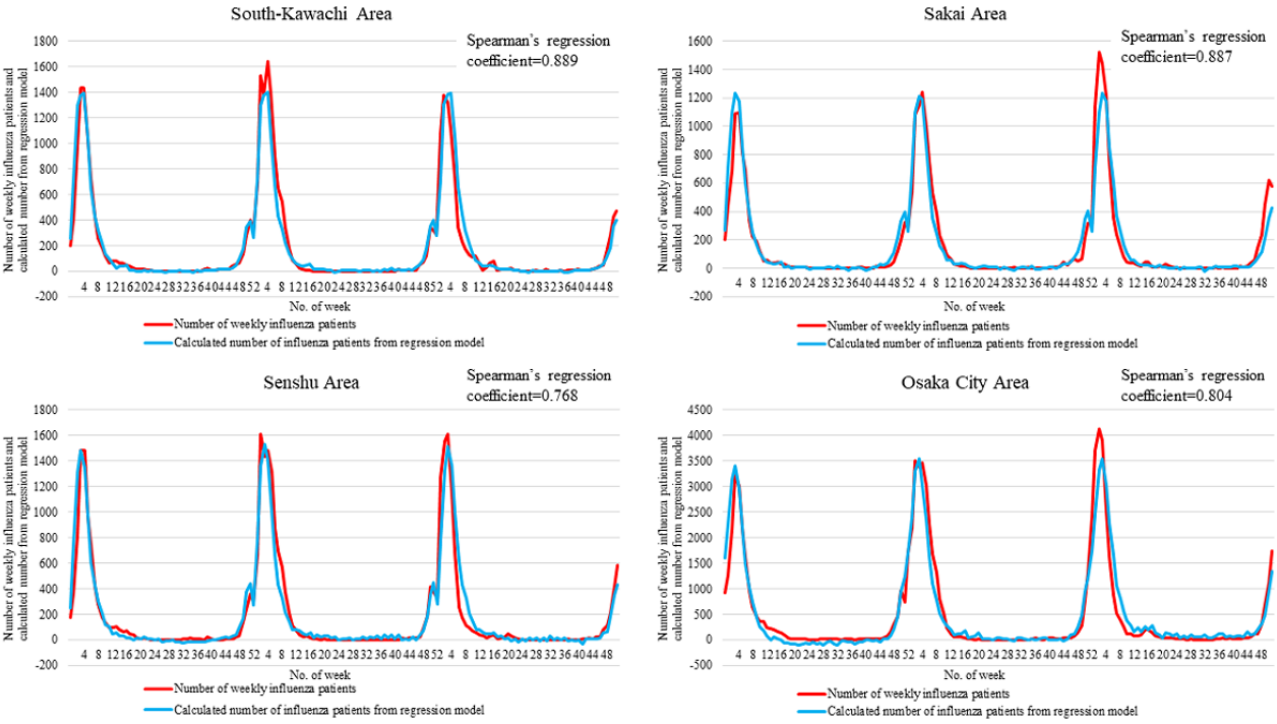
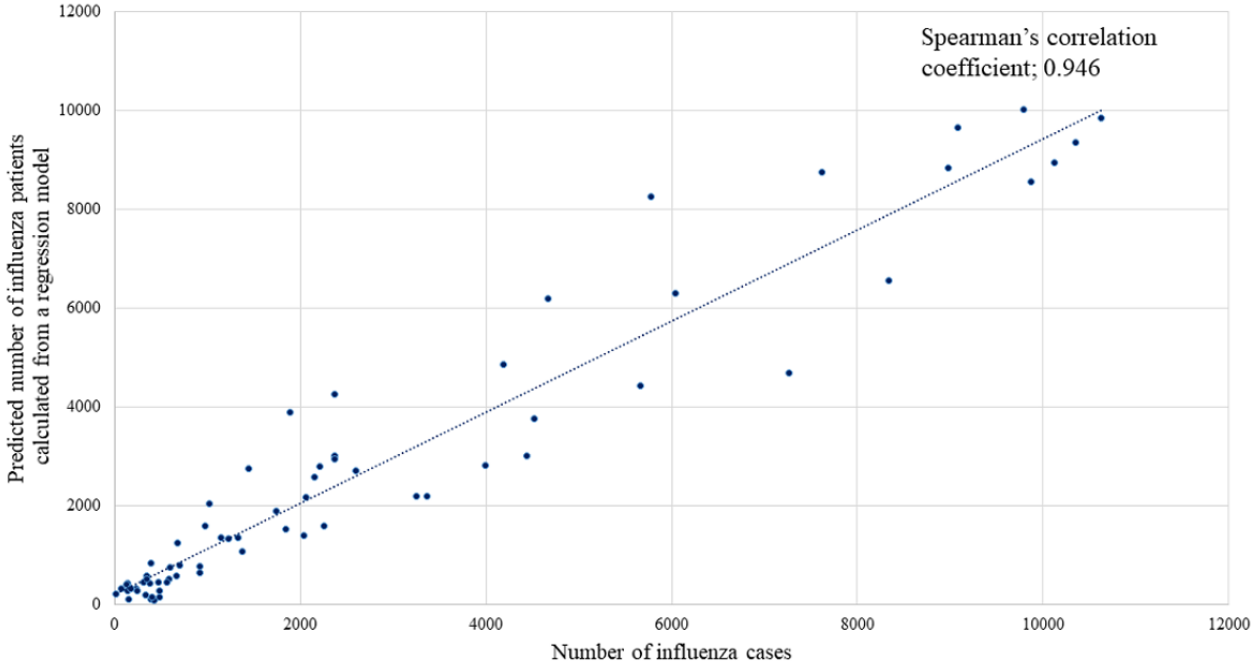


Figure 5. A scatter plot of the number of influenza patients and number of patients calculated using the liner regression model in December-April.



Discussion

Principal Findings

This study revealed that the number of times that "fever" was selected in the mobile app was positively associated with the number of pediatric influenza patients in large urban areas of Japan. We also revealed that the validity of the linear regression model did not differ among the regions. In the analysis by age group, there was also no difference in the relationship between the number of times that "fever" was selected in the mobile app and the number of influenza cases. In the analysis by age group, there was also no difference in the relationship between the number of times that "fever" was selected in the mobile app and the number of influenza cases. The number of pediatric influenza patients and the calculated number from the linear regression model were highly correlated during the influenza epidemic periods (December–April) in Japan. Thus, the number of pediatric influenza patients was related to the number of times that the mobile app was used, and the contribution of the linear regression model was high. Although traditional infectious disease reports require 10 days to 2 weeks, there is a relationship between the number of times that the mobile app was used and the number of influenza patients during the influenza season, and comparing various models using this parameter will lead to the construction of an optimal prediction model in the future.

Comparison With Prior Work

First, the number of pediatric influenza patients was strongly correlated with the calculated number estimated using the linear regression model based on the number of times that "fever" was selected in the mobile app. There have been several studies on syndromic surveillance of influenza using telephone triage data [24-26], and we also revealed that the number of calls for telephone triage was related to the number of influenza patients in Japan [13]. There have also been some syndromic surveillance studies of influenza using mobile app data. A study in Guatemala reported that the Pearson correlation coefficient between the number of ILI patients and the data of mobile apps that members of the general public used to record symptoms was –0.412 [27]. On the other hand, a study in South Korea reported that the Spearman correlation coefficient between the number of influenza patients and the data from a mobile app that supports parents and caregivers when young children have fever was 0.878 [28]. Spearman correlation coefficient between the number of influenza patients and the calculated number estimated using the regression model was 0.946 in this study. Although the number of participants was 189 and the maximum number of influenza patients was approximately 40 per week in the study by Guatemala. The number of participants in this study was approximately 15,000, while that in a study from South Korea was approximately 5 million. The number of infected patients and the data used in the analysis would affect the validity of the results. In order to accurately predict infectious disease epidemics, in addition to obtaining a large amount of data for the analysis, it is also necessary for the disease (eg, influenza) to infect a large number of people.

Next, we divided Osaka Prefecture into eight regions and analyzed the data, and found no difference in the validity of the linear regression model in any of the regions. These results were similar in the sub-group analysis divided by age group. Perry and colleagues compared the predictive performance of N4SID (a numerical algorithm for subspace state space system identification), exponentially weighted moving average, fast orthogonal search (FOS), and regression models in predicting the number of patients visiting the emergency department for respiratory diseases using the number of telephone consultations in Canada [7]. They found that the FOS model had better prediction accuracy than the regression model when the population was large but that the regression model had the best prediction performance in regions with small populations. In this study, we used a linear regression model to assess the relationship between the use of the mobile app and the number

of influenza patients. The volume of data for which the linear regression model is suitable is not known; however, the fact that a good correlation was found between the number of influenza patients and the calculated number of influenza patients estimated using the regression model in all regions may indicate that this model was suitable for the analysis of the amount of data in this data set.

Third, the number of influenza patients and the calculated number of influenza patients estimated using the regression model were well correlated in the analysis of the influenza epidemic season from December to April. The same result was obtained in our previous study using telephone triage data [13]. In a previous study, the data set used in the regression model included the data from 17,000 individuals per year, while the number in this study was approximately 5000 per year. The relationship in the noninfluenza season was not as good as in the influenza season. In addition, there was an interaction between influenza season and the number of times “fever” was selected in the mobile app, and it is likely that the relationship is higher during the influenza season. This is because the number of times that “fever” was selected in the mobile app was used in this study. During the noninfluenza season, “fever” was selected due to other infectious diseases other than influenza, so the relationship between the number of pediatric influenza patients and the value calculated from the regression model might be low.

Limitations

The present study was associated with some limitations. First, reports on influenza in Japan are fixed-point observations based on the Infectious Diseases Control Law, and the survey of influenza is not a survey of all cases. Second, regarding the criteria for reporting influenza patients in Japan, patients who are diagnosed with influenza based on clinical symptoms are included [29]. In Japan, it includes cases reported as influenza based on clinical symptoms alone without the use of an influenza diagnosis kit, so some of these patients may not truly have influenza. Third, as a result of the global spread of the novel COVID-19 in 2020, there has been a change in people’s attitudes toward the prevention of infectious diseases. Therefore, it may affect the frequency of mobile app usage and the number of influenza patients. Fourth, since this study included pediatric patients, it is unclear whether it is appropriate for predicting epidemics in adult and elderly populations. Fifth, as this study followed the assumptions of linear regression analysis and it was a population-based study in a metropolitan area of Japan, the generalizability of these results may be high. However, it has not been assessed for validity in other regions, so external validity needs to be assessed, and we will assess external validity in the future.

Conclusions

In this study, the number of times that mobile apps were used was positively associated with the number of influenza patients. In particular, the relationship between the number of times that mobile apps were used and the number of influenza patients was good during the influenza epidemic season.

Acknowledgments

We thank our colleagues from Osaka University Center of Medical Data Science and Advanced Clinical Epidemiology Investigator’s Research Project for providing their insight and expertise for our research.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Map of medical districts in Osaka prefecture.
[PDF File (Adobe PDF File), 118 KB - formative_v6i2e31131_app1.pdf]

Multimedia Appendix 2
The weekly number of times that "fever" was selected in the mobile app and the weekly number of influenza patients in Osaka, Japan between 2017 and 2019.
[XLSX File (Microsoft Excel File), 24 KB - formative_v6i2e31131_app2.xlsx]

References

1. Kara EO, Elliot AJ, Bagnall H, Foord DGF, Pnaiser R, Osman H, et al. Absenteeism in schools during the 2009 influenza A(H1N1) pandemic: a useful tool for early detection of influenza activity in the community? Epidemiol Infect 2011 Oct 21;140(7):1328-1336. [doi: 10.1017/s0950268811002093]
2. Fan Y, Yang M, Jiang H, Wang Y, Yang W, Zhang Z, et al. Estimating the effectiveness of early control measures through school absenteeism surveillance in observed outbreaks at rural schools in Hubei, China. PLoS One 2014 Sep 24;9(9):e106856 [FREE Full text] [doi: 10.1371/journal.pone.0106856] [Medline: 25250786]
3. Ma T, Englund H, Bjelkmar P, Wallensten A, Hulth A. Syndromic surveillance of influenza activity in Sweden: an evaluation of three tools. Epidemiol Infect 2014 Dec 04;143(11):2390-2398. [doi: 10.1017/s0950268814003240]

4. Vergu E, Grais RF, Sarter H, Fagot J, Lambert B, Valleron A, et al. Medication sales and syndromic surveillance, France. Emerg Infect Dis 2006 Mar;12(3):416-421 [FREE Full text] [doi: 10.3201/eid1203.050573] [Medline: 16704778]
5. Socan M, Erculj V, Lajovic J. Early detection of influenza like illness through medication sales. Cent Eur J Public Health 2012 Jun 1;20(2):156-162 [FREE Full text] [doi: 10.21101/cejph.a3735] [Medline: 22966744]
6. Hall G, Krahn T, Van Dijk A, Evans G, Moore K, Maier A, et al. Emergency department surveillance as a proxy for the prediction of circulating respiratory viral disease in Eastern Ontario. Can J Infect Dis Med Microbiol 2013;24(3):150-154 [FREE Full text] [doi: 10.1155/2013/386018] [Medline: 24421826]
7. Perry AG, Moore KM, Levesque LE, Pickett CWL, Korenberg MJ. A Comparison of Methods for Forecasting Emergency Department Visits for Respiratory Illness Using Telehealth Ontario Calls. Can J Public Health 2010 Nov 1;101(6):464-469. [doi: 10.1007/bf03403965]
8. Hulth A, Rydevik G, Linde A. Web queries as a source for syndromic surveillance. PLoS One 2009 Feb 6;4(2):e4378 [FREE Full text] [doi: 10.1371/journal.pone.0004378] [Medline: 19197389]
9. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. Nature 2009 Feb 19;457(7232):1012-1014. [doi: 10.1038/nature07634] [Medline: 19020500]
10. Davidson MW, Haim DA, Radin JM. Using networks to combine "big data" and traditional surveillance to improve influenza predictions. Sci Rep 2015 Jan 29;5:8154 [FREE Full text] [doi: 10.1038/srep08154] [Medline: 25634021]
11. Dong X, Boulton ML, Carlson B, Montgomery JP, Wells EV. Syndromic surveillance for influenza in Tianjin, China: 2013-14. J Public Health (Oxf) 2017 Jun 01;39(2):274-281. [doi: 10.1093/pubmed/fdw022] [Medline: 26968483]
12. Katayama Y, Kiyohara K, Komukai S, Kitamura T, Ishida K, Hirose T, et al. Relationship between the number of pediatric patients with rotavirus and telephone triage for associated symptoms. Am J Emerg Med 2021 Jan;39:6-10 [FREE Full text] [doi: 10.1016/j.ajem.2020.03.039] [Medline: 32241629]
13. Katayama Y, Kiyohara K, Komukai S, Kitamura T, Ishida K, Hirose T, et al. The relationship between seasonal influenza and telephone triage for fever: A population-based study in Osaka, Japan. PLoS One 2020 Aug 6;15(8):e0236560 [FREE Full text] [doi: 10.1371/journal.pone.0236560] [Medline: 32760164]
14. Katayama Y, Kiyohara K, Hirose T, Matsuyama T, Ishida T, Nakao S, et al. A Mobile App for Self-Triage for Pediatric Emergency Patients in Japan: 4 Year Descriptive Epidemiological Study. JMIR Pediatr Parent 2021 Jun 30;4(2):e27581 [FREE Full text] [doi: 10.2196/27581] [Medline: 34255709]
15. The Census in Osaka Prefecture in 2015. Osaka Prefectural Government. URL: https://www.pref.osaka.lg.jp/attach/1891/00210094/27jinkoutoukihon.pdf [accessed 2021-05-13]
16. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. The Lancet 2007 Oct;370(9596):1453-1457. [doi: 10.1016/S0140-6736(07)61602-X]
17. Nakamura Y, Sugawara T, Kawanohara H, Ohkusa Y, Kamei M, Oishi K. Evaluation of Estimated Number of Influenza Patients from National Sentinel Surveillance Using the National Database of Electronic Medical Claims. Jpn J Infect Dis 2015;68(1):27-29. [doi: 10.7883/yoken.jjid.2014.092]
18. Murakami Y, Hashimoto S, Kawado M, Ohta A, Taniguchi K, Sunagawa T, et al. Estimated Number of Patients with Influenza A(H1)pdm09, or Other Viral Types, from 2010 to 2014 in Japan. PLoS One 2016 Jan 19;11(1):e0146520 [FREE Full text] [doi: 10.1371/journal.pone.0146520] [Medline: 26784031]
19. Okabe N, Yamashita K, Taniguchi K, Inouye S. Influenza surveillance system of Japan and acute encephalitis and encephalopathy in the influenza season. Pediatr Int 2000 Apr 20;42(2):187-191. [doi: 10.1046/j.1442-200x.2000.01206.x] [Medline: 10804740]
20. Shimada T, Sunagawa T, Taniguchi K, Yahata Y, Kamiya H, Yamamoto KU, et al. Jpn J Infect Dis 2015;68(2):151-158. [doi: 10.7883/yoken.jjid.2014.125]
21. Infectious Disease Surveillance System in Japan. National institute of infectious diseases. URL: https://www.niid.go.jp/niid/images/epi/nesid/nesid_en.pdf.(in [accessed 2021-05-13]
22. The Information Center of Infectious Disease in Osaka Prefecture. URL: http://www.iph.pref.osaka.jp/infection/2-old.html [accessed 2021-05-13]
23. The Health and Medical Care Plan by Osaka Prefectural Government. Osaka Prefectural Government. URL: https://www.pref.osaka.lg.jp/attach/2502/00281086/compactban%20kizonbyousyousyuusei.pdf [accessed 2021-05-13]
24. Lucero-Obusan C, Winston CA, Schirmer PL, Oda G, Holodniy M. Enhanced Influenza Surveillance Using Telephone Triage and Electronic Syndromic Surveillance in the Department of Veterans Affairs, 2011-2015. Public Health Rep 2017 Jul 10;132(1_suppl):16S-22S [FREE Full text] [doi: 10.1177/0033354917709779] [Medline: 28692402]
25. Yih WK, Teates KS, Abrams A, Kleinman K, Kulldorff M, Pinner R, et al. Telephone triage service data for detection of influenza-like illness. PLoS One 2009 Apr 17;4(4):e5260 [FREE Full text] [doi: 10.1371/journal.pone.0005260] [Medline: 19381342]
26. Moore K. Real-time syndrome surveillance in Ontario, Canada: the potential use of emergency departments and Telehealth. Eur J Emerg Med 2004 Feb;11(1):3-11. [doi: 10.1097/00063110-200402000-00002] [Medline: 15167186]

27. Prieto JT, Jara JH, Alvis JP, Furlan LR, Murray CT, Garcia J, et al. Will Participatory Syndromic Surveillance Work in Latin America? Piloting a Mobile Approach to Crowdsource Influenza-Like Illness Data in Guatemala. JMIR Public Health Surveill 2017 Nov 14;3(4):e87 [FREE Full text] [doi: 10.2196/publichealth.8610] [Medline: 29138128]
28. Kim M, Yune S, Chang S, Jung Y, Sa SO, Han HW. The Fever Coach Mobile App for Participatory Influenza Surveillance in Children: Usability Study. JMIR Mhealth Uhealth 2019 Oct 17;7(10):e14276 [FREE Full text] [doi: 10.2196/14276] [Medline: 31625946]
29. Zaraket H, Saito R. Japanese Surveillance Systems and Treatment for Influenza. Curr Treat Options Infect Dis 2016 Oct 10;8(4):311-328 [FREE Full text] [doi: 10.1007/s40506-016-0085-5] [Medline: 28035195]

Abbreviations

FOS: fast orthogonal search
ILI: influenza-like illness

Edited by A Mavragani; submitted 10.06.21; peer-reviewed by E Lau, M Salimi, C Lynch; comments to author 04.07.21; revised version received 28.08.21; accepted 22.12.21; published 10.02.22.

Please cite as:
Katayama Y, Kiyohara K, Hirose T, Ishida K, Tachino J, Nakao S, Noda T, Ojima M, Kiguchi T, Matsuyama T, Kitamura T
An Association of Influenza Epidemics in Children With Mobile App Data: Population-Based Observational Study in Osaka, Japan
JMIR Form Res 2022;6(2):e31131
URL: https://formative.jmir.org/2022/2/e31131
doi:10.2196/31131
PMID:35142628

Original Paper

Determining Priorities in the Aboriginal and Islander Mental Health Initiative for Youth App Second Phase Participatory Design Project: Qualitative Study and Narrative Literature Review

Josie Povey[1], BOT, MPH; Michelle Sweet[2], PhD; Tricia Nagel[1], MBBS, PhD; Anne Lowell[3], PhD; Fiona Shand[4], PhD; Jahdai Vigona[1], Cert III (ComS); Kylie M Dingwall[5], PhD

[1]Menzies School of Health Research, Charles Darwin University, Casuarina, Australia
[2]Menzies School of Health Research, Charles Darwin University, Adelaide, Australia
[3]Northern Institute, Charles Darwin University, Darwin, Australia
[4]Black Dog Institute, University of New South Wales, Sydney, Australia
[5]Menzies School of Health Research, Charles Darwin University, Alice Springs, Australia

Corresponding Author:
Josie Povey, BOT, MPH
Menzies School of Health Research
Charles Darwin University
Red 9 CDU Campus
Casuarina, 0820
Australia
Phone: 61 8 8946 8487
Email: josie.povey@menzies.edu.au

Abstract

Background: Digital mental health tools can promote access to culturally safe early intervention mental health services for Aboriginal and Torres Strait Islander young people. Participatory design methodology facilitates user engagement in the co-design of digital resources. However, several challenges have been identified that limit the methodological rigor of this approach.

Objective: This paper aims to present an in-depth account of the second phase of participatory design in the development of the Aboriginal and Islander Mental Health Initiative for Youth (AIMhi-Y) app.

Methods: A *first idea* storyboard, generated from a formative phase of the AIMhi-Y project, was refined through a series of youth co-design workshops and meetings. A narrative review of the literature, 6 service provider interviews, and engagement with an expert reference group also informed the design process. Generative design activities, storyboarding, discussions, and voting strategies were used.

Results: The participatory design process identified the app features preferred by young people and service providers and assessed their alignment with current recommendations from the scientific literature. Findings from the co-design process are presented across 9 app characteristic domains. Integration of findings into app design proved complex. Although most preferred features identified by young people were included to some degree, other inclusions were restricted by budget, time, and the need to integrate best practice recommendations. A process of prioritization was required.

Conclusions: Participatory design is often cited in the development of digital mental health resources; however, methods are diverse and often lack detailed descriptions. This study reports the outcomes and strategies used to determine priorities in the second phase of the development of the AIMhi-Y app. We provide an example and the key learnings to inform others seeking to use participatory design with a similar cohort.

(JMIR Form Res 2022;6(2):e28342) doi:10.2196/28342

KEYWORDS

Aboriginal and Torres Strait Islander; young people; digital mental health; app; participatory design; decision-making; mobile phone

Introduction

Background

Aboriginal and Torres Strait Islander young people possess many strengths that stem from an ancient and resilient culture [1]. However, they also experience negative influences on well-being, mental illness, and suicide at higher rates than their non-Indigenous counterparts [2,3]. Help-seeking is impeded by stigma, shame, differences in language and culture, geographical isolation, and cost [4-7]. Mental health services that integrate Aboriginal and Torres Strait Islander people's holistic health worldviews are recommended [8] but remain limited [4]. Culturally responsive technological innovations offer the potential to bridge mental health service gaps and address increasing demand [9-12].

Digital mental health (dMH) resources include "mental health, suicide prevention, and alcohol and other drugs services delivered via a digital platform" [13]. These can be stand-alone, clinician-supported, or used alongside face-to-face services [13]. Evidence supporting the acceptability and effectiveness of dMH solutions for Aboriginal and Torres Strait Islander young people is emerging [9,14-17]. Nevertheless, culturally adapted dMH resources remain limited [18,19]. End user involvement in the design of dMH interventions is recommended [20,21] and assists in valuing Aboriginal and Torres Strait Islander expertise [11] and decolonizing research processes [22,23].

Participatory design is often used in dMH development to engage users. It is a multistage process (defining the problem, identifying solutions, and generating and evaluating prototypes), with end user involvement throughout [21,24,25]. Research tools can include brainstorming, storyboarding, scenarios, prototyping, role plays, 2D sketching, group discussion, and user journeys [24,26]. Participatory design upskills and engages end users in experiential action-based activities with shared informed decision-making throughout [21,27]. Integration of best practice evidence and expert advice [21,26], visibility of outcomes [28,29], adequate time and resourcing [30], and a flexible iterative approach [30] are recommended to ensure success.

Although considered best practice, several challenges exist in undertaking participatory design. Tensions in preferences among designers; end users; and therapy and pedagogy experts, whose input is vital in designing resources that are appealing, usable, and address therapeutic goals, can prove challenging [31,32]. In addition, differences in user preferences based on age, gender, video game familiarity, and mental health need often arise [19,33-35]. Strategies to address these predictable tensions between and within stakeholder groups and manage selection bias include the following: defining roles, upskilling participants, developing a positive team approach, and supporting meaningful engagement [30,36,37]. Such strategies value end user input and help overcome power differentials [19,33-35,38]. In addition, co-design in cross-cultural contexts requires specific engagement strategies that embrace language and worldview differences and empower Indigenous young people throughout [39,40]. In summary, successful participatory design characteristics include the following:

- End user participation throughout
- Finding balance between tailoring-to-context and a one-size-fits-all approach
- Generation of resources through experiential and playful action-based activities
- Shared decision-making throughout, which is transparent and iterative
- Ensuring the empowerment, upskilling, and safety of participants
- Strategies to address predictable tensions between user preferences and other stakeholder groups (eg, app developers, researchers, clinicians, and educators)
- Strategies to address differences in end user preferences related to gender, age, mental health need, and technological familiarity
- Strategies to balance stakeholder expectations and resource availability (eg, time and budget)

Although recommended [24,26], in-depth reporting of processes used throughout the participatory design of dMH tools remains scarce [41], with some notable exceptions [36,42]. This omission inhibits the development of participatory design as a sound methodological approach and limits the understanding of *how* and *if* participatory design increases effectiveness, reach, adoption, and implementation of dMH resources [33,43,44].

The Aboriginal and Islander Mental health initiative (AIMhi) program of research develops, tests, and implements dMH resources for Aboriginal and Torres Strait Islander people [45,46]. Resources such as the Stay Strong Plan, a culturally adapted low-intensity cognitive behavioral therapy (CBT) intervention (available in digital and paper-based formats) have demonstrated effectiveness in reducing psychological distress and substance misuse among adults [16,47]. Developed and implemented over a decade within the Greater Darwin and Tiwi Island regions, the initiative is recognized and trusted within the community. Community feedback, suggesting changes to appeal to young people [9,10], led to the commencement of a youth-focused program of research (Aboriginal and Islander Mental Health Initiative for Youth [AIMhi-Y]). Phase 1, reported elsewhere [48], aimed to understand Aboriginal and Torres Strait Islander young people's lived experience of mental health and well-being, views on dMH resources, and design and content preferences. Preferred features included a smartphone-based app that incorporates strengths-based mental health information presented through relatable storytelling and a fun, appealing, easy-to-use interface, which encouraged app progression [48]. Phase 1 resulted in the first 2D storyboard, which was used as a basis for this second design phase.

Objective

Using a participatory design approach, this study (phase 2) aims to iteratively enhance the AIMhi-Y app storyboard and finalize the first release of the AIMhi-Y app. This was done through collaboration with Aboriginal and Torres Strait Islander young people, service providers, integration of evidence and best practice recommendations, and iterative usability testing with the app developer. In this paper, we provide an in-depth report

of our findings and experiences in developing the AIMhi-Y app.

Methods

Study Design

This 2-year qualitative participatory design project conducted youth co-design activities (workshops and meetings), service provider interviews, a narrative synthesis of the literature, and app developer meetings. Several activities occurred concurrently as shown in Figure 1.

Figure 1. Timeline of phase 2 Aboriginal and Islander Mental Health Initiative for Youth (AIMhi-Y) app development project activities. IYRG: Indigenous youth reference group.

Participant Selection and Setting

Purposive sampling was used to recruit young people and service providers with experience in working with young people across 3 sites (2 schools and 1 residential drug rehabilitation facility). All 3 sites that engaged in phase 1 [48] continued their engagement in phase 2. Youth participants invited to the co-design workshops were nominated by the teaching or support staff at their site. Inclusion criteria were as follows: Aboriginal or Torres Strait Islander descent, aged 8-18 years, and willing to talk in a group setting in English (at urban and remote sites) or Tiwi language (at the remote site). Service providers were key contact persons at each site and their engagement with the project ranged from 6 months to 2 years.

After completing the co-design workshops at the school and residential rehabilitation facility, a separate group of young people was recruited to form an Indigenous youth reference group (IYRG). This smaller group provided prompt feedback throughout the app development stage. Recruitment occurred via email and social media advertising, disseminated through project networks. Individuals identifying as Aboriginal or Torres Strait Islander, aged 15-25 years, and willing to attend meetings via videoconference (owing to the COVID-19 social distancing regulations) were recruited to the IYRG. Participants involved in the phase 1 or phase 2 co-design workshops were prioritized throughout recruitment, aiming for diverse age and gender representation.

Informed oral consent was obtained (face-to-face or via phone) from all participants and from guardians of those aged <16 years before participation. A pictorial consent flip chart and participant information handouts were used to assist communication. Interpreters were available for those who spoke a language other than English at home. The first author (JP) led the project, supervised by 5 senior researchers with clinical and dMH expertise. In addition, guidance from 4 Indigenous researchers (n=2, 50% youth and n=2, 50% women) played a vital role in establishing rapport, developing and refining data collection activities, and interpreting during face-to-face discussions. Research team members were trained in distress management and suicide awareness. Site-specific risk management processes were agreed upon by each organization before commencement. They included the provision of immediate safety measures by researchers, notification of senior research team members, and follow-up and referral to treatment services provided by the nominated site contact person.

An existing expert reference group (established in phase 1) included 12 service providers and researchers with relevant expertise from mental health, drug and alcohol, child protection, and education sectors (n=3, 25% men and n=5, 42% Aboriginal). The expert reference group met or were contacted via email biyearly to advise on research processes, app design, and content [48].

Ethical Approval

Ethical approval was obtained from Menzies School of Health Research Human Research Ethics Committee (HREC 2017-2991), including the Indigenous subcommittee and the Northern Territory Department of Education Research Ethics Committee (reference number: 13417).

Data Collection

Co-design Workshops and IYRG Meetings

A total of 16 co-design workshops and 5 IYRG meetings were held with youth participants. Workshops and meetings used generative and discussion methods to design and refine app attributes [24]. Participants were presented with previous findings, mock-ups, storyboards, and app prototypes. Voting (show of hands) and discussion strategies were used to reach a consensus. Co-design workshop and meeting activities (Textboxes 1 and 2) were developed, informed by the literature and research team consultation.

Textbox 1. Co-design workshop activities guide.

Workshop 1

- User flow diagrams (paper-based 2D visual representations; A0 size) integrating the app storyboard drafted in phase 1; oriented and educated young people to purpose, features, and content through a minilecture.
- Interactive group discussion; elicited participants' preferred features, interactive components, and usability preferences.
- Introduction of mock-ups (Figure 2) using printed worksheets; initiated group discussion and written feedback on design styles, storylines, and character attributes.

Workshop 2

- Indigenous researcher-drawn home page mock-ups; initiated group discussions about metaphors, reward systems, and progress displays.
- Young people generated their own home page ideas through individual and small group sketch activities.
- Revised design styles (based on workshop 1) were introduced and discussed. Voting (show of hands) was conducted to reach an agreement on look and feel.

Workshop 3

- Video scripts based on lived experience examples provided from phase 1 [48] were developed by the research team and presented to the group through interactive discussion. Young people were asked to draw a series of scenes to complete the short video. Sketching and discussion explored participant's understanding of information, preferred perspectives (first person, third person, or noncharacter delivered), imagery, storylines, character strengths, challenges, and personal attributes.

Workshop 4

- A total of 3 youth-generated ideas for app narratives and metaphors (generated within workshop 2) were presented. Young people sketched and described a mock-up of their preferred narrative or generated a new one.

Textbox 2. Indigenous youth reference group meetings activities guide.

Meeting 1

- Group members were presented with relevant findings from phase 1 and the co-design workshops through a minilecture. A revised app storyboard oriented the participants to overall aim, app narrative, structure, and characters.
- Indigenous youth researcher–generated logo mock-ups prompted group discussion.
- Visual aids prompted discussion of preferred reward systems, background images, color schemes, and design element preferences.

Meeting 2

- Iterated versions of the logo, backgrounds, and reward images were presented for discussion.
- Researcher-generated digital home page mock-ups prompted group discussion.
- Persona development activities prompted group discussion and explored the character's attributes such as family, strengths, cultural identity, and language.

Meeting 3

- Completed backgrounds and logo were presented for confirmation by group members.
- Storyboard activities and group discussion completed a detailed review of narrative, wording, reward metaphors, and images for Ramone quest.
- Character names and support person roles were reviewed and discussed.

Meeting 4

- Draft app prototype was presented using screen-sharing software. Initial impressions and feedback were elicited through group discussion.
- A researcher-generated mock-up of rewards page prompted group discussion.

Meeting 5

- Aboriginal and Islander Mental Health Initiative for Youth app prototype 1.0 presented with feedback elicited for integration at a later stage.
- Pilot study procedures were presented and discussed. Strategies for recruitment and consent were discussed, with changes made in line with recommendations from participants.

Figure 2. Prototypes used to elicit discussion on design attributes.

Workshops were facilitated by the first author (JP) and 2 youth Indigenous researchers (JV), promoting group rapport and enabling an insider perspective [49]. A senior Indigenous researcher, who was an interpreter and community elder, was present at all workshops in the remote school, offering participants the option to engage using their first language. Several unsuccessful attempts to involve an interpreter at the residential rehabilitation center were made. This was considered important as few of these participants spoke English as their first language, in contrast to the youth participants from the urban schools. The research team acknowledges that the presence of interpreters would have likely deepened discussion and improved rapport [6]. Workshops began with an icebreaker activity, introductions, sharing of food, and discussing kinship and family relationships. Workshops were 40 to 90 minutes in duration. Groups were divided by gender to follow local cultural protocols [50]. Young people who identified as gender diverse chose the group with which they felt most comfortable to work. Larger groups (>10 people) were divided into self-selected peer groups of 3 to 5 people. Groups rotated around activities in a classroom environment. Workshops were audio recorded and complemented by field notes.

The IYRG meetings were facilitated by the first author (JP), a youth Indigenous researcher (JV), and a senior researcher (MS) and held via videoconference after school. Meetings started with introductions, icebreaker activities, and discussing kinship and family relationships. Meetings were 90 minutes in duration and members were reimbursed for their time (Aus $75 [US $54] sitting fee).

Service Provider Interviews

A total of 6 semistructured in-depth interviews of 40 to 60 minutes duration with service providers were conducted concurrently with youth co-design activities. Interviews aimed to gain feedback on the app storyboard, refine features, and explore implementation considerations. Interviews were conducted at the study site or over the phone by the first author (JP), with an Indigenous researcher (JV) present on 2 occasions. Service provider interview guides (Multimedia Appendix 1) were informed by the literature and refined within the research team.

Narrative Literature Synthesis

A narrative literature synthesis was undertaken to inform app development. Searches used a combination of search terms, including mental health, technology use, adolescence, Indigenous populations, therapeutic interventions, and efficacy. Inclusion criteria were broad, aiming to capture evidence for dMH interventions across multiple population groups (youth, adult, and Indigenous). Exclusion criteria included articles not published in English owing to limited resources. Searches were conducted in multiple web-based databases, including EBSCOhost, Scopus, CINAHL, PubMed, Google Scholar, Informit, and organization websites (Multimedia Appendix 2) and included literature published from 2000-2020. Several dMH development guidelines and safety standards were included in the review. Multiple searches were conducted between July 2019 and July 2020, and the findings were used to inform the design process. Therefore, the specific number of articles found and excluded was not recorded. The first author (JP) conducted these searches and formulated the synthesis of findings from the review of relevant literature. Other authors (KMD and TN) provided feedback and guidance for the identification of additional relevant literature. References were uploaded to the Endnote referencing system (Clarivate) and categorized into groups.

Analysis

The first author (JP) transcribed all the sections of the co-design workshop recordings that were relevant to the design process. The first and second authors independently thematically coded data from workshops, meetings, and interviews by combining an inductive process, generating codes from the data (MS) [51], with deductive analysis, based on the app preferences identified in phase 1 (JP) [52]. Following consensus, further discussions

were conducted with the full research team. Scientific literature was concurrently synthesized into table format by the first author (JP) and thematically analyzed according to the recommended dMH attributes. The themes emerging from youth co-design workshops, meetings, and service provider interviews were then integrated into the themes identified through the literature review. Findings from all analyses informed the design and development of the AIMhi-Y app (version 1.0). The dMH safety standards and youth preferences were prioritized throughout development. Analysis occurred before and concurrent to meetings between the app developer and the research team to translate design preferences into prototypes. Demographic data were analyzed for descriptive statistics to describe the sample.

Results

Participant Demographics

Co-design Workshops and Meetings

A total of 16 co-design workshops (5 groups with a series of 2-4 workshops) and 5 IYRG meetings involved 75 young people (Table 1). Participants were primarily female (40/75, 53%) and currently engaged in school (66/75, 89%; Table 2). A total of 10 young people (n=6, 60% remote and n=4, 40% urban) continued engagement from phase 1 [48]. In all, 38% (6/16) IYRG participants were involved in previous co-design workshops. Researcher observations of participant body language and interaction within the group indicated that engagement in workshop activities ranged from disinterest to enthusiastic participation in consultation and co-design and was influenced by multiple factors, including participant mood, research activity, and facilitator. Flexibility in the workshop design allowed participants to engage and reengage at will. On one occasion, a young person expressed distress unrelated to workshop content and was provided immediate support by the research team, with follow-up by the identified site contact person, as per the agreed risk management processes.

Table 1. Number of youth participants at each co-design workshop or meeting (N=75).

Workshop or meeting	Remote School, n (%)		Urban School, n (%)		Rehabilitation service, n (%)	IYRG[a], n (%)
	Male	Female	Male	Female	Mixed gender	Mixed gender
1	7 (9)	6 (8)	12 (16)	19 (25)	4 (5)	16 (21)
2	6 (8)	5 (7)	11 (15)	15 (20)	5 (7)	14 (19)
3	7 (9)	6 (8)	8 (11)	16 (21)	N/A[b]	8 (11)
4	9 (12)	7 (9)	N/A	N/A	N/A	6 (8)
5	N/A	N/A	N/A	N/A	N/A	5 (7)

[a]IYRG: Indigenous youth reference group.

[b]N/A: not applicable.

Table 2. Co-design workshop and Indigenous youth reference group participant demographics.

Variable	Values
Gender (N=75), n (%)	
Female	40 (53)
Male	33 (44)
Gender diverse	2 (3)
Age (years; n=56[a]), mean (SD; range)	15.14 (1.74; 8-18)
Reside in a very remote community (n=65[a]), n (%)	24 (37)
English not the main language spoken at home (n=56[a]), n (%)	21 (38)
Not currently engaged in school (n=65[a]), n (%)	7 (11)
Owned a smartphone (n=56[a]), n (%)	41 (73)

[a]Owing to missing data.

Service Provider Interviews

A total of 6 service provider interviews (4, 67% men and 2, 33% Aboriginal) across 3 sites included teachers (2/6, 33%), school or service support personnel (3/6, 50%), and a manager (1/6, 17%).

Youth, IYRG, and Service Provider App Preferences

Overview

Most youth participants held strong preferences for activity types, engagement strategies, interface, graphic design, and language preferences. Safety, security, psychological approach, evidence, implementation, and accessibility features, although valued by some young people, were more prominent throughout service provider preferences. We present the findings from co-design activities in the following sections.

Activity Types

Young people's preferred activity types included *videos, minigames, and self-monitoring* of well-being and goals. Activities that incorporate *strengths-based mental health information* and *skill development* were preferred:

> *...It's looking like a good app...Add more videos...[make them] short, talking about mental health and wellbeing...[I] like the depth of the animal game with the mindfulness part...* [Female participants, aged 15-17 years, IYRG members]

Activity types revealed many diverse opinions from youth and service providers. Participants suggested that drivers of differences in preferences were related to age, gender, video game familiarity, community of origin, school attendance, level of distress, and cultural connectedness:

> *...if it's like a really young one, they will be looking onto the [mini] game, then they feel more comfortable...if a female client, [who is] really worried; really down; very lonely...she may be looking onto something else...kids who're having significant absence from school, they don't [respond to trophies or points]...[they] need a reward in that they're getting more - more games, more activities...* [Service provider]

Engagement

Participants valued options to *customize, personalize, represent progress, gamify, prompt real-world action, allow exploration,* and *provide unlimited use*:

> *...is this one of the characters you get to customize...choose what they are wearing?* [Female participant, aged 13 years, co-design workshop]
>
> *...as you went down the path...have footsteps to show you have moved...the sound of feet crunching on dirt...things that...reflect the decision you made...* [Female participant, aged 15 years, co-design workshop]
>
> *...[include] motivational things that make you get up and do something...like star jumps...to get you moving...or like 'judging on the questions you answered today...it would be helpful for you to do...'* [Female participant, aged 15 years, co-design workshop]
>
> *...if they're really enjoying it, and they finish it...Is there something that they can keep on going back to? Like their goals – that keep changing or updating?* [Service provider]

Participants valued *storytelling* with relatable characters, which helped build trust and prompt reflection. Stories told from different perspectives, including peers and family members, were necessary. The character's gender influenced some young people's willingness to listen:

> *[I] like that he is telling his own story...because that would help other people overcome what they are going through as well...* [Female participant, aged 16 years, co-design workshop]
>
> *Male or female I would listen because that...is their experience in life...I reckon girls are more interesting than boys...we already know everything about boys, but understanding girls would be better...* [Male participant, aged 16 years, co-design workshop]
>
> *Nah, I wouldn't listen to it as much [if the character was female]...[Include] two different stories - for girl - like teasing and gossip and anxiety - or some thing more serious like for boys...* [Male participant, aged 15 years, co-design workshop]

Figure 3 provides selected youth-drawn storyboards and descriptions highlighting story content and imagery suggestions.

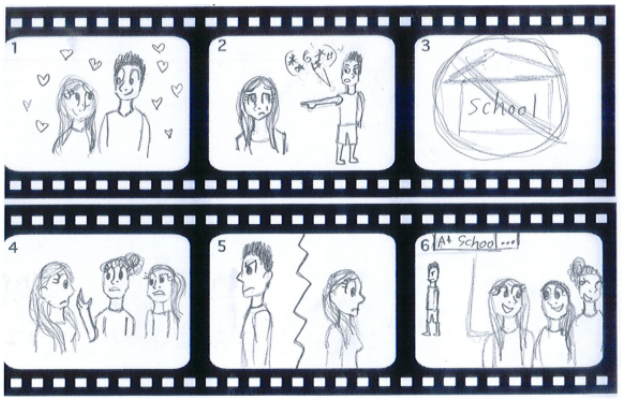
The description of the storyboard, *body stressing*, was as follows:

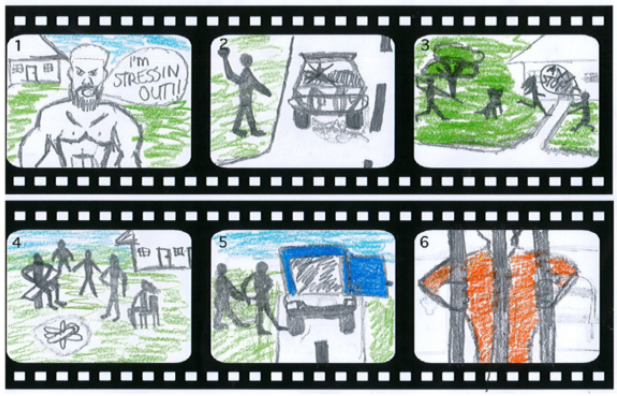
> *[He is] feeling stressed can lead to sickness, drugs and alcohol and car crash, [he is an] angry man...getting in trouble with police – gettin' sent to prison...he is trying to kill himself...he gotta stop gunja (marijuana), be with family...recovery...he [character] should tell the story - he should tell it after [it's happened]...Mum as well - yeah she can tell her part of the story...[and] uncle for taking him out hunting, like see the difference.* [Male participant, aged 18 years, co-design workshop]

The description of the storyboard, *relationship problems*, was as follows:

> *...there is a girl, she has worries with [her] boyfriend and gossip...not going to school, fighting with her mum, focusing on her boyfriend too much and texting all the time, staying up late, or fake going to sleep early...[she is feeling] sad, depressed. He is stressing her out, he want money, she is jealous...he is jealous of her, he wants to fight, about another boy, he is really clingy...[Then] she dumps the boyfriend and has her own life...talks to her best friend...she helps her out to dump the boy, she takes her back to school, forget about that boy, she makes different excuses to get away from him...she talks to grandmother and grandmother talks to boyfriends parents and says keep away from my grandaughter.* [Female participant, aged 18 years, co-design workshop]

Figure 3. Participant-drawn storyboards: 'Relationship problems' and 'body stressing'.





Interface and Graphic Design

Participants valued *aesthetic, relevant,* and *intuitive designs,* with diverse opinions on styles and graphics noted:

> *...some people might like to see like a groovy animation and some might like a more real one...everyone is different.* [Female participant, aged 15 years, co-design workshop]

> *...it was [interesting] at first and then it just started drag on a bit...if I was a fisherman I would find it interesting but I don't fish [we] are from [Central Australia]...make it look like it is in the outback, like in community...* [Male participant, aged 16 years, co-design workshop]

IYRG members often referred to valuing and including diversity of all those involved and incorporated this into their app content and design suggestions:

> *I like how the desert and bush intertwine. We don't all face same situation [but] we're [Indigenous youth] still interconnected. The trouble that we face is similar no matter where we are.* [Female participant, aged 16 years, IYRG Member]

> *Get [images] from all over NT and put in the shapes – incorporate salt water and desert...[include] stories from different places incorporating the [Aboriginal] languages from that place...We live in open places both Top End and Central Desert – needs opening perspective, looking up and out e.g. sunset...* [Multiple IYRG members]

Language

Youth and service providers identified preferences for *audio options* and *integration of Indigenous Languages*:

> *Is it something you only read, or is there someone going to be reading it?* [Male participant, aged 17 years, co-design workshop]

Youth participants highlighted preferences for gender-neutral *familiar metaphors and imagery* relating to nature, a journey, making or nurturing something, sport, ceremony, and lore:

> *Make it look more alive with bird life and plants.* [Female participant, aged 16 years, co-design workshop]

> *Going out bush, fishing...driving...over time you can get a campsite, fishing rod...rewards like that...could you like, pick the upgrade...[a] hot chocolate...* [Female participant, aged 15 years, co-design workshop]

> *Rewards need to be gender neutral.* [Female participant, aged 17 years, IYRG]

Figure 4 provides youth-drawn images and descriptions of proposed metaphors.

The description of the proposed metaphor, *we all have feelings,* was as follows:

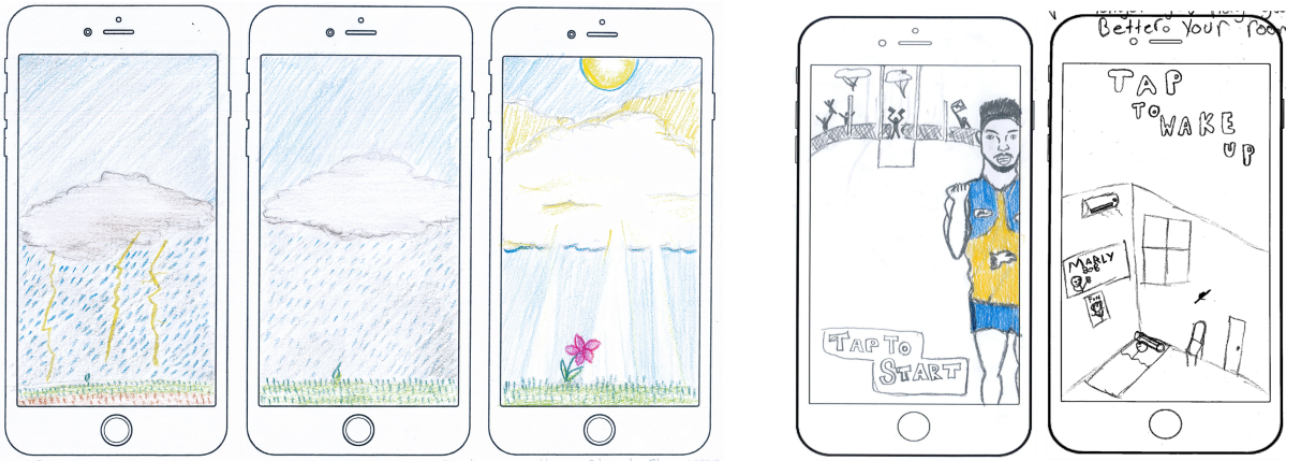
> *The rainy, stormy, cranky cloud is angry and it wants to burn everything with lightening. But then it realises that it shouldn't be destroying nature, but helping nature grow. The grumpy cloud calms down and stops striking lightening because it can see a little seed trying to grow. The cloud pours rain on the sapling. The cloud stops raining and moves onto the side so the sun could shine on the little sapling. The sapling grows into a beautiful flower. The rain bought nature to life.* [Female participant, aged 14 years, co-design workshop]

The descriptions of the proposed metaphors, *home page ideas,* were as follows:

> *Tap game to start...kick the footy to score.* [Male participant, aged 18 years, co-design workshop]

> *The longer you play the game the better your room looks, like upgrades...aircon...music...* [Male participant, aged 17 years, co-design workshop]

Figure 4. Participants proposed metaphors: 'We all have feelings' and 'Homepage ideas'.



Safety

Young people and service providers highlighted the importance of *crisis support features*:

> *I liked on the get help page, you took in consideration different ways people get help...text, phone...* [Female participant, aged 17 years, IYRG Member]

Psychological Approach

Young people and service providers valued strengths-based brief interventions, identifying prototypes that integrated a culturally responsive low-intensity CBT intervention as appealing [48]:

> *Sometimes we're using that [Stay Strong app] here in our program...You know the yellow colour (representing worries), the green colour (representing strengths) – when it comes, then they [clients] can feel it...if we can consistently use the same thing, they feel like, "Okay. This is the one. We know how to do that..."* [Service provider]

Security, Privacy, and Reliability

Most young people and service providers preferred security features to be optional:

> *...Do they really know which password they put on last month?...It's not necessary...[But] if they want to, then yeah [include the option].* [Service provider]

Evidence and Implementation

The potential to promote *early intervention* and *continuity of care* was also highlighted by service providers as important; however, it required consideration about the transfer of information among devices:

> *...We speak to those boys about...going to the school counsellor, and they just don't want to be there...this would be better for us to say, "Have a look at this app"...So, that preventative side of things would be really good.* [Service provider]
>
> *We are trying to connect these people back to the community. So, if...we can just give them [other service providers] – this is what this client was working on...[if] they can use the same password and they already went half-way and they can continue.* [Service provider]

Accessibility

All participants highlighted *free availability on a smartphone* to allow real-time anonymous use, with features that accommodate *data use concerns* and *platform compatibility* as preferred:

> *That's what I like...a lot of them are ashamed but if they're using this [app], it's discreet...no one else knows they are using it...They've all got their phones, it kind of gives them that power...a sense of independence...Is it for free?* [Service provider]

Service providers also noted that *availability on other devices* was necessary for some settings:

> *If we can get the app on an iPad?...no one can use phones at school.* [Service provider]

Narrative Literature Synthesis

Overview

Few papers specific to Indigenous youth populations and dMH apps were identified, suggesting that the field is still emerging. Publications relevant to this synthesis included certification guidelines [13]; evaluative tools [31,53,54]; and systematic reviews relating to Indigenous psychotherapy [55-57], dMH approaches for adolescents [58-62], digital health solutions for Indigenous people [11,63-66], and app features [20,44,64,67,68]. Information about Aboriginal and Torres Strait Islander people's technology use [12,69] was also included. A total of 3 mental health treatment apps designed to meet the needs of Aboriginal and Torres Strait Islander people were identified, with 1 (33%) app specifically designed for youth [10,15,16,70,71]. In the following section, we present recommendations from the literature across multiple population groups, including youth, adults, and Indigenous populations, recognizing the varied depth of understanding in each field. Multimedia Appendix 3 summarizes recommendations for the development of dMH tools for Indigenous young people [6,9,11-14,16,17,18,19,31,44,46-48,53-55,57-60,62,63,65,66,68,69,71-100].

Therapeutic Approach

Current guidelines recommend that mental health apps incorporate effective therapeutic approaches such as CBT [13,44]. Although there is limited empirical evidence for the effectiveness of psychotherapy for Indigenous people, culturally responsive approaches are preferred [9,57,72,73]. Strengths-based, culturally adapted, low-intensity CBT and acceptance commitment therapy have been shown to be effective with Aboriginal and Torres Strait Islander populations through randomized controlled trials [16,47].

Activity Types

Current dMH tool recommendations suggest that the inclusion of mental health information can prompt early identification and help-seeking [101]. Young people often prefer videos with limited text [62]. Behavioral activation techniques, such as goal-setting, planning, mindfulness, and encouragement of real-world activities, are frequently used strategies in digital health design [44]. Self-monitoring of thoughts, feelings, or behaviors has also demonstrated some evidence of favorable acceptability and effectiveness with adults and young people [44,62,74,102] despite other reports that it is *boring* and *feels like punishment* [103].

Engagement

Engagement strategies, such as notifications, social sharing, gamification, and personalization, can improve adherence and intrinsic motivation to dMH apps, if they are delivered and timed appropriately [44,48,62,68,75]. The ability to share information or progress can be particularly salient for cultures with holistic health worldviews; however, it can demotivate some if not aligned with the personal aims of the user [75,76,103]. Notifications should remind users of the benefits and promote opportunities to engage (ie, when feeling stressed or low mood), which can assist in habit formation and motivation [19,35,44,74]. Storytelling through fictional characters aids immersion, models behavior, and allows users to learn skills in a low-risk environment, which can prompt help-seeking [31,77].

Interface and Graphic Design

Culturally relevant, intuitive, aesthetic, and minimalist designs increase the acceptability and usability of dMH tools [53,75,78], particularly for those whose first language is not English [79]. Furthermore, mental health apps designed for use by nonclinical populations allow greater accessibility, reduce stigma, and facilitate preventative use [44,61].

Language

Language, which is simple, concrete, confident, hopeful, nonclinical, nonbiased, and matched to the literacy level of intended end users, promotes access and adherence to dMH tools [44]. Mechanisms that support and promote literacy, such as audio options, visual prompts, and metaphors, can improve understanding [75,80]. In addition, integration of Indigenous languages can aid communication and engagement for Indigenous people [71,75].

Security, Privacy, and Reliability

Apps should include password protection, which can aid the protection of privacy when devices are shared [9,13,81]. Privacy policies should be available, understandable, and regularly updated, with clear reporting of privacy features embedded in the design [13]. Security should also be monitored regularly for breaches.

Safety

The Australian National Safety and Quality dMH standards highlight several critical inclusions and considerations for crisis support [13]. The dMH services should include mechanisms for recognizing and responding to acute deterioration in mental state. Links to crisis support should be presented in attractive, easy-to-access ways throughout the app [13].

Evidence and Implementation

Although some dMH programs have demonstrated effectiveness with culturally diverse [16,82] and other young people [58,83], overall, there is a paucity of evidence of the effectiveness of mental health apps for adolescents [59,84]. It is also unclear how dMH solutions perform outside trial conditions, with reach, uptake, and adherence remaining problematic, highlighting the need for continued focus on implementation and sustainability [59,85]. Culturally responsive early intervention mental health services are lacking in most regions of Australia, especially in regional and remote areas [73,104]. The dMH interventions can promote access to early intervention, complement existing services, and prompt help-seeking [62,105].

Accessibility

Smartphones are the most accessible digital tools for mental health treatment worldwide [44] and are increasingly available to Aboriginal and Torres Strait Islander young people [69,86,87]. However, digital tool design should consider data credit, ongoing data use, offline use, availability free of charge, and platform compatibility [9,13,48].

Integrating Participant Feedback and the Scientific Literature

Overview

Feedback on storyboards and prototypes was sought from a diverse group of participants and sources, generating a range of perspectives and recommendations. Most young people demonstrated unique and strong preferences for specific activity types, storytelling, strengths-based information, character attributes, and gamification features. Service providers showed similar preferences, also highlighting the importance of features to support the psychological approach, accessibility, implementation, and continuity of care. The literature strongly supports CBT approaches, safety and quality features, and the necessity for evidence, which was not disputed by the young people or service providers but was also not prioritized. We reflect on the processes undertaken to engage, integrate multiple views, and prioritize features for inclusion in the following sections.

Engagement

One of the key challenges throughout was maintaining young people's interest and engagement through culturally safe processes over a 2-year time frame. The ongoing engagement of all sites and several young people throughout multiple phases of this project and the rich in-depth data collected suggest that the strategies used were reasonably successful. Strategies to maintain engagement included working within supportive schools and services and employing youth and senior Indigenous researchers with established connections to the community. Continually reviewing and revising co-design activities to meet the needs of the young people and frequently recruiting throughout the data collection process further aided success.

Integrating Multiple Views

Another key challenge was the differing youth preferences highlighted among individuals across and within co-design workshop groups. For example, the inclusion of storytelling through characters was a recurring preference. However, differing preferences for the narrative, context, character identities, and depth of information presented were identified. These challenges were compounded by the staggered timing and nonstatic group of participants in co-design workshops. The IYRG engaged a group that was diverse in terms of geographical location, gender, and age. This group was provided the findings from the co-design workshops to ensure they were informed by previous work undertaken. This improved efficiency and led to improved decision-making capacity of young people throughout development.

Prioritization

Throughout this participatory design project, we identified app features preferred by participants and assessed their alignment with current recommendations and scientific literature. Although we included many preferred features to some degree, the *available budget and prioritization of best practice recommendations* influenced what was integrated into the prototype. A prioritization process was required, with some features earmarked for later inclusion (Table 3). Prioritization was based on the consideration of young people's preferences, recommendations from the literature, and study protocols developed for a feasibility study (commenced in August 2020). Selected features, such as notifications and prompted mood monitoring, were popular among young people and recommended in the literature. Given the resource limitations, these were deprioritized in initial development in preference for the diversity of content (ie, variety of strengths, challenges, strategies, and rewards), activities (ie, minigames, videos, and pictorial selection or upload), and appeal (ie, characters of differing age, gender, and geographical location).

Table 3. Aboriginal and Islander Mental Health Initiative for Youth (version 1.0) prioritization of app features for development.

App features/Included	Not included owing to resource limitations; considered for next version	Justification
Therapeutic approach		
Low-intensity culturally adapted CBT[a] intervention (evidence-based)	N/A[b]	G[c], L[d], Y[e], SP[f]
Stories and prompts to address anxiety and low mood	N/A	L, Y
Activity types		
Videos	N/A	L, Y
Mental health information with links to further information	N/A	G, L, Y, SP
Minigames to prompt relaxation and real-world activities	N/A	L, Y
Self-monitoring of personal goals	Self-monitoring of mood and behavior	L, Y, SP
Engagement		
Storytelling through characters	N/A	L, Y
Real-time engagement	N/A	L, Y, SP
Gamification (eg, levels and rewards)	N/A	L, Y
Progress and summary page	N/A	L, Y
Ability to complete a personalized quest—add names and photos throughout (customization)	Ability to customize characters and notifications and select voice or language settings	G, L, Y, SP
N/A	Notifications to remind users of program benefits and encourage use	L, Y
Interface and design		
Simple and intuitive interface	N/A	G, L, Y, SP
Smooth, easy to use, accurate, and logical flow	N/A	G, L, Y
Aesthetic and minimalist design	N/A	G, L, Y, SP
Culturally relevant graphics tailored to target group	N/A	G, L, Y, SP
Language		
Simple, concrete, confident, hopeful, nonclinical, nonbiased language, and matched to the literacy level of users	N/A	G, L, Y, SP
Mechanisms to support literacy (eg, visual prompts and cues, use of metaphors, and minimal written content)	Mechanisms to support literacy (eg, audio of all app content)	L, Y, SP
Indigenous language words used within English text in line with stories	Full integration of multiple Indigenous languages throughout	L, Y
Privacy		
Up-to-date, available, and understandable privacy policy	N/A	G, L
N/A	Regularly monitored security	G, L
N/A	Password protection—optional	G, L, Y, SP
Safety		
Easily accessible links to crisis support	N/A	G, L, Y, SP
Evidence and implementation		
Experimental trials to examine efficacy	N/A	G, L
Analytics (use data)	N/A	G, L
N/A	The ability for users to provide feedback	G, L
N/A	Assessments embedded with prompts at different time points	G, L
Accessibility		
Compatible with mobile phone—currently only Android	Compatible with all platforms	G, L, Y, SP

App features/Included	Not included owing to resource limitations; considered for next version	Justification
Suitable for supported or self-driven use	N/A	L, SP
Offline use available	N/A	L, Y
N/A	Users made aware of data use	G
N/A	Multi-user tablet version	SP

[a]CBT: cognitive behavioral therapy.

[b]N/A: not applicable.

[c]G: guidelines.

[d]L: recommended in the literature.

[e]Y: youth preference.

[f]SP: service provider preference.

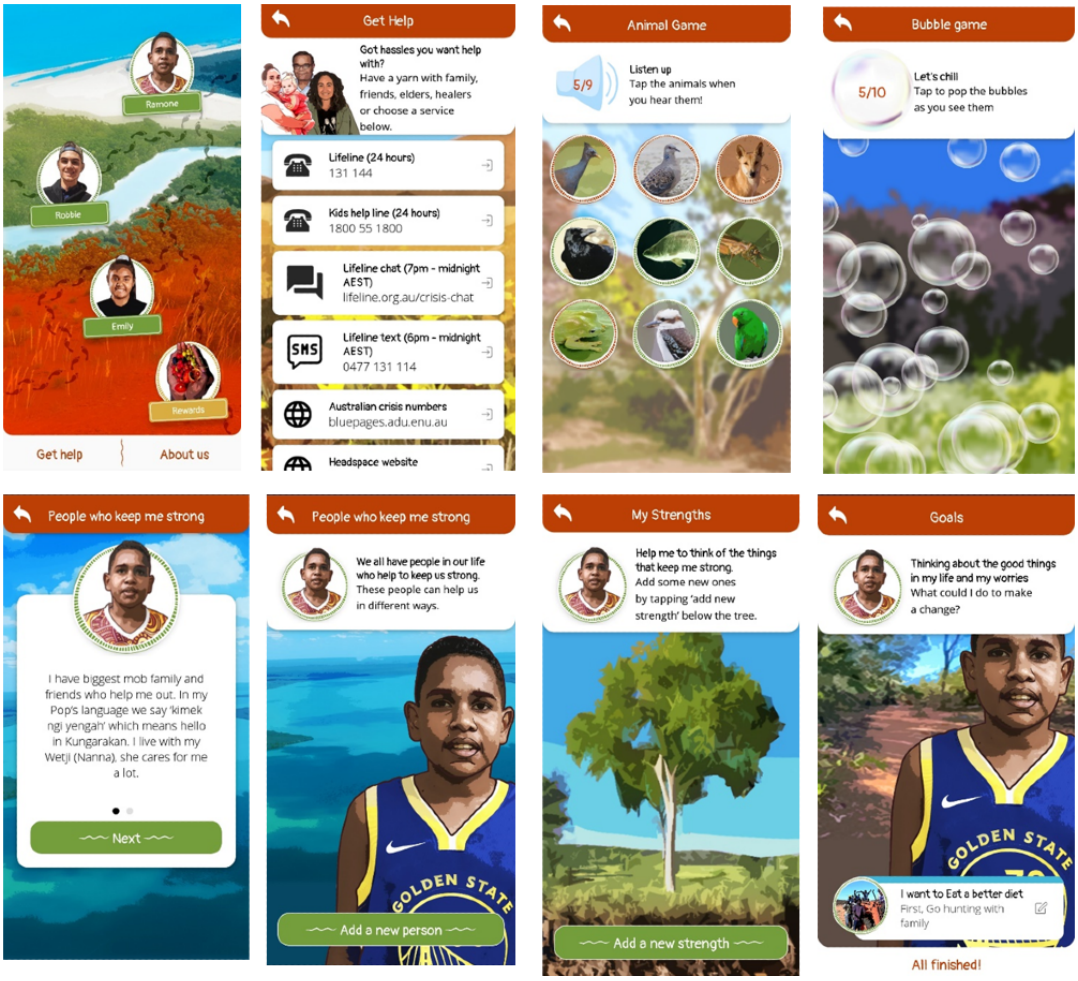
App Development

App development occurred through a series of research team meetings with the app developers, involving iterative user testing of developed prototypes. Input from the IYRG was sought by the research team throughout development for important design decisions.

Outcome: AIMhi-Y App (Version 1.0)

The product developed is a smartphone-based AIMhi-Y app (version 1.0) that integrates culturally adapted low-intensity CBT, psychoeducation, and mindfulness-based activities into a universal early intervention (Figure 5). The app aims to increase mental health literacy, self-management, and help-seeking for Aboriginal and Torres Strait Islander young people. Users assist fictional characters through a series of quest levels, aiming to become familiar with the content before beginning their own quest. The integrated AIMhi Stay Strong therapy follows a 4-step process: identifying people who keep them strong, strengths, worries, and supported goal-setting, which is interspersed with psychoeducational videos [63]. Activities and information target both anxiety and low mood. A summary page collates user or character information and presents their progress. Minigames promote relaxation; encourage real-world activities; and provide fun, engaging, and immersive sensory experiences. An example is a game in which users listen to an immersive soundtrack and are asked to identify familiar animal sounds by selecting related images. Storytelling aims to develop relationships and facilitate skill development in a safe, nonthreatening environment. The app aims to be easy to use with intuitive designs and nonclinical youth-friendly language. Aboriginal language words relevant to specific characters are integrated throughout. Vibrant colors and design elements reflect the natural landscapes of different Northern Territory (Australia) regions. Help contacts and links for further information are available throughout. A deidentified database captures app use and user interaction with app features.

Figure 5. Selected screenshots from Aboriginal and Islander Mental Health Initiative for Youth app (version 1.0).



Discussion

Principal Findings

This second phase participatory design project integrated Aboriginal and Torres Strait Islander young people's and service providers' preferences with recommendations from the scientific literature to develop the AIMhi-Y app (version 1.0). Young people expressed diverse and strong desires for app features, which were generally well aligned with findings in the scientific literature. Throughout this project, young people have been involved in an iterative design, development, and review process. Several challenges and key learnings have been noted. We reflect on how our processes align with the attributes of successful participatory design.

This project engaged a large, diverse group of young people with varied demographic, cultural, and linguistic backgrounds, many of whom may not usually be involved in co-design processes [38,106]. The planning, reflection, and refinement of culturally responsive co-design activities with the input of 4 Indigenous researchers led to increased engagement and acceptability of our processes. However, maintaining young people's participation for 2 years during adolescence, where social, emotional, and vocational needs are changing, was difficult: a challenge highlighted by others [106,107]. New participants were recruited throughout the process and they shared new ideas, suggestions, and opinions frequently. Our findings align with other studies that have identified differing preferences based on age, gender, and mental health status [88], suggesting that dMH tools need to be designed and tailored with specific target audiences in mind [33,48,68,108].

The staggered timing of diverse groups across multiple sites and the iterative nature of participatory design challenged our democratic processes: difficulties others have identified [20,36,38]. Each co-design workshop was iteratively modified in response to previous findings, limiting the use of voting or consensus methods [38,37]. Therefore, informed by previous findings, literature, and safety standards, researchers assumed a decision-making role, which shifted power differentials and impacted the extent of young people's influence in the co-design process. The sustained and regular inclusion of 2 youth and 2 senior Indigenous researchers and an experienced research team optimized youth, clinician, and Aboriginal and Torres Strait Islander influence in decision-making. The inclusion of the IYRG throughout development improved the decision-making capacity of young people in some domains. However, the information and decisions presented to the group were prioritized by the research team to keep within time frame constraints. Systematically taking findings and decisions from one group to the other and back again (eg, through decision logs) might have improved the transparency of our decision-making.

Furthermore, although young people provided many innovative and creative suggestions, budget and timeline restrictions limited what was incorporated into this first prototype app design. This is consistent with the challenges highlighted by others in facilitating a process allowing participants to be creative and innovative while providing rules and guidelines to keep work within the project scope and budget [109]. Finally, determining the degree to which upskilling of participants occurred in a culturally and linguistically diverse, iterative situation, where knowledge from all stakeholders remained both tacit and latent, presented a significant challenge [24]. In addition, the knowledge presented, the responses sought, and the interpretation of feedback by the research team were inevitably influenced by the researchers' experience, availability, paradigms, and worldviews [110].

Overall, the integration of findings into the app prototype proved complex, requiring consideration of young people's and service providers' preferences, recommendations from the literature, and budget. Future iterations will be informed by a feasibility study (commenced in August 2020) and might include additional activities; greater accessibility across platforms and devices; options for customization; self-monitoring; full integration of Aboriginal languages; and additional characters representing diverse ages, genders, language groups, and geographic areas facing differing age-appropriate strengths and challenges. Key learnings throughout this project include the importance of understanding and upskilling all those involved in the co-design project on previous work undertaken, the practical limitations, findings from previous research, and safety recommendations early in the co-design process. Although seeking diverse opinions undoubtedly strengthened our approach, a detailed plan of how to integrate the information across stakeholder groups would have aided our process. Identifying stakeholder roles, opportunities for input, responsibilities, influence, and decision-making strategies across app characteristic domains may have allowed more decision-making opportunities to be presented to young people.

Limitations

All of the small number of service providers interviewed had existing relationships and interest in the project, likely resulting in favorable comments. However, their involvement ensured the consideration of context and supported seeding ideas for further exploration and testing, thereby strengthening rather than detracting from the process. Furthermore, young people's engagement in this project was influenced by organizational, cultural, and linguistic contexts; group dynamics; facilitator experience; and availability. The difficulties in engaging an interpreter at the drug rehabilitation site undoubtedly impacted young people's engagement. The IYRG members were reimbursed for their participation, which may have influenced their comments. However, others have suggested that reimbursement might aid participation by promoting a sense of pride and responsibility [38]. Although sustained engagement facilitated trust and rapport over time, the transient nature of this population group proved challenging as new relationships needed to be established. Finally, the research team took a lead role in decision-making and prioritization of app features owing to resource limitations, impacting the extent of young people's influence in the co-design process.

Conclusions

By using a participatory design approach, we engaged a large and diverse group of young people to develop a new culturally informed early intervention mental health treatment app. We identified challenges in reaching consensus, upskilling participants, and allowing equal representation of views in a genuinely participatory process. By describing our processes, we aim to provide transparency in reporting of participatory design approaches. Although the process proved challenging, we successfully created a new app integrating youth preferences and best practice recommendations within a limited time frame and budget. A current feasibility study is underway to evaluate the usability, feasibility, and appropriateness of the AIMhi-Y app and its potential for improving the well-being of Aboriginal and Torres Strait Islander young people.

Acknowledgments

The authors would like to acknowledge all the young people, parents, teachers, support people, and members of our expert reference group who made this research possible. The authors would like to thank Senior Indigenous Researchers Cathy Pumuralimawu Stassi and Patj Patj Janama Robert Mills, for their assistance with interpreting, youth engagement, and refining app content. Tiwi Land Council provided approval for this project to occur. Funding was provided by Channel 7 Children's Research Foundation, South Australia (reference number 181700). The first author (JP) was supported by Australian Government Research Training Program Scholarship and an Australian Rotary Health, Ian Scott Scholarship.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Service provider interview guide.
[DOCX File , 13 KB - formative_v6i2e28342_app1.docx]

Multimedia Appendix 2
Search strategy.

[DOCX File , 15 KB - formative_v6i2e28342_app2.docx]

Multimedia Appendix 3
Recommended features for digital mental health tools for Indigenous young people.
[DOCX File , 27 KB - formative_v6i2e28342_app3.docx]

References

1. Salmon M, Doery K, Dance P, Chapman J, Gilbert R, Williams R. Defining the Indefinable: Descriptors of Aboriginal and Torres Strait Islander Peoples' Cultures and Their Links to Health and Wellbeing. Canberra: Aboriginal and Torres Strait Islander Health Team, Research School of Population Health, The Australian National University; 2018.
2. Causes of Death, Australia: statistics on the number of deaths, by sex, selected age groups, and cause of death classified to the International Classification of Diseases (ICD). Australian Bureau of Statistics. URL: https://www.abs.gov.au/statistics/health/causes-death/causes-death-australia/latest-release#intentional-self-harm-suicides-key-characteristics; [accessed 2022-01-19]
3. Aboriginal and Torres Strait Islander Adolescent and Youth Health and Wellbeing. Sydney, New South Wales: Australian Institute of Health and Welfare; 2018.
4. Price M, Dalgleish J. Help-seeking among indigenous Australian adolescents: exploring attitudes, behaviours and barriers. Youth Stud Australia 2013 Mar;32(1) [FREE Full text]
5. Walker BF, Stomski NJ, Price A, Jackson-Barrett E. Perspectives of Indigenous people in the Pilbara about the delivery of healthcare services. Aust Health Rev 2014;38(1):93. [doi: 10.1071/ah13074]
6. Lowell A, Maypilama E, Yikaniwuy S, Rrapa E, Williams R, Dunn S. "Hiding the story": indigenous consumer concerns about communication related to chronic disease in one remote region of Australia. Int J Speech Lang Pathol 2012 Jun;14(3):200-208. [doi: 10.3109/17549507.2012.663791] [Medline: 22443611]
7. Vicary D, Westerman T. That's just the way he is': some implications of Aboriginal mental health beliefs. Australian E J Advancement Mental Health 2014 Dec 17;3(3):103-112 [FREE Full text] [doi: 10.5172/jamh.3.3.103]
8. Dudgeon P, Bray A, D'costa B, Walker R. Decolonising psychology: validating social and emotional wellbeing. Aust Psychol 2020 Nov 12;52(4):316-325. [doi: 10.1111/ap.12294]
9. Leckning B, Li S, Cunningham T, Guthridge S, Robinson G, Nagel T, et al. Trends in hospital admissions involving suicidal behaviour in the Northern Territory, 2001-2013. Australas Psychiatry 2016 Jun;24(3):300-304. [doi: 10.1177/1039856216629838] [Medline: 26858241]
10. Povey J, Mills PP, Dingwall KM, Lowell A, Singer J, Rotumah D, et al. Acceptability of mental health apps for aboriginal and Torres Strait Islander Australians: a qualitative study. J Med Internet Res 2016 Mar 11;18(3):e65 [FREE Full text] [doi: 10.2196/jmir.5314] [Medline: 26969043]
11. Dingwall K, Puszka S, Sweet M, Nagel T. "like drawing into sand": acceptability, feasibility, and appropriateness of a new e‐mental health resource for service providers working with Aboriginal and Torres Strait Islander people. Australian Psychol 2020 Nov 12;50(1):60-69 [FREE Full text] [doi: 10.1111/ap.12100]
12. Hobson GR, Caffery LJ, Neuhaus M, Langbecker DH. Mobile health for first nations populations: systematic review. JMIR Mhealth Uhealth 2019 Oct 07;7(10):e14877 [FREE Full text] [doi: 10.2196/14877] [Medline: 31593537]
13. National safety and quality digital mental health standards. Australian Commission on Safety and Quality in Health Care. URL: https://www.safetyandquality.gov.au/standards/national-safety-and-quality-digital-mental-health-standards [accessed 2022-01-19]
14. Tighe J, Shand F, McKay K, Mcalister TJ, Mackinnon A, Christensen H. Usage and acceptability of the iBobbly app: pilot trial for suicide prevention in aboriginal and Torres Strait islander youth. JMIR Ment Health 2020 Dec 01;7(12):e14296 [FREE Full text] [doi: 10.2196/14296] [Medline: 33258782]
15. Shand F, Ridani R, Tighe J, Christensen H. The effectiveness of a suicide prevention app for indigenous Australian youth: study protocol for a randomized controlled trial. Trials 2013;14(1):396. [doi: 10.1186/1745-6215-14-396]
16. Tighe J, Shand F, Ridani R, Mackinnon A, De La Mata N, Christensen H. Ibobbly mobile health intervention for suicide prevention in Australian Indigenous youth: a pilot randomised controlled trial. BMJ Open 2017 Jan 27;7(1):e013518 [FREE Full text] [doi: 10.1136/bmjopen-2016-013518] [Medline: 28132007]
17. Dingwall K, Nagel T, Hughes JT, Kavanagh DJ, Cass A, Howard K, et al. Wellbeing intervention for chronic kidney disease (WICKD): a randomised controlled trial study protocol. BMC Psychol 2019 Jan 08;7(1):2 [FREE Full text] [doi: 10.1186/s40359-018-0264-x] [Medline: 30621791]
18. Orji R, Moffatt K. Persuasive technology for health and wellness: state-of-the-art and emerging trends. Health Informatics J 2018 Mar;24(1):66-91 [FREE Full text] [doi: 10.1177/1460458216650979] [Medline: 27245673]
19. Orji R, Mandryk R. Developing culturally relevant design guidelines for encouraging healthy eating behavior. Int J Human Comput Stud 2014 Feb;72(2):207-223 [FREE Full text] [doi: 10.1016/j.ijhcs.2013.08.012]

20. Orlowski SK, Lawn S, Venning A, Winsall M, Jones GM, Wyld K, et al. Participatory research as one piece of the puzzle: a systematic review of consumer involvement in design of technology-based youth mental health and well-being interventions. JMIR Hum Factors 2015 Jul 09;2(2):e12 [FREE Full text] [doi: 10.2196/humanfactors.4361] [Medline: 27025279]
21. Hagen P, Collin P, Metcalf A, Nicholas M, Rahilly K, Swainston N. Participatory Design of Evidence-based Online Youth Mental Health Promotion, Prevention, Early Intervention and Treatment. Melbourne: Young and Well Cooperative Research Centre; 2012.
22. Engaging Young People for Health and Sustainable Development: Strategic Opportunities for the World Health Organization and Partners. Geneva: World Health Organization; 2018.
23. Baum F, MacDougall C, Smith D. Participatory action research. J Epidemiol Community Health 2006 Oct;60(10):854-857 [FREE Full text] [doi: 10.1136/jech.2004.028662] [Medline: 16973531]
24. Vandekerckhove P, de Mul M, Bramer WM, de Bont AA. Generative participatory design methodology to develop electronic health interventions: systematic literature review. J Med Internet Res 2020 Apr 27;22(4):e13780 [FREE Full text] [doi: 10.2196/13780] [Medline: 32338617]
25. Allin S, Shepherd J, Tomasone J, Munce S, Linassi G, Hossain SN, et al. Participatory design of an online self-management tool for users with spinal cord injury: qualitative study. JMIR Rehabil Assist Technol 2018 Mar 21;5(1):e6 [FREE Full text] [doi: 10.2196/rehab.8158] [Medline: 29563075]
26. Moore G, Wilding H, Gray K, Castle D. Participatory methods to engage health service users in the development of electronic health resources: systematic review. J Particip Med 2019 Feb 22;11(1):e11474 [FREE Full text] [doi: 10.2196/11474] [Medline: 33055069]
27. Co-designing with young people: the fundamentals. Orygen. URL: https://tinyurl.com/5eczhpks [accessed 2022-02-01]
28. Maher LM, Hayward B, Hayward P, Walsh C. Increasing patient engagement in healthcare service design: a qualitative evaluation of a co-design programme in New Zealand. Patient Experience J 2017 Apr 24;4(1):23-32. [doi: 10.35680/2372-0247.1149]
29. Bowen S, McSeveny K, Lockley E, Wolstenholme D, Cobb M, Dearden A. How was it for you? Experiences of participatory design in the UK health service. CoDesign 2013 Dec;9(4):230-246 [FREE Full text] [doi: 10.1080/15710882.2013.846384]
30. Pallesen K, Rogers L, Anjara S, De Brún A, McAuliffe E. A qualitative evaluation of participants' experiences of using co-design to develop a collective leadership educational intervention for health-care teams. Health Expect 2020 Apr;23(2):358-367 [FREE Full text] [doi: 10.1111/hex.13002] [Medline: 31999883]
31. Cheek C, Fleming T, Lucassen MF, Bridgman H, Stasiak K, Shepherd M, et al. Integrating health behavior theory and design elements in serious games. JMIR Ment Health 2015;2(2):e11 [FREE Full text] [doi: 10.2196/mental.4133] [Medline: 26543916]
32. Sundram F, Hawken SJ, Stasiak K, Lucassen MF, Fleming T, Shepherd M, et al. Tips and traps: lessons from codesigning a clinician e-monitoring tool for computerized cognitive behavioral therapy. JMIR Ment Health 2017 Jan 11;4(1):e3 [FREE Full text] [doi: 10.2196/mental.5878] [Medline: 28077345]
33. Fleming T, Merry S, Stasiak K, Hopkins S, Patolo T, Ruru S, et al. The importance of user segmentation for designing digital therapy for adolescent mental health: findings from scoping processes. JMIR Ment Health 2019 May 08;6(5):e12656 [FREE Full text] [doi: 10.2196/12656] [Medline: 31066705]
34. Orji R, Vassileva J, Mandryk R. Modeling the efficacy of persuasive strategies for different gamer types in serious games for health. User Model User-Adap Inter 2014 Jul 14;24(5):453-498 [FREE Full text] [doi: 10.1007/s11257-014-9149-8]
35. Oyibo K, Orji R, Vassileva J. Developing culturally relevant design guidelines for encouraging physical activity: a social cognitive theory perspective. J Healthc Inform Res 2018 May 31;2(4):319-352 [FREE Full text] [doi: 10.1007/s41666-018-0026-9]
36. Frauenberger C, Good J, Fitzpatrick G, Iversen O. In pursuit of rigour and accountability in participatory design. Int J Hum Comput Stud 2015 Feb;74:93-106 [FREE Full text] [doi: 10.1016/j.ijhcs.2014.09.004] [Medline: 26109833]
37. Witteman HO, Presseau J, Nicholas Angl E, Jokhio I, Schwalm JD, Grimshaw JM, et al. Negotiating tensions between theory and design in the development of mailings for people recovering from acute coronary syndrome. JMIR Hum Factors 2017 Mar 01;4(1):e6 [FREE Full text] [doi: 10.2196/humanfactors.6502] [Medline: 28249831]
38. Thabrew H, Fleming T, Hetrick S, Merry S. Co-design of eHealth interventions with children and young people. Front Psychiatry 2018;9:481 [FREE Full text] [doi: 10.3389/fpsyt.2018.00481] [Medline: 30405450]
39. Verbiest ME, Corrigan C, Dalhousie S, Firestone R, Funaki T, Goodwin D, et al. Using codesign to develop a culturally tailored, behavior change mHealth intervention for indigenous and other priority communities: a case study in New Zealand. Transl Behav Med 2018 Nov 02;9(4):720-736. [doi: 10.1093/tbm/iby093] [Medline: 30388262]
40. Liebenberg L, Sylliboy A, Davis-Ward D, Vincent A. Meaningful engagement of indigenous youth in PAR. Int J Qual Methods 2017 Jul 24;16(1):160940691770409 [FREE Full text] [doi: 10.1177/1609406917704095]
41. Wiewiora A, Keast R, Brown K. Opportunities and challenges in engaging citizens in the co-production of infrastructure-based public services in Australia. Public Manag Rev 2015 Jan 14;18(4):483-507 [FREE Full text] [doi: 10.1080/14719037.2014.999820]
42. Cockbill S, May A, Mitchell V. The assessment of meaningful outcomes from co-design: a case study from the energy sector. She Ji J Design Econ Innov 2019;5(3):188-208 [FREE Full text] [doi: 10.1016/j.sheji.2019.07.004]

43. DeSmet A, Thompson D, Baranowski T, Palmeira A, Verloigne M, De BI. Is participatory design associated with the effectiveness of serious digital games for healthy lifestyle promotion? A meta-analysis. J Med Internet Res 2016 Apr 29;18(4):e94 [FREE Full text] [doi: 10.2196/jmir.4444] [Medline: 27129447]
44. Bakker D, Kazantzis N, Rickwood D, Rickard N. Mental health smartphone apps: review and evidence-based recommendations for future developments. JMIR Ment Health 2016 Mar 01;3(1):e7 [FREE Full text] [doi: 10.2196/mental.4984] [Medline: 26932350]
45. Nagel T. Relapse prevention in remote indigenous mental health. Charles Darwin University. URL: https://researchers.cdu.edu.au/en/studentTheses/relapse-prevention-in-remote-indigenous-mental-health [accessed 2022-02-04]
46. Nagel T, Thompson C. AIMHI NT 'Mental Health Story Teller Mob': developing stories in mental health. Aust E J Advance Mental Health 2014 Dec 17;6(2):119-124. [doi: 10.5172/jamh.6.2.119]
47. Nagel T, Robinson G, Condon J, Trauer T. Approach to treatment of mental illness and substance dependence in remote Indigenous communities: results of a mixed methods study. Aust J Rural Health 2009 Aug;17(4):174-182. [doi: 10.1111/j.1440-1584.2009.01060.x] [Medline: 19664081]
48. Povey J, Sweet M, Nagel T, Mills P, Stassi C, Puruntatameri A, et al. Drafting the Aboriginal and Islander Mental Health Initiative for Youth (AIMhi-Y) app: results of a formative mixed methods study. Internet Interv 2020 Sep;21:100318 [FREE Full text] [doi: 10.1016/j.invent.2020.100318] [Medline: 32477884]
49. Kelly J, Saggers S, Taylor K, Pearce G, Massey P, Bull J, et al. "Makes you proud to be black eh?": reflections on meaningful indigenous research participation. Int J Equity Health 2012 Aug 08;11:40 [FREE Full text] [doi: 10.1186/1475-9276-11-40] [Medline: 22873769]
50. Engaging with Aboriginal children and young people toolkit. Commissioner for Children and Young People. URL: https://tinyurl.com/mvmz9ut7 [accessed 2022-02-01]
51. Thomas D. A general inductive approach for analyzing qualitative evaluation data. Am J Eval 2016 Jun 30;27(2):237-246 [FREE Full text] [doi: 10.1177/1098214005283748]
52. Thorne S. Data analysis in qualitative research. Evidence Based Nurs 2000;3(3):68-70 [FREE Full text] [doi: 10.1136/ebn.3.3.68]
53. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
54. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the User Version of the Mobile Application Rating Scale (uMARS). JMIR Mhealth Uhealth 2016;4(2):e72 [FREE Full text] [doi: 10.2196/mhealth.5849] [Medline: 27287964]
55. Antonio M, Chung-Do J. Systematic review of interventions focusing on Indigenous adolescent mental health and substance use. Am Indian Alsk Native Ment Health Res 2015;22(3):36-56. [doi: 10.5820/aian.2203.2015.36] [Medline: 26425865]
56. Leske S, Harris MG, Charlson FJ, Ferrari AJ, Baxter AJ, Logan JM, et al. Systematic review of interventions for Indigenous adults with mental and substance use disorders in Australia, Canada, New Zealand and the United States. Aust N Z J Psychiatry 2016 Nov;50(11):1040-1054. [doi: 10.1177/0004867416662150] [Medline: 27514405]
57. Pomerville A, Burrage RL, Gone JP. Empirical findings from psychotherapy research with indigenous populations: a systematic review. J Consult Clin Psychol 2016 Dec;84(12):1023-1038. [doi: 10.1037/ccp0000150] [Medline: 27736112]
58. Richardson T, Stallard P, Velleman S. Computerised cognitive behavioural therapy for the prevention and treatment of depression and anxiety in children and adolescents: a systematic review. Clin Child Fam Psychol Rev 2010 Sep;13(3):275-290. [doi: 10.1007/s10567-010-0069-9] [Medline: 20532980]
59. 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]
60. Clarke A, Kuosmanen T, Barry M. A systematic review of online youth mental health promotion and prevention interventions. J Youth Adolesc 2015 Jan;44(1):90-113. [doi: 10.1007/s10964-014-0165-0] [Medline: 25115460]
61. Bergin A, Vallejos E, Davies E, Daley D, Ford T, Harold G, et al. Preventive digital mental health interventions for children and young people: a review of the design and reporting of research. NPJ Digit Med 2020;3:133 [FREE Full text] [doi: 10.1038/s41746-020-00339-7] [Medline: 33083568]
62. Liverpool S, Mota CP, Sales CM, Čuš A, Carletto S, Hancheva C, et al. Engaging children and young people in digital mental health interventions: systematic review of modes of delivery, facilitators, and barriers. J Med Internet Res 2020 Jun 23;22(6):e16317 [FREE Full text] [doi: 10.2196/16317] [Medline: 32442160]
63. Hensel J, Ellard K, Koltek M, Wilson G, Sareen J. Digital health solutions for indigenous mental well-being. Curr Psychiatry Rep 2019 Jul 01;21(8):68 [FREE Full text] [doi: 10.1007/s11920-019-1056-6] [Medline: 31263971]
64. Antezana G, Bidargaddi N, Blake V, Schrader G, Kaambwa B, Quinn S, et al. Development of an online well-being intervention for young people: an evaluation protocol. JMIR Res Protoc 2015 Apr 30;4(2):e48 [FREE Full text] [doi: 10.2196/resprot.4098] [Medline: 25929201]
65. Brusse C, Gardner K, McAullay D, Dowden M. Social media and mobile apps for health promotion in Australian Indigenous populations: scoping review. J Med Internet Res 2014 Dec 10;16(12):e280 [FREE Full text] [doi: 10.2196/jmir.3614] [Medline: 25498835]

66. Jones L, Jacklin K, O'Connell ME. Development and use of health-related technologies in indigenous communities: critical review. J Med Internet Res 2017 Jul 20;19(7):e256 [FREE Full text] [doi: 10.2196/jmir.7520] [Medline: 28729237]
67. Cowpertwait L, Clarke D. Effectiveness of web-based psychological interventions for depression: a meta-analysis. Int J Ment Health Addiction 2013 Jan 18;11(2):247-268 [FREE Full text] [doi: 10.1007/s11469-012-9416-z]
68. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
69. Indigenous media and communications. McNair Innovations. URL: https://mcnair.com.au/wp-content/uploads/Indigenous-Media-Infographic.pdf [accessed 2022-02-04]
70. Kurdiji App: using technology to save young Indigneous lives. Kurdiji. URL: https://kurdijiapp.wordpress.com/ [accessed 2022-02-04]
71. Nagel T, Sweet M, Dingwall K, Puszka S, Hughes J, Kavanagh D, et al. Adapting wellbeing research tools for Aboriginal and Torres Strait Islander people with chronic kidney disease. BMC Nephrol 2020 Apr 15;21(1):130 [FREE Full text] [doi: 10.1186/s12882-020-01776-y] [Medline: 32293331]
72. MacLean S, Ritte R, Thorpe A, Ewen S, Arabena K. Health and wellbeing outcomes of programs for Indigenous Australians that include strategies to enable the expression of cultural identities: a systematic review. Aust J Prim Health 2017 Sep;23(4):309-318. [doi: 10.1071/PY16061] [Medline: 28619126]
73. Westerman T. Engaging Australian Aboriginal youth in mental health services. Aust Psychol 2010 Sep 02;45(3):212-222 [FREE Full text] [doi: 10.1080/00050060903451790]
74. Huguet A, Rao S, McGrath P, Wozney L, Wheaton M, Conrod J, et al. A systematic review of cognitive behavioral therapy and behavioral activation apps for depression. PLoS One 2016;11(5):e0154248 [FREE Full text] [doi: 10.1371/journal.pone.0154248] [Medline: 27135410]
75. Shepherd M, Merry S, Lambie I, Thompson A. Indigenous adolescents' perception of an eMental health program (SPARX): exploratory qualitative assessment. JMIR Serious Games 2018 Jul 05;6(3):e13 [FREE Full text] [doi: 10.2196/games.8752] [Medline: 29980495]
76. Kappen D, Orji R. Gamified and persuasive systems as behavior change agents for health and wellness. XRDS 2017 Sep 14;24(1):52-55 [FREE Full text] [doi: 10.1145/3123750]
77. Heilemann MV, Martinez A, Soderlund PD. A mental health storytelling intervention using transmedia to engage Latinas: grounded theory analysis of participants' perceptions of the story's main character. J Med Internet Res 2018 May 02;20(5):e10028 [FREE Full text] [doi: 10.2196/10028] [Medline: 29720357]
78. Reilly R, Stephens J, Micklem J, Tufanaru C, Harfield S, Fisher I, et al. Use and uptake of web-based therapeutic interventions amongst Indigenous populations in Australia, New Zealand, the United States of America and Canada: a scoping review. Syst Rev 2020 May 31;9(1):123 [FREE Full text] [doi: 10.1186/s13643-020-01374-x] [Medline: 32475342]
79. Kral IB. Plugged in: remote Australian indigenous youth and digital culture. SSRN J 2010 [FREE Full text] [doi: 10.2139/ssrn.2244615]
80. Toombs E, Kowatch KR, Dalicandro L, McConkey S, Hopkins C, Mushquash CJ. A systematic review of electronic mental health interventions for Indigenous youth: results and recommendations. J Telemed Telecare 2021 Oct;27(9):539-552. [doi: 10.1177/1357633X19899231] [Medline: 31937199]
81. Larsen M, Nicholas J, Christensen H. A systematic assessment of smartphone tools for suicide prevention. PLoS One 2016;11(4):e0152285 [FREE Full text] [doi: 10.1371/journal.pone.0152285] [Medline: 27073900]
82. Fleming T, Dixon R, Frampton C, Merry S. A pragmatic randomized controlled trial of computerized CBT (SPARX) for symptoms of depression among adolescents excluded from mainstream education. Behav Cogn Psychother 2011 Dec 05;40(5):529-541 [FREE Full text] [doi: 10.1017/s1352465811000695]
83. Merry S, Stasiak K, Shepherd M, Frampton C, Fleming T, Lucassen M. 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]
84. March S, Donovan C, Baldwin S, Ford M, Spence S. Using stepped-care approaches within internet-based interventions for youth anxiety: three case studies. Internet Interv 2019 Dec;18:100281 [FREE Full text] [doi: 10.1016/j.invent.2019.100281] [Medline: 31890629]
85. 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]
86. Taylor A. Information communication technologies and new Indigenous mobilities? Insights from remote Northern Territory Communities. J Rural Community Develop 2012;7(1):59-73.
87. Media usage amongst Aboriginal and Torres Strait people. McNair Innovations. URL: https://mcnair.com.au/media-release-media-usage-amongst-aboriginal-and-torres-strait-islander-people-infographic/ [accessed 2022-02-01]
88. Fortes Tondello G, Valtchanov D, Reetz A, Wehbe R, Orji R, Nacke L. Towards a trait model of video game preferences. Int J Human Comput Interact 2018 Apr 19;34(8):732-748 [FREE Full text] [doi: 10.1080/10447318.2018.1461765]

89. Johnson D, Deterding S, Kuhn K, Staneva A, Stoyanov S, Hides L. Gamification for health and wellbeing: a systematic review of the literature. Internet Interv 2016 Nov;6:89-106 [FREE Full text] [doi: 10.1016/j.invent.2016.10.002] [Medline: 30135818]
90. Kelders SM, Sommers-Spijkerman M, Goldberg J. Investigating the direct impact of a gamified versus nongamified well-being intervention: an exploratory experiment. J Med Internet Res 2018 Jul 26;20(7):e247 [FREE Full text] [doi: 10.2196/jmir.9923] [Medline: 30049669]
91. Christie G, Shepherd M, Merry S, Hopkins S, Knightly S, Stasiak K. Gamifying CBT to deliver emotional health treatment to young people on smartphones. Internet Interv 2019 Dec;18:100286 [FREE Full text] [doi: 10.1016/j.invent.2019.100286] [Medline: 31890633]
92. Titov N, Dear BF, Johnston L, McEvoy PM, Wootton B, Terides MD, et al. Improving adherence and clinical outcomes in self-guided internet treatment for anxiety and depression: a 12-month follow-up of a randomised controlled trial. PLoS One 2014;9(2):e89591 [FREE Full text] [doi: 10.1371/journal.pone.0089591] [Medline: 24586897]
93. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile App Rating Scale: A New Tool for Assessing the Quality of Health Mobile Apps. JMIR mHealth uHealth 2015 Mar 11;3(1):e27. [doi: 10.2196/mhealth.3422]
94. Orji R, Reilly D, Oyibo K, Orji F. Deconstructing persuasiveness of strategies in behaviour change systems using the ARCS model of motivation. Behav Inf Technol 2018 Sep 17;38(4):319-335 [FREE Full text] [doi: 10.1080/0144929x.2018.1520302]
95. Shand F, Mackinnon A, O'Moore K, Ridani R, Reda B, Hoy M, et al. The iBobbly Aboriginal and Torres Strait Islander app project: study protocol for a randomised controlled trial. Trials 2019 Apr 05;20(1):198 [FREE Full text] [doi: 10.1186/s13063-019-3262-2] [Medline: 30953556]
96. Cass A, Lowell A, Christie M, Snelling PL, Flack M, Marrnganyin B, et al. Sharing the true stories: improving communication between Aboriginal patients and healthcare workers. Medical Journal of Australia 2002 May 20;176(10):466-470. [doi: 10.5694/j.1326-5377.2002.tb04517.x]
97. Davies J, Bukulatjpi S, Sharma S, Caldwell L, Johnston V, Davis JS. Development of a culturally appropriate bilingual electronic app about hepatitis b for Indigenous Australians: towards shared understandings. JMIR Res Protoc 2015 Jun 10;4(2):e70 [FREE Full text] [doi: 10.2196/resprot.4216] [Medline: 26063474]
98. Kral I. Youth media as cultural practice: remote indigenous youth speaking out loud. Aust Aborig Stud 2011:1 [FREE Full text]
99. Kral I. Shifting perceptions, shifting identities: communication technologies and the altered social, cultural and linguistic ecology in a remote indigenous context. Aust J Anthropol 2014 Jun 01;25(2):171-189 [FREE Full text] [doi: 10.1111/taja.12087]
100. van Ballegooijen W, Cuijpers P, van Straten A, Karyotaki E, Andersson G, Smit JH, et al. Adherence to internet-based and face-to-face cognitive behavioural therapy for depression: a meta-analysis. PLoS One 2014;9(7):e100674 [FREE Full text] [doi: 10.1371/journal.pone.0100674] [Medline: 25029507]
101. Titov N, Dear B, Nielssen O, Staples L, Hadjistavropoulos H, Nugent M, et al. ICBT in routine care: a descriptive analysis of successful clinics in five countries. Internet Interv 2018 Sep;13:108-115 [FREE Full text] [doi: 10.1016/j.invent.2018.07.006] [Medline: 30206525]
102. Hetrick SE, Robinson J, Burge E, Blandon R, Mobilio B, Rice SM, et al. Youth codesign of a mobile phone app to facilitate self-monitoring and management of mood symptoms in young people with major depression, suicidal ideation, and self-harm. JMIR Ment Health 2018 Jan 23;5(1):e9 [FREE Full text] [doi: 10.2196/mental.9041] [Medline: 29362208]
103. Orji R, Lomotey R, Oyibo K, Orji F, Blustein J, Shahid S. Tracking feels oppressive and 'punishy': exploring the costs and benefits of self-monitoring for health and wellness. Digit Health 2018;4:2055207618797554 [FREE Full text] [doi: 10.1177/2055207618797554] [Medline: 30202544]
104. Gulliver A, Griffiths KM, Christensen H. Perceived barriers and facilitators to mental health help-seeking in young people: a systematic review. BMC Psychiatry 2010 Dec 30;10(1):e10. [doi: 10.1186/1471-244x-10-113]
105. Youth mental health report: youth survey 2012-16. Mission Australia, Black Dog Institute. URL: https://apo.org.au/node/75809 [accessed 2022-02-01]
106. Powers JL, Tiffany JS. Engaging youth in participatory research and evaluation. J Public Health Manag Pract 2006 Nov;Suppl:S79-S87. [doi: 10.1097/00124784-200611001-00015] [Medline: 17035908]
107. Jacquez F, Vaughn L, Wagner E. Youth as partners, participants or passive recipients: a review of children and adolescents in community-based participatory research (CBPR). Am J Community Psychol 2013 Mar;51(1-2):176-189. [doi: 10.1007/s10464-012-9533-7] [Medline: 22718087]
108. Jessen S, Mirkovic J, Ruland CM. Creating gameful design in mHealth: a participatory co-design approach. JMIR Mhealth Uhealth 2018 Dec 14;6(12):e11579. [doi: 10.2196/11579] [Medline: 30552080]
109. Fossey E, Harvey C, McDermott F, Davidson L. Understanding and evaluating qualitative research. Aust N Z J Psychiatry 2002 Dec;36(6):717-732. [doi: 10.1046/j.1440-1614.2002.01100.x] [Medline: 12406114]
110. Titov N, Schofield C, Staples L, Dear BF, Nielssen O. A comparison of Indigenous and non-Indigenous users of MindSpot: an Australian digital mental health service. Australas Psychiatry 2019 Aug;27(4):352-357. [doi: 10.1177/1039856218789784] [Medline: 30058351]

Abbreviations

AIMhi-Y: Aboriginal and Islander Mental Health Initiative for Youth
CBT: cognitive behavioral therapy
dMH: digital mental health
IYRG: Indigenous youth reference group

Edited by A Mavragani; submitted 16.03.21; peer-reviewed by S Darling, J Tighe, M Webb, M Lucassen, M O'Connell; comments to author 27.05.21; revised version received 14.10.21; accepted 15.12.21; published 18.02.22.

Please cite as:
Povey J, Sweet M, Nagel T, Lowell A, Shand F, Vigona J, Dingwall KM
Determining Priorities in the Aboriginal and Islander Mental Health Initiative for Youth App Second Phase Participatory Design Project: Qualitative Study and Narrative Literature Review
JMIR Form Res 2022;6(2):e28342
URL: https://formative.jmir.org/2022/2/e28342
doi:10.2196/28342
PMID:35179498

Original Paper

Use of Social Networks in the Context of the Dietitian's Practice in Brazil and Changes During the COVID-19 Pandemic: Exploratory Study

Jackson Sbardelotto[1*]; Bárbara Birck Martins[1*]; Caroline Buss[2*], PhD, RD, MSc

[1]Nutrition Course, Federal University of Health Sciences of Porto Alegre, Porto Alegre, Brazil

[2]Department of Nutrition, Graduate Study Program in Health Sciences, Federal University of Health Sciences of Porto Alegre, Porto Alegre, Brazil

[*]all authors contributed equally

Corresponding Author:
Caroline Buss, PhD, RD, MSc
Department of Nutrition
Graduate Study Program in Health Sciences
Federal University of Health Sciences of Porto Alegre
R. Sarmento Leite, 245
Porto Alegre, 90050-170
Brazil
Phone: 55 513303 8830
Email: carolinebuss@ufcspa.edu.br

Abstract

Background: Social networks have been pointed out as 1 of the greatest means of spreading information. A large part of the population is already present on these platforms, looking up subjects such as health, nutrition, and food. To reach this audience, it may be important for dietitians to explore social networks. However, there is a gap in scientific studies on exploring the ways in which these platforms are used by dietitians in Brazil, and the roles they play in the profession have not been well defined.

Objective: This study aims to describe the roles that social networks play in dietitians' practice in Brazil and their mode of use of social networks. This study also aims to identify professionals' perceptions and opinions regarding the use of these tools, as well as changes in behavior on social network usage caused by the COVID-19 pandemic.

Methods: We carried out a quantitative cross-sectional study, collecting data through an online questionnaire, submitted between October 2020 and January 2021 to dietitians registered on the Federal Council of Dietitians. All participants included in the study answered questions about the use of social networks in their professional context.

Results: In total, 264 (91.7%) of the 288 participants reported using social networks for professional practice. Instagram was the social network most often used by professionals (224/264, 84.8%). Dietitians (N=288) related to the use of social networks (always to almost always) for sharing information about their services (n=114-72 [39.6%-25%], respectively), following the work of other dietitians (n=172-64 [59.7%-22.2%], respectively), and writing about topics related to food and nutrition (n=166-53 [57.6%-18.4%], respectively). The roles played by social networks in the professional context of dietitians were attracting more clients (210/289, 72.7%) and keeping in touch with them (195/289, 67.5%). Furthermore, 227 (78.5%) of the 289 dietitians strongly agreed that social networks are good tools to promote their services. During the COVID-19 pandemic, 216 (74.7%) of the 289 participants noticed changes in their behavior, feelings, or beliefs on the use of social networks related to professional practice, and 149 (51.6%) have increased the frequency of sharing information about nutrition and health in general on social networks.

Conclusions: The main roles of social networks in the professional context of dietitians are to attract clients and to facilitate the contact between professional and client. The modes of use reported by the professionals included sharing information about their services, following the work of professional colleagues, and writing about topics related to nutrition. Most of them reported believing that social networks are an effective way to disseminate their services. Moreover, most professionals claimed to have noticed changes in their behaviors or beliefs on social media during the COVID-19 pandemic.

(JMIR Form Res 2022;6(2):e31533) doi:10.2196/31533

KEYWORDS

dietitian; social networks; nutrition; health communication; COVID-19; social media; Brazil; perception; health information; usage; behavior

Introduction

Social networks have been pointed out as 1 of the greatest means of spreading information, with 4.2 billion active users in the world, which corresponds to 53.6% of the world population [1]. In 2019, in a prepandemic scenario, the percentage of active users around the world was 42%, representing about 3 billion users [2]. The year 2020 marked the biggest growth in the number of users in the past 3 years. The internet has become 1 of the main sources for seeking information about health-related matters, especially with the restriction of face-to-face contact due to the COVID-19 pandemic [3].

About 72% of adult internet users search for information about diagnoses, treatments, and reports of people in the same health condition [4]. Social networks can have many benefits for health professionals' practice, by providing greater possibility to interact, share experiences, and communicate in real time with other health professionals as well as their clients/patients [5]. One health-related topic with high interest on social networks is regarding nutrition and healthy eating. Evidence shows that possible impacts of social networks can be related to facilitating eating behavior changes, as they allow users to acquire some knowledge and understanding about healthy eating [6].

Social networks have reach to a large audience and are low-cost tools. Therefore, social networks are considered tools with potential to offer great benefits to dietitians. They can assist professionals in their online nutritional interventions, disseminating evidence-based information, and can work as an extension of their service [7,8]. A large part of the population is already present on these platforms, looking up subjects such as health, nutrition, and food [9]. Many people started to do so when there was a restriction of face-to-face consultations during the period of social isolation due to the COVID-19 pandemic. At the same time, researchers are using social media to share recommendations and decisions being made in times of the COVID-19 pandemic. In addition, dietitians have had to shift from in-person client interactions to telenutrition consultations [3,10].

The estimated average time spent on social media per user around the world is currently 2 hours and 25 minutes a day [1]. Therefore, to reach this audience, it may be important for professionals to explore social networks to provide adequate information about nutrition and food, to implement the practice of educational actions, and to offer nutrition services. In Brazil, the average time spent per day by users of social networks is 3 hours and 42 minutes. Brazil is in third place in this global ranking [1]. Thus, conducting studies on this subject in the country is valid and necessary to make possible a better understanding of the role of social networks in the profession and, later, to find opportunities of its use.

In the light of the increasing use of social media around the world among dietitians and the general population, Dumas et al [11] designed a review to map existing evidence on the effects of social media on nutritional practice. The authors found 4 interventional studies with dietitians and social media users. From the users' perspective, online interaction with professionals can be favorable for changing behavior. However, further studies isolating variables will be necessary to measure the effectiveness of networks in this process. We also identified studies with expert opinions about the behavior of professionals on social networks, dealing with advantages and disadvantages of the use of networks in the profession and guides on ethical conduct on digital platforms [6,7,12]. Few of these have been studied so far (ie, discussion forums, blogs, and Facebook).

Saboia et al [13] observed the influences on the eating behavior of the followers of dietitians and nongraduates on Instagram in Brazil and Portugal. They were able to describe some usage characteristics of the professionals and how they interact with the public. However, it is necessary to further explore the use of networks in the profession to identify their modes of use.

Another review conducted by Saboia et al [14] also evidenced the lack of studies that can describe how dietitians are using social networks. In a 2020 publication [15], the same authors designed a questionnaire to understand the dynamics of social media use by dietitians for future application with dietitians in Portugal.

Given the low number of studies that have actually explored the ways that dietitians use social media, and how the COVID-19 pandemic may have altered such use, we developed this study. The aim of this study was to identify how dietitians in Brazil use social networks and to describe the professionals' perceptions and opinions on the use of these tools, as well as changes in behavior on social network usage due to the COVID-19 pandemic.

Methods

Study Design

We conducted a quantitative cross-sectional study, collecting data through an online questionnaire, submitted between October 2020 and January 2021. People participated in the research only after reading and being aware of the free and informed consent terms, which provided information about the nature of the study.

Ethics Approval

The research ethics committee of the Federal University of Health Sciences of Porto Alegre approved the study (#3981023).

Sample and Participants

According to data from the Federal Council of Dietitians (FCD), there were 150,892 active dietitians in Brazil in 2019 [16]. Adopting a conservative approach to estimating the prevalence of the use of social media by dietitians, a ratio of 50% was used to maximize the sample size. A sampling error of 6% and a 95% CI were adopted. The minimum number of responses for the sample to be representative of this population was calculated as 267 participants.

The criterion for being included as participants was to be registered with the Regional Council of Dietitians (RCD). We performed recruitment in 3 stages: (1) We sent emails to a contact list of registered dietitians available on the RCD website, inviting them to take part in the research and answer the questionnaire; (2) in the same way, we sent invitations by direct messages to the professionals' Instagram accounts; and (3) professionals with more than 10,000 followers released the questionnaire link, inviting other dietitians to respond through the stories tool on Instagram.

Data Collection

All participants in the study answered an online questionnaire about how they use social networks in their professions. The study data were collected and organized using REDCap electronic data capture tools hosted by the Federal University of Health Sciences of Porto Alegre [17,18].

To describe the use of social media, we based and adapted the questions included in the questionnaire through the Global Digital Report of 2021 [1]. We included additional questions to identify professionals' social network usage modes based on the study by Dunne et al [19]. We also created more questions to explore professionals' usage modes, opinions, and perceptions. We divided the questions into 6 categories: (1) sociodemographic information, (2) characterization of the sample, (3) negative and positive points perceived using social networks, (4) ways of using social networks, (5) professionals' perceptions on the use of social media, and (6) changes in perceptions and behavior on these platforms due to the COVID-19 pandemic.

Data Analysis

Categorical results were presented by frequency and percentage. Symmetric quantitative results were presented by average and SD, or the median and the 25th percentile (P25) and the 75th percentile (P75) when asymmetric. Associations were analyzed using the chi-square test with the aid of adjusted standardized residues. The software used for the analyzes was SPSS Statistics v25 (IBM), and the significance level was .05.

Results

Participant Characteristics

In total, 334 dietitians were included in the study. Of the 334 dietitians who started, 289 (86.5%) participants reached the end of the questionnaire. Of these, 283 (97.9%) dietitians were female, and the median age was 29 (25-34) years. Information about the professional practice of the participants is shown in Table 1.

Table 1. Demographic and professional practice of dietitians who responded about their use of social networks (N=289), Brazil, 2021.

Characteristics	Value
Demographics	
Gender (female), n (%)	283 (97.9)
Age (years), median (P25-P75)	29 (25-34)
Region, n (%)	
North	7 (2.4)
Northeast	53 (18.3)
Midwest	16 (5.5)
Southeast	111 (38.4)
South	102 (35.3)
Professional practice time (years), median (P25[a]-P75[b])	5 (2-10)
Professional field, n (%)	
Collective feeding nutrition	15 (5.2)
Clinical nutrition	208 (72)
Sports and exercise nutrition	34 (11.8)
Public health nutrition	13 (4.5)
Production chain: industry and food trade nutrition	4 (1.4)
Teaching: research and extension nutrition	15 (5.2)

[a]P25: 25th percentile.

[b]P75: 75th percentile.

The use of social networks in the daily lives of dietitians was identified in 288 (99.7%) of the 289 participants, and of these, 264 (91.7%) claimed to use these platforms for their professional practice. Of the 24 (8.3%) professionals who stated that they did not use social networks linked to professional practice, 8 (33%) worked in the clinical nutrition field, 6 (25%) in public health nutrition, 5 (21%) in nutrition in collective feeding, and 5 (21%) in superior nutrition education and research. In a comparison of the number of dietitians who answered no to this question with the total number of professionals in the fields of

professional activity, there were 8 (3.8%) of 208 in the clinical nutrition field, 6 (39%) of 13 in public health nutrition, 5 (33%) of 15 in nutrition in collective feeding, and 5 (33%) of 15 in superior nutrition education and research.

When evaluating modes of use of social networks by dietitians according to their working field, 133 (63.9%) of the 208 participants in the clinical nutrition field reported the habit of writing about topics such as food, feeding, and related ones, and 67 (32.2%) of these professionals reported the practice of recording videos on these subjects and sharing the videos on their social platforms. In contrast, 3 (20%) of the 15 professionals in the collective feeding field and 4 (21.1%) of the 19 professionals in the areas of superior nutrition education and research (n=15, 78.9%) and in the food industry (n=4, 21.1%) stated that they never wrote about these topics in their social networks. Regarding the videos related to these themes, 8 (53%) of the 15 nutrition professionals in the collective feeding field said they never recorded one, and 9 (47.4%) of the 19 nutrition professionals in the superior nutrition education and research field and in the food industry stated the same. The professionals in the clinical nutrition field stated the habit of following a schedule for organizing the content to be published on social networks. We detailed the use of these platforms in their professions as well as the modes of use in Tables 2 and 3, respectively.

Table 2. Characterization of the use of social networks by dietitians who responded about their use of social networks, Brazil, 2021.

Use of social networks	n (%)
Use of social networks in their daily lives (N=289)	288 (99.7)
Use of social networks related to professional practice (N=288)	264 (91.7)
Most frequently used social network (N=264)	
Instagram	224 (84.8)
WhatsApp	29 (11)
Facebook	4 (1.5)
LinkedIn	2 (0.8)
Others	5 (1.9)
Snapchat/Twitter	0
Frequency of use of social networks (N=288)	
≥5 times a day	197 (68.4%)
2-4 times a day	71 (24.7%)
Once a day	10 (3.5%)
3-6 times a week	9 (3.1%)
<3 times a week	1 (0.3%)
Time spent in organizing, creating, and posting materials (N=275)	
>6 hours per week	33 (12%)
5-6 hours per week	38 (13.8%)
3-4 hours per week	98 (35.6%)
1-2 hours per week	59 (21.5%)
<1 hour per week	23 (8.4%)
I pay someone to do these tasks.	3 (1.1%)
I do not do it.	21 (7.6%)

Table 3. Modes of use of social networks in professional practice by dietitians who responded about their use of social networks (N=288), Brazil, 2021.

Modes of use	Never, n (%)	Almost never, n (%)	Sometimes, n (%)	Almost always, n (%)	Always, n (%)
I share information about my services on social networks (eg, localization, specializations, and previous work).	17 (5.9)	24 (8.3)	61 (21.2)	72 (25)	114 (39.6)
I share third-party content related to health and food on my social media profiles.	31 (10.8)	81 (28.1)	123 (42.7)	33 (11.5)	20 (6.9)
I follow the work of other nutrition professionals through social networks.	2 (0.7)	4 (1.4)	46 (15.9)	64 (22.2)	172 (59.7)
I write about topics related to food/nutrition and share the content on my social networks.	13 (4.5)	18 (6.3)	38 (13.2)	53 (18.4)	166 (57.6)
I record food/nutrition videos and share them on my social networks.	66 (22.9)	38 (13.2)	60 (20.8)	45 (15.6)	79 (27.5)
Before sharing content on nutrition and health in general, I make sure that what I am posting is based on evidence and comes from a safe source.	2 (0.7)	3 (1.0)	9 (3.1)	19 (6.6)	255 (88.6)
Whenever I share any material, I make sure to credit the source.	15 (5.2)	28 (9.7)	58 (20.1)	52 (18.1)	135 (46.9)
I have an organized schedule for my social networks' posts and content.	75 (26.0)	47 (16.3)	65 (22.6)	42 (14.6)	59 (20.5)

Dietitians identified the positive roles of social network usage in professional practice in the questionnaire. "I attract more clients/patients through social networks" was the most cited by 210 (72.7%) of the 289 participants, followed by "I keep in touch more easily with my clients" cited by 195 (67.5%) of the participants. Negative roles of social networks were also indicated, such as excessive time browsing social media, mentioned by 140 (48.4%) of the 289 participants. The second negative point, mentioned by 139 (48.1%) of the 289 professionals, was that the comparison with other professionals' accounts triggers a feeling of inferiority. Displayed in Multimedia Appendix 1 are all the data about the roles social networks play in the professional context of dietitians. Table 4 describes the professionals' general opinions and perceptions on the use of social media.

Regarding the changes in the use of social networks during the COVID-19 pandemic by dietitians, 217 (75.1%) of the 289 participants reported that they were performing online consultations according to the authorization given by the FCD (Brazil 2020). Of these 217, 179 (82.5%) professionals declared to be satisfied or very satisfied with this type of service. The changes related to feelings, beliefs, and behavior on the social networks of professionals during the COVID-19 pandemic are described in Table 5.

Table 4. Opinions and perceptions on the use of social networks by dietitians who responded about their use of social networks (N=289), Brazil, 2021.

Opinions and perceptions	Strongly disagree, n (%)	Somewhat disagree, n (%)	Neither agree nor disagree, n (%)	Somewhat agree, n (%)	Strongly agree, n (%)
Social networks are good tools to promote your services.	0	4 (1.4)	8 (2.8)	50 (17.3)	227 (78.5)
Licensed professionals are the ones who write most health-related content on social media.	80 (27.7)	108 (37.4)	33 (11.4)	60 (20.7)	8 (2.8)
Social networks play an important role regarding health promotion.	0	17 (5.9)	32 (11)	199 (41.2)	121 (41.9)
Being active on social media might be dangerous for professionals.	99 (34.2)	78 (20.1)	67 (23.2)	54 (18.7)	11 (3.8)
In social networks, most professionals behave according to the practical aspects established by the RCD[a].	57 (19.7)	120 (41.5)	58 (20.1)	49 (17)	5 (1.7)
Social networks are unnecessary for dietitians.	210 (72.7)	31 (10.7)	19 (6.6)	15 (5.2)	14 (4.8)
Dietitians who post frequently on social media are at a commercial advantage when compared to those who do not post frequently or who do not have social networks accounts at all.	6 (2.1)	24 (8.3)	47 (16.2)	82 (28.4)	130 (45)

[a]RCD: Regional Council of Dietitians.

Table 5. Prevalence of behavior changes during the COVID-19 pandemic on social network usage related to the professional practice of dietitians who responded about their use of social networks (N=289), Brazil, 2021.

Behavior changes	n (%)
I noticed changes in my behavior, feelings, or beliefs regarding the use of social networks related to professional practice.	216 (74.7)
I started spending more time checking out the work of other nutrition professionals on social networks.	96 (33.2)
During the pandemic, I started to share or increased the frequency of sharing information about nutrition and health in general on social networks.	149 (51.6)
I felt more inspired to present information about my services on social networks.	127 (43.9)
I started using new tools (such as live video streaming, interactive question stickers, and polls, among others) on social networks.	125 (43.3)
During this period, I changed my belief regarding the relevance of using social networks for the professional practice of dietitians.	75 (26.0)
I felt some kind of concern for not feeling qualified enough to talk about subjects related to nutrition on social networks.	46 (15.9)

Discussion

Principal Findings

This study identified the roles that social networks play in the professional context of dietitians in Brazil and their modes of use, as well as the professionals' perceptions and opinions on the use of these tools and the changes in behavior on social network usage due to the COVID-19 pandemic. During this period, many professionals increased the frequency of their use of social media and changed their opinions on how relevant these tools are in their profession. The results showed that almost all the participants used social media for professional practice, which we expected since a large part of the population is already present on these platforms and uses them for professional reasons [1].

We identified that the use of social networks is more prevalent in clinical nutrition than in other fields. This result is relevant, as it shows there is a specific group of professionals who use these social platforms more frequently. It is possible to make a direct association between social networks and professional practice since internet users are making their occupations and activities public, sharing photographic records of their job duties, and searching for profiles that could be interested in their work. Therefore, we expected the association of social media with professional practice. Dietitians use these platforms to attract clients/patients, facilitate contact, and make it more frequent.

The large percentage of statements about using social networks to attract clients/patients is related to the fact that professionals post information about their services on these platforms, making it possible for the content to be visible to many internet users. As a positive point, professionals can quantify this number through metrics from the social networks themselves, allowing them to identify the numbers of views and interactions of posts. These metrics are not viable for offline media, which can be an advantage of social media compared to traditional media. Another almost exclusive feature of social networking, considered a positive point, is that internet users can contact professionals through direct and instant messaging.

Social networks also allow the professional to take an active role in the process of exchanging messages, identifying the profiles of potential clients/patients, and getting an opportunity to contact them. These actions are not viable in traditional media, such as banners, magazines, newspapers, or radio and television. Professionals can then use social media to write about topics they have mastered and spread that content to other users. When compiling information about a particular subject in their profile, other users start to perceive that professional as someone who shows authority on the subject, which brings credibility, thus facilitating more people from social networks to become interested in their services.

The participants in this study also pointed out the role of social networks as a means of facilitating contact with patients and making these contacts more frequent, which is relevant since professionals can closely monitor their patients/clients. Thus, the professional can send information at a faster speed and can use the tools to instruct their patients and answer questions. One of the advantages of social networks in this sense is that they have a wealth of visual tools that can facilitate the educational process. Sometimes, the professional can also opt for a more informal approach, which can influence the relationship with the patient. Still, social media has other types of tools that can make interactions between users more attractive and more frequent. The interactions provided by these platforms can also strengthen the bond between professional and patient. During the COVID-19 pandemic, a considerable part of the studied professionals reported starting using tools such as live video streams, polls, and question stickers on social networks, which may be related to the search for better interaction with the public, which was modified by the social distancing during the pandemic.

We expected the results about the indicated modes of use (sharing information about their services, checking the work of colleagues, and writing about topics such as food) since these behaviors are in line with the role social networks play in attracting clients/patients. Checking on colleagues' work was also expected since it is common for internet users to follow pages that are attractive to them, such as people who share common interests or colleagues with whom they previously had personal contact.

Still on the modes of use, the publication on topics related to food can bring credibility, as it provides “samples” of the professionals' knowledge. At the same time, attractive layouts

or more casual content can also draw the attention of the social media public. We identified that publishing texts and videos on topics related to food is more frequent among professionals in the clinical nutrition field. In addition, the organization of content to be published using a schedule is an almost exclusive behavior of these professionals. In contrast, writing and recording videos on these topics is an uncommon behavior for professionals working in the areas such as collective deeding, superior nutrition education and research, and the food industry and commerce. These data suggest more frequent behaviors in a specific group of professionals within social networks, being less frequent in other areas. In this niche of work, there is a greater possibility of attracting clients, so these professionals dedicate more time to using social networks and have a variety of behaviors on these platforms.

Regarding the COVID-19 pandemic, some professionals reported adopting new behaviors on social networks or that certain existing behaviors have become more frequent. As an example, more than half of the participants stated that they started to publish or increased the frequency of publishing information about nutrition and health in general on their social networks. This was not surprising, since the social isolation measures adopted by the government in the country of study required more time indoors. When in isolation, it is expected that individuals will spend more hours of their day on the internet, including the use of social networks, since many activities started to occur virtually during the pandemic. It was expected that some professionals would start to explore these platforms as a possible way of extending their service, when prevented from conducting consultations and attracting clients in person. In addition, 2020, the year of isolation, was the year of the greatest growth for users of social networks in the past 3 years, which may have influenced the results mentioned earlier.

Most respondents cited the action of following the work of other professionals on social networks, which serves for dietitians to have references from other colleagues. Thus, it is possible to address topics the other professional discusses, becoming a source of ideas for their publications. Additionally, it can be useful to gather more knowledge about some subject that that colleague has mastered, so the interaction between professionals can be beneficial in these cases.

However, the frequent visit to other professionals' accounts can cause a feeling of inferiority, as reported as 1 of the negative points in using social networks by part of the dietitians participating in the study. Possibly, some facts caused this feeling, such as some professionals having more followers, appearing to be more successful than others, dominating specific subjects, or having a higher interaction rate in their accounts. Still, a third of the participants reported that they spent more time checking on the work of other professionals during the COVID-19 pandemic. This result could also be expected, considering the increase in the number of users and the time spent on social networks. Thus, it is reasonable that all reported behaviors grow in frequency, rather than just some of them exclusively.

Instagram was reported as the most used social network when linked to professional practice, which may be related to the platform's popularity in the country of the study. In Kemp [20], Instagram was placed in the 11th position in the ranking of the most accessed websites and in 3rd place as an application in the category of social networks in the country, with an average time of use per user of 14 hours monthly in the mobile version.

In a similar study conducted in 2019 with sports dietitians in the United Kingdom [19], different results were found regarding the most used social networks in the professional practice of dietitians. Twitter was the most popular social network among participants, followed by Facebook and WhatsApp. In the year of the study (2019), Twitter had a higher number of users (46% of the UK population) compared to the current year (44.3%). Meanwhile, Instagram was 1 of the fastest-growing social networks in these years (from 47% of the population in 2019 to 52.5% in 2021) in the region [2,20]. The popularity of Twitter in the year of the study may be related to this result. Furthermore, 100% of the study participants had a Twitter account, while only 68% were present on Instagram. The percentage of professionals who used social platforms was high but lower than the results found in our study. This may be linked to the difference in percentages of the number of users and time spent on social networks in each region. In Brazil, it is estimated that 70.3% of the population are users of social media, with an average daily use of 3 hours and 42 minutes [20]; in the United Kingdom, in 2019, that number was around 67% of the population, and the average daily usage estimated was 1 hour and 50 minutes.

Most of the UK study participants pointed out that using social media is beneficial in their practice, which also occurred in our study. In the study mentioned, the ways of using these platforms were as follows: to update on scientific research, to search for recipes for meals, and to give information to their clients [19]. The results of our study showed that for the most part, professionals use social networks to share information about their services, to check on the work of professional colleagues, and to write about topics related to nutrition. The difference in the results of each study on the modes of use is notable, as it shows that social media can have different usage for professionals in the same field of activity. Thus, its use may vary according to the country and the social network used in each location.

An interesting result observed in this study was that the participants disagreed that accredited professionals are responsible for most of the content published on social networks, which corroborates with a systematic review that identified that issues of quality and reliability are 1 of the main limitations of the use of social media for communication in health [21]. Respondents also cited to believe that most professionals do not act according to the ethical aspects established by the RCD. As examples of violated behaviors in the Code of Ethics and Conduct of Dietitians, we can mention promotions or drawings of lots of procedures carried out by dietitians, dissemination of body images of themselves or third parties (even with authorization), and preference or promotion of products or company brands related to food and nutrition activities [22].

It is important to note that during the COVID-19 pandemic, some respondents reported they changed their views regarding the relevance of using social networks for their work practices. Still, a considerable part of the professionals started conducting online consultations with the permission of the FCD. The committee exceptionally suspended Article 36 of CFN Resolution No. 599, of February 25, 2018, which established that nutritional assessment and diagnosis should be carried out in person and that only nutritional guidance and monitoring could be conducted remotely. Thus, dietitians can carry out online consultations until the end of the COVID-19 pandemic [23,24], which shows the importance of the professionals' ability to adapt to the context in which they are inserted. Social networks became relevant for professionals when their reality changed, making these platforms a means to get closer to clients/patients.

A large part of the professionals participating in the study reported that they felt more motivated to present the services by digital means and that they changed their opinions and perceptions on social networks during the pandemic. The need to continue exercising their professional activities competitively and the inability to carry out face-to-face consultations may have contributed to a change in the way professionals perceive social networks and influenced their motivations to present their services on these platforms.

Limitations

This study had limitations. Many of the responses to the questionnaire originated from invitations made through the Instagram platform itself, which reflected in the data of a greater number of users of this social network in our study. However, according to statistics on the use of social networks in Brazil, Instagram is 1 of the most popular ones [20], which also provides greater scope for researching this topic.

Conclusion

In conclusion, this study explored the use of social networks by dietitians in Brazil and the changes in their use during the COVID-19 pandemic, identifying a wide adoption of these platforms for professional usage. Attracting clients and facilitating contact between professional and client were the main reported roles of social networks for dietitians. The modes of use reported by the professionals included sharing information about their services, following the work of professional colleagues, and writing about topics related to nutrition. Regarding the professionals' perceptions, most of them reported believing that social networks are an effective way to disseminate their services. Moreover, most professionals claimed to have noticed changes in their behaviors or beliefs on social media during the COVID-19 pandemic. Future research in this area could work with open questions to explore more in-depth different ways of using social media and what other benefits dietitians and other health professionals can obtain through them.

Acknowledgments

All the authors of this paper have directly participated in the planning, execution, or analysis of this study, having read and approved the final version submitted. The authors also have agreed to authorship in the indicated order.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The roles social networks play in the professional context of dietitians who responded about their use of social networks (N=289), Brazil, 2021.
[DOCX File , 14 KB - formative_v6i2e31533_app1.docx]

References

1. Kemp S. Digital 2021: The Latest Insights into the 'State of Digital'. We are social. URL: https://wearesocial.com/uk/blog/2021/01/digital-2021-the-latest-insights-into-the-state-of-digital [accessed 2022-02-10]
2. Kemp S. Digital, Social Media, Mobile et E-Commerce en 2019. We are social. URL: https://wearesocial.com/global-digital-report-2019 [accessed 2022-02-10]
3. Farid D. COVID-19 and telenutrition: remote consultation in clinical nutrition practice. Curr Dev Nutr 2020 Dec;4(12):nzaa124 [FREE Full text] [doi: 10.1093/cdn/nzaa124] [Medline: 33409442]
4. Barthel M, Shearer E, Gottfried J, Mitchell A. The Evolving Role of News on Twitter and Facebook. Pew Research Center. 2015. URL: https://www.journalism.org/2015/07/14/the-evolving-role-of-news-on-twitter-and-facebook/ [accessed 2022-02-10]
5. Chan W, Leung A. Use of social network sites for communication among health professionals: systematic review. J Med Internet Res 2018 Mar 28;20(3):e117 [FREE Full text] [doi: 10.2196/jmir.8382] [Medline: 29592845]
6. McGloin AF, Eslami S. Digital and social media opportunities for dietary behaviour change. Proc Nutr Soc 2015 May;74(2):139-148. [doi: 10.1017/S0029665114001505] [Medline: 25319345]

7. Ahmed OH, Weiler R, Schneiders AG, McCrory P, Sullivan SJ. Top tips for social media use in sports and exercise medicine: doing the right thing in the digital age. Br J Sports Med 2015 Jul;49(14):909-910 [FREE Full text] [doi: 10.1136/bjsports-2014-094395] [Medline: 25940634]
8. Pizzuti AG, Patel KH, McCreary EK, Heil E, Bland CM, Chinaeke E, et al. Healthcare practitioners' views of social media as an educational resource. PLoS One 2020 Feb 6;15(2):e0228372 [FREE Full text] [doi: 10.1371/journal.pone.0228372] [Medline: 32027686]
9. Pollard CM, Pulker CE, Meng X, Kerr DA, Scott JA. Who uses the internet as a source of nutrition and dietary information? An Australian population perspective. J Med Internet Res 2015 Aug 26;17(8):e209 [FREE Full text] [doi: 10.2196/jmir.4548] [Medline: 26310192]
10. Cuello-Garcia C, Pérez-Gaxiola G, van Amelsvoort L. Social media can have an impact on how we manage and investigate the COVID-19 pandemic. J Clin Epidemiol 2020 Nov;127:198-201 [FREE Full text] [doi: 10.1016/j.jclinepi.2020.06.028] [Medline: 32603686]
11. Dumas A, Lapointe A, Desroches S. Users, uses, and effects of social media in dietetic practice: scoping review of the quantitative and qualitative evidence. J Med Internet Res 2018 Feb 20;20(2):e55 [FREE Full text] [doi: 10.2196/jmir.9230] [Medline: 29463487]
12. Helm J, Jones RM. Practice Paper of the Academy of Nutrition and Dietetics: social media and the dietetics practitioner: opportunities, challenges, and best practices. J Acad Nutr Diet 2016 Nov;116(11):1825-1835. [doi: 10.1016/j.jand.2016.09.003] [Medline: 27788767]
13. Saboia I, Pisco Almeida AM, Sousa P, Pernencar C. I am with you: a netnographic analysis of the Instagram opinion leaders on eating behavior change. Procedia Comput Sci 2018;138:97-104. [doi: 10.1016/j.procs.2018.10.014]
14. Saboia I, Almeida A, Sousa P, Pernencar C. Dietitians and nutritionists behaviour on social media: a scoping literature review. In: Proceedings of the 13th International Joint Conference on Biomedical Engineering Systems and Technologies. Setúbal: SciTePress - Science and Technology Publications; 2020:530-538.
15. Saboia I, Pisco Almeida AM, Sousa P, Pernencar C, Silva R. What are dietitians and nutritionists doing on social media? A proposal of an online survey. Procedia Comput Sci 2021;181(19):793-802. [doi: 10.1016/J.PROCS.2021.01.232] [Medline: 22489612]
16. Estatística. CFN CF de N. 2019. URL: https://www.cfn.org.br/index.php/estatistica/ [accessed 2022-02-10]
17. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap Consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
18. 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]
19. Dunne DM, Lefevre C, Cunniffe B, Tod D, Close GL, Morton JP, et al. Performance nutrition in the digital era: an exploratory study into the use of social media by sports nutritionists. J Sports Sci 2019 Nov 26;37(21):2467-2474. [doi: 10.1080/02640414.2019.1642052] [Medline: 31345110]
20. Kemp S. Digital 2021: Brazil. URL: https://datareportal.com/reports/digital-2021-brazil [accessed 2022-02-10]
21. Moorhead S, Hazlett D, Harrison L, Carroll J, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. J Med Internet Res 2013 Apr 23;15(4):e85 [FREE Full text] [doi: 10.2196/jmir.1933] [Medline: 23615206]
22. Código de Ética e de Conduta do Nutricionista. CFN CF de N. 2018. URL: https://www.cfn.org.br/wp-content/uploads/2018/04/codigo-de-etica.pdf [accessed 2022-02-10]
23. Resolução CFN no 646, de 18 de março de 2020. CFN CF de N. URL: https://www.cfn.org.br/wp-content/uploads/2020/03/Resol-CFN-646-codigo-etica.pdf [accessed 2022-02-10]
24. Resolução CFN no 600, de 25 de Fevereiro De 2018 Internet. CFN CF de N. 2018. URL: http://www.cfn.org.br/wp-content/uploads/resolucoes/Res_600_2018.htm [accessed 2022-02-10]

Abbreviations

FCD: Federal Council of Dieticians
P25: 25th percentile
P75: 75th percentile
RCD: Regional Council of Dieticians

Edited by A Mavragani; submitted 06.07.21; peer-reviewed by C Hao, L Chirchir; comments to author 23.10.21; revised version received 18.11.21; accepted 12.01.22; published 25.02.22.

Please cite as:
Sbardelotto J, Martins BB, Buss C
Use of Social Networks in the Context of the Dietitian's Practice in Brazil and Changes During the COVID-19 Pandemic: Exploratory Study
JMIR Form Res 2022;6(2):e31533
URL: https://formative.jmir.org/2022/2/e31533
doi:10.2196/31533
PMID:35023837

Viewpoint

Psychiatry on Twitter: Content Analysis of the Use of Psychiatric Terms in French

Sarah Delanys[1,2], MD; Farah Benamara[3], PhD; Véronique Moriceau[3], PhD; François Olivier[1,2], MD; Josiane Mothe[3], PhD

[1]Fédération Régionale de Recherche en Psychiatrie et santé mentale d'Occitanie, Toulouse, France

[2]Centre Hospitalier de Montauban, Montauban, France

[3]Institut de Recherche en Informatique de Toulouse, Université de Toulouse, Toulouse, France

Corresponding Author:
Sarah Delanys, MD
Fédération Régionale de Recherche en Psychiatrie et santé mentale d'Occitanie
134 route d'Espagne
Toulouse, 31000
France
Phone: 33 5 61 43 78 52
Email: s.delanys@ch-montauban.fr

Abstract

Background: With the advent of digital technology and specifically user-generated contents in social media, new ways emerged for studying possible stigma of people in relation with mental health. Several pieces of work studied the discourse conveyed about psychiatric pathologies on Twitter considering mostly tweets in English and a limited number of psychiatric disorders terms. This paper proposes the first study to analyze the use of a wide range of psychiatric terms in tweets in French.

Objective: Our aim is to study how generic, nosographic, and therapeutic psychiatric terms are used on Twitter in French. More specifically, our study has 3 complementary goals: (1) to analyze the types of psychiatric word use (medical, misuse, or irrelevant), (2) to analyze the polarity conveyed in the tweets that use these terms (positive, negative, or neural), and (3) to compare the frequency of these terms to those observed in related work (mainly in English).

Methods: Our study was conducted on a corpus of tweets in French posted from January 1, 2016, to December 31, 2018, and collected using dedicated keywords. The corpus was manually annotated by clinical psychiatrists following a multilayer annotation scheme that includes the type of word use and the opinion orientation of the tweet. A qualitative analysis was performed to measure the reliability of the produced manual annotation, and then a quantitative analysis was performed considering mainly term frequency in each layer and exploring the interactions between them.

Results: One of the first results is a resource as an annotated dataset. The initial dataset is composed of 22,579 tweets in French containing at least one of the selected psychiatric terms. From this set, experts in psychiatry randomly annotated 3040 tweets that corresponded to the resource resulting from our work. The second result is the analysis of the annotations showing that terms are misused in 45.33% (1378/3040) of the tweets and that their associated polarity is negative in 86.21% (1188/1378) of the cases. When considering the 3 types of term use, 52.14% (1585/3040) of the tweets are associated with a negative polarity. Misused terms related to psychotic disorders (721/1300, 55.46%) were more frequent to those related to depression (15/280, 5.4%).

Conclusions: Some psychiatric terms are misused in the corpora we studied, which is consistent with the results reported in related work in other languages. Thanks to the great diversity of studied terms, this work highlighted a disparity in the representations and ways of using psychiatric terms. Moreover, our study is important to help psychiatrists to be aware of the term use in new communication media such as social networks that are widely used. This study has the huge advantage to be reproducible thanks to the framework and guidelines we produced so that the study could be renewed in order to analyze the evolution of term usage. While the newly build dataset is a valuable resource for other analytical studies, it could also serve to train machine learning algorithms to automatically identify stigma in social media.

(JMIR Form Res 2022;6(2):e18539) doi:10.2196/18539

KEYWORDS

social media analysis; psychiatric term use; social stigma; Twitter; social media; mental health

Introduction

Stigma of Mental Disorders

Mental health stigma finds its roots in the history of psychiatry, in its connection to madness representations. Throughout history, the mentally ill patient has been given a pejorative label that induces social rejection. The term "stigma" comes from the ancient Greek "stitzein," which means "to tattoo or mark with a red iron." Jean-Yves Giordana [1], psychiatrist at the Nice hospital, rightly defines stigmatization as "a general attitude, a prejudicial one induced by low knowledge or ignorance of a situation or a state."

Stigma and discriminatory behaviors have multiple negative impacts. Stigma in mental health leads the individual away from society, which often causes social isolation. Indeed, stigmatized people confront difficulties in their daily life such as integration into the professional world [1,2], access to housing [1], and interpersonal relationships [3]. Difficulties also concern the treatment itself, including delay in initial medical consultation, difficulty in accepting the illness, and tenuous therapeutic alliance, etc [1].

Many studies analyzing newspaper articles point out a major diversion in the use of psychiatric terms [4,5]. A French survey conducted by the L'Observatoire Société et Consommation [6] found that the French terms "schizophrène" (schizophrenic) and "schizophrénie" (schizophrenia) are particularly used in the context of violent news items and are often used metaphorically, which may lead to dangerousness, contradictory behavior, or negative connotation. Pignon et al [7], on the other hand, focused on the impact of the change of the terminology of bipolar disorder. The authors observed that substituting the term "manic-depressive psychosis" for the term "bipolar disorder" reduces stigma by disassociating this disorder from the representation of madness and dangerousness leading to the social exclusion classically associated with psychotic disorders.

With the rise of the internet and social media, it has become important to analyze how psychiatric terms are used by people in general to act effectively against stigmatization. Indeed, from the internet users' point of view, Berry et al [8] showed that tweeting about mental health helps reduce isolation, fight stigmatization, and raise awareness of mental health by improving knowledge, promoting free expression, and strengthening coping and empowerment strategies.

In this paper, we focus on tweets in French as Twitter is one of the most used social media platforms in France [9], and the tweets are publicly available with some conditions.

Twitter as a Resource to Analyze the Usage of Psychiatric Terms

More than 500 million tweets are posted daily in more than 40 languages [10]. In March 2019, Twitter had 321 million active users worldwide (at least one use per month) among which 10.3 million were in France [11], making Twitter third in popularity behind Facebook and YouTube, the other two most popular social networks, with 35 million and 19 million active users, respectively. The sociodemographic profile of Twitter users in France is more male, younger, and more educated than the general population. They are mainly students, with some managers and intellectual professions [12-14].

Twitter offers its users the opportunity to post short messages named "tweets" (140 characters maximum in our study although since we collected the data, the maximum length has doubled), making possible analyses of a large number of tweets in a short time. In addition, Twitter provides 1% of tweets posted each day worldwide, allowing free access to a large database accessible for various purposes including research.

Since 2014, many studies have addressed discourse content about psychiatry on Twitter, suggesting that social networks convey stigmatizing representations of mental health and people with mental health conditions. To our knowledge, existing studies deal only with English and Greek languages. Moreover, they focus on a limited number of psychiatric disorder terms such as depression, schizophrenia, and autism. Lachmar et al [15] created the hashtag #MDLL (#mydepressionlookslike) and analyzed 3225 tweets highlighting 7 topics when Twitter users talk about depression: dysfunctional thoughts, impact on daily life, social difficulties, hiding behind a mask, sadness and apathy, suicidal behaviors/ideas, and seeking support/help. Reavley et al [16] analyzed a corpus of tweets about schizophrenia and depression in English. This corpus was collected from the 1% database of Twitter using two keywords: #schizophrenia and #depression. They found that 5% of tweets related to schizophrenia convey stigmatizing remarks while less than 1% are related to depression. In addition, in their dataset, they found the polarity is mostly positive (65% of the tweets analyzed) when writing about depression while it is rather neutral (43%) for schizophrenia. Joseph et al [17] found that tweets containing the hashtag #schizophrenia convey a negative sentiment more frequently than tweets containing #diabetes (21% vs 12.6%, respectively). Similarly, Athanasopoulou et al [18] showed that tweets about schizophrenia in Greek tend to be more negative, medically inappropriate, sarcastic, and used in a nonmedical way than tweets about diabetes. Robinson et al [19] analyzed and compared messages about 5 psychiatric disorders (autism, depression, eating disorders, obsessive-compulsive disorder, and schizophrenia) and 5 physical diseases (AIDS, asthma, cancer, diabetes, and epilepsy). In their corpus, schizophrenia and HIV were the most stigmatized diseases. These diseases are perceived as dangerous and with an uncontrollable and unpredictable nature. The authors found more than 40% of stigmatizing tweets were about schizophrenia compared to less than 5% of those about depression. Finally, Alvarez-Mon et al [20] recently studied the use of the term "psychosis" and compared it to some medical terms from the field of somatic medicine (diabetes, HIV, Alzheimer disease, and breast cancer). The results showed a predominance of nonmedical content (33.3%) with a high frequency of misuse and pejorative opinion tone (36.2%) in the tweets related to psychosis compared to the tweets related to the physical diseases studied.

Toward the First Analysis of Psychiatric Terms in French Tweets

As far as we know, this is the first study that proposes an in-depth analysis of psychiatric term usage in tweets in French. In particular, we propose the following:

- Analysis of a wide range of psychiatric terms going beyond a small set of nosographic terms. Our study considers a wide range of nosographic terms but also generic and therapeutic psychiatric terms.
- Multilayer annotation scheme that includes the type of word use (medical usage, misuse, or irrelevant usage) and the opinion orientation of the tweet (positive, negative, neutral, or mixed).
- New dataset of about 22,579 tweets containing the selected terms among which 3040 are manually annotated by clinical psychiatrists. The dataset will be made available to the research community.
- Qualitative analysis of the annotated data in terms of interannotator agreement along with quantitative analysis considering mainly term frequency in each layer and exploring the interactions between them.
- Comparison of our results to those obtained by analyzing tweets in English. Our results constitute a first important step toward an automatic detection of stigma in social media.

Methods

Objectives

The multidisciplinary study reported in this paper has been conducted by clinical psychiatrists and computer scientist experts in natural language processing and information extraction from social media. The main objective of the study is to analyze how psychiatric terms are used on Twitter, in particular whether they are used in a medical use. The other goal is to analyze the opinion polarity of these terms and thus highlight the main stereotypes they convey. Our assumption is that psychiatric terms are often misused and these misusages probably have a negative polarity.

Tweet Collection

Our corpus is new and composed of tweets in French that contain selected terms relative to psychiatry. To guarantee a wide lexical convergence of the extracted tweets, we grouped terms according to 3 dimensions:

- Generic terms indicating different morphological variations of the French stem “psychiatr” (psychiatr) such as “psychiatrie” (psychiatry), “psychiatrique” (psychiatric) and “psychiatre” (psychiatrist)
- Nosographic terms relative to psychiatric disorders. Following the *Diagnostic and Statistical Manual of Mental Disorder* taxonomy [21] that classifies mental disorders in order to improve diagnoses, treatment, and research, we grouped terms in 5 main categories: schizophrenia spectrum and other psychotic disorders, bipolar and depressive disorders, autism spectrum disorder, anxiety disorders, and other disorders.
- Therapeutic terms relative to the most used drugs in the psychiatry field.

In each dimension, we selected a set of representative terms experts considered as the most important for this study. For each term, we also considered its slang versions, such as schizo for schizophrenia. We selected a total of 120 psychiatric terms (see Table 1 for examples and frequencies and Multimedia Appendix 1 for the detailed list).

Our dataset is composed of tweets collected via the OSIRIM platform that hosts a Twitter stream representing the 1% of global tweets since 2015, with a total of 73,345,245 tweets. From this collection, we selected tweets in French—using the tag provided by Twitter on tweets—that were posted from January 1, 2016, to December 31, 2018, and contain at least one psychiatric term from our list. After removing retweets and duplicates, we got at a total of 22,579 tweets (Table 2).

Table 1. Examples of terms for each dimension along with their frequencies and English translation (n=120).

Psychiatric terms	Example terms	terms, n
Generic	Psychiat-*	1
Diagnostic		31
Schizophrenia spectrum	Psychose (*psychosis*), Psychotique (*psychotic*), Schizophrène (*schizophrenic*), Schizo (*schizo*)	6
Bipolar and depressive disorders	Maniaque (*manic*), Bipolaire (*bipolar*), Hypomaniaque (*hypomanic*)	7
Autism spectrum	Autisme (*autism*), Autiste (*autistic*)	2
Anxiety disorders	Phobie (*phobia*), TOC[a] (*obsessive compulsive disorder*)	6
Other disorders	Hyperactif (hyperactive), borderline	10
Therapeutic	Neuroleptique (*neuroleptic*), Xanax (*alprazolam*), Theralite (*lithium*)	88

[a]TOC: trouble obsessionnel compulsif.

Table 2. Number of tweets containing the selected terms (a tweet may contain several keywords).

Psychiatric terms	tweets, n
Generic	6993
Diagnostic	12,149
Schizophrenia spectrum	1304
Bipolar and depressive disorders	3500
Autism spectrum	4389
Anxiety disorders	5855
Other disorders	101
Therapeutic	1853

Annotation Guidelines

We designed an annotation guideline to analyze the use of the 120 selected psychiatric terms in tweets in French. To this end, two clinical psychiatrists first analyzed a small subset of 157 tweets randomly selected in order to define the annotation guidelines. These tweets were then removed from the initial collection and never used again in the study.

Our annotation scheme is multilayered and aims at answering 2 main questions: Do the psychiatric terms used in the tweet convey a medical use or not? What is the overall opinion given in the tweet? We detail each layer and illustrate them by example tweets extracted from our corpus. In these examples, psychiatric terms are in bold font. All examples are given in French along with their English translation. Note that translations may not perfectly reflect the initial writing, as tweets often use slang, abbreviations, and contain grammatical errors.

Types of Term Use

We define three possible types of term use: medical use, misuse or irrelevant use.

Medical use corresponds to the medical definition of the term. The term is used to refer to a medical pathology or to the domain of psychiatry, as in Textbox 1.

Misuse occurs when a psychiatric term is used in a figurative or metaphoric way. These misuses often convey prejudices, stereotypes, or humor and thus make psychiatry commonplace and strengthen the stigma of psychiatry and people suffering from psychiatric disorders, as in Textbox 2.

Irrelevant use occurs when the tweet is not understandable (lack of context, link to a URL, advertising, etc) or not relevant to psychiatry (use of synonyms), as in Textbox 3.

Textbox 1. Examples of medical uses (psychiatric terms are in bold font).

- *Tellement dégueulasse le **valium** en gouttes (Oral **valium** is so disgusting)*
- *Tout à l'heure g écouter une vidéo des voix qu'les **schizo** entendent dans leurs têtes g pas pu tenir + de 30sec g cru devenir folle (I listened to a video of voices heard by **schizophrenic** people I couldn't hold more than 30 sec I thought I was going insane)*

Textbox 2. Examples of misuse (psychiatric terms are in bold font).

- *Là j'suis en colère tu changes toutes les minutes, à croire que t'es **bipolaire**. (I'm angry you're changing your mind every minute, I'd think you're **bipolar**)*
- *Tu viens d faire quoi sale **autiste** (What have you just done, you f*** **autistic**)*

Textbox 3. Examples of irrelevant use (psychiatric terms are in bold font).

- *qd t une **schizo** https://t.co/SB3Z1DR7cX (when a **schizophrenic** https://t.co/SB3Z1DR7cX)*
- ***Psychose**, C'est un peu vieux mais c'est trop cool (**Psycho**, it's a little bit old but it's so cool)*

Overall Opinion of the Tweet

As usually defined in sentiment analysis [22], polarity or orientation indicates whether the opinion is positive, negative, or neutral. We consider these 3 possible values and also include mixed opinion to account for cases where the opinion can be positive and negative at the same time. We consider opinion orientation of the author at the tweet level regardless of whether the expressed opinion is related to a psychiatric term.

A tweet is annotated as having positive polarity when the writer expresses a positive personal opinion on facts, events, or on a quote (1); the general idea of the tweet is in favor of psychiatry (2); the writer defends the proper medical use of psychiatric

terms regardless of their valence (3); or with the presence of positive terms or smileys (4), as in Textbox 4.

A tweet has a negative polarity when the writer expresses a negative personal opinion on facts, events, or on a quote (1); with the presence of terms that are basically negative (2); the tweet includes ironic or sarcastic comments (3); the tweet reports negative facts connected to psychiatry (4); the tweet contains a positive smiley linked to a negative content (5); the tweet marks a derogatory or insulting positioning (6); or the psychiatric term is used in the tweet to refer to an inconvenient situation or to a topic releasing a negative emotion (7), as in Textbox 5.

Mixed or neutral polarity orientation mainly covers cases where the opinion of the writer is not clearly expressed (1) or the writer's opinion is mixed, both positive and negative (2), as in Textbox 6.

See Multimedia Appendix 2 for other examples of types of term use and polarity orientation.

Textbox 4. Examples of positive polarity (psychiatric terms are in bold font).

1. *C'est trop top la **psychiatrie** tu vas t'éclater! (**Psychiatry** is so great, you'll have so much fun!)*
2. *Mon Rdv **psychiatrie** de demain tombe à la perfection. Pour une fois je l'avoue, j'en ai grandement besoin. (Tomorrow is the perfect timing for my **psychiatric** appointment. To be honest, for once, I really need it)*
3. ***Bipolaire** c'est un vrai trouble **psychiatrique**, mesdames arrêtez de le mettre en TN vous n'êtes pas **bipolaires** vous êtes juste mal éduquées. (**Bipolar** disorder is a real **mental health** condition. Ladies, stop using this term as tweet name. You are not **bipolar**, you are just poorly-educated)*
4. *ça va mieux t'inquiète pas merci, j'ai pris 3 **Xanax** et ils commencent à faire effet (Feeling better, thanks, don't worry. I took 3 **Xanax** tablets and it has started to work)*

Textbox 5. Examples of negative uses (psychiatric terms are in bold font).

1. *La **psychiatrie** ça brise encore plus les gens. (**Psychiatry** breaks people down even more)*
2. *Il vend sa mère au diable se marie avec une chetana et Il finit en **psychiatrie**. Le pacte 666 l'a détruit. (He sells his mother to the devil, he gets married to a she-devil and he ends up in **psychiatric** hospital. He has been wiped out by pact 666)*
3. *La France est une terre d'asile... **psychiatrique** ! (France is a land of asylum… **psychiatric** asylum!)*
4. *Paris: la **psychiatre** vendait de faux certificats médicaux aux envahisseurs sans-papiers (Paris: **psychiatry** used to sell fake medical certificates to paperless invaders)*
5. *Les artistes finissent presque tous en hôpital **psychiatrique** / (Almost all artists end up in **psychiatric** hospital)*
6. *Selon une grosse conne **psychiatrique** le harcèlement d'activité est une loi de France (According to a **dumb lunatic** woman, harassing is a custom in France)*
7. *Et franchement les garçons radins c'est grave ma **phobie** (Sincerely, stingy boys are basically my greatest **phobia***

Textbox 6. Examples of mixed or neutral use (psychiatric terms are in bold font).

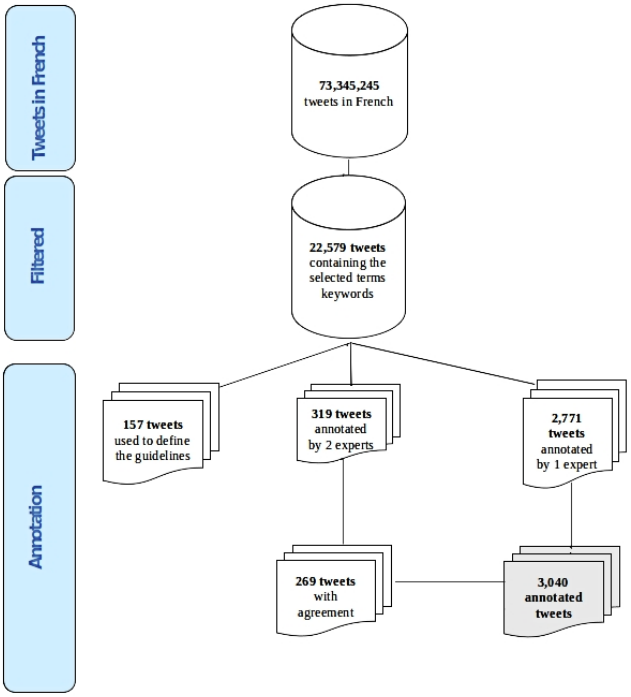
1. *Croyez-vous qu'un **psychiatre** prendrait les médicaments qu'il prescrit (Do you think a **psychiatrist** would take the medicine he prescribes?)*
2. *La psychiatrie c'est cool, Faire ça dans un lieu de stage où ils te harcèlent jusqu'à la dernière heure de tout tout ton stage par contre moins. (**Psychiatry** is fun but throughout the internship they badger you, is less fun)*

Annotation Procedure and Ethics

Our data were manually annotated by two French native speakers, both clinical psychiatrists, using the Brat tool. We performed a 3-step annotation where an intermediate analysis of agreement and disagreement between annotators was completed. Annotators were first trained on 157 tweets that helped them better understand the task and adjust the annotation guidelines. Annotators were then asked to separately annotate 319 tweets (around 10% of the annotated corpus) so that interannotator agreements could be computed. Before moving to the real annotation, annotators were asked to discuss main cases of disagreement, which resulted in a set of 269 tweets. After adjudication, a total of 2771 tweets were manually annotated by one expert. In the end, our dataset is composed of 3040 tweets (269 + 2771) annotated according to our multilevel annotation scheme (Figure 1).

Regarding ethics, we did not request validation with the research ethics board since this study does not involve patients and does not use personal digital data. In addition, our data are composed of textual content from the public domain. Finally, as we will make the dataset publicly available to the research community and conform to the Twitter Developer Agreement and Policy that allows unlimited distribution of the numeric identification number of each tweet.

Figure 1. Annotation procedure.



Interannotator Agreement

Interannotator agreement allows assessment of the amount of agreement between annotators beyond chance and provides a measure of the reliability to the annotation guide. We used the Cohen kappa statistical measure defined as follows [23]:

Where p_o and p_e are probabilities that correspond to the observed and the expected agreements, respectively. The latter probability measures the possible agreement by chance when each annotator randomly selects a given category. Kappa measure ranges from −1 to +1 where $K \leq 0$ indicates no agreement, $0.6 \leq K \leq 0.8$ a high agreement, and $K = 1$ a perfect agreement. We used Microsoft Excel to compute Cohen kappa from the contingency table of frequencies with the rows and columns indicating the categories (agreement frequencies are in the diagonal cells whereas disagreements are in the off-diagonal cells).

Results

Tweet Collection

We conducted a descriptive analysis that relies on the tweet collection we built . This does not require statistical tests.

From the initial collection of about 73 million tweets in French, 22,579 contain at least one term from our list of 120 terms. We observed that 25 terms out of 120 are not present in the dataset. They are mainly therapeutic terms such as international nonproprietary name of neuroleptics or antidepressants. The remaining 95 terms are diagnostic or generic terms referring to psychiatry.

From these 22,579 tweets, 3040 (13.46%) were manually annotated. Annotating tweets is time consuming and requires a high level of expertise in psychiatry. For this reason, we could annotate a limited number of tweets only. In future work, we will consider automatic annotation using supervised machine learning based on these examples, but this is out of the scope of this study.

Table 3 provides the overall frequency of annotated tweets for each dimension of psychiatric terms. Note that a tweet may contain several terms and hence the total is greater than 3040. We observe that tweets with diagnostic terms are the most frequent and that schizophrenia spectrum terms are dominant followed by generic and then bipolar disorders terms.

Table 3. Overall frequency of annotated tweets in each psychiatric term dimension.

Psychiatric terms	tweets, n
Total	3850
Generic	1086
Diagnostic	2604
Schizophrenia	1300
Bipolar/depressive	647
Autism	232
Anxiety	400
Other disorders	25
Therapeutic	160

Interannotator Agreement

In this section, we report on interannotator agreement on both the type of use and the overall opinion levels of the annotation scheme.

Kappa values were computed on the set of 319 tweets independently annotated by the two experts. For the types of term uses (that is to say medical use, misuse, or irrelevant use), we got a K=0.829 whereas for the opinion level (ie, positive, negative, or neutral), we got K=0.817. Interannotator agreement being very high (over 0.80), the guideline is considered reliable and the annotation reproducible.

We note that it is consistent with the percentage agreement, 84.3% (269/319).

Analysis of Psychiatric Terms in the Annotated Dataset

Among the annotated tweets: 12.30% (374/3040) are annotated as irrelevant, 142.37% (1288/3040) as medical use 45.33% (1370/3040) as misuse. Concerning polarity, we could observe that 52.14% (1585/3040) are annotated as having a negative polarity, and 86.21% (1188/3040) of the tweets annotated as misuse are negative while 0.65% (9/3040) are positive. Furthermore, 19.02% (245/1288) of the tweets annotated as medical use have a positive polarity. It is interesting to note that most tweets annotated as medical use are neutral whereas tweets annotated as misuse are mostly negative (Table 4).

We can observe that misuses related to the spectrum of psychotic disorders are more frequent (721/1300, 55.46%) than those related to the spectrum of depression (15/280, 5.36%; Table 5).

Finally, term misuse with a positive polarity is rare and a lot of psychiatric terms such as “schizo” (schizophrenic), “bipolaire” (bipolar), and “autiste” (autistic) are used as insults. See the statistics in Multimedia Appendix 3.

In the remainder of this section, we provide a deeper inspection of each of our two annotation levels for both generic and diagnostic terms focusing on the 13 terms that are present in more than 50 tweets. We fixed this threshold to draw solid conclusions and reliable comparison with related studies. Therapeutic terms being rare in the annotated corpus, we did not include them in this detailed analysis.

Table 4. Distribution of annotated tweets according to the type of term use.

Psychiatric term	Types of term use, n						
	Medical			Misuse			Irrelevant
	Pos[a]	Neg[b]	Neut[c]	Pos	Neg	Neut	
Generic	109	269	376	0	252	17	63
Schizophrenia	79	85	214	2	633	89	198
Bipolar/depressive	125	102	196	7	117	47	53
Autism	63	22	74	0	43	7	23
Anxiety	26	51	94	0	176	20	33
Other disorders	13	3	6	0	3	0	0
Therapeutic	11	33	37	3	19	19	38

[a]Pos: positive.

[b]Neg: negative.

[c]Neut: neutral.

Table 5. Frequency of term misuse according to the most frequent psychiatric terms in the annotated dataset.

Psychiatric terms	Frequency
Total	439
Schizo	70
Maniaque (manic)	65
Psychose (psychosis)	56
Phobie (phobia)	50
Autiste (autistic)	41
Bipolaire (bipolar)	41
Psychotique (psychotic)	40
Psychiatr-	25
Schizophrene (schizophrenic)	25
Schizophrenie (schizophrenia)	11
Depressif/ve (depressed)	7
Depression (depressive disorder)	5
Autisme (autism)	3

Generic Terms

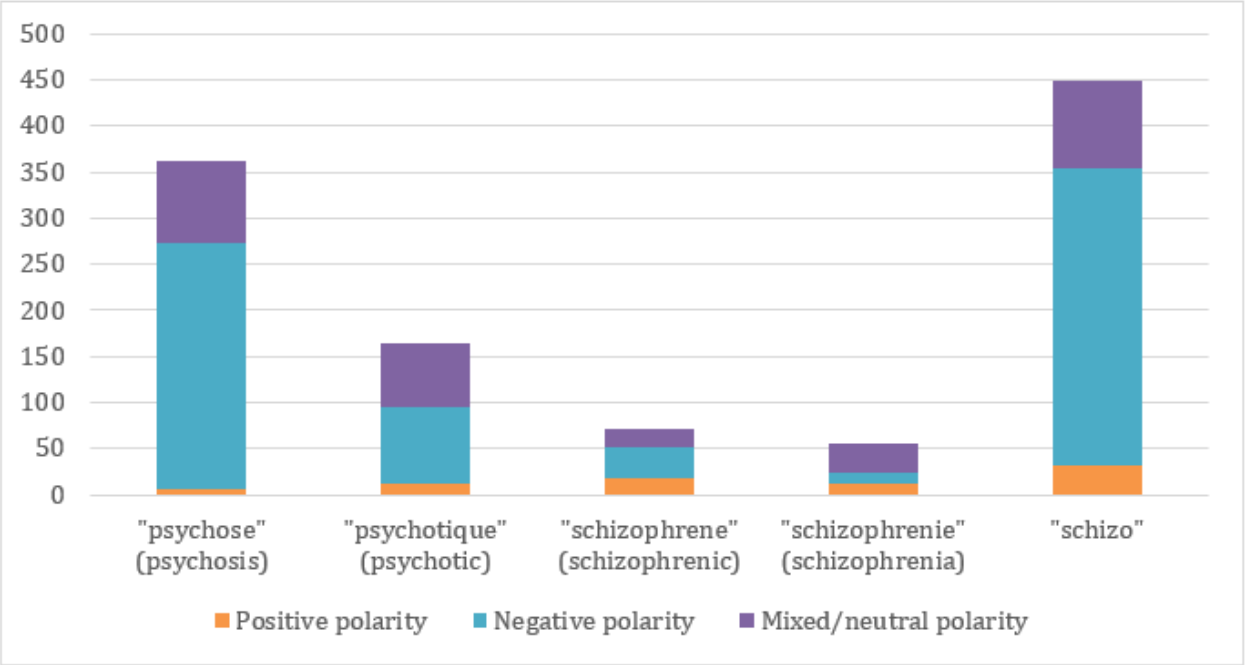
The terms “psychiatrie” (psychiatry), “psychiatrique” (psychiatric), and “psychiatre” (psychiatrist) have been collected using the stem “psychiatry.” These generic terms are mainly used in a medical use (754/1086, 69.43%), but only 14.5% (109/754) of these tweets have a positive polarity.

We observe that generic terms are often used with expressions like “ça relève de la psychiatrie” (It’s a matter of psychiatry) and “finir dans un hôpital psychiatrique” (to end up in a psychiatric hospital). The first expression occurs in 11 tweets and is related to a behavior or a person that is not understood, that seems different or out of the norm. In the same way, the second expression is used as a synonym of irrecoverable.

Schizophrenia Spectrum and Other Psychotic Disorders

We annotated all the tweets of the corpus that contain the terms “schizo,” “schizophrène” (schizophrenic), “schizophrénie” (schizophrenia), “psychose” (psycho), and “psychotique” (psychotic). Among the tweets containing these terms (Table 3), 48.69% (633/1300) are annotated as “misuse” with a negative polarity. All the terms from the spectrum of psychotic disorders have a negative polarity except for “schizophrenie” (schizophrenia), which has mainly a mixed or neutral polarity (Table 4). In particular, 70.1% (354/505) of the tweets containing “schizo” have been annotated as “misuse” versus only 25% (18/73) of the tweets containing “schizophrène.” On the other hand, “psychotique” (psychotic) is more misused (68/172, 39.5%) than “schizophrène” (18/73, 24.66%). In the same way, “psychose” (psycho) is frequently misused (278/494, 56.3%), meaning an excessive fear, a collective terror maintained by a stressful environment or by media. It is often related to violent events, terrorist attacks, or in a context of political tensions. It is more misused than “schizophrénie” (shizophrenia) (6/56, 11%; see Figure 2).

Figure 2. Polarity conveyed by any use according to the terms belonging to the spectrum of psychotic disorders.

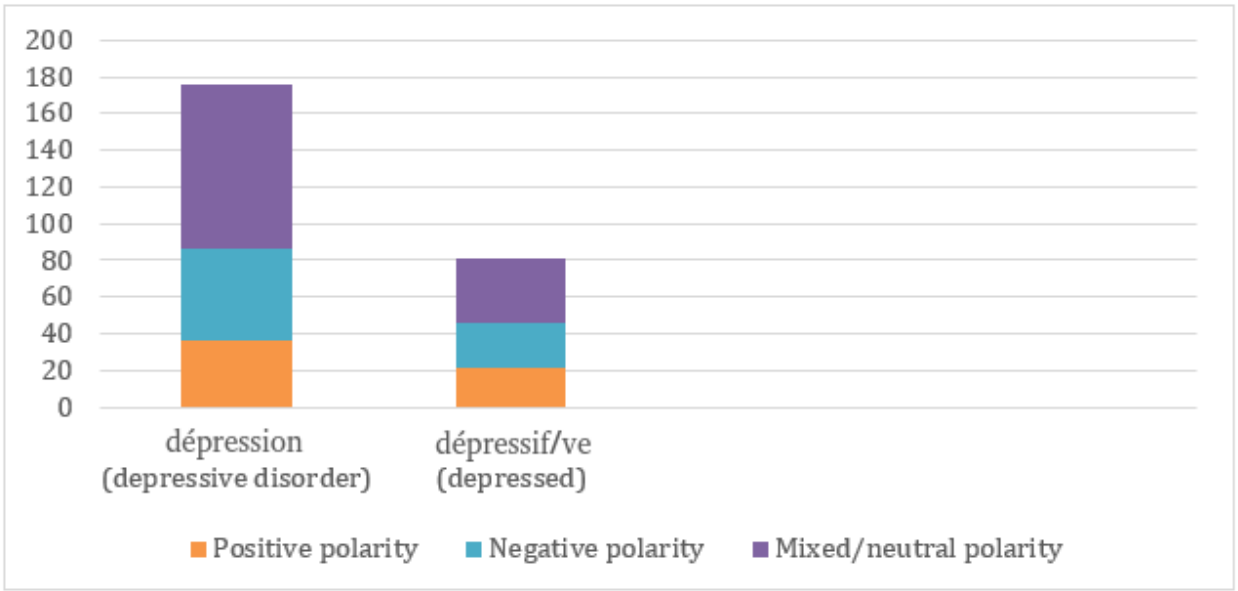


Depressive and Bipolar Disorders

In the domain of depression, all tweets containing the terms “depression” (nervous breakdown) and “dépressif/ive” (depressive) have been annotated, and 5.4% (15/280) are annotated as “misuse.”

From Figure 3, we see that 26.2% (50/191) to 28% (25/89) of the tweets only (depending on the keyword) have a negative polarity. When depression is concerned, tweets could be sorted into 5 categories: personal testimony or experience, recourse to care, defense of medically correct use, qualification of a state of mind, and misused to evoke something other than psychological behavior.

Figure 3. Polarity conveyed by any use according to the depression spectrum terms in the annotated dataset.



When analyzing the distribution of terms relative to bipolar disorders (mainly the term “bipolaire” [bipolar]), 40.6% (112/276) of the tweets are annotated as “misuse” and 24.2% (58/240) of those annotated as “medical use” have a positive polarity. These tweets often describe a person who changes her or his mind or emotions (going from tears to laughter) or who is difficult to understand (as in “Là j’suis en colère tu changes toutes les minutes, à croire que t’es bipolaire” [I’m angry you’re changing your mind every minute, it seems like you’re bipolar]). This term is sometimes related to objects, animals, weather, or political parties that have changing behaviors or contradict themselves.

Finally, 65% (34/52) of the tweets containing the term "maniaque" (maniac) are annotated as "misuse," the term being most often a synonym of obsession. This is one of the few terms being misused but having a positive polarity (4/34, 12%).

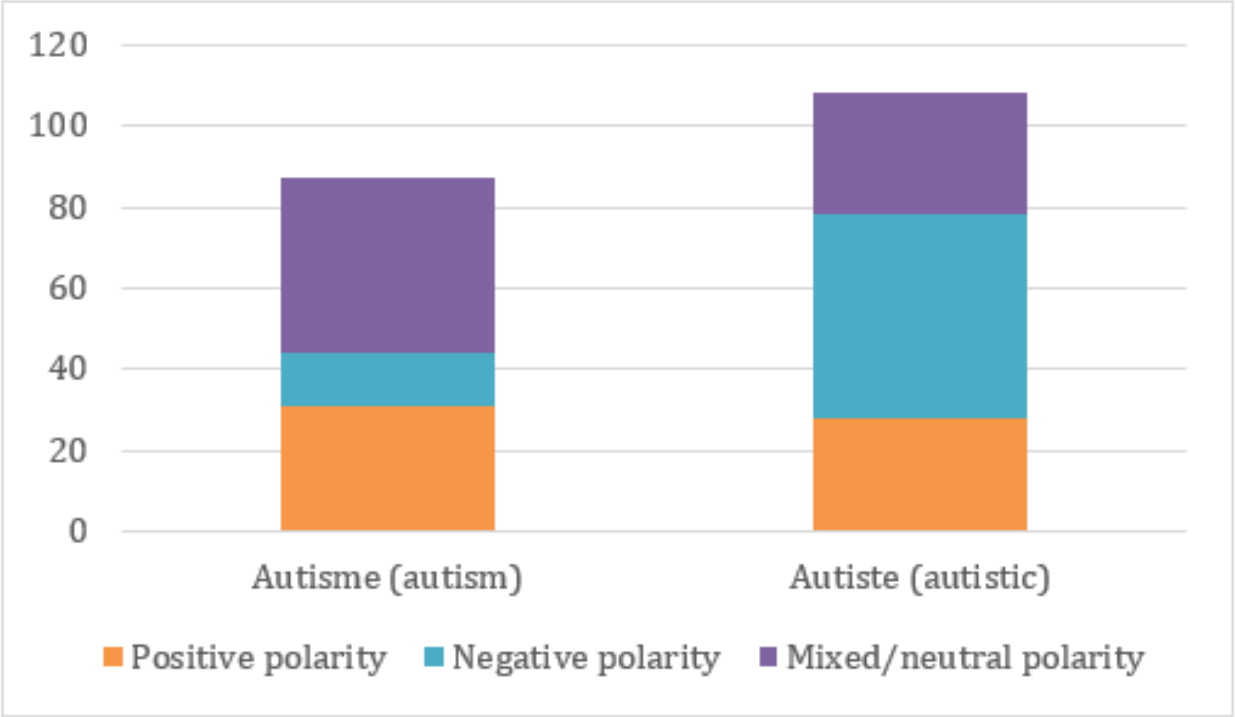
Anxiety Disorders

The term "phobie" (phobia) is the most frequent specific term in our corpus (5494/22,579, 24.33%) and is related to something that is hated or causes anger. Among the annotated tweets containing this term, 50.1% (180/359) are annotated as "misuse" and have mostly a negative polarity (163/180, 90.56%).

Autism Spectrum Disorder

In the domain of autism, we found that 40.9% (47/115) of the tweets containing the term "autiste" (autistic) are annotated as "misuse" and 43.5% (50/115) have a negative polarity. This is not the case for the term "autism" (autism), which is annotated as "medical use" in 91% (84/92) of the tweets and has a negative polarity in only 14% (13/92) of the tweets. The term "autiste" seems to be close to an insult meaning idiot or referring to a person with less adaptation and strange behaviors. On the contrary, the term "autisme" has most often a positive polarity: we noticed that a lot of tweets defend the medical use of this term by conveying information, testimonials from autistic people and families or references to articles about autism. These messages allow one to fight against prejudices (Figure 4).

Figure 4. Polarity conveyed by any use according to the autism spectrum disorder terms in the annotated dataset.



Discussion

Principal Findings

The main goal of this study was to analyze how psychiatric terms are used on Twitter. A descriptive analysis of our annotated corpus shows that terms from the psychiatric domain are often misused (1378/3040, 45.33%) and that most of the tweets about psychiatry convey a negative polarity (1585/3040, 52.14%).

There are some differences in psychiatric term distribution in our corpus. Indeed, there are far fewer tweets about depression than psychosis. This difference may also be due to the fact that there are many more terms related to psychiatry than to depression and thus potentially more tweets. The most frequent psychiatric terms in our corpus are "phobie," "bipolaire," and "autiste," probably meaning that they now belong to everyday language.

Comparison With Related Work in Social Media

Our analysis reveals that depression is less prone to stigmatization, in the sense of term misuse, than schizophrenia. These results confirm the existence of stigmatization and negative prejudices related to psychotic disorders. In particular, thus, the terms "dépression" and "autisme" are less prone to misuse than the terms " psychose" and "psychotique" as well as "schizophrène" and "schizophrénie." This trend toward stigmatization is consistent with the results of previous work in English social networks [16,19]. For example, Robinson et al [19] found stigmatizing behaviors in less than 6% of tweets about depression and autism versus more than 40% of tweets about schizophrenia. Regarding polarity, Joseph et al [17] found that 33% and 21.1% of tweets containing the words "schizophrenic" and "schizophrenia," respectively, convey a negative opinion, which is consistent with our results of 47% (34/73) and 18% (10/56), respectively. On the other hand, the term "depression" in our study is less often associated with a

positive polarity (36/191, 18.9%) versus 65% according to Reavley et al [16]). This difference may reflect the lack of consensus in how to define the positive polarity but also the variability in the sample size or way the tweets were collected.

Limitations

Our results cannot be directly generalized to the discourse conveyed on Twitter or to the representations conveyed in the general population for the following reasons. First, the Twitter user community is not representative of the general population. Indeed, no sociodemographic data were collected to describe the characteristics of the tweet authors in our corpus. Second, the link between thought and written discourse remains complex, which is why it is impossible to extrapolate the ideas conveyed on Twitter to social representations. Third, the method we used to collect tweets also contributes to the difficulty of generalizing our results since Twitter does not provide information concerning the representativeness of the 1% of tweets compared to all the tweets posted daily. Thus, the results of this study apply to our corpus of tweets but nothing can be said on their generalization, especially since our annotated dataset is relatively small. An alternative way to collect tweets could have been to use the Twitter Filter Realtime Streaming API, although because of limitations, we are not fully sure a replication of this can be guaranteed to collect the same tweets (eg, removed tweets), while since we are storing them, we can guarantee replication.

Another limitation concerns the manual annotation itself as the dataset is naturally biased toward annotators' subjectivity mainly because tweets are short and the lack of context forces annotators to rely on cultural and background knowledge to better understand the tweet content. This is particularly true in case of irony and humor. Overall, this is a more general problem when manually building linguistic resources [24], and the reliable interannotator agreement we obtained (0.829 for the types of use and 0.817 for the opinion level) guarantees a good quality of the resource produced in this study.

Conclusions

The analysis of the messages posted on Twitter in French in this pilot descriptive study highlights the existence of a misuse of most of the psychiatric terms studied in our corpus of tweets and the preponderance of a negative polarity conveyed by tweets when talking about psychiatry. It is important to mention that this work, despite some limitations, is part of an international research landscape that is expanding in recent years.

We believe it is necessary to pursue this research on digital social networks in order to improve the quality of discourse analysis and to work toward a better representativeness. Such study is also useful to evaluate the impact of antistigma campaigns on the content of social networks. To this end, future research could focus on a larger corpus of tweets and/or to other popular social networks. Another future direction is the automatic detection of stigma relying on natural language processing techniques. The annotated dataset that we built can then serve as a valuable source to train machine learning techniques to jointly identify the type of use and the opinion conveyed by each tweet.

Acknowledgments

This work has been supported by FERREPSY Occitanie (Fédération Régionale de Recherche en Psychiatrie et santé mentale d'Occitanie), Centre Hospitalier Gérard Marchant. We also thank Dr Dejean Sébastien, statistician at the Institut de Mathématiques, Univeristé de Toulouse, France, for his valuable comments and guidance in this study. We also thank Dr Rouyre Pauline, Centre Hospitalier Gérard Marchant, for his guidance in this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
List of terms used in our study grouped in three dimensions: generic, diagnostic and therapeutic.
[DOCX File , 15 KB - formative_v6i2e18539_app1.docx]

Multimedia Appendix 2
Additional tables.
[DOCX File , 50 KB - formative_v6i2e18539_app2.docx]

Multimedia Appendix 3
Term frequencies.
[DOCX File , 19 KB - formative_v6i2e18539_app3.docx]

References

1. Giordana J. La Stigmatisation en Psychiatrie et en Santé Mentale. Masson: Elsevier; 2010.
2. Crisp AH, Gelder MG, Rix S, Meltzer HI, Rowlands OJ. Stigmatisation of people with mental illnesses. Br J Psychiatry 2000 Jul;177:4-7. [doi: 10.1192/bjp.177.1.4] [Medline: 10945080]

3. Lampropoulos D, Fonte D, Apostolidis T. La stigmatisation sociale des personnes vivant avec la schizophrénie: une revue systématique de la littérature. L'Évolution Psychiatrique 2019 Apr;84(2):346-363. [doi: 10.1016/j.evopsy.2018.09.002]
4. Athanasopoulou C, Välimäki M. 'Schizophrenia' as a metaphor in greek newspaper websites. Stud Health Technol Inform 2014;202:275-278. [Medline: 25000070]
5. Magliano L, Read J, Marassi R. Metaphoric and non-metaphoric use of the term "schizophrenia" in Italian newspapers. Soc Psychiatry Psychiatr Epidemiol 2011 Oct 18;46(10):1019-1025. [doi: 10.1007/s00127-010-0274-3] [Medline: 20640568]
6. L'image de la schizophrénie à travers son traitement médiatique. L'OBSoCo. 2015. URL: https://tinyurl.com/yckujj37 [accessed 2022-01-28]
7. Pignon B, Tebeka S, Leboyer M, Geoffroy P. De « Psychose maniaco-dépressive » à « Troubles bipolaires » : une histoire des représentations sociales et de la stigmatisation en rapport avec la nosographie. Annales Médico-psychologiques, revue psychiatrique 2017 Jul;175(6):514-521. [doi: 10.1016/j.amp.2016.02.015]
8. Berry N, Lobban F, Belousov M, Emsley R, Nenadic G, Bucci S. #WhyWeTweetMH: understanding why people use Twitter to discuss mental health problems. J Med Internet Res 2017 Apr 05;19(4):e107 [FREE Full text] [doi: 10.2196/jmir.6173] [Medline: 28381392]
9. Les 50 chiffres à connaître sur les médias sociaux en 2019. URL: https://www.blogdumoderateur.com/50-chiffres-medias-sociaux-2019/ [accessed 2022-01-27]
10. Twitter. Wikipédia. 2019 Oct 31. URL: https://fr.wikipedia.org/w/index.php?title=Twitter&oldid=163661816 [accessed 2020-02-01]
11. Tauzin A. Classement des réseaux sociaux en France et dans le monde en 2019. Agence Tiz. 2019. URL: https://www.tiz.fr/utilisateurs-reseaux-sociaux-france-monde/ [accessed 2022-01-27]
12. Infographie #WhoUsesTwitter. GlobalWebIndex. 2014. URL: https://guillaume-dardier.fr/utilisateurs-twitter-france.html [accessed 2020-02-01]
13. Link BG, Struening EL, Rahav M, Phelan JC, Nuttbrock L. On stigma and its consequences: evidence from a longitudinal study of men with dual diagnoses of mental illness and substance abuse. J Health Soc Behav 1997 Jun;38(2):177-190. [Medline: 9212538]
14. Depla MFIA, de Graaf R, van Weeghel J, Heeren TJ. The role of stigma in the quality of life of older adults with severe mental illness. Int J Geriatr Psychiatry 2005 Feb;20(2):146-153. [doi: 10.1002/gps.1264] [Medline: 15660405]
15. Lachmar EM, Wittenborn AK, Bogen KW, McCauley HL. #MyDepressionLooksLike: examining public discourse about depression on twitter. JMIR Ment Health 2017 Oct 18;4(4):e43 [FREE Full text] [doi: 10.2196/mental.8141] [Medline: 29046270]
16. Reavley NJ, Pilkington PD. Use of Twitter to monitor attitudes toward depression and schizophrenia: an exploratory study. PeerJ 2014;2:e647 [FREE Full text] [doi: 10.7717/peerj.647] [Medline: 25374786]
17. Joseph AJ, Tandon N, Yang LH, Duckworth K, Torous J, Seidman LJ, et al. #Schizophrenia: use and misuse on Twitter. Schizophr Res 2015 Jul;165(2-3):111-115. [doi: 10.1016/j.schres.2015.04.009] [Medline: 25937459]
18. Athanasopoulou C, Sakellari E. 'Schizophrenia' on Twitter: content analysis of greek language tweets. Stud Health Technol Inform 2016;226:271-274. [Medline: 27350523]
19. Robinson J, Bailey E, Hetrick S, Paix S, O'Donnell M, Cox G, et al. Developing social media-based suicide prevention messages in partnership with young people: exploratory study. JMIR Ment Health 2017 Oct 04;4(4):e40 [FREE Full text] [doi: 10.2196/mental.7847] [Medline: 28978499]
20. Alvarez-Mon MA, Llavero-Valero M, Sánchez-Bayona R, Pereira-Sanchez V, Vallejo-Valdivielso M, Monserrat J, et al. Areas of interest and stigmatic attitudes of the general public in five relevant medical conditions: thematic and quantitative analysis using Twitter. J Med Internet Res 2019 May 28;21(5):e14110 [FREE Full text] [doi: 10.2196/14110] [Medline: 31140438]
21. DSM-5 : Manuel diagnostique et statistique des troubles mentaux. Masson: Elsevier; 2015.
22. Benamara F, Taboada M, Mathieu Y. Evaluative language beyond bags of words: linguistic insights and computational applications. Computational Linguistics 2017 Apr;43(1):201-264. [doi: 10.1162/coli_a_00278]
23. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas 2016 Jul 02;20(1):37-46. [doi: 10.1177/001316446002000104]
24. Wynne M. Developing Linguistic Corpora: A Guide to Good Practice (AHDS Guides to Good Practice). Oxford: Oxbow Books; 2005.

Edited by G Eysenbach; submitted 03.03.20; peer-reviewed by D Carvalho, R Zowalla; comments to author 13.04.20; revised version received 06.06.20; accepted 11.06.20; published 14.02.22.

Please cite as:
Delanys S, Benamara F, Moriceau V, Olivier F, Mothe J
Psychiatry on Twitter: Content Analysis of the Use of Psychiatric Terms in French
JMIR Form Res 2022;6(2):e18539
URL: https://formative.jmir.org/2022/2/e18539
doi:10.2196/18539
PMID:35156925

Original Paper

Loss of Weight Gained During the COVID-19 Pandemic: Content Analysis of YouTube Videos

Hao Tang[1], MS; Sungwoo Kim[2], MS; Priscila E Laforet[3], MPH; Naa-Solo Tettey[4], MPH, EdD; Corey H Basch[4], MPH, EdD

[1]Department of Health and Behavior Studies, Teachers College, Columbia University, New York, NY, United States

[2]Department of Human Development, Teachers College, Columbia University, New York, NY, United States

[3]Mailman School of Public Health, Columbia University, New York, NY, United States

[4]Department of Public Health, William Paterson University, Wayne, NJ, United States

Corresponding Author:
Hao Tang, MS
Department of Health and Behavior Studies
Teachers College
Columbia University
528 West 121st Street
New York, NY, 10027
United States
Phone: 1 2126783964
Email: ht2489@tc.columbia.edu

Abstract

Background: Many people experienced unintended weight gain during the COVID-19 pandemic, which has been discussed widely on social media.

Objective: This study aims to describe the content of weight loss videos on YouTube (Google LLC) during the COVID-19 pandemic.

Methods: By using the keywords *weight loss during quarantine*, the 100 most viewed English-language videos were identified and coded for content related to losing weight gained during the COVID-19 pandemic.

Results: In total, 9 videos were excluded due to having non-English content or posting data before the COVID-19 pandemic. The 91 videos included in the study sample acquired 407,326 views at the time of study and were roughly 14 minutes long. A total of 48% (44/91) of the sample videos included graphic comparisons to illustrate weight change. Videos that included a graphic comparison were more likely to have content related to trigger warnings (χ^2_1=6.05; *P*=.01), weight loss (χ^2_1=13.39; *P*<.001), negative feelings during quarantine (χ^2_1=4.75; *P*=.03), instructions for losing weight (χ^2_1=9.17; *P*=.002), self-love (χ^2_1=6.01; *P*=.01), body shaming (χ^2_1=4.36; *P*=.04), and special dietary practices (χ^2_1=11.10; *P*<.001) but were less likely to include food recipes (χ^2_1=5.05; *P*=.03). Our regression analysis results suggested that mentioning quarantine (*P*=.05), fat-gaining food (*P*=.04), self-care and self-love (*P*=.05), and body shaming (*P*=.008) and having presenters from both sexes (*P*<.001) are significant predictors for a higher number of views. Our adjusted regression model suggested that videos with content about routine change have significantly lower view counts (*P*=.03) than those of videos without such content.

Conclusions: The findings of this study indicate the ways in which YouTube is being used to showcase COVID-19–related weight loss in a pre-post fashion. The use of graphic comparisons garnered a great deal of attention. Additional studies are needed to understand the role of graphic comparisons in social media posts. Further studies that focus on people's attitudes and behaviors toward weight change during the COVID-19 pandemic and the implications of social media on these attitudes and behaviors are warranted.

(JMIR Form Res 2022;6(2):e35164) doi:10.2196/35164

KEYWORDS

COVID-19; quarantine; weight loss; weight gain; social media; YouTube

Introduction

COVID-19 is a contagious respiratory illness caused by the novel coronavirus (SARS-CoV-2). The World Health Organization declared the COVID-19 outbreak a global pandemic due to its rapid spread and alarming severity in March 2020 [1]. In an early attempt to reduce the transmission of SARS-CoV-2 (ie, to "flatten the curve") [2], over 100 countries have implemented self-quarantine at different points since January 2020 (referred to as *lockdown*) [2,3]. In March 2020, different regions in the United States enforced lockdowns of varying stringency, with most regions restricting outdoor activities and shutting down schools and other nonessential businesses [4,5]. Some states announced stay-at-home orders with a mandate that all nonessential work was to be conducted from home [4]. After the COVID-19 vaccines became available to the general public in December 2020, businesses gradually reopened, but social distancing was largely still encouraged [6]. The unprecedented shutdown has effectively slowed the spread of COVID-19 and averted an estimated 531 million coronavirus infections around the world, including 60 million infections in the United States [7]. Despite this however, the negative consequences of the COVID-19 pandemic persist. These include job losses [8] and the increased prevalence of mental health issues (eg, depression and anxiety) [9,10], as well as a myriad of additional societal losses [11]. The COVID-19 quarantine has also resulted in significant lifestyle changes, such as variations to customary eating habits and physical activity [12,13]. These changes have resulted in widespread concerns about weight gain and body image [14]. Reflecting this, the term *quarantine 15* is highly discussed on social media [15,16]. For example, a recent search yielded approximately 50,600 videos on YouTube (Google LLC) that include *quarantine 15* in their titles. Further, a cursory search of the term *quarantine 15* yielded more than 619,000 Instagram posts, not including thousands of posts using related terms and those found in other social media outlets, such as Twitter and Facebook.

Social media has long been a popular tool for sharing and disseminating prompt, health-related information and a cost-effective information and education platform that can help with intervening in health behaviors [17], including weight management. For example, weight loss is one of the most searched topics on the internet [17]. However, not all social media use is positive, especially when it comes to promoting health behaviors such as safe and effective weight loss. For example, content that stigmatizes weight gain or triggers eating disorders is present on social media [16,18]. Moreover, recent evidence suggests that image-centric social media platforms have a greater impact on body image dissatisfaction and eating disorder behaviors than non–image-centric social media platforms [19]. As one of the most popular video-sharing platforms around the globe [20-22], YouTube has been the focal point of a range of studies related to COVID-19. Given the widespread concerns about undesired weight gain, the popularity of YouTube, and the long-existing risk of social media content [14,16,17], it is important to characterize the trending weight-related social media posts during the COVID-19 quarantine to investigate the effect of quarantine on individuals' behaviors and health [16]. Therefore, the purpose of this study was to describe the content of weight loss videos on YouTube during the COVID-19 pandemic.

Methods

This was a qualitative, content analysis study that adapted methods from prior studies that analyzed YouTube videos on COVID-19 vaccination [23,24]. By using "visitor" mode, we conducted searches with the keywords *quarantine weight loss* and *quarantine weight gain* and found that the results were similar. To keep the view count as a valid measure, we used 1 key term—*quarantine weight loss*—to filter the first 100 videos by view count. However, 6 of the most viewed videos were not presented in English, and 3 videos were posted before the pandemic (ie, before 2019). As such, only 91 videos were coded and analyzed.

A total of 23 coding categories were used to code each video, and only 7 basic information categories [25], including (1) URL, (2) upload date, (3) view count, (4) thumbs-up, (5) thumbs-down, (6) video length (in minutes and seconds), and (7) presenter sex, were coded and recorded on the same day. Thumbs-up and thumbs-down counts were subsequently used to calculate the like to dislike ratio. Additional categories were extracted from related articles, a World Health Organization report, and the first 10 most viewed videos [14,25-27]. These characteristic categories were (1) including a trigger warning or disclaimer, (2) mentioning quarantine, (3) including a graphic comparison of pre- and post-pandemic weight (cover, picture, or video), (4) mentioning weight gain during quarantine, (5) mentioning an exact amount of quarantine weight gain, (6) mentioning weight loss during quarantine, (7) mentioning an exact amount of quarantine weight loss, (8) mentioning personal causes of quarantine weight gain, (9) mentioning negative feelings during quarantine, (10) highlighting food of low nutritional quality (dessert or ultraprocessed food), (11) mentioning exercise, (12) mentioning how to lose weight during quarantine, (13) highlighting weight loss pills or products, (14) mentioning self-love or self-care, (15) mentioning body shaming, (16) mentioning a specific diet, and (17) including recipes.

Each video was coded as "1" (yes) or "0" (no) for whether the video mentioned these characteristic categories. Means, SDs, and ranges were calculated for the view count, thumbs-up, thumbs-down, and video length variables. Frequencies and percentages were calculated for all categorical variables. Chi-square tests were used to investigate associations between the inclusion of a graphic comparison of pre- and postpandemic weight and other video characteristics. In addition, a correlation analysis was conducted to determine if there were significant relationships among various video characteristics. Finally, a regression analysis was conducted, with view count (in thousands) as a dependent variable and different video characteristics as independent variables, to test for significant predictors of higher view counts. After the initial interpretation of results, a stepwise regression was performed to reduce the complexity of our model and produce a more efficient model. The descriptive analyses (correlation and chi-square tests) were

performed by using SPSS version 27 (IBM Corporation), and the regression analysis with subset selection was conducted with RStudio 1.4.1717. This study was not reviewed by the institutional review boards of Columbia University and William Paterson University because it did not involve human subjects, per their policies.

Results

The 91 YouTube videos on weight loss during quarantine had an average of 407,326 views; this value had a high SD of 836,478 views. Most (65/91, 71%) of the videos were uploaded between January 1 and August 2, 2020, and only 29% (26/91) of videos were uploaded between August 2, 2020, and March 2, 2021. The majority of videos were presented by females (65/91, 71%). The videos were roughly 14 minutes long on average, though the length ranged from 2 minutes to 1 hour. Most videos were very positively rated; the average like to dislike ratio was 98%. A complete list of the characteristics included in these videos is outlined in Table 1.

Table 1. Characteristics of the most viewed YouTube videos on quarantine weight loss (N=91).

Characteristics	Values
Upload date of videos, n (%)	
Between January 1 and August 2, 2020	65 (71)
Between August 2, 2020, and March 2, 2021	26 (29)
View count, mean (SD; range)	407,326,69 (836,478.02; 44,181-5,396,499)
Thumbs-up (like) count, mean (SD; range)	13,004.82 (24,237.7; 639-176,000)
Thumbs-down (dislike) count, mean (SD; range)	203.31 (297.4; 16-1500)
Video length (seconds), mean (SD; range)	847.19 (614.8; 141-4063)
Presenter sex, n (%)	
Female	65 (71)
Male	22 (24)
Both	4 (4)
Includes a trigger warning or disclaimer, n (%)	
Yes	14 (15)
No	77 (85)
Mentions quarantine, n (%)	
Yes	83 (91)
No	8 (9)
Includes a graphic comparison of pre- and postpandemic weight (cover, picture, or video), n (%)	
Yes	44 (48)
No	47 (52)
Mentions the exact amount of quarantine weight gain, n (%)	
Yes	16 (18)
No	75 (82)
Mentions the exact amount of quarantine weight loss, n (%)	
Yes	37 (42)
No	52 (58)
Mentions weight gain during quarantine, n (%)	
Yes	29 (32)
No	62 (68)
Mentions weight loss during quarantine, n (%)	
Yes	57 (63)
No	34 (37)
Mentions personal causes of quarantine weight gain, n (%)	
Yes	17 (19)
No	74 (81)
Mentions negative feelings during quarantine, n (%)	
Yes	39 (43)
No	52 (57)
Highlights fat-gaining food (dessert or ultraprocessed food), n (%)	
Yes	17 (19)
No	74 (81)
Mentions exercise, n (%)	

Characteristics	Values
Yes	62 (68)
No	29 (32)
Mentions how to lose weight during quarantine, n (%)	
Yes	38 (42)
No	53 (58)
Highlights weight loss pills or products, n (%)	
Yes	24 (26)
No	67 (74)
Mentions self-love or self-care, n (%)	
Yes	30 (33)
No	61 (67)
Mentions body shaming, n (%)	
Yes	10 (11)
No	81 (89)
Mentions a specific diet, n (%)	
Yes	27 (30)
No	64 (70)
Includes recipes, n (%)	
Yes	33 (36)
No	58 (64)

Table 2 features a compilation of chi-square tests of independence for testing the relationship between the inclusion of a graphic pre-post weight comparison and various other video characteristics. Videos that included a trigger warning were more likely to feature a graphic comparison (χ^2_1=6.05; *P*=.01). Such videos that included a graphic comparison of pre- and postpandemic weight also mentioned weight loss more often than videos without a graphic comparison (χ^2_1=13.39; *P*<.001) and more often mentioned negative feelings during quarantine (χ^2_1=4.75; *P*=.03). In addition, videos with a graphic comparison more frequently included how-to instructions (χ^2_1=9.17; *P*=.002) and more frequently mentioned self-love (χ^2_1=6.01; *P*=.01), body shaming (χ^2_1=4.36; *P*=.04), and special dietary practices (χ^2_1=11.10; *P*<.001). However, videos with graphic comparisons significantly less often included food recipes (χ^2_1=5.05; *P*=.03).

Table 2. Associations between the inclusion of a graphic comparison of pre- and postpandemic weight and video characteristics.

Categories	Includes a graphic comparison			Chi-square (*df*)	*P* value
	No, n	Yes, n	Total, N		
Presenter sex	47	44	91	0.47 (2)	.79
Female	35	30	65		
Male	10	12	22		
Both	2	2	4		
Includes a trigger warning or disclaimer	47	44	91	6.05 (1)	.01[a]
No	44	33	77		
Yes	3	11	14		
Mentions quarantine	47	44	91	0.01 (1)	.92
No	4	4	8		
Yes	43	40	83		
Mentions weight gain	47	44	91	0.79 (1)	.37
No	34	28	62		
Yes	13	16	29		
Mentions weight loss	47	44	91	13.39 (1)	<.001[a]
No	26	8	34		
Yes	21	36	57		
Includes a COVID-19 weight change–related term	47	43	90	9.56 (1)	.002[a]
No	45	31	76		
Yes	2	12	14		
Mentions weight gain cause	47	44	91	0.92 (1)	.34
No	40	34	74		
Yes	7	10	17		
Mentions negative feelings during quarantine	47	44	91	4.75 (1)	.03[a]
No	32	20	52		
Yes	15	24	39		
Mentions fat-gaining food	47	44	91	2.24 (1)	.14
No	41	33	74		
Yes	6	11	17		
Mentions exercise	47	44	91	3.28 (1)	.07
No	19	10	29		
Yes	28	34	62		
Mentions routine or life change	47	44	91	0.07 (1)	.79
No	19	19	38		
Yes	28	25	53		
Mentions how to lose weight	47	44	91	9.17 (1)	.002[a]
No	34	18	52		
Yes	13	26	39		
Highlights weight loss pills or products	47	44	91	0.04 (1)	.85
No	35	32	67		

Categories	Includes a graphic comparison			Chi-square (*df*)	*P* value
	No, n	Yes, n	Total, N		
Yes	12	12	24		
Mentions self-love or self-care	47	44	91	6.01 (1)	.01[a]
No	37	24	61		
Yes	10	20	30		
Mentions body shaming	47	44	91	4.51 (1)	.03[a]
No	45	36	81		
Yes	2	8	10		
Mentions a specific diet (eg, keto diet, etc)	47	44	91	10.17 (1)	.001[a]
No	40	24	64		
Yes	7	20	27		
Includes recipes	47	44	91	3.68 (1)	.03[a]
No	25	33	58		
Yes	22	11	33		

[a]Statistically significant at a $P<.05$ significance level.

A regression analysis with a full list of 21 factors was conducted to create a statistically significant model ($F_{21,68}=3.223$; $P<.001$) with moderate model fit ($R^2=34.4\%$). Videos that mentioned COVID-19 quarantine had approximately 580,000 more views than those that did not mention COVID-19 quarantine ($P=.05$). In addition, talking about fat-gaining food ($P=.04$), self-love ($P=.05$), and body shaming ($P=.008$) significantly contributed to higher view counts. Further, videos with both male and female presenters had, on average, 1.8 million more views than videos with male presenters and 1.5 million more views compared to videos with female presenters ($P<.001$). Table 3 outlines the full list of regression coefficients.

Table 3. Regression analysis with full factors.

Characteristics	B (SE)	*t* test (*df*)[a]	*P* value
Intercept	−3420.035 (2411.484)	−1.418 (68)	.16
Like to dislike ratio	3537.054 (2449.459)	1.444 (68)	.15
Length (seconds)	−0.241 (0.141)	−1.714 (68)	.09
Includes a trigger warning	278.817 (236.567)	1.179 (68)	.24
Mentions quarantine	580.534 (290.706)	1.997 (68)	.05[b]
Mentions a graphic comparison	−17.898 (205.592)	−0.087 (68)	.93
Mentions weight gain	−111.256 (205.264)	−0.542 (68)	.59
Mentions weight loss	−295.420 (174.849)	−1.690 (68)	.10
Includes a COVID-19 weight change–related term	−217.298 (229.365)	−0.947 (68)	.35
Mentions weight gain cause	−535.506 (264.648)	−2.023 (68)	.047[b]
Mentions negative feelings	128.832 (188.775)	0.682 (68)	.50
Mentions fat-gaining food	470.720 (218.611)	2.153 (68)	.04[b]
Mentions exercise	28.272 (176.027)	0.161 (68)	.87
Mentions a routine or life change	−392.194 (209.015)	−1.876 (68)	.07
Mentions how to lose weight	350.152 (182.994)	1.913 (68)	.06
Highlights weight loss pills or products	16.195 (190.823)	0.085 (68)	.93
Mentions self-love or self-care	389.582 (195.430)	1.993 (68)	.05[b]
Mentions body shaming	700.194 (255.749)	2.738 (68)	.008[c]
Mentions a specific diet (eg, keto, etc)	−56.951 (198.153)	−0.287 (68)	.78
Includes recipes	−113.440 (179.326)	−0.633 (68)	.53
Presenter is female	269.498 (187.232)	1.439 (68)	.16
Includes both male and female presenters	1848.469 (394.523)	4.685 (68)	<.001[c]

[a]Values are from a regression analysis.

[b]Statistically significant at a $P<.05$ level.

[c]Statistically significant at a $P<.01$ level.

To reduce the number of parameters and improve model fit, forward and backward stepwise regressions were performed. Among the two models, the backward stepwise regression model was selected because it had a lower Akaike Information Criteria value (1195.92) compared to that of the forward regression model (1208.15). The selected model was, overall, significantly similar to the full model ($F_{11,79}=5.506$; $P<.001$) and had a slightly improved model fit ($R^2=35.5\%$). The reduced model included 11 out of the 21 variables, and the complete list of the coefficients can be found in Table 4. The effects of mentioning quarantine ($P=.049$) and body shaming ($P=.009$) and presenter sex ($P<.001$) stayed significant. In addition, mentioning changes in life routine was selected as a significant predictor of view count ($P=.03$), though videos that included content about routine change had approximately 380,000 fewer views.

Table 4. Backward stepwise regression analysis with reduced variables[a].

Characteristics	B (SE)	*t* test (*df*)[b]	*P* value
Intercept	−3323.7 (2291.4)	−1.451 (79)	.15
Like to dislike ratio	3329.7 (2304.5)	1.445 (79)	.15
Length (seconds)	−0.2 (0.1)	−1.778 (79)	.08
Mentions quarantine	556.1 (278.4)	1.997 (79)	.049[c]
Mentions weight gain	−271.6 (167.1)	−1.626 (79)	.11
Highlights fat-gaining food	342.1 (192.4)	1.778 (79)	.08
Mentions a routine or life change	−385.6 (173.4)	−2.212 (79)	.03[c]
Mentions how to lose weight	296.8 (157.1)	1.889 (79)	.06
Mentions self-love or self-care	221.5 (165.5)	1.339 (79)	.18
Mentions body shaming	630.9 (237.1)	2.660 (79)	.009[d]
Presenter is female	246.0 (171.9)	1.432 (79)	.16
Includes both male and female presenters	1889.4 (369.1)	5.100 (79)	<.001[d]

[a]The Akaike Information Criteria and Bayesian Information Criteria values of the model were 1195.92 and 1195.92, respectively.

[b]Values are from a regression analysis.

[c]Statistically significant at a $P<.05$ level.

[d]Statistically significant at a $P<.01$ level.

Discussion

The findings of this study are important in that they indicate the ways in which YouTube is being used to showcase weight loss in a pre-post fashion. Further, videos that used graphic comparisons garnered the most attention and included less of the studied content compared to those that did not use such comparisons. The power of graphic depiction has long been recognized in many fields, including research, education, and business [28-30]. One example is the social comparison theory, which posits that self-worth is often determined through the assessment of differences and similarities with others [31]. This highlights why exposure to body images on social media can result in both positive and negative consequences [32,33].

Many studies have found that social media posts with pictures usually induce higher levels of engagement with eating disorder behaviors than those induced by posts without pictures [34,35]. In this study, videos that included a graphic comparison of pre- and postpandemic weight usually talked about weight loss (mentions weight loss, mentions how to lose weight, and mentions the exact amount of quarantine weight loss) but not weight gain. It is possible that weight loss has always been a popular topic in social media [36]. As such, video makers wanted to make their video content stand out by showing the efficacy of their weight loss journeys with graphic comparisons.

On the other hand, it should be noted that graphic posts can magnify the risk of social media use with regard to body objectification, body dissatisfaction, and eating disorders [19,37-40]. Graphic comparison is a strategy that is used to motivate participants by demonstrating the potential results of following a suggested regimen or advice. However, instead of promoting body positivity, some graphic comparisons have the opposite effect because they are based on the erroneous belief that fat shaming or weight stigma can serve as a motivator for weight loss [41,42]. Transformational graphic images often focus on decreased size and high amounts of weight loss as determinants of good health. The linking of weight to health can lead to negative body image and decreased self-esteem for those who do not meet the criteria in the posted images or videos [43]. In addition, this strategy does not often lead to motivation but instead can discourage and decrease the self-efficacy of viewers who do not believe that they can achieve the weight loss goals portrayed in such images or videos [43]. Furthermore, considering racial and ethnic representation and cultural body image standards within these images is important for reaching a diverse audience and achieving body inclusivity [44].

Including trigger warnings could be an effective strategy for limiting exposure to content that can distort body image. We found that videos that include a pre-post graphic comparison are more likely to include trigger warnings or disclaimers and mention self-love, self-care, and body shaming compared to those without a graphic comparison. Only 15% (14/91) of these popular videos discussing weight change included trigger warnings or disclaimers. Social media platforms and content creators should be more aware of the potential risk of content related to body image and promote policies to reduce this risk. Future studies should seek to develop best practices for developing graphic images in a way that promotes health and body positivity instead of just weight loss and a thin body ideal.

The literature indicates that several lifestyle changes during the COVID-19 quarantine have resulted in weight gain and increased the risk for obesity [45,46], which is a primary public health concern [47,48]. Studies have confirmed that weight gain was commonplace around the globe during the COVID-19 quarantine [14,45,46,49]. The interruption of usual routines and

restriction of social behaviors due to the COVID-19 quarantine can result in increased boredom [12] and stress [50], which in turn induce emotional eating and food craving [51-53]. Additionally, higher energy intake; the higher consumption of sugar, fat, and alcohol; and the limited availability of fresh fruits and vegetables [13,54] during quarantine increased the risk for overweight and many metabolic diseases [55]. In addition, the decrease in physical activity and increase in sedentary behaviors, such as screen time, during the COVID-19 quarantine may also contribute to weight and fat gain [56].

This study is limited by its cross-sectional design, the sole inclusion of English-language videos, and the search term being limited to 1 phrase. There is no indication of how our results may have differed at other points in time during the pandemic, as quarantine rules varied greatly over time. Despite these limitations, this study does contribute to a gap in the literature and may encourage researchers to conduct studies related to the loss of weight gained during the COVID-19 pandemic.

Social media use increased significantly during the COVID-19 pandemic, since it has helped people feel connected with others and has kept them updated with news and entertained while staying at home [16,57]. More than half of US adults reported the increased use of social media platforms after the pandemic [58]. Social media sites like YouTube provide an opportunity for lay content creators, as well as public health organizations, to reach large audiences and provide content that can promote improved body image and increased focus on health rather than weight [59]. The lessons learned from the COVID-19 pandemic should serve as a catalyst for public health practitioners to develop evidence-based tools that people can use to remain healthy should an extended quarantine occur again. Many creators of social media content that focuses on weight loss are not trained in health education or public health and may not be using evidence-based strategies to develop content. It would be beneficial for trusted organizations to develop evidence-based social media education and training with guidelines for maintaining a healthy weight and establishing healthy behaviors during times of quarantine or other public health emergencies. Further studies that focus on people's attitudes and behaviors toward weight change during the COVID-19 pandemic and the implications of social media on these attitudes and behaviors are warranted.

Authors' Contributions

CHB and HT conceptualized the study. HT collected the data. SK and PEL conducted the data analysis. All authors contributed to manuscript production.

Conflicts of Interest

None declared.

References

1. Cucinotta D, Vanelli M. WHO declares COVID-19 a pandemic. Acta Biomed 2020 Mar 19;91(1):157-160 [FREE Full text] [doi: 10.23750/abm.v91i1.9397] [Medline: 32191675]
2. Perra N. Non-pharmaceutical interventions during the COVID-19 pandemic: A review. Phys Rep 2021 May 23;913:1-52 [FREE Full text] [doi: 10.1016/j.physrep.2021.02.001] [Medline: 33612922]
3. Levenson M. Scale of China's Wuhan shutdown is believed to be without precedent. The New York Times. 2020 Jan 22. URL: https://www.nytimes.com/2020/01/22/world/asia/coronavirus-quarantines-history.html [accessed 2021-08-10]
4. Chappell B, Romo V. New York, Illinois governors issue stay at home orders, following California's lead. National Public Radio. 2020 Mar 20. URL: https://www.npr.org/sections/coronavirus-live-updates/2020/03/20/818952589/coronavirus-n-y-gov-cuomo-says-100-of-workforce-must-stay-home [accessed 2022-01-11]
5. Ravani S. Bay Area coronavirus decision: Behind the scenes of nation's first shelter-in-place order. San Francisco Chronicle. 2020 Mar 21. URL: https://www.sfchronicle.com/bayarea/article/Bay-Area-coronavirus-decision-Behind-the-scenes-15148425.php [accessed 2022-01-11]
6. COVID-19 restrictions. Map of COVID-19 case trends and restrictions. USA Today. URL: https://www.usatoday.com/storytelling/coronavirus-reopening-america-map/ [accessed 2021-08-19]
7. Hsiang S, Allen D, Annan-Phan S, Bell K, Bolliger I, Chong T, et al. The effect of large-scale anti-contagion policies on the COVID-19 pandemic. Nature 2020 Aug;584(7820):262-267. [doi: 10.1038/s41586-020-2404-8] [Medline: 32512578]
8. Unemployment rates during the COVID-19 pandemic. Congressional Research Service. 2021. URL: https://fas.org/sgp/crs/misc/R46554.pdf [accessed 2021-08-19]
9. Schippers MC. For the greater good? The devastating ripple effects of the Covid-19 crisis. Front Psychol 2020 Sep 29;11:577740 [FREE Full text] [doi: 10.3389/fpsyg.2020.577740] [Medline: 33132987]
10. Holmes EA, O'Connor RC, Perry VH, Tracey I, Wessely S, Arseneault L, et al. Multidisciplinary research priorities for the COVID-19 pandemic: a call for action for mental health science. Lancet Psychiatry 2020 Jun;7(6):547-560 [FREE Full text] [doi: 10.1016/S2215-0366(20)30168-1] [Medline: 32304649]
11. Lopez-Leon S, Wegman-Ostrosky T, Perelman C, Sepulveda R, Rebolledo PA, Cuapio A, et al. More than 50 long-term effects of COVID-19: a systematic review and meta-analysis. Sci Rep 2021 Aug 09;11(1):16144 [FREE Full text] [doi: 10.1038/s41598-021-95565-8] [Medline: 34373540]

12. Di Renzo L, Gualtieri P, Pivari F, Soldati L, Attinà A, Cinelli G, et al. Eating habits and lifestyle changes during COVID-19 lockdown: an Italian survey. J Transl Med 2020 Jun 08;18(1):229 [FREE Full text] [doi: 10.1186/s12967-020-02399-5] [Medline: 32513197]
13. Bennett G, Young E, Butler I, Coe S. The impact of lockdown during the COVID-19 outbreak on dietary habits in various population groups: A scoping review. Front Nutr 2021 Mar 04;8:626432 [FREE Full text] [doi: 10.3389/fnut.2021.626432] [Medline: 33748175]
14. Jordan AK, Barnhart WR, Studer-Perez EI, Kalantzis MA, Hamilton L, Musher-Eizenman DR. 'Quarantine 15': Pre-registered findings on stress and concern about weight gain before/during COVID-19 in relation to caregivers' eating pathology. Appetite 2021 Nov 01;166:105580. [doi: 10.1016/j.appet.2021.105580] [Medline: 34186158]
15. Keel PK, Gomez MM, Harris L, Kennedy GA, Ribeiro J, Joiner TE. Gaining "The Quarantine 15:" Perceived versus observed weight changes in college students in the wake of COVID-19. Int J Eat Disord 2020 Nov;53(11):1801-1808 [FREE Full text] [doi: 10.1002/eat.23375] [Medline: 32856752]
16. Pearl RL. Weight stigma and the "Quarantine-15". Obesity (Silver Spring) 2020 Jul;28(7):1180-1181 [FREE Full text] [doi: 10.1002/oby.22850] [Medline: 32324954]
17. Jane M, Hagger M, Foster J, Ho S, Pal S. Social media for health promotion and weight management: a critical debate. BMC Public Health 2018 Jul 28;18(1):932 [FREE Full text] [doi: 10.1186/s12889-018-5837-3] [Medline: 30055592]
18. Das L, Mohan R, Makaya T. The bid to lose weight: impact of social media on weight perceptions, weight control and diabetes. Curr Diabetes Rev 2014;10(5):291-297. [doi: 10.2174/1573399810666141010112542] [Medline: 25311196]
19. Griffiths S, Murray SB, Krug I, McLean SA. The contribution of social media to body dissatisfaction, eating disorder symptoms, and anabolic steroid use among sexual minority men. Cyberpsychol Behav Soc Netw 2018 Mar;21(3):149-156 [FREE Full text] [doi: 10.1089/cyber.2017.0375] [Medline: 29363993]
20. Digital 2021 - Social media marketing and management dashboard. Hootsuite. URL: https://www.hootsuite.com/resources/digital-trends [accessed 2021-08-10]
21. YouTube: Statistics and facts. Statista. 2021 Jul 12. URL: https://www.statista.com/topics/2019/youtube/ [accessed 2021-08-10]
22. Newberry C. 25 YouTube statistics that may surprise you: 2021 edition. Hootsuite. 2021 Feb 02. URL: https://blog.hootsuite.com/youtube-stats-marketers/ [accessed 2021-08-10]
23. Basch CE, Basch CH, Hillyer GC, Meleo-Erwin ZC, Zagnit EA. YouTube videos and informed decision-making about COVID-19 vaccination: Successive sampling study. JMIR Public Health Surveill 2021 May 06;7(5):e28352 [FREE Full text] [doi: 10.2196/28352] [Medline: 33886487]
24. Basch CH, Hillyer GC, Zagnit EA, Basch CE. YouTube coverage of COVID-19 vaccine development: implications for awareness and uptake. Hum Vaccin Immunother 2020 Nov 01;16(11):2582-2585 [FREE Full text] [doi: 10.1080/21645515.2020.1790280] [Medline: 32701403]
25. Basch CH, Mohlman J, Fera J, Tang H, Pellicane A, Basch CE. Community mitigation of COVID-19 and portrayal of testing on TikTok: Descriptive study. JMIR Public Health Surveill 2021 Jun 10;7(6):e29528 [FREE Full text] [doi: 10.2196/29528] [Medline: 34081591]
26. Kriaucioniene V, Bagdonaviciene L, Rodríguez-Pérez C, Petkeviciene J. Associations between changes in health behaviours and body weight during the COVID-19 quarantine in Lithuania: The Lithuanian COVIDiet Study. Nutrients 2020 Oct 13;12(10):3119 [FREE Full text] [doi: 10.3390/nu12103119] [Medline: 33065991]
27. Social media and COVID-19: A global study of digital crisis interaction among Gen Z and millennials. World Health Organization. 2021 Dec 01. URL: https://www.who.int/news-room/feature-stories/detail/social-media-covid-19-a-global-study-of-digital-crisis-interaction-among-gen-z-and-millennials [accessed 2021-08-10]
28. Dunlap JC, Lowenthal PR. Getting graphic about infographics: design lessons learned from popular infographics. Journal of Visual Literacy 2016 Sep 08;35(1):42-59 [FREE Full text] [doi: 10.1080/1051144x.2016.1205832]
29. Guo D, Zhang S, Wright KL, McTigue EM. Do you get the picture? A meta-analysis of the effect of graphics on reading comprehension. AERA Open 2020 Jan 24;6(1):233285842090169 [FREE Full text] [doi: 10.1177/2332858420901696]
30. Khoury CK, Kisel Y, Kantar M, Barber E, Ricciardi V, Klirs C, et al. Science-graphic art partnerships to increase research impact. Commun Biol 2019 Aug 06;2:295 [FREE Full text] [doi: 10.1038/s42003-019-0516-1] [Medline: 31396575]
31. Festinger L. A theory of social comparison processes. Hum Relat 1954 May 01;7(2):117-140. [doi: 10.1177/001872675400700202]
32. Bennett BL, Whisenhunt BL, Hudson DL, Wagner AF, Latner JD, Stefano EC, et al. Examining the impact of social media on mood and body dissatisfaction using ecological momentary assessment. J Am Coll Health 2020 Jul;68(5):502-508. [doi: 10.1080/07448481.2019.1583236] [Medline: 30908131]
33. Sherlock M, Wagstaff DL. Exploring the relationship between frequency of Instagram use, exposure to idealized images, and psychological well-being in women. Psychol Pop Media Cult 2019;8(4):482-490. [doi: 10.1037/ppm0000182]
34. Newberry C, McLachlan S. Social media advertising 101: How to get the most out of your ad budget. Hootsuite. 2020 Sep 09. URL: https://blog.hootsuite.com/social-media-advertising/ [accessed 2021-09-07]
35. Li Y, Xie Y. Is a picture worth a thousand words? An empirical study of image content and social media engagement. J Mark Res 2019 Nov 18;57(1):1-19 [FREE Full text] [doi: 10.1177/0022243719881113]

36. Park MB, Wang JM, Bulwer BE. Global dieting trends and seasonality: Social big-data analysis may be a useful tool. Nutrients 2021 Mar 25;13(4):1069 [FREE Full text] [doi: 10.3390/nu13041069] [Medline: 33806069]
37. Sidani JE, Shensa A, Hoffman B, Hanmer J, Primack BA. The association between social media use and eating concerns among US young adults. J Acad Nutr Diet 2016 Sep;116(9):1465-1472 [FREE Full text] [doi: 10.1016/j.jand.2016.03.021] [Medline: 27161027]
38. Wilksch SM, O'Shea A, Ho P, Byrne S, Wade TD. The relationship between social media use and disordered eating in young adolescents. Int J Eat Disord 2020 Jan;53(1):96-106. [doi: 10.1002/eat.23198] [Medline: 31797420]
39. Teo NSY, Collinson SL. Instagram and risk of rumination and eating disorders: An Asian perspective. Psychol Pop Media Cult 2019;8(4):491-508. [doi: 10.1037/ppm0000205]
40. Walker M, Thornton L, De Choudhury M, Teevan J, Bulik CM, Levinson CA, et al. Facebook use and disordered eating in college-aged women. J Adolesc Health 2015 Aug;57(2):157-163 [FREE Full text] [doi: 10.1016/j.jadohealth.2015.04.026] [Medline: 26206436]
41. Puhl RM, Heuer CA. Obesity stigma: important considerations for public health. Am J Public Health 2010 Jun;100(6):1019-1028. [doi: 10.2105/AJPH.2009.159491] [Medline: 20075322]
42. Vartanian LR, Pinkus RT, Smyth JM. Experiences of weight stigma in everyday life: Implications for health motivation. Stigma Health 2018;3(2):85-92. [doi: 10.1037/sah0000077]
43. Hunger JM, Major B, Blodorn A, Miller CT. Weighed down by stigma: How weight-based social identity threat contributes to weight gain and poor health. Soc Personal Psychol Compass 2015 Jun;9(6):255-268 [FREE Full text] [doi: 10.1111/spc3.12172] [Medline: 29225670]
44. Lazuka RF, Wick MR, Keel PK, Harriger JA. Are we there yet? Progress in depicting diverse images of beauty in Instagram's body positivity movement. Body Image 2020 Sep;34:85-93. [doi: 10.1016/j.bodyim.2020.05.001] [Medline: 32534269]
45. Zachary Z, Brianna F, Brianna L, Garrett P, Jade W, Alyssa D, et al. Self-quarantine and weight gain related risk factors during the COVID-19 pandemic. Obes Res Clin Pract 2020;14(3):210-216 [FREE Full text] [doi: 10.1016/j.orcp.2020.05.004] [Medline: 32460966]
46. Zeigler Z. COVID-19 self-quarantine and weight gain risk factors in adults. Curr Obes Rep 2021 Sep;10(3):423-433 [FREE Full text] [doi: 10.1007/s13679-021-00449-7] [Medline: 34251647]
47. Hales CM, Fryar CD, Carroll MD, Freedman DS, Ogden CL. Trends in obesity and severe obesity prevalence in US youth and adults by sex and age, 2007-2008 to 2015-2016. JAMA 2018 Apr 24;319(16):1723-1725 [FREE Full text] [doi: 10.1001/jama.2018.3060] [Medline: 29570750]
48. Adult obesity facts. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/obesity/data/adult.html [accessed 2021-08-10]
49. One year later, a new wave of pandemic health concerns. American Psychological Association. 2021 Mar 11. URL: https://www.apa.org/news/press/releases/stress/2021/one-year-pandemic-stress [accessed 2021-08-25]
50. 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]
51. Torres SJ, Nowson CA. Relationship between stress, eating behavior, and obesity. Nutrition 2007;23(11-12):887-894. [doi: 10.1016/j.nut.2007.08.008] [Medline: 17869482]
52. Yau YHC, Potenza MN. Stress and eating behaviors. Minerva Endocrinol 2013 Sep;38(3):255-267 [FREE Full text] [Medline: 24126546]
53. Moynihan AB, van Tilburg WAP, Igou ER, Wisman A, Donnelly AE, Mulcaire JB. Eaten up by boredom: consuming food to escape awareness of the bored self. Front Psychol 2015 Apr 01;6:369 [FREE Full text] [doi: 10.3389/fpsyg.2015.00369] [Medline: 25883579]
54. Deschasaux-Tanguy M, Druesne-Pecollo N, Esseddik Y, de Edelenyi FS, Alles B, Andreeva VA, et al. Diet and physical activity during the COVID-19 lockdown period (March-May 2020): results from the French NutriNet-Santé cohort study. medRxiv. Preprint posted online on June 5, 2020 [FREE Full text] [doi: 10.1101/2020.06.04.20121855]
55. Mattioli AV, Sciomer S, Cocchi C, Maffei S, Gallina S. Quarantine during COVID-19 outbreak: Changes in diet and physical activity increase the risk of cardiovascular disease. Nutr Metab Cardiovasc Dis 2020 Aug 28;30(9):1409-1417 [FREE Full text] [doi: 10.1016/j.numecd.2020.05.020] [Medline: 32571612]
56. 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]
57. Jiang Y. Problematic social media usage and anxiety among university students during the COVID-19 pandemic: The mediating role of psychological capital and the moderating role of academic burnout. Front Psychol 2021 Feb 05;12:612007 [FREE Full text] [doi: 10.3389/fpsyg.2021.612007] [Medline: 33613391]
58. Samet A. How the coronavirus is changing US social media usage. Insider Intelligence. 2020 Jul 29. URL: https://www.emarketer.com/content/how-coronavirus-changing-us-social-media-usage [accessed 2021-08-10]
59. Webb JB, Vinoski ER, Bonar AS, Davies AE, Etzel L. Fat is fashionable and fit: A comparative content analysis of Fatspiration and Health at Every Size Instagram images. Body Image 2017 Sep;22:53-64. [doi: 10.1016/j.bodyim.2017.05.003] [Medline: 28624756]

Edited by G Eysenbach; submitted 28.11.21; peer-reviewed by J Harris; comments to author 23.12.21; revised version received 23.12.21; accepted 28.12.21; published 09.02.22.

Please cite as:
Tang H, Kim S, Laforet PE, Tettey NS, Basch CH
Loss of Weight Gained During the COVID-19 Pandemic: Content Analysis of YouTube Videos
JMIR Form Res 2022;6(2):e35164
URL: https://formative.jmir.org/2022/2/e35164
doi:10.2196/35164
PMID:34978534

Original Paper

The Accessibility of YouTube Fitness Videos for Individuals Who Are Disabled Before and During the COVID-19 Pandemic: Preliminary Application of a Text Analytics Approach

Shevali Kadakia[1*]; Catherine Stratton[2*], BA, MPH; Yinfei Wu[3], BA; Josemari Feliciano[3], BSc, MPH; Yetsa A Tuakli-Wosornu[2,4], MPH, MD

[1]Department of Computing and Mathematical Sciences, California Institute of Technology, Pasadena, CA, United States

[2]Department of Chronic Disease Epidemiology, Yale School of Public Health, New Haven, CT, United States

[3]Department of Biostatistics, Yale School of Public Health, New Haven, CT, United States

[4]Department of Physical Medicine and Rehabilitation, University of Pittsburgh Medical Center, Pittsburgh, PA, United States

[*]these authors contributed equally

Corresponding Author:
Yetsa A Tuakli-Wosornu, MPH, MD
Department of Physical Medicine and Rehabilitation
University of Pittsburgh Medical Center
3471 Fifth Ave
Pittsburgh, PA, 15213
United States
Phone: 1 4436104322
Email: yetsa.tuakli-wosornu@yale.edu

Abstract

Background: People with disabilities face barriers to in-person physical activity (PA), including a lack of adaptive equipment and knowledgeable instructors. Given this and the increased need for digital resources due to widespread COVID-19 lockdowns, it is necessary to assess the accessibility of digital fitness resources for people with disabilities. To investigate whether YouTube fitness content creators have made videos accessible to people with disabilities would be informative about access to PA during COVID-19 and could also provide insight into the feasibility of individuals who are disabled relying on YouTube for PA in a post–COVID-19 world.

Objective: This study aims to ascertain if disability-friendly PA videos on YouTube are accessible through searching general fitness terms and whether a change in the availability of accessible fitness resources for people with disabilities occurred on YouTube between before and during the COVID-19 pandemic on "Hospital/Medical Institutions," "Individual(s)," and "Other(s)" channels. Secondary aims are to investigate if different categories of YouTube channels produce more accessible fitness content and highlight any disparities in disability-friendly PA content on YouTube.

Methods: A cross-sectional text analysis of exercise-related YouTube videos was conducted. The authors used Python (version 3.0) to access the YouTube database via its data application programming interface. Terms pertaining to PA that were searched on YouTube were *at-home exercise*, *exercise at home*, *exercise no equipment*, *home exercise*, *home-based exercise*, *no equipment workout*, and *workout no equipment*. Various elements (eg, view count and content generation) of the videos published between January 1 and June 30, 2019 (n=700), were compared to the elements of videos published between January 1 and June 30, 2020 (n=700). To capture a broad idea of disability-friendly videos on YouTube, videos were labeled "accessible" if they were found in the first 100 video results and if their title, description, or tags contained the following terms: *para*, *paralympic*, *adaptive*, *adapted*, *disabled*, *disability*, *differently abled*, *disability-friendly*, *wheelchair accessible*, and *inclusive*. Each video and channel were categorized as "Hospitals/Medical Institutions," "Individuals," or "Other(s)."

Results: The analysis revealed a statistically significant increase in viewership of fitness content on YouTube (P=.001) and in fitness content generated by Hospitals/Medical Institutions (P=.004). Accessible terms applicable to people with disabilities had minimal appearances in 2019 (21 videos) and 2020 (19 videos). None of the top viewed fitness videos that populated on YouTube from 2019 or 2020 were accessible.

Conclusions: The proportion of accessible disability-friendly videos remains diminutive relative to the prevalence of disability in the general population, revealing that disability-friendly videos are seldom findable on YouTube. Thus, the need for disability-friendly fitness content to be easily searched and found remains urgent if access to digital fitness resources is to improve.

(JMIR Form Res 2022;6(2):e34176) doi:10.2196/34176

KEYWORDS

persons with disabilities; disability; exercise; physical activity; digital health; YouTube; accessibility; fitness; COVID-19; text analysis; social media; video

Introduction

Background

Physical activity (PA) is a critical health strategy for the prevention and maintenance of strong physical and mental health as well as upholding a high quality of life [1-4]. There is strong evidence that people with disabilities report markedly lower rates of PA than their abled-bodied peers [1-4]. For example, according to one study, only 45% of Americans adults with a mobility disability participated in aerobic PA [5]. This lower rate of PA, in part, explains how people with disabilities present with serious illnesses such as obesity, heart disease, stroke, diabetes, and cancer at higher rates than the general population. Therefore, strategies to address these barriers should be developed [1].

People who are disabled have often struggled to access exercise trainers and equipment due to a lack of social support in the fitness and sports sectors, insufficient knowledge of disability among fitness instructors, and a shortage of adaptive fitness resources in gyms [3,4,6]. For example, Richardson et al [4] conducted semistructured interviews with individuals with disabilities about their experiences with the gym. Although several participants indicated that they believed the gym could have the power to improve their physical wellness and social engagement, they also noted that their experiences were often at odds with the gym's culture [4]. This proves to be an extreme setback, as it has been documented that people with disabilities tend to be more willing to participate in PA if the gym instructor has medical knowledge of their particular diagnosis or disability [3,4,7]. In fact, one study conducted by the Lakeshore Foundation in collaboration with Degree found that 81% of people with disabilities feel uncomfortable using traditional gym and fitness spaces and resources [8]. The reasons for this include, but are not limited to, having greater trust in the source of instruction and greater comfort in the safety of PA if it is being led by someone who would understand the manifestations and possible limitations of a particular diagnosis [7]. These barriers will require systemic change. In the interim, it is possible that people with disabilities might consider alternative modes of PA, such as accessing digital fitness resources.

Understanding the accessibility of fitness resources for people with disabilities on social media platforms such as YouTube could be beneficial for them since some people with disabilities who have discomfort toward in-person fitness settings might be more inclined to use online resources. Thus, considering the need for at-home PA resources due to the social deterrents associated with in-person PA for some people with disabilities, investigations into the accessibility of digital fitness resources as an alternative for people with disabilities are timely and warranted.

Furthermore, the COVID-19 pandemic has increased our dependencies on digital options for activities such as fitness [9-13]. Though increases in PA content on YouTube have not been widely reported, individuals and groups of people creating fitness content on YouTube have seen significant spikes in metrics of engagement with digital resources, such as their number of subscribers and views [14]. For example, more patients have begun relying on hospitals' or medical institutions' online fitness sessions to improve their stress and anxiety [15]. Research has found that not only can engaging in PA online be effective in providing the same benefits of more traditional modes of PA [16] but also transitioning to the virtual space has resulted in some benefits, one being the larger audience with whom fitness instructors interact [17,18]. An example of this includes The University of Milan's *#StayHomeStayFit* movement. This movement reached over 21,000 people, which is a 100-fold increase compared to their prior in-person fitness classes [19]. Given the increasing popularity of digital PA resources, it is important to determine if their accessibility has extended to the disability community as well. Moreover, the unique barriers to in-person PA resources for people with disabilities make investigations into the accessibility of digital fitness resources for people with disabilities, and how these resources could be extended in a postpandemic world, important.

Although there could be a shortage of digital PA resources for people with disabilities, the authors acknowledge that such content may exist on YouTube as well. Our principal concern is how easily this content can be discovered for use when general terms related to fitness are used. If a person with a disability spends a disproportionate amount of time searching for accessible videos or cannot successfully identify it, they are not having an equitable experience to that of their abled-bodied peers. Therefore, the existence of disability-friendly content—content that is created for or adaptable to people with disabilities—is not the focus of this study. Instead, the authors are investigating whether disability-friendly content can be easily found using common search terms.

Prior Work

Many studies have been conducted on YouTube videos, but few have analyzed the accessibility of YouTube videos for people with disabilities. Most prior work concerns the accessibility of physical fitness centers for people who are disabled, not the accessibility of online fitness content [20-22]. Thus, to the authors' best knowledge, no studies have analyzed this matter.

Objectives

This study has two primary objectives. The first is to assess how frequently disability-friendly accessibility terms are used in YouTube fitness videos when users search general PA terms. The second primary objective is to determine if there were changes in the accessibility of disability-friendly PA resources on YouTube between before and during the COVID-19 pandemic. The secondary aims are to ascertain if certain types of channels produce more accessible fitness content and to highlight disparities in accessing fitness opportunities on YouTube for people with disabilities, if any exist.

Methods

Video Collection for Study Analysis

A cross-sectional text analysis of exercise- and fitness-related YouTube videos was conducted. Data about videos published from January 1 to June 30, 2019 ("pre-pandemic") and from January 1 to June 30, 202l ("during COVID-19 pandemic"), were collected using the following search PA terms: *at-home exercise*, *exercise at home*, *exercise no equipment*, *home exercise*, *home-based exercise*, *no equipment workout*, and *workout no equipment*. The authors selected these terms to capture broad exercise content (eg, "exercise" and "workout") that could be used without the need for equipment most people only can access at a gym. The authors recognize that other fitness terms could be used, such as terms referring to a specific sport (eg, "basketball"), but this investigation aims to capture the experience of using YouTube for PA for the general public rather than for smaller groups of individuals who play specific sports. General terms related to fitness were searched instead of specific terms related to people with disabilities, as the study's purpose is to determine whether mainstream PA videos include accommodations for people with disabilities. January 2020 was deemed a starting date for COVID-19 videos given that this is when the first lockdown in the world was reported [23].

In lieu of a random sample, we sought to replicate the YouTube video search process to make practical conclusions about the experience of finding accessible videos on YouTube. Prior work suggests that 95% of YouTube traffic is on the first page of search results, which contains 20 to 30 videos [24]. Thus, to establish a rigorous sample size that captures the videos most viewers access, this study collected the first 100 videos that populated on YouTube for each PA search term in both 2019 and 2020. Videos were eligible for inclusion regardless of the country of its creation if the video was created in the English language so that they would be searchable to the study investigators. The authors used Python (version 3.0; Python Software Foundation) to access the YouTube database via its data application programming interface [25]. To deidentify the collected data, YouTube channels and videos were labelled as channel or video as "1, 2, 3..."

Defining "People With Disabilities" and "Disabled Individuals"

These two terms are used interchangeably to reflect a balanced use of disability-friendly language. For the purpose of this study, people with disabilities refers to any individual who self-identifies with a physical, psychological, or intellectual disability. Since this is a preliminary investigation into disability-friendly content on YouTube, the authors are not framing the definition of disabled individuals around a specific diagnosis or criteria. The authors are more concerned with whether the yielded content includes accessibility terms, and less so with whether the people searching for these terms have a particular disability.

Measuring "Accessibility"

Although there may be videos on YouTube that contain disability-friendly content, if they cannot be efficiently found, their utility to people with disabilities diminishes. Therefore, for the purposes of this preliminary assessment of accessible PA videos on YouTube, a video was deemed "accessible" if it was found in the first 100 results from the PA search terms and if its title, tags, or description contained one of the following accessibility terms: *paraparalympic*, *disabled*, *disability*, *differently abled*, *disability-friendly*, *wheelchair-accessible*, *adaptive*, *adapted*, or *inclusive*. These terms were not used in the initial search for PA content since the investigators wanted to ascertain how common accessibility terms are used within commonly searched and viewed PA videos.

The authors acknowledge that videos meeting these criteria still may not be accessible to all users and that additional terms may be appropriate. However, the authors agreed that the aforementioned terms were appropriate and should be analyzed for the following reasons. *Para* or *paralympic* have an implication that combines disability-identifying individuals with sport or fitness [26]. The terms *disabled*, *disability*, *differently abled*, and *disability-friendly* were selected since they are all centered around the word "disability." It should be noted that although the authors will not use the term *differently abled* to refer to people with disabilities since the term is generally opposed within the disability community [27], it is still frequently used and therefore should be searched to better ascertain what terms content generators might use to describe disability-friendly content [28]. The term *wheelchair-accessible* was included since many individuals who are disabled use wheelchairs, and all these individuals face similar challenges in accessing fitness resources [21]. *Adaptive*, *adapted*, or *inclusive* were analyzed because, although as stand-alone terms they do not necessarily denote disability-friendliness, when combined with words associated with PA, the connotation becomes stronger. The concept of inclusive sport and fitness has shaped an association with disability [29,30].

Parameters Collected From Videos

Since video titles and YouTube channel names alone often do not provide comprehensive descriptions of video content, video tags (words or phrases creators choose with which to associate their videos) and video descriptions were also gathered and analyzed. Frequencies of the appearance of "accessible" terms in video tags, descriptions, and titles were recorded.

Data Analysis

Words with the greatest frequencies of appearance in the video titles, tags, and descriptions were collected to assess potential differences in how content generators were describing and

tailoring their videos, and if accessibility terms were among the words with greatest frequency. Other collected metrics were compared between 2019 and 2020 content, including the view counts of the generated videos. Frequent consecutive wording pairs ("bigrams") were compiled within the included video using the tidytext package (version 0.3.1) in R (R Foundation for Statistical Computing) because bigrams can be programmed to remove extraneous words such as "or" and "the" that do not speak to unique context or PA YouTube videos, thereby giving better insight into the video's specific content. For example, the words "Yoga for neck pain" would generate "yoga" and "neck" as one bigram and "neck" and "pain" as another bigram. If a hyperlink or name was generated in bigrams, it was replaced with the annotation "[hyperlink]," "[omitted first name]," or "[omitted last name]."

Categorization of Channels

Each video and channel was categorized based on whether the channel is run by "Hospitals/Medical Institutions," "Individuals," or "Other(s)." Videos published on a hospital's or medical institution's channel were categorized separately from videos created by individuals—a singular person unaffiliated with an established hospital, medical institution, or practice. Remaining videos were placed in the "Other(s)" category. Examples include a certified hospital being placed in the "Hospitals/Medical Institutions" category, a singular person creator being placed in the "Individual" category, and a group creator placed in the "Other" category. This categorization provided greater insight into which content generators are producing the most accessible disability-friendly exercise content on YouTube before and during COVID-19. Attention was given to the "Hospitals/Medical Institutions" channels given that people with disabilities report greater comfort with receiving PA instruction from medical professionals [4]. The creditability of the "Individual(s)" channels, however, cannot be verified with as much ease. Therefore, people with disabilities may be reluctant to use them as frequently for PA. Further, given their expertise with people with disabilities, hospitals and medical institutions might be more inclined to implement accessible terms when posting PA resources during COVID-19 [8,15]. The initial text analysis (eg, search terms such as *exercise no equipment*) and the subsequent text analysis for disability-friendly terms (eg, search terms such as *adaptive*) were compared across the categories including a chi-square test on the change in video generation by "Hospitals/Medical Institutions" channels between 2019 and 2020.

Results

Video titles, descriptions, tags, and transcriptions were collected for a total of 1400 PA videos between 2019 and 2020. The 1400 videos include the first 100 videos to populate for each of the seven PA search terms created in each year (2019 and 2020). Removing duplicate videos resulted in 1038 unique videos (508 in 2019 and 530 in 2020). Viewership in 2020 of content created in 2020 increased significantly when compared to the viewership in 2019 of content created in 2019 with median video view counts of 52,288 (IQR 2891-401,879) and 122,837 (IQR 7257-728,854) for 2019 and 2020, respectively (P=.001).

The analysis revealed that accessible terms applicable to people with disabilities had minimal appearances in 2019 (21 videos) and 2020 (19 videos) among the 1038 unique videos.

Considering the three domains of interest, "Individuals" channels generated more exercise-related videos between January and June 2020 when compared to 2019. The videos created by each category are summarized in Figure 1. Among the top 10 fitness content generators on YouTube with the most views on their videos, none were "Hospitals/Medical Institutions" channels in 2019, and one was a "Hospitals/Medical Institutions" channel in 2020 (Table 1). After accounting for 29 of the 508 (6%) unique videos from 2019, "Hospitals/Medical Institutions" channels generated 60 of the 530 (11%) videos from 2020 (Figure 1). To investigate whether the proportion of the most viewed videos created by "Hospitals/Medical Institutions" channels had increased significantly in 2020, a chi-square test was conducted that revealed a moderate but statistically significant increase (5%) in PA content generated on YouTube by "Hospitals/Medical Institutions" channels (χ^2=8.1476; P=.004).

Figure 1. Histogram of total videos published from January to June in 2019 and 2020.

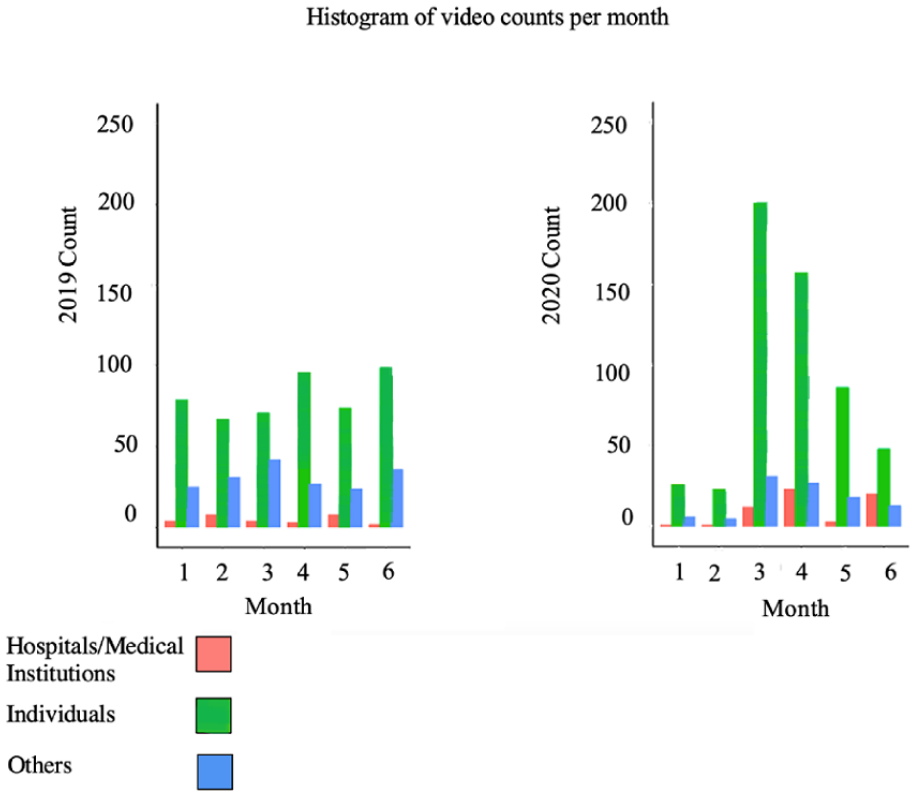


Table 1. Top 10 fitness channels on YouTube in 2019 and 2020 (based on video views).

Deidentified channels (2019)	Channel view count as of June 2020 (rounded to nearest million)	Category (2019)	Deidentified channels (2020)	Channel view count as of June 2020 (rounded to nearest million)	Category (2020)
1	29	Individuals	1	34	Individuals
2	28	Others	2	28	Individuals
3	25	Individuals	3	22	Individuals
4	17	Individuals	4	21	Hospitals/Medical Institutions
5	14	Individuals	5	18	Individuals
6	13	Others	6	15	Individuals
7	13	Individuals	7	13	Individuals
8	13	Others	8	12	Others
9	13	Individuals	9	12	Individuals
10	13	Individuals	10	10	Others

When the PA terms were searched, none of the study's accessible terms populated in the 20 words with the largest aggregate word counts or the top 10 most frequently used word pairs (bigrams) for 2019 and 2020 (Tables 2 and 3). In 2020, two bigrams were tied in 10th place resulting in 11 bigrams being reported.

All the top five viewed videos created by "Hospitals/Medical Institutions," "Individuals," and "Other(s)" channels in 2019 and 2020 were inaccessible (Table 4).

Table 2. Top 20 words with largest aggregate word count of 1400 video descriptions in 2019 and 2020 (excludes filler words).

Year and word	Word appearances, n
2019	
workout	139
home	88
https[a]	46
body	42
exercise	42
video	38
minute	30
5	27
Fat	26
equipment	25
exercises	24
abs	23
10	20
free	20
http	20
hiit	18
min	18
cardio	15
download	15
visit	15
2020	
workout	150
home	109
https	56
body	41
video	35
minute	34
5	32
exercise	32
equipment	27
fat	26
free	26
abs	25
ready	23
hiit	22
min	22
exercises	21
10	20
burn	19
visit	19
join	18

[a]Before tokenization (text parsing), symbols are converted into white space. Accordingly, *http* was kept in the word tally after symbol removal from any embedded link in the video description.

Table 3. Top 10 bigram counts of 1400 video descriptions in 2019 and 2020.

Year and bigram	Bigram appearances, n
2019	
home workout	19
visit https	15
https [hyperlink]	13
body workout	12
home exercise	12
abs workout	11
5 minute	9
10 minute	8
cardio workout	8
[omitted first name] [omitted last name]	8
2020	
home workout	21
visit https	19
https [hyperlink]	18
abs workout	17
[omitted first name] [omitted last name]	14
5 minute	13
join [omitted first name]	13
body workout	12
home exercise	10
10 minute	8
body home	8

Table 4. Top five fitness viewed videos by Hospitals/Medical Institutions, Individuals, and Others in 2019 and 2020.

Year, category, and deidentified video titles	Accessible or not accessible	View count as of June 2020, n
2019		
Hospitals/Medical Institutions		
1	Not accessible	857,433
2	Not accessible	400,604
3	Not accessible	325,860
4	Not accessible	320,632
5	Not accessible	308,392
Individuals		
1	Not accessible	31,972,886
2	Not accessible	23,421,718
3	Not accessible	23,421,342
4	Not accessible	23,418,009
5	Not accessible	23,292,783
Others		
1	Not accessible	29,604,801
2	Not accessible	29,600,357
3	Not accessible	29,414,241
4	Not accessible	12,927,417
5	Not accessible	12,924,144
2020		
Hospitals/Medical Institutions		
1	Not accessible	3,108,295
2	Not accessible	1,708,937
3	Not accessible	58,240
4	Not accessible	44,078
5	Not accessible	14,616
Individuals		
1	Not accessible	24,521,523
2	Not accessible	22,420,881
3	Not accessible	19,890,094
4	Not accessible	19,484,086
5	Not accessible	15,835,194
Others		
1	Not accessible	18,331,311
2	Not accessible	3,589,580
3	Not accessible	2,900,454
4	Not accessible	2,251,818
5	Not accessible	2,174,462

Discussion

Significance of Findings

Although there was a statistically significant increase in the number of videos created by "Hospitals/Medical Institutions" channels in the top viewed videos over the study period, none of the videos by these creators were accessible by our study's definition. The authors created their definition of accessible by considering that *disability* encompasses visible impairments such as amputations and invisible disabilities like chronic pain and disease [31-33]. Before the pandemic, many people with disabilities would have benefitted from digital resources since they struggled to find adequate fitness programs due to barriers such as inaccessible buildings, classes, and equipment, or cost and limited social inclusion [6]. During the pandemic, the need for digital resources was heightened due to lockdowns. The statistically significant increase in views on PA videos on YouTube during COVID-19 reflects the increased dependency on digital resources during the pandemic. The absence of a proportional increase in videos using the study's accessibility terms, however, reinforces the need for YouTube content to be more accessible for people with disabilities and that higher viewership does not necessarily correlate with greater utility.

As full participants in this active social media platform [34], persons with diverse disabilities and ailments could benefit from popular YouTube channels including accommodations for people with disabilities. Although more content curated for or adaptable to people with disabilities may exist, it is unfair that disabled individuals have barriers to accessing these digital resources with comparable ease, especially given that over 1 billion people have a disability globally [35]. YouTube content creators are encouraged to include some accommodations for people with disabilities in their PA videos to make fitness a more inclusive environment.

The analysis of the most frequently used 20 words and word pairs in video descriptions showed no words applicable to people with disabilities, even after reviewing the list for potentially relevant terms not part of the study's accessible terms. This suggests that most of the content created for at-home exercises were either not inclusive of people who are disabled or would be quite difficult for people with disabilities to find. When terms that could be relevant to disabled users are used by creators less frequently, the question of whether disability-friendly content exists becomes less significant than the question of if such content can be found for use. Not being in the first 100 results shown by YouTube means that the content will rarely be accessed because users rarely look beyond the first page of results.

This analysis shows that the standards for giving everyone equal access opportunities are not being met. In this sense, the COVID-19 crisis has further exposed and exacerbated pre-existing social inequities such as disability stigma and ableist attitudes [36-38]. A particularly damaging form of ableism is the reality that people who are disabled are often invisible to mainstream citizens, programs, and policies. Despite the global burden of disability, for instance, even sweeping international policies have been called out for omitting and failing to consider the experiences of people who are disabled [39].

Beyond YouTube, however, it is encouraging that the COVID-19 pandemic has accelerated action in grassroots and international advocacy groups, as they increasingly recognize the imperative need for digital inclusiveness—including with exercise, health, and fitness content. Mooven, an online resource center, was created in response to the stay-at-home orders. With the help of the International Federation of Adapted Physical Activity, Mooven offers guidance and feedback on exercises [40]. Additionally, the nonprofit Inter Campus uses sports to develop resilience in children and help them cope through the pandemic. On the European front, many programs are taking action to adequately prepare trainers to work with people with disabilities [41]. For digital media access, the Universal Fitness Innovation & Transformation organization created a repository of fitness content specifically for people who are disabled and persons with chronic pain [42]. Finally, in regard to overall connectivity, a United Nations Children's Fund (UNICEF) program increases internet connection for children in 11 different countries [43]. These programs' work to increase outreach provides a positive outlook on the increased accessibility of sports.

Limitations and Next Steps

This preliminary assessment of the availability and searchability of disability-friendly fitness videos on YouTube has several limitations that could be addressed through subsequent studies.

The first limitation is selection bias of search terms. The authors sought to select sensible terms that were conducive to both exercise and disability-friendly content. This approach did populate videos focused on specific disabilities. The investigators are aware, however, there are other terms that could fulfill the same purpose and, therefore, all potentially relevant terms were not included in this study. As acknowledgement of this, the investigators consider this study a preliminary study focusing on accessible fitness content on YouTube for disability in general. Similar studies in the future could search for fitness content targeted for patients with specific conditions such as "stroke," "cerebral palsy (CP)," "multiple sclerosis (MS)," and "rheumatoid arthritis" to provide the opportunity to search for more targeted results. Additionally, cross-sectional methodology introduces inherent limitations related to generalizability. To manage this bias, the investigators selected a point in time where there was a reasonably suspected shift in video curation on YouTube given the COVID-19 pandemic.

This study solely sought to assess how a generic YouTube user could find disability-friendly content. In other words, the investigators cannot conclude from this investigation if people with disabilities find the videos useful or the proportion of viewers that people who are disabled represent. A future study that would be merited would be a qualitative study into the experiences of people with disabilities using YouTube as a fitness resource. Given the cited preference among some people with disabilities for content created by relevant experts [8] and the fact that the overall increase in fitness content created by "Hospitals/Medical Institutions" channels on YouTube during

COVID-19 did not translate to a meaningful increase in accessible content, another future study would include analyzing accessibility of disability-friendly videos only created by Hospitals/Medical Institutions. Members of the disability community have previously identified safety as an important variable in their decision to engage in PA [44,45]. Although it is possible for “Individual” or “Other” channels to provide safe PA options for people with disabilities, not all “Individuals” will have experience working with people with disabilities unlike "Hospitals/Medical Institutions." If safety is a key determinant of engagement, people with disabilities may be dissuaded from participating. Thus, understanding the amount of content that "Hospitals/Medical Institutions" channels create can provide further insight into the accessibility of exercise videos available on YouTube.

Another limitation is that the official beginning of lockdowns varied per country. This may have affected the quantity of content created during each month analyzed in this study. Despite this, there were lockdowns as of January 2020, which justifies this choice for the study. A future study could analyze fitness videos created only in certain countries and limit the search to the dates where lockdowns were in place in that country.

Videos analyzed for this project were limited to being in English so they could be searchable to the study investigators. This may have resulted in videos created in different languages that could be useful for people with disabilities being excluded from this study. To get a more comprehensive understanding of all PA videos, a future study could analyze videos created in multiple languages.

Conclusions

This study concludes that current YouTube general fitness content is either lacking in disability-friendly content or the content is not easily accessible for YouTube users who are disabled. Despite COVID-19 galvanizing a broader appreciation for seeking PA digitally, this YouTube analysis found there were no increases in accessible disability-friendly exercise content since the pandemic, including Hospitals/Medical Institutions, which have been a source of trust and engagement for the disability community. Increasing disability-friendly fitness content will be important for improving barriers to digital fitness resources within the disability community in a post–COVID-19 era.

Conflicts of Interest

None declared.

References

1. Physical activity for people with disability. Centers for Disease Control and Prevention. 2020 May 21. URL: https://www.cdc.gov/ncbddd/disabilityandhealth/features/physical-activity-for-all.html [accessed 2021-12-30]
2. Jaarsma E, Smith B. Promoting physical activity for disabled people who are ready to become physically active: a systematic review. Psychol Sport Exerc 2018 Jul;37:205-223. [doi: 10.1016/j.psychsport.2017.08.010]
3. Nikolajsen H, Sandal LF, Juhl CB, Troelsen J, Juul-Kristensen B. Barriers to, and facilitators of, exercising in fitness centres among adults with and without physical disabilities: a scoping review. Int J Environ Res Public Health 2021 Jul 09;18(14):7341 [FREE Full text] [doi: 10.3390/ijerph18147341] [Medline: 34299792]
4. Richardson EV, Smith B, Papathomas A. Disability and the gym: experiences, barriers and facilitators of gym use for individuals with physical disabilities. Disabil Rehabil 2017 Sep;39(19):1950-1957. [doi: 10.1080/09638288.2016.1213893] [Medline: 27626359]
5. Hollis N, Zhang Q, Cyrus A, Courtney-Long E, Watson K, Carroll D. Physical activity types among US adults with mobility disability, Behavioral Risk Factor Surveillance System, 2017. Disabil Health J 2020 Jul;13(3):100888 [FREE Full text] [doi: 10.1016/j.dhjo.2020.100888] [Medline: 32061542]
6. Rimmer JH, Riley B, Wang E, Rauworth A, Jurkowski J. Physical activity participation among persons with disabilities: barriers and facilitators. Am J Prev Med 2004 Jun;26(5):419-425. [doi: 10.1016/j.amepre.2004.02.002] [Medline: 15165658]
7. Shirazipour CH, Evans MB, Leo J, Lithopoulos A, Martin Ginis KA, Latimer-Cheung AE. Program conditions that foster quality physical activity participation experiences for people with a physical disability: a systematic review. Disabil Rehabil 2020 Jan;42(2):147-155. [doi: 10.1080/09638288.2018.1494215] [Medline: 30324815]
8. Trainers for hire. Degree. URL: https://www.degreedeodorant.com/us/trainersforhire/ [accessed 2021-09-24]
9. Su Y, Venkat A, Yadav Y, Puglisi LB, Fodeh SJ. COVID-19 twitter-based analysis reveals differential concerns across areas with socioeconomic disparities. medRxiv. Preprint posted online on November 18, 2020. [doi: 10.1101/2020.11.18.20233973]
10. Conversano C, Di Giuseppe M, Miccoli M, Ciacchini R, Gemignani A, Orrù G. Mindfulness, age and gender as protective factors against psychological distress during COVID-19 pandemic. Front Psychol 2020;11:1900. [doi: 10.3389/fpsyg.2020.01900] [Medline: 33013503]
11. Walker G, McCabe T. Psychological defence mechanisms during the COVID-19 pandemic: a case series. Eur J Psychiatry 2021;35(1):41-45 [FREE Full text] [doi: 10.1016/j.ejpsy.2020.10.005] [Medline: 33191967]
12. Suciu P. Fitness goes to social media during COVID-19 outbreak. Forbes. 2020 Mar 19. URL: https://www.forbes.com/sites/petersuciu/2020/03/19/fitness-goes-to-social-media-during-covid-19-outbreak/?sh=5c18630238ea [accessed 2022-02-02]

13. Dwyer MJ, Pasini M, De Dominicis S, Righi E. Physical activity: benefits and challenges during the COVID-19 pandemic. Scand J Med Sci Sports 2020 Jul;30(7):1291-1294 [FREE Full text] [doi: 10.1111/sms.13710] [Medline: 32542719]
14. Hoffower H. The couple behind a home workout channel with 6 million YouTube followers says they've seen a spike in subscribers amid the coronavirus pandemic, and it shows the effect social distancing is having on people's routines. Business Insider. URL: https://www.businessinsider.com/coronavirus-increase-at-home-workouts-youtube-fitness-blender-2020-3 [accessed 2021-09-24]
15. Trevino KM, Raghunathan N, Latte-Naor S, Polubriaginof FCG, Jensen C, Atkinson TM, et al. Rapid deployment of virtual mind-body interventions during the COVID-19 outbreak: feasibility, acceptability, and implications for future care. Support Care Cancer 2021 Feb;29(2):543-546 [FREE Full text] [doi: 10.1007/s00520-020-05740-2] [Medline: 32902712]
16. McDonough DJ, Helgeson MA, Liu W, Gao Z. Effects of a remote, YouTube-delivered exercise intervention on young adults' physical activity, sedentary behavior, and sleep during the COVID-19 pandemic: Randomized controlled trial. J Sport Health Sci 2021 Jul 24:S2095-2546(21)00091-0 [FREE Full text] [doi: 10.1016/j.jshs.2021.07.009] [Medline: 34314877]
17. Claus E, Feliciano J, Benz L, Calvocoressi L. Social media partnerships with patient organizations for neuro-oncology patient recruitment. Neurooncol Pract 2020 Mar;7(2):143-151 [FREE Full text] [doi: 10.1093/nop/npz049] [Medline: 32626583]
18. 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]
19. Lucini D, Gandolfi C, Antonucci C, Cavagna A, Valzano E, Botta E, et al. #StayHomeStayFit: UNIMI's approach to online healthy lifestyle promotion during the COVID-19 pandemic. Acta Biomed 2020 Sep 07;91(3):e2020037 [FREE Full text] [doi: 10.23750/abm.v91i3.10375] [Medline: 32921731]
20. Johnson M, Stoelzle H, Finco K, Foss S, Carstens K. ADA compliance and accessibility of fitness facilities in western Wisconsin. Top Spinal Cord Inj Rehabil 2012;18(4):340-353 [FREE Full text] [doi: 10.1310/sci1804-340] [Medline: 23459713]
21. Rimmer JH, Padalabalanarayanan S, Malone LA, Mehta T. Fitness facilities still lack accessibility for people with disabilities. Disabil Health J 2017 Apr;10(2):214-221 [FREE Full text] [doi: 10.1016/j.dhjo.2016.12.011] [Medline: 28143707]
22. Calder A, Sole G, Mulligan H. The accessibility of fitness centers for people with disabilities: a systematic review. Disabil Health J 2018 Oct;11(4):525-536. [doi: 10.1016/j.dhjo.2018.04.002] [Medline: 29936142]
23. Wuhan lockdown: a year of China's fight against the Covid pandemic. BBC News. URL: https://www.bbc.com/news/world-asia-china-55628488 [accessed 2021-12-30]
24. Do people scroll past page one on YouTube results? Capital & Growth. URL: https://capitalandgrowth.org/answers/2981187/Do-people-scroll-past-page-one-on-YouTube-results [accessed 2022-02-02]
25. Perrin A, Anderson M. Share of U.S. adults using social media, including Facebook, is mostly unchanged since 2018. Pew Research Center. URL: https://www.pewresearch.org/fact-tank/2019/04/10/share-of-u-s-adults-using-social-media-including-facebook-is-mostly-unchanged-since-2018/ [accessed 2022-02-02]
26. Paralympics history. International Paralympic Committee. URL: https://www.paralympic.org/ipc/history [accessed 2021-12-29]
27. Disability Language Style Guide. National Center on Disability and Journalism. URL: https://ncdj.org/style-guide/ [accessed 2021-12-29]
28. Ahlquist S. Negotiating human engagement and the fixity of computational design: Toward a performative design space for the differently-abled bodymind. International Journal of Architectural Computing 2020 May 18;18(2):174-193. [doi: 10.1177/1478077120919850]
29. Inclusive sports and fitness. Special Olympics Unified Champion Schools. URL: https://media.specialolympics.org/resources/community-building/youth-and-school/unified-champion-schools/ESP6-Inclusive-Sports-Yellow.pdf [accessed 2021-12-30]
30. Inclusive Sport. URL: https://inclusivesport.net [accessed 2022-02-04]
31. Gow MA, Mostert Y, Dreyer L. The promise of equal education not kept: specific learning disabilities - the invisible disability. Afr J Disabil 2020;9:647 [FREE Full text] [doi: 10.4102/ajod.v9i0.647] [Medline: 32158643]
32. Visagie S, Swartz L. "There is nothing wrong with me": disability invisibility in a rural South African town. Disabil Rehabil 2018 Jul;40(15):1799-1807. [doi: 10.1080/09638288.2017.1313909] [Medline: 28399720]
33. Clemens L, Langdon D. How does cognition relate to employment in multiple sclerosis? A systematic review. Mult Scler Relat Disord 2018 Nov;26:183-191. [doi: 10.1016/j.msard.2018.09.018] [Medline: 30268039]
34. Shelton K. The value Of search results rankings. Forbes. URL: https://www.forbes.com/sites/forbesagencycouncil/2017/10/30/the-value-of-search-results-rankings/?sh=7aae9f9d44d3 [accessed 2022-02-02]
35. Disability and health. WHO. 2021 Nov 24. URL: https://www.who.int/news-room/fact-sheets/detail/disability-and-health [accessed 2022-02-04]
36. Krouse HJ. COVID-19 and the widening gap in health inequity. Otolaryngol Head Neck Surg 2020 Jul;163(1):65-66. [doi: 10.1177/0194599820926463] [Medline: 32366172]
37. Chakraborty J. Social inequities in the distribution of COVID-19: an intra-categorical analysis of people with disabilities in the U.S. Disabil Health J 2021 Jan;14(1):101007 [FREE Full text] [doi: 10.1016/j.dhjo.2020.101007] [Medline: 32981853]

38. Bambra C, Riordan R, Ford J, Matthews F. The COVID-19 pandemic and health inequalities. J Epidemiol Community Health 2020 Nov;74(11):964-968 [FREE Full text] [doi: 10.1136/jech-2020-214401] [Medline: 32535550]
39. Wolbring G, Burke B. Reflecting on education for sustainable development throught two lenses: ability studies and disability studies. Sustainability 2013 May 24;5(6):2327-2342. [doi: 10.3390/su5062327]
40. Boursier C. Mooven APA video feedback. International Federation of Adapted Physical Activity. 2020 Mar 27. URL: https://ifapa.net/mooven-apa-video-feedback/ [accessed 2022-02-02]
41. Carty C. Project information sheet European fitness inclusion training for work. European Observatoire of Sport and Employment. 2013. URL: http://eose.org/wp-content/uploads/2013/10/EIFTW-Project-Information-Sheet.pdf [accessed 2022-02-02]
42. UFIT homexercise. UFIT. URL: https://justdoufit.com/ufit-homexercise/ [accessed 2022-02-02]
43. Giga. UNICEF. URL: https://www.unicef.org/innovation/giga [accessed 2022-02-02]
44. Salt E, Wiggins AT, Hooker Q, Crofford L, Rayens MK, Segerstrom S. The effects of pain severity, pain catastrophizing, depression, and exercise on perceived disability in acute low back pain patients. Res Theory Nurs Pract 2018 Nov;32(4):436-448. [doi: 10.1891/1541-6577.32.4.436] [Medline: 30567914]
45. Reynolds ER, Ashbaugh AD, Hockenberry BJ, McGrew CA. Multiple sclerosis and exercise: a literature review. Curr Sports Med Rep 2018 Jan;17(1):31-35. [doi: 10.1249/JSR.0000000000000446] [Medline: 29315107]

Abbreviations

CP: cerebral palsy
MS: multiple sclerosis
PA: physical activity
UNICEF: United Nations Children's Fund

Edited by A Mavragani; submitted 09.10.21; peer-reviewed by R Shimkhada, L Herrmann; comments to author 23.11.21; revised version received 31.12.21; accepted 14.01.22; published 15.02.22.

Please cite as:
Kadakia S, Stratton C, Wu Y, Feliciano J, Tuakli-Wosornu YA
The Accessibility of YouTube Fitness Videos for Individuals Who Are Disabled Before and During the COVID-19 Pandemic: Preliminary Application of a Text Analytics Approach
JMIR Form Res 2022;6(2):e34176
URL: https://formative.jmir.org/2022/2/e34176
doi:10.2196/34176
PMID:35044305

Original Paper

A Brief, Daily, Online Mental Health and Well-being Intervention for University Staff During the COVID-19 Pandemic: Program Description and Outcomes Using a Mixed Methods Design

Alexandra Parker[1,2], BA, MPsych, PhD; Sarah Dash[1], PhD; Matthew Bourke[1], BSc; Rhiannon Patten[1], MSc, PhD; Melinda Craike[1,3], PhD; Peter Baldwin[1,4], BA, MPsych, PhD; Warwick Hosking[1], BA, PhD; Itamar Levinger[1,5], BSc, MSc, PhD; Vasso Apostolopoulos[1], BSc, PhD; Maximilian de Courten[3], MPH, MD; Jenny Sharples[1], MPsych, PhD; Monika Naslund[1], MPsych; Vasileios Stavropoulos[1], MA, PhD; Mary Woessner[1], BSc, MSc, PhD; Christopher Sonn[1], PhD; Caroline Stansen[1], BA, MPsych; Michaela Pascoe[1,6], BSc, MSc, PhD

[1]Institute for Health and Sport, Victoria University, Melbourne, Australia

[2]Orygen, Centre for Youth Mental Health, University of Melbourne, Melbourne, Australia

[3]Mitchell Institute, Victoria University, Melbourne, Australia

[4]Black Dog Institute, Sydney, Australia

[5]Australian Institute for Musculoskeletal Science, Department of Medicine, Western Health, Melbourne Medical School, University of Melbourne, Melbourne, Australia

[6]Peter MacCallum Cancer Centre, Melbourne, Australia

Corresponding Author:
Alexandra Parker, BA, MPsych, PhD
Institute for Health and Sport
Victoria University
Ballarat Road
Melbourne, 3011
Australia
Phone: 61 399195874
Email: alex.parker@vu.edu.au

Abstract

Background: The unprecedented changes and isolation measures to contain COVID-19 have had multiple psychological and social impacts, with implications for professional and personal functioning. Evidence-informed interventions that can be rapidly implemented under pandemic conditions to support mental health during such times are urgently needed.

Objective: The aim of this study was to determine the acceptability and preliminary outcomes of a daily online mental health promotion program for tertiary education staff during the COVID-19 pandemic.

Methods: The "Victoria University (VU) Elevenses" program was delivered as an uncontrolled intervention at Victoria University (VU) in the western metropolitan region of Melbourne, Australia. In April 2020, an email invitation was sent to all academic and professional staff inviting them to: (1) participate in the program and (2) opt-in to the research component. The "VU Elevenses" program provided 10-15–minute microinterventions comprising lifestyle and well-being strategies to promote mental health via an online meeting platform at 11 AM each weekday. A mixed methods approach was used to evaluate the program, combining structured questionnaires with semistructured interviews to investigate the experiences of staff who participated in the program.

Results: Between 16 and 90 participants provided weekly program feedback. A total of 106 university staff opted into the longitudinal research component and 10 staff participated in the interviews. Participants reported high levels of satisfaction with sessions and perceived benefits for mental health. Approximately one quarter of participants reported moderate to severe symptoms of depression, anxiety, and stress at baseline, with significant reductions in these symptoms in the first 7 weeks of the program, corresponding with easing in mandatory isolation ("lockdown") restrictions. Symptoms of depression, anxiety, and stress all increased when lockdown measures were reintroduced, but not to the same levels as found during the initial lockdown period. Overall changes in depression and anxiety from baseline to the end of the program were explained by changes in COVID-19–related distress, whereas changes in self-compassion explained changes in stress.

Conclusions: We show that it is feasible and acceptable to develop and deliver a program of brief interventions in a timely manner, using a simple and accessible online platform. Although participation in the program was initially associated with reduced symptoms of depression, anxiety, and stress, participants' mental health worsened with the reintroduction of a "lockdown" period. However, as symptoms of depression, anxiety, and stress did not return to levels observed at the start of the VU Elevenses program, participation in the uncontrolled intervention may have offered a protective benefit against the impact of the second significant lockdown period.

(JMIR Form Res 2022;6(2):e35776) doi:10.2196/35776

KEYWORDS

workplace mental health; well-being; mental health promotion; online intervention; telehealth; COVID-19 pandemic; COVID-19; pandemic; health promotion

Introduction

The Australian national response to the global pandemic of COVID-19 has been consistent with international approaches recommending or mandating physical distancing and self-isolation. Psychological and social risks associated with this period of rapid change and uncertainty included the inability to access workplaces; educational institutions; and social, cultural, and sporting events. Psychological stressors include concerns about income, job insecurity, and changes to work practices (eg, remote working, reduced hours, or loss of job), with nearly one third of Australians reporting worsening household finances due to the COVID-19 pandemic during 2020 [1]. Additional health-related anxiety for self and loved ones, the complexity of caring responsibilities (including remote learning/schooling from home), potentially unsafe home environments, and the impacts of social isolation [2] contributed to significant increases in rates of psychological distress compared with Australian national data from 2017 to 2018 [1,3]. Findings from a survey from the beginning of the pandemic of almost 1500 Australian adults indicated that more than 1 in 4 respondents reported depression scores that were moderately to extremely severe, and approximately 1 in 5 had moderately to extremely severe stress during the period of peak COVID-19 self-isolation and physical distancing requirements [4].

Structure and consistency in everyday life are important in preserving mental well-being during enforced periods of self-isolation and uncertainty [5]. Using helpful coping strategies, maintaining daily structure, and staying connected were consistently endorsed by national (eg, Beyond Blue [6]) and international (eg, [5]) health organizations as critical to maintaining mental well-being during the COVID-19 crisis. Further, the delivery of psychological interventions and interactions via online meeting platforms [7], and brief, targeted interventions for well-being have been shown to be effective in promoting mental health in clinical and nonclinical adult populations [8-10]. To promote and maintain physical and mental well-being for staff of a tertiary education institution during the COVID-19 pandemic, we delivered an evidence-informed, timely, accessible, responsive online intervention comprising brief, daily (weekday) microinterventions and strategies targeting six essential lifestyle areas for well-being (healthy eating, physical activity, reducing alcohol intake, improving sleep, healthy relationships and social connection, and stress management) [11]. In deference to the warmth, comfort, and connection shared over a cup of tea and sticky bun (ie, morning tea referred to as "elevenses") between Paddington Bear and his dear friend, Mr Gruber [12], our program was titled the "Victoria University (VU) Elevenses." The aim of this study was to determine the acceptability of the program; assess changes in stress, depressive, and anxiety symptoms in program participants over time; and determine mediators of change in symptoms of stress, depression, and anxiety from baseline to the endpoint of the program.

Methods

Setting

The study was conducted at Victoria University (VU) in the western metropolitan region of Melbourne, Australia. VU has eight Victorian campuses, with 1994 academic (teaching and research) and professional staff, 988 sessional staff (academic and polytechnic teachers), and 598 casual professional staff. Among the total VU staff, 59.64% (2135/3580) identify as female.

Participants

All staff received an email invitation to participate in the online program and to opt-in to the research component. All staff were eligible to participate in the program, with no exclusion criteria or minimum attendance requirements. Using the general university mailing list, invitations to join the daily online sessions were issued over the first 2 weeks of the program, with staff asked to opt-in via email to receive ongoing meeting invitations. At the end of the program, all those who opted in to meeting invitations (17.93%, 642/3580) were invited by email to participate in the semistructured interviews.

Ethics Approval

All procedures involving human participants were approved by the Victoria University Human Research Ethics Committee (HRE20-054) and complied with the ethical standards of this institutional committee.

Research Design

This study followed a mixed methods design. A repeated-measures within-subjects design was used to collect data on mental health outcomes and potential mediators of change from baseline to 32 weeks after program commencement. Initially, it was planned that follow-up assessments would occur every 3 weeks. However, given that the program ran longer than

anticipated due to Melbourne's extended lockdown period, the repeated-measures time points became longer as the program progressed to reduce participant burden. Additionally, brief weekly questionnaires were used to determine the acceptability of the program over the first 13 weeks of the program. At the end of the program, we also conducted in-depth semistructured interviews in a small sample of participants to explore their experiences with VU Elevenses.

Intervention Program

The "VU Elevenses" program focused on mental health promotion, including universal, selected, and indicated prevention strategies [13]. Participants were provided information on how to access mental health treatment from existing internal (ie, Employee Assistance Program) and external (ie, Beyond Blue, headspace, Black Dog Institute, Lifeline, Relationships Australia, Family Relationships Advice Line) sources, as the "VU Elevenses" program was not designed to be a treatment for mental health conditions. The program was aligned with the VU Employee Wellbeing Policy [14].

The "VU Elevenses" program comprised brief (10-15 minutes), evidence-informed microinterventions and strategies to promote physical and mental well-being, utilizing an online meeting platform at 11 AM each weekday. The intervention had three main phases: (1) managing immediate concerns and stressors, (2) adjusting to working and studying remotely, and (3) preparing to return to work and study on campus. The design and implementation of the program were guided by the principles of inclusivity and accessibility, responsiveness, consistency, and connectedness. The sessions were live and unpolished to maintain authenticity and connection, and to allow for the rapid sharing of evidence-based clinical content for mental health and well-being support.

A collective of VU practitioners and researchers rotated and delivered interventions relevant to one of six identified lifestyle areas (healthy eating, physical activity, reducing alcohol intake, improving sleep, healthy relationships and social connection, and stress management) [15]. Each microintervention aimed to promote skill-building through simple mindfulness strategies, deep breathing exercises, relaxation exercises, time-management and routine-setting strategies, self-compassion strategies, physical activity guidance, sleep tips, nutrition advice, and fun activities for community connection (eg, quizzes and group singing sessions; see Table 1).

Members of the program team planned the following week of microintervention content, resources, and presenters. Resources for further information (ie, fact sheets, websites, apps, YouTube clips, online interventions) and links to recordings were provided to participants for each daily session via VU's intranet. The program was evaluated for 32 weeks, from April 2020 to November 2020, with sessions each weekday for the first 24 weeks, followed by three times per week for the final 8 weeks of the program.

Table 1. Session content overview for the first three weeks of the "Victoria University (VU) Elevenses" program.

Day[a]	Lifestyle intervention theme	Content	Presenter	Resources shared
1	Stress management	Introduction to "VU Elevenses," psychoeducation on sympathetic and parasympathetic nervous systems, deep breathing exercise (box breathing)	Clinical psychologist/research academic	Links to breathing apps, box breathing jpeg
2	Stress management	Simple mindfulness strategies (Notice 5 Things and Leaves on a Stream)	Clinical psychologist/research academic	Links to mindfulness and meditation apps
3	Relationships	Managing roles and responsibilities	Clinical psychologist/teaching and research academic	Links to resources for maintaining mental health and relationships when working from home (ie, Lifeline and Relationships Australia fact sheets)
4	Physical activity	Stretching exercises for shoulder, neck, and back	Accredited exercise physiologist/research academic	Links to web pages of accredited organizations for home-based stretching exercises
5	Relationships (community connection)	Online quiz (1908s music)	Clinical psychologist/research academic	None
6	Relationships	Schooling from home	Clinical and community psychologist/teaching academic	Links to popular media articles with expert opinions on challenges of remote education
7	Stress management	Resetting workload expectations	Clinical psychologist/teaching and research academic	Links to popular media articles with expert opinions on academic expectations during COVID-19
8	Stress management	Self-compassion meditation	Probationary psychologist/teaching and research academic	Links to website for additional self-compassion resources
9	Physical activity	Exercise "snacking": 5 minutes of light aerobic activities	Accredited exercise physiologist/research academic	Link to YouTube clip of 5-minute cardio work out
10	Relationships (community connection)	Online quiz (general knowledge)	Clinical psychologist/research academic	None
11	Stress management	Progressive muscle relaxation	Clinical psychologist/research academic	Link to online audio clip of recorded progressive muscle relaxation
12	Sleep	Guidelines and tips to improve sleep quality	Teaching and research academic	Links to sleep guidelines and national Sleep Health Foundation
13	Stress management	Managing time and work commitments during change	Clinical psychologist/research academic	Links to sprint technique ("Pomodoro") and popular media article on time management
14	Physical activity	Exercise "snacking": 10-15 minutes of resistance activities	Accredited exercise physiologist/research academic	Links to web pages of accredited organizations for home-based stretching exercises and YouTube home workout clip
15	Relationships (community connection)	Introduction to Indigenous mindfulness practices and connection to country	Indigenous Community Liaison Officer and licensed Wayapa practitioner	Links to Wayapa and YouTube clip from Victorian Aboriginal Heritage Council

[a]Days 16-20 included healthy eating (nutrition and mental health), "desk yoga," pleasant activity scheduling, exercise "snacking" (light aerobic activities), and a music sing-along.

Outcomes

Program Acceptability

All participants were provided with a link to a brief survey at the end of each week for the first 13 weeks of the intervention. Participants were asked to provide weekly feedback on the acceptability of the intervention by answering the following questions on a scale from 1 (strongly disagree) to 7 (strongly agree): (1) "I am generally satisfied with the sessions this week," (2) "I found the content helpful for my mental health and well-being," and (3) "I felt more connected with my colleagues." To measure engagement, each respondent was asked to report the number of live sessions they attended that week. Attendance at the daily sessions was assessed at week 1, week 16, and week 32 using the online meeting platform's participant logs.

Participant Experiences

One-on-one semistructured interviews were conducted with a sample (n=10) of VU staff members who attended at least one VU Elevenses session, as previous research shows that data saturation for the most part occurs between 6 and 12 interviews [16]. Participants were prompted about their level of engagement, their overall perceptions of the program, and how useful they found the program. The full interview schedule can be seen in Multimedia Appendix 1.

Mental Health, Well-being, and Health Behaviors

The mental health of participants was measured using the Short Form Depression, Anxiety and Stress Scale (DASS-21) [17]. The Pittsburgh Sleep Quality Index [18] was used to measure sleep quality; the WHO Alcohol, Smoking & Substance Involvement Screening Test [19] was used to measure alcohol intake; the Coping Scale [20] and Self-Compassion Short Form (SCS-SF) [21] were used to measure stress management; the Multidimensional Scale of Perceived Social Support (MSPSS) [22] and Social Connectedness and the Social Assurance Scales (SCS) [23] were used to measure healthy relationships and social connection; the Starting the Conversation Diet Instrument [24] was used to measure diet quality; and the Active Australia Survey was used to measure current leisure-time physical activity [25]. Additionally, COVID-19–specific distress was measured using the Coronavirus Anxiety Scale [26].

Analysis of Acceptability and Participants' Experiences

To determine acceptability of the program, percentages of participants who agreed or strongly agreed they were satisfied with the content of the session, felt the content helped with their mental health and well-being, and thought the program made them feel more connected with their colleagues during the first 13 weeks of the program were calculated. Additionally, the average number of live daily sessions attended by respondents in a week was calculated.

The semistructured interviews were digitally recorded and then transcribed verbatim. Transcripts were read and reread while listening to the recordings to ensure accuracy. Braun and Clarke's [27] thematic analysis guide was used to examine the personal experiences and meanings of the program. Data were organized in a meaningful and systematic way, with coding used to reduce data into smaller chunks of meaning. Codes were examined and organized into broader themes and were reviewed and modified. The final themes were defined and named.

Analyses of Mental Health Outcomes

Quantitative data analysis was performed using SPSS (v.25) and the R package. Trajectories of change in symptoms of depression, anxiety, and stress were estimated using multilevel growth curve models. First, an unconditional model with a random intercept was run for each outcome (null model). Subsequently, a series of models were run with a fixed linear effect of time (Model A), quadratic effect of time (Model B), and cubic effect of time (Model C) added. Because the time between follow-ups differed, time was entered as the number of weeks from baseline centered on the date of baseline collection. All models were fitted with a random intercept. Additionally, models were run with random slopes. However, the inclusion of random slopes did not significantly improve the fit of any model; thus, the results are reported for models with a random intercept only. The likelihood ratio test was used to compare model fits and determine which model fit the data best. Before analysis, models were checked for normality. The results indicated that depression and anxiety were not normally distributed, and therefore these outcomes were log-transformed before analysis. Because logs of zeroes cannot be computed, 1 was added to all scores of depression and anxiety before transformations were made.

To determine mediators of changes in depression symptoms from baseline to the end of the program, 1-1-1 multilevel multiple mediation models were estimated using the MLmed macro tool [28]. Models were estimated with random intercepts, fixed slopes, and uncorrelated residuals. The significance of indirect effects was estimated via Monte Carlo 95% CIs based on 10,000 resamples. Variables for each of the six identified lifestyle areas (healthy eating, physical activity, reducing alcohol intake, improving sleep, healthy relationships and social connection, and stress management) were tested as possible mediators. Additionally, to account for potential changes due to fluctuating responses to COVID-19, COVID-19–related distress was also tested as a potential mediator. To keep models parsimonious, the three variables with the strongest within-person correlations [29] with each of the outcomes were included in a single multiple mediation model. Within-person correlations between variables are provided in Multimedia Appendix 2.

Missing Data

There were large amounts of missing data. Only six participants provided data at the fourth follow-up, and therefore data from this time point were excluded from analyses. Of the remaining five assessments, 43 participants only completed the baseline assessment, 18 participants completed two assessments, 8 participants completed three assessments, 15 participants completed four assessments, and 20 participants completed all five assessments. Data were available for 106 participants at baseline, 51 at the first follow-up, 42 at the second follow-up, 36 at the third follow-up, and 31 at the fifth follow-up. One-way analysis of variance and Kruskal–Wallis tests were used to determine if there was a significant difference in study variables at baseline between participants with complete data and those missing data at follow-up. Missing value analysis indicated no significant differences in symptoms of depression, anxiety, or stress, or in demographics between participants with different amounts of missing data, indicating that the data were missing completely at random. Missing data for symptoms of depression, anxiety, and stress were handled using the full information maximum-likelihood estimation method. This has been shown to reduce parameter errors compared to ad hoc approaches to handling missing data, and provides appropriate standard errors of estimations, even with substantial amounts of missing data and small sample sizes [30,31].

For the mediation analysis, 31 participants completed both the baseline and follow-up assessments. Results indicated that participants who completed both baseline and follow-up assessments reported significantly less COVID-19–related

distress at baseline than participants who did not complete the follow-up assessment, indicating that the data were not missing completely at random. Because multilevel models cannot handle missing data for independent variables using maximum likelihood, missing values for mediator variables were imputed using multiple imputation. A total of 10 imputed data sets were created using Markov Chain Monte Carlo with predictive mean matching. Imputation models included all mediators and outcomes as predictors of missing values. Results from the multiple imputed data sets were combined using Rubin's Rule [32].

Results

Program Acceptability

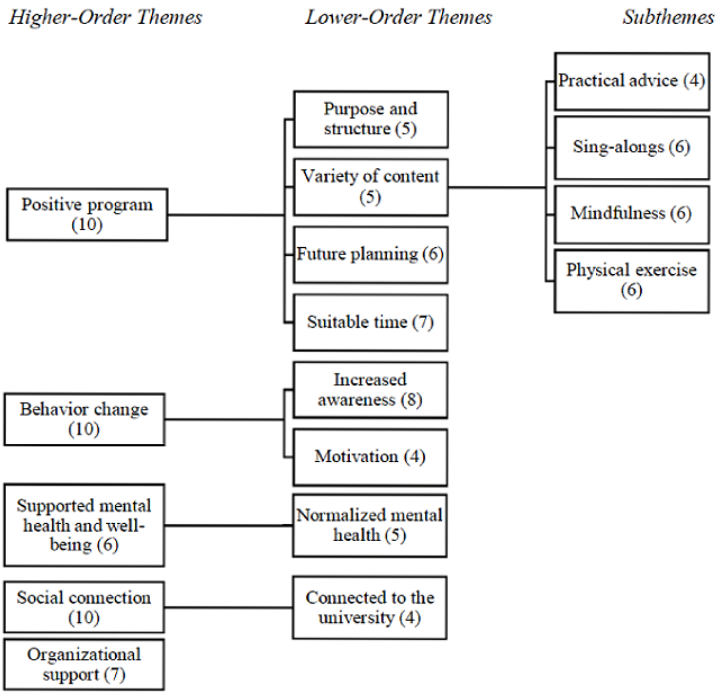
Attendance at the daily sessions ranged from 174 to 240 participants in the first week of the program, 63-84 at week 16, and 21-24 in the final week of the program. The number of responses to weekly feedback ranged from 16 to 90 participants. The average weekly attendance for participants who responded to the weekly feedback questionnaires ranged between 2.38 and 3.22 sessions. Participants reported high levels of satisfaction with the sessions during the first 13 weeks of the study (81.0% and 97.7% of weekly respondents agreeing or strongly agreeing that they were satisfied with the sessions they attended that week), and found the sessions they attended helpful for their mental health and well-being (80.9% and 94.2% of weekly respondents agreeing or strongly agreeing). Increased connection with colleagues was moderately endorsed (46.1% and 94.1% of weekly respondents agreed or strongly agreed).

Program Experiences

Overview of Themes

Participants in the semistructured interviews were aged, on average, 54 years (range 29-69 years), were predominantly professional staff (90%), and predominantly female (90%). Thematic analysis identified five higher-order themes: a positive program, facilitating behavior change, supporting mental health and well-being, providing social connection, and organizational support (Figure 1). Exploration of the higher-order themes, lower-order themes, and subthemes formed the basis of the findings.

Figure 1. Themes from semistructured interviews with participants.



A Positive Program

Four lower-order themes emerged from the higher-order theme of a positive program: (1) provided purpose and structure, (2) variety of relevant content, (3) future planning, and (4) suitable time. The VU Elevenses program was reported as a positive experience for participants during the COVID-19 pandemic. The program provided structure and purpose by supporting participants to have something to look forward to during extended periods of remote working and mandated isolation. The 11 AM time was suitable and the 15-minute duration was sufficient. Further, the program provided a variety of content, ranging from practical advice, sing-a-longs, mindfulness, and physical activities. This variety was reported as positive, as participants could opt in and out of the program depending on what suited their needs. Participants reported that the mix of content kept the program refreshing and engaging. Further, participants reported that they wanted the program to continue after the immediate response to the COVID-19 pandemic.

> *… It's been mind-blowing. It's been inspiring. It's been educational. It's taught me strategies. It's given me physical, mental, and different perspectives to think about how to do things better during this*

> *lockdown COVID…it's the best thing since sliced bread. It should continue. It should have happened sooner. So I think COVID has really pushed us to thinking, doing, thinking better in how we plan our days and to take care of ourselves* [Participant 6]
>
> *…bringing us back and reminding us to set some structure, have things to look forward to in the day, schedule in things that you want, and not just leave them to chance* [Participant 3]
>
> *I think there is so much value and benefit of what has happened with Elevenses that could continue on beyond, you know, what we are facing right now* [Participant 4]

Facilitating Behavior Change

Two lower-order themes emerged from the higher-order theme of facilitating behavior change: (1) increased awareness and (2) motivation. All participants described the VU Elevenses program as contributing to changing well-being behaviors during the COVID-19 pandemic. For example, some participants changed how they engaged with news and social media or incorporated breathing exercises into their daily lives. Participants felt the program created awareness and motivation for change, including regular prompts and reminders of strategies and skills they could implement, which facilitated implementing these changes.

> *…one of the strategies was about minimizing access to social media, particularly around the news. Not constantly looking at the news 'cause it is what it is. And so I think that was one that I found really useful to just think about where am I getting that information… finding one or two [trusted] sources… and that helped* [Participant 2]
>
> *…without really consciously thinking about making behavior change, I think that just having that kind of reminder was helpful* [Participant 1]
>
> *…overall I think the Elevenses, was very motivational. Every time I went into it, I left feeling very positive. So if there was a day where I felt unmotivated and I felt a bit like, "Ah, I can't be bothered," when I'd go into the session, I'd leave feeling a bit more motivated* [Participant 4]

Supporting Mental Health and Well-being

One lower-order theme emerged from the higher-order theme of supporting mental health and well-being: normalizing mental health. Participants discussed how the program provided them with more confidence, insight, and understanding in managing their mental health and well-being during the COVID-19 pandemic. The online group program appeared to normalize any difficulties they were experiencing during the extended lockdown.

> *… my mental health has definitely adjusted and moved along the journey of the pandemic. So again, I think the program was very attuned to the situation and supporting mental health... I feel strongly that it did promote wellbeing and mental health in a very accessible way* [Participant 3]
>
> *… I think the breathing one, I sort of kept in my mind when I was feeling stressed, to do that. And then just the recognition that this has been identified as something that maybe other people might be going through as well helped* [Participant 8]

Providing Social Connection

One lower-order theme emerged from the higher-order theme of social connection: connected to the university. Participants reported that the program offered a space for staff to connect socially every day. Further, participants discussed feeling connected to the university and staff during the program, which seemed to provide a sense of comfort and familiarity.

> *… and it was also really great to see some of my colleagues that I haven't been, physically been able to connect with, this made things really much better…* [Participant 1]
>
> *…made me feel a little bit more connected to the wider university because it made me feel more connected to what people were doing within the university* [Participant 8]

Organizational Support

Having access to the daily VU Elevenses program and feeling valued at work appeared to be beneficial for participants during the COVID-19 pandemic: "…that sense of connection to the organization and feeling like the organization actually values the staff has been the key thing for me." [Participant 6]

Longitudinal Changes in Depression, Anxiety, and Stress

A total of 106 participants completed the baseline questionnaire. At baseline, the average age of the 106 participants was 47.34 (SD 11.82) years and 92 (86.8%) identified as female. Overall, 79 (74.5%) participants were administrative or executive staff and 26 (24.5%) were teaching or research staff. The majority (70/106, 66.0%) were full-time staff. At baseline, 22.9%, 15.2%, and 25.2% of participants reported moderate-to-severe symptoms of depression, anxiety, and stress, respectively.

Results from the multilevel linear growth models are shown in Table 2. The results showed that symptoms of depression, anxiety, and stress all followed a cubic trajectory decreasing from baseline to June 8, increasing from June 8 to July 31 (corresponding with a return to strict lockdown), and decreasing from July 31 to November 27 as lockdown restrictions were eased (Figures 2-4).

Table 2. Results from multilevel growth curve models for depression, anxiety, and stress.

Effects	Depression[a]	Anxiety[a]	Stress
Model A			
Fixed effect: time			
B (95% CI)	–0.003 (–0.009 to 0.003)	–0.010 (–0.016 to –0.004)	–0.081 (–0.116 to 0.047)
P value	.39	.002	<.001
Model fit			
–2LL	569.179	546.004	1408.421
Likelihood ratio (null model), χ^2 (*df*=1)	0.752	9.501	20.581
P value	.39	.002	<.001
Model B			
Fixed Effects			
Time			
B (95% CI)	–0.017 (–0.039 to 0.006)	–0.019 (–0.041 to 0.004)	–0.154 (–0.278 to –0.031)
P value	.15	.02	.02
Time^2			
B (95% CI)	0.000 (–0.000 to 0.001)	0.000 (–0.000 to 0.001)	0.002 (–0.001 to 0.005)
P value	.21	.42	.23
Model fit			
–2LL	567.637	545.385	1406.950
Likelihood ratio (Model A), χ^2 (*df*=1)	1.542	0.619	1.471
P value	.21	.43	.23
Model C			
Fixed effects			
Time			
B (95% CI)	–0.137 (–0.194 to –0.080)	–0.090 (–0.149 to –0.031)	–0.678 (–1.000 to –0.356)
P value	<.001	.003	<.001
Time^2			
B (95% CI)	0.013 (0.007 to 0.018)	0.007 (0.002 to 0.013)	0.055 (0.024 to 0.085)
P value	<.001	.009	<.001
Time^3			
B (95% CI)	–0.000 (–0.000, to –0.000)	–0.000 (–0.000 to –0.000)	–0.001 (–0.001 to –0.000)
P value	<.001	.01	.001
Model fit			
–2LL	548.578	538.925	1395.580
Likelihood ratio (Model A), χ^2 (*df*=2)	20.601	7.079	12.841
P value	<.001	.03	.002
Likelihood ratio (Model B), χ^2 (*df*=1)	19.059	6.460	11.370
P value	<.001	.006	<.001

[a]Outcomes are log-transformed.

Figure 2. Change in depression scores during the program. The subscale indicates if outcomes were log-transformed (ln). Higher subscale numbers equate to greater symptoms. Time points shaded in grey indicate data collection periods during which participants were in strict lockdowns. DASS-21: Short Form Depression, Anxiety and Stress Scale.

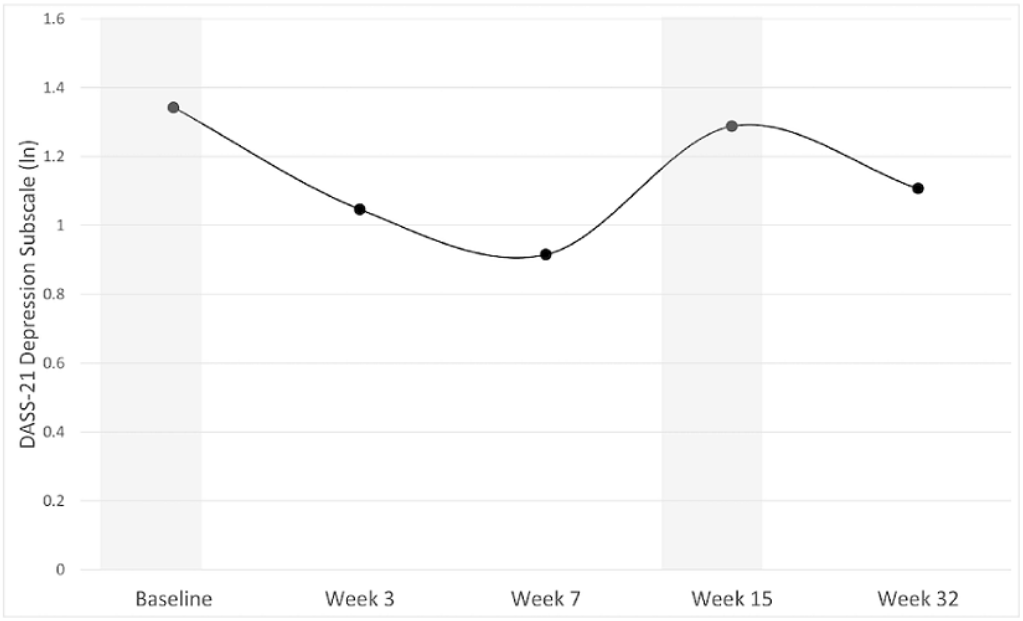


Figure 3. Change in anxiety scores during the program. The subscale indicates if outcomes were log-transformed (ln). Higher subscale numbers equate to greater symptoms. Time points shaded in grey indicate data collection periods during which participants were in strict lockdowns. DASS-21: Short Form Depression, Anxiety and Stress Scale.

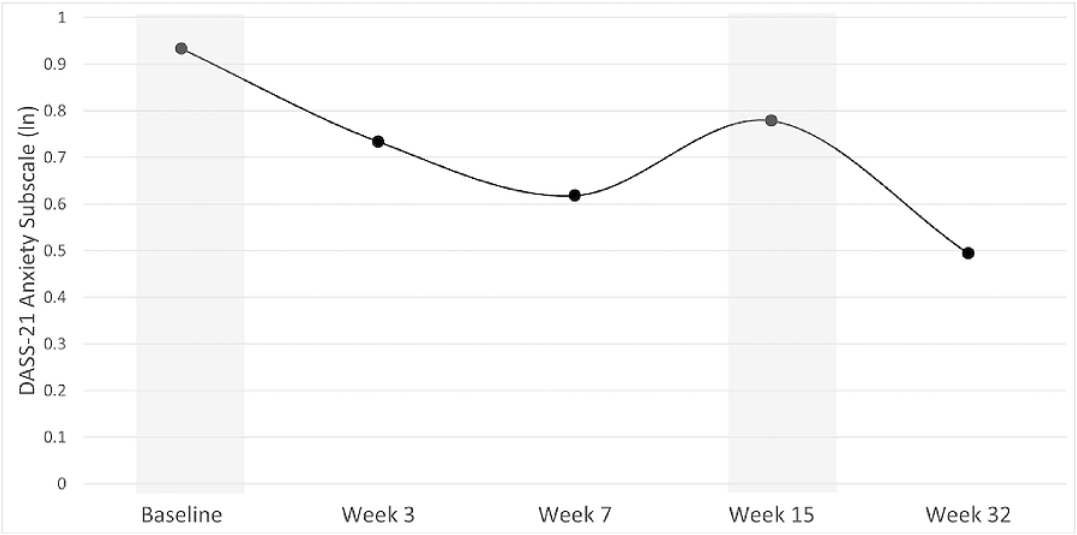
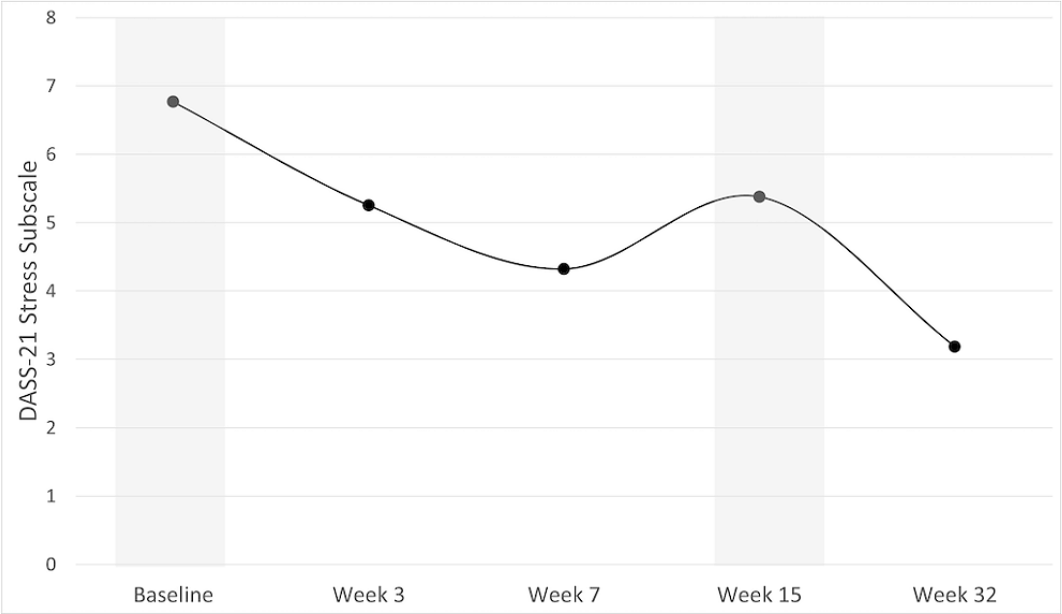


Figure 4. Change in stress scores during the program. DASS-21: Short Form Depression, Anxiety and Stress Scale.



Protective Factors Against Poor Mental Health

Results from the repeated-measures correlation (Multimedia Appendix 2) showed that COVID-19–related distress, self-compassion, and social connectedness had the strongest within-person correlations with symptoms of depression. The results from the mediation model for depression are presented in Table 3, showing a significant negative indirect effect of time through COVID-19–related distress on depression. There were no significant indirect effects through self-compassion or social connectedness. When controlling for mediators, there was no direct effect of time on symptoms of depression, indicating that changes in symptoms of depression were mediated by changes in COVID-19–related distress.

Table 3. Results from the multilevel mediation model on depression.

Effects	Path	B (95% CI)
Time through COVID-related distress	a1	–0.206 (–0.285 to –0.127)
COVID-related distress through depression	b1	6.228 (1.213 to 11.242)
Indirect effect	b1×a1	–1.299 (–2.558 to –0.040)
Time through self-compassion	a2	4.297 (2.228 to 6.366)
Self-compassion through depression	b2	–0.304 (–0.590 to –0.018)
Indirect effect	b2×a2	–1.284 (–2.579 to 0.011)
Time through social connectedness	a3	3.686 (0.755 to 6.617)
Social connectedness through depression	b3	0.001 (–0.200 to 0.202)
Indirect effect	b3×a3	–0.022 (–0.761 to 0.717)
Time through depression	c	1.144 (–0.450 to 2.737)

Repeated-measures correlation showed that COVID-19–related distress, self-compassion, and sleep quality had the strongest within-person correlations with symptoms of anxiety. Results from the mediation analysis for anxiety (Table 4) showed that there was also a significant negative indirect effect of time through COVID-19–related distress. There were no significant indirect effects through self-compassion or sleep quality. When controlling for mediators, there was no direct effect of time on symptoms of anxiety, indicating that changes in symptoms of anxiety were mediated by changes in COVID-19–related distress.

Table 4. Results from the multilevel mediation model on anxiety.

Effects	Path	B (95% CI)
Time through COVID-related distress	a1	–0.239 (–0.309 to –0.170)
COVID-related distress through anxiety	b1	7.659 (4.957 to 10.360)
Indirect effect	b1×a1	–1.828 (–2.585 to –1.071)
Time through self-compassion	a2	4.718 (2.972 to 6.464)
Self-compassion through anxiety	b2	–0.103 (–0.210 to 0.004)
Indirect effect	b2×a2	–0.480 (–1.005 to 0.044)
Time through sleep quality	a3	1.451 (0.741 to 2.162)
Sleep quality through anxiety	b3	–0.14 (–0.425 to 0.133)
Indirect effect	b3×a3	–0.215 (–0.675 to 0.246)
Time through anxiety	c	0.27 (–0.523 to 1.067)

Repeated-measures correlation showed that COVID-19–related distress, self-compassion, and diet quality had the strongest within-person correlations with stress. Results for the mediation analysis for stress (Table 5) showed that there was a significant indirect effect of time through self-compassion. There were no significant indirect effects through COVID-19–related distress or diet quality. When controlling for mediators, there was no direct effect of time, indicating that changes in stress were mediated by changes in self-compassion.

Table 5. Results from the multilevel mediation model on stress.

Effects	Path	B (95% CI)
Time through COVID-related distress	a1	–0.221 (–0.293 to –0.148)
COVID-related distress through stress	b1	5.189 (0.532 to 9.846)
Indirect effect	b1×a1	–1.156 (–2.335 to 0.024)
Time through self-compassion	a2	3.822 (1.702 to 5.942)
Self-compassion through stress	b2	–0.438 (–0.641 to –0.234)
Indirect effect	b2×a2	–1.662 (–2.745 to –0.580)
Time through diet quality	a3	0.831 (–0.075 to 1.738)
Diet quality through stress	b3	–0.341 (–0.907 to 0.225)
Indirect effect	b3×a3	–0.260 (–0.780 to 0.261)
Time through stress	c	–0.500 (–2.154 to 1.154)

Discussion

Principal Findings

This study examined the acceptability, outcomes, and potential mediators of daily, evidence-informed, online microinterventions to promote the mental health of staff of a tertiary education institution during the COVID-19 pandemic. Regarding program acceptability, results from weekly questionnaires showed that the vast majority of participants were satisfied with the program content, and believed that the content was helpful for their mental health and well-being. Additionally, qualitative interviews showed that participants had positive perceptions about the program's structure and content, believed it supported their mental health and well-being, helped them implement positive behavioral change, and made them feel more socially connected and supported by the university. Results from the longitudinal analysis of mental health symptoms showed that participants had moderate-to-severe symptoms of depression, anxiety, and stress at baseline, and that symptoms declined over the first 7 weeks of the intervention. Additionally, although participants' symptoms of depression, anxiety, and stress increased with the commencement of a second, stricter lockdown period, symptoms did not return to the levels reported at the beginning of the initial COVID-19 lockdown. This suggests that participation in the "VU Elevenses" program may have protected against worsening symptoms of depression, anxiety, and stress associated with mass lockdown, fears of an outbreak, and economic recession [33]. Additionally, results from the multilevel mediation analysis suggested that reductions in depression and anxiety were explained by reductions in COVID-19–related distress, and that participants who experienced greater increases in self-compassion over the study experienced greater reductions in symptoms of stress.

The moderate to severe symptoms of depression, anxiety, and stress reported by approximately 15%-25% of our participants at baseline are expected responses to the unknown and stressful circumstances of a health epidemic [2,34] and mandated self-isolation [35,36]. As stated by international mental health

experts, a rise in these symptoms was predicted to occur during the COVID-19 pandemic, as well as an increase in the prevalence of diagnosed depression and anxiety disorders, self-harm behaviors, and suicide [34]. Indeed, the mental health impacts of the COVID-19 pandemic are apparent, with reported increases in the experience of psychological distress, symptoms of depression and anxiety, and incidence of mental disorders [37-39], and highlight the need for effective mental health supports during this time of crisis.

Results from this study indicate that a brief daily online intervention may be an acceptable way to provide mental health support during a time of crisis. Staff experienced the VU Elevenses program as a positive program during the COVID-19 pandemic, and the vast majority of staff agreed that they found the content helpful for their mental health and well-being. Similarly, high levels of satisfaction have been reported in multiple other online interventions during the COVID-19 pandemic [40,41], indicating that it may be viable to deliver mental health interventions online. Additionally, many participants agreed that they felt more connected with their peers and the university due to the VU Elevenses program. It is promising that our program, delivered during the period of mandated self-isolation, increased the sense of connection with colleagues. An adverse consequence of COVID-19 restrictions is the impact of social isolation and loneliness on mental health and well-being [35,36]. This is likely a consequence of the intervention being delivered in a group setting and allowing participants to connect with their peers daily, which can increase perceptions of social support during social isolation [42]. Additionally, by delivering the intervention in groups, the program may have normalized any difficulties staff may have been experiencing during the lockdown.

The longitudinal changes in mental health reported are similar to those reported in previous studies. Several studies reported worse mental health status at the beginning of the COVID-19 pandemic compared to prepandemic levels; however, mental health improved during the early months of the pandemic [43-45]. This may be because the greatest burden of lockdowns occurs early, and people may be able to adapt to their new situation over time [46]. Other studies have also reported a nonlinear trajectory in mental health, with mental health problems reemerging after the initial improvement [44]. This is likely to follow increases in case numbers and reintroductions of local lockdown measures, contributing to worsening mental health [44]. Although previous online mental health interventions reported short-term improvements in mental health during the COVID-19 pandemic [47-49], the results from this study indicate that these improvements may not be sustained long-term, especially if local cases began to rise and stricter restrictions are reintroduced. Therefore, rapid-response health promotion interventions need to be adapted to changing participant needs throughout a crisis period to be able to provide the most appropriate supports when they are required the most.

Results from the mediation analysis indicated that preintervention to postintervention changes in depression and anxiety were explained by changes in COVID-19–related distress. This suggests that the program may have provided participants with the strategies and skills to adapt their behaviors and regulate their emotions to reduce the distress associated with the pandemic. By contrast, changes in stress were explained by changes in self-compassion over time. The Elevenses intervention used several strategies to improve self-compassion, including self-compassion meditation, resetting expectations, and simple mindfulness strategies (Table 1). Self-compassion may promote a healthier and more adaptive style of thinking and decrease rumination [50], which may explain why it was associated with changes with stress over time.

In line with recommendations to share information on managing stress, depression, and anxiety using telehealth [2] or digital responses [34], our program delivered psychological and lifestyle strategies and was regularly accessed by a group of VU staff. The weekly evaluation demonstrated a high level of program acceptability and indicated that participants who engaged in the program perceived real benefits to their mental health and well-being. The uncontrolled intervention was offered each workday, and there was no requirement to attend each day, providing flexibility for individual circumstances.

Strengths and Limitations

The content of our program was evidence-informed and was grounded in the six identified lifestyle intervention areas for well-being [11], including stress management, physical activity, nutrition, reducing substance use, sleep, and relationships. These influence coping, resilience, and mental health generally as well as in the context of stressful experiences [34]. The program's content can be delivered in any workplace that has practitioners or clinicians on staff, can facilitate access to allied health professionals, or source and share appropriate existing online resources and clinical content. Despite the strength of having a flexible and responsive approach to content, aspects of the program that were based on general clinical content such as self-compassion, relaxation, and mindfulness practices could readily be reproduced. In addition, our findings demonstrate that it is possible to develop and deliver an acceptable program of online microinterventions in a timely manner, utilizing simple and accessible low-cost online meeting platforms, offered to an entire workplace at one time point. Until May 15, 2020, Australia's federal government response was focused on managing health and financial needs rather than on the mental health implications of the COVID-19 crisis [2,51]. Our program was able to address this gap within our workplace, as it was developed and launched within 4 weeks from when staff were required to work remotely, and delivered wholly online to adhere to physical distancing and self-isolation requirements. This program allowed for the rapid collection of feedback from participants as well as timely sharing among the research team to adapt to anticipated changes in social regulations and restrictions. To the best of our knowledge, this is the first live, brief, weekday mental health promotion program in a workplace during a crisis.

Several limitations should be considered in the interpretation of these results. First, participants who opted into the program and the evaluation, including the semistructured interviews, may be subject to self-selection bias. Participants who value mental health promotion initiatives may be more likely to participate. Despite our intention for this program to be a

universal mental health promotion and prevention intervention, only 18% of invited staff participated in the program, indicating modest reach into the target population [52]. Indeed, we do not have data to determine if our sample represents the larger staff body. The gender disparity in participation in the research component (females: 80% research opt-in vs 60% overall VU staff) may reflect the well-documented gender differences in help-seeking behaviors related to mental health, or reflect the attitudinal or societal barriers reported by men in accessing mental health support [53]. This should be considered for future program marketing and content development. As this program was developed as a rapid mental health promotion response, we could not develop an exhaustive internal promotion and engagement strategy. Inviting staff to participate via a general announcement email list can be problematic as many staff may not read these emails in detail. Follow-up responses to the opt-in research component were lower than expected and it is possible that the length of the survey may have contributed to this low response rate. As an uncontrolled intervention without a comparison group, it is possible that the changes in psychological symptoms reflect natural adaptation to new routines and life circumstances over time. This is reflected in the results of the mediation models, which showed that changes in COVID-19–related distress explained changes in depression and anxiety from baseline to the study endpoint.

Conclusions

This study demonstrates the acceptability of a rapidly developed program of microinterventions, delivered in a timely manner utilizing a simple and accessible low-cost online meeting platform during a time of crisis when mental health is highly likely to deteriorate. Participation in the program may have protected against worsening symptoms of depression, anxiety, and stress associated with extended COVID-19 lockdown periods, and improved symptoms of stress by increasing self-compassion.

Acknowledgments

The authors would like to thank the contribution of our Victoria University advisory team members: Prof Corinne Reid (former Deputy Vice Chancellor, Research), Prof Anne-Marie Hede (Dean, Graduate Research), Sharon Jenner (Senior Advisor, Health and Wellbeing), Margaret Theologou (Manager, Counselling and Accessibility), Amanda Rea (Executive Officer, Institute for Health and Sport), and Prof Jeannie Rea (Senior Project Manager, Planetary Health). This work was supported by the Place-Based Planetary Health Grant (PH099), Victoria University, Australia. The funding source did not have a role in study design; data collection, analysis, or interpretation; writing of the report; or the decision to submit this paper for publication.

Authors' Contributions

AP conceived the study with all authors contributing to formulating the research question, designing the study, carrying it out, and writing the publication. Data analysis was completed by MB, PB, CS, and AP. The corresponding author (AP) had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Participant semistructured interviews schedule.
[DOCX File , 17 KB - formative_v6i2e35776_app1.docx]

Multimedia Appendix 2
Repeated-measures correlation between factors associated with mental health and symptoms of depression, anxiety, and stress.
[DOCX File , 16 KB - formative_v6i2e35776_app2.docx]

References

1. Household impacts of COVID-19 survey, 14-17 April. Australian Bureau of Statistics. 2021 Jul 14. URL: https://www.abs.gov.au/statistics/people/people-and-communities/household-impacts-covid-19-survey/latest-release [accessed 2022-02-14]
2. Zhou X, Snoswell CL, Harding LE, Bambling M, Edirippulige S, Bai X, et al. The role of telehealth in reducing the mental health burden from COVID-19. Telemed J E Health 2020 Apr 01;26(4):377-379. [doi: 10.1089/tmj.2020.0068] [Medline: 32202977]
3. Dodd RH, Dadaczynski K, Okan O, McCaffery KJ, Pickles K. Psychological wellbeing and academic experience of university students in Australia during COVID-19. Int J Environ Res Public Health 2021 Jan 20;18(3):866 [FREE Full text] [doi: 10.3390/ijerph18030866] [Medline: 33498376]
4. Stanton R, To QG, Khalesi S, Williams SL, Alley SJ, Thwaite TL, et al. Depression, anxiety and stress during COVID-19: associations with changes in physical activity, sleep, tobacco and alcohol use in Australian adults. Int J Environ Res Public Health 2020 Jun 07;17(11):4065 [FREE Full text] [doi: 10.3390/ijerph17114065] [Medline: 32517294]

5. Mental health and psychosocial considerations during the COVID-19 outbreak. World Health Organization. 2020 Mar 18. URL: https://www.who.int/docs/default-source/coronaviruse/mental-health-considerations.pdf [accessed 2022-02-14]
6. Ways to look after your mental health amid the coronavirus pandemic. Beyond Blue. URL: https://coronavirus.beyondblue.org.au/Managing-my-daily-life/Coping-with-isolation-and-being-at-home/ways-to-look-after-your-mental-health-amid-the-covid-19-pandemic [accessed 2022-02-14]
7. Sucala M, Schnur JB, Constantino MJ, Miller SJ, Brackman EH, Montgomery GH. The therapeutic relationship in e-therapy for mental health: a systematic review. J Med Internet Res 2012 Aug 02;14(4):e110 [FREE Full text] [doi: 10.2196/jmir.2084] [Medline: 22858538]
8. Bunge EL, Williamson RE, Cano M, Leykin Y, Muñoz RF. Mood management effects of brief unsupported internet interventions. Internet Interv 2016 Sep;5:36-43 [FREE Full text] [doi: 10.1016/j.invent.2016.06.001] [Medline: 30135805]
9. Ivandic I, Freeman A, Birner U, Nowak D, Sabariego C. A systematic review of brief mental health and well-being interventions in organizational settings. Scand J Work Environ Health 2017 Mar 01;43(2):99-108 [FREE Full text] [doi: 10.5271/sjweh.3616] [Medline: 28042963]
10. Fischer R, Bortolini T, Karl JA, Zilberberg M, Robinson K, Rabelo A, et al. Rapid review and meta-meta-analysis of self-guided interventions to address anxiety, depression, and stress during COVID-19 social distancing. Front Psychol 2020 Oct 28;11:563876. [doi: 10.3389/fpsyg.2020.563876] [Medline: 33192837]
11. Egger GJ, Binns AF, Rossner SR. The emergence of "lifestyle medicine" as a structured approach for management of chronic disease. Med J Aust 2009 Feb 02;190(3):143-145. [doi: 10.5694/j.1326-5377.2009.tb02317.x] [Medline: 19203313]
12. Bond M. A Bear Called Paddington. Glasgow, Scotland: HarperCollins Children's Books; 2003.
13. Mrazek P, Haggerty R. Reducing risks for mental disorders: Frontiers for preventive intervention research. Washington, DC: National Academy Press; 1994:19-20.
14. Employee Wellbeing Policy. Victoria University. 2019. URL: https://policy.vu.edu.au/document/view.php?id=460 [accessed 2022-02-14]
15. Evidence Overwhelmingly Supports Efficacy of Lifestyle Medicine. The American College of Lifestyle Medicine. URL: https://www.lifestylemedicine.org/ACLM/Resources/Scientific_Evidence/ACLM/About/What_is_Lifestyle_Medicine_/Scientific_Evidence.aspx?hkey=ed4b4130-6ce9-41bb-8703-211bc98eed7f [accessed 2022-02-14]
16. 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]
17. Lovibond PF, Lovibond SH. The structure of negative emotional states: comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. Behav Res Ther 1995 Mar;33(3):335-343. [doi: 10.1016/0005-7967(94)00075-u] [Medline: 7726811]
18. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989 May;28(2):193-213. [doi: 10.1016/0165-1781(89)90047-4] [Medline: 2748771]
19. Humeniuk R, Henry-Edwards S, Ali R, Poznyak V, Monteiro M. The Alcohol, Smoking and Substance Involvement Screening Test (ASSIST): Manual for use in Primary Care. Geneva: World Health Organization; 2010:42-47.
20. Billings AG, Moos RH. The role of coping responses and social resources in attenuating the stress of life events. J Behav Med 1981 Jun;4(2):139-157. [doi: 10.1007/bf00844267]
21. Raes F, Pommier E, Neff KD, Van Gucht D. Construction and factorial validation of a short form of the self-compassion scale. Clin Psychol Psychother 2011 Jun 08;18(3):250-255. [doi: 10.1002/cpp.702] [Medline: 21584907]
22. Zimet GD, Dahlem NW, Zimet SG, Farley GK. The multidimensional scale of perceived social support. J Pers Assess 1988 Mar;52(1):30-41. [doi: 10.1207/s15327752jpa5201_2]
23. Lee RM, Draper M, Lee S. Social connectedness, dysfunctional interpersonal behaviors, and psychological distress: testing a mediator model. J Counsel Psychol 2001;48(3):310-318. [doi: 10.1037/0022-0167.48.3.310]
24. Paxton AE, Strycker LA, Toobert DJ, Ammerman AS, Glasgow RE. Starting the conversation performance of a brief dietary assessment and intervention tool for health professionals. Am J Prev Med 2011 Jan;40(1):67-71. [doi: 10.1016/j.amepre.2010.10.009] [Medline: 21146770]
25. Australian Institute of Health and Welfare. The Active Australia Survey: A guide and manual for implementation, analysis and reporting. Canberra: Australian Government; 2003:2-4.
26. Lee SA. Coronavirus Anxiety Scale: a brief mental health screener for COVID-19 related anxiety. Death Stud 2020 Apr 16;44(7):393-401. [doi: 10.1080/07481187.2020.1748481] [Medline: 32299304]
27. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
28. Rockwood NJ. Advancing the formulation and testing of multilevel mediation and moderated mediation models. Thesis. The Ohio State University. 2017. URL: https://etd.ohiolink.edu/apexprod/rws_olink/r/1501/10?clear=10&p10_accession_num=osu1489578419777238 [accessed 2022-02-14]
29. Bakdash JZ, Marusich LR. Repeated Measures Correlation. Front Psychol 2017;8:456. [doi: 10.3389/fpsyg.2017.00456] [Medline: 28439244]

30. Shin T, Davison ML, Long JD. Maximum likelihood versus multiple imputation for missing data in small longitudinal samples with nonnormality. Psychol Methods 2017 Sep;22(3):426-449. [doi: 10.1037/met0000094] [Medline: 27709974]
31. Newman DA. Longitudinal modeling with randomly and systematically missing data: a simulation of ad hoc, maximum likelihood, and multiple imputation techniques. Organ Res Methods 2016 Jun 29;6(3):328-362. [doi: 10.1177/1094428103254673]
32. Rubin D. Multiple imputation for nonresponse in survey. Hoboken, NJ: John Wiley & Sons; 1987:19-21.
33. Robinson E, Sutin A, 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]
34. Holmes EA, O'Connor RC, Perry VH, Tracey I, Wessely S, Arseneault L, et al. Multidisciplinary research priorities for the COVID-19 pandemic: a call for action for mental health science. Lancet Psychiatry 2020 Jun;7(6):547-560 [FREE Full text] [doi: 10.1016/S2215-0366(20)30168-1] [Medline: 32304649]
35. Nkire N, Mrklas K, Hrabok M, Gusnowski A, Vuong W, Surood S, et al. COVID-19 pandemic: demographic predictors of self-isolation or self-quarantine and impact of isolation and quarantine on perceived stress, anxiety, and depression. Front Psychiatry 2021 Feb 1;12:553468. [doi: 10.3389/fpsyt.2021.553468] [Medline: 33597900]
36. Pancani L, Marinucci M, Aureli N, Riva P. Forced social isolation and mental health: a study on 1,006 Italians under COVID-19 lockdown. Front Psychol 2021 May 21;12:663799. [doi: 10.3389/fpsyg.2021.663799] [Medline: 34093358]
37. Campion J, Javed A, Sartorius N, Marmot M. Addressing the public mental health challenge of COVID-19. Lancet Psychiatry 2020 Aug;7(8):657-659 [FREE Full text] [doi: 10.1016/S2215-0366(20)30240-6] [Medline: 32531299]
38. O'Connor RC, Wetherall K, Cleare S, McClelland H, Melson AJ, Niedzwiedz CL, et al. Br J Psychiatry 2021 Jun;218(6):326-333 [FREE Full text] [doi: 10.1192/bjp.2020.212] [Medline: 33081860]
39. Kola L, Kohrt BA, Hanlon C, Naslund JA, Sikander S, Balaji M, et al. COVID-19 mental health impact and responses in low-income and middle-income countries: reimagining global mental health. Lancet Psychiatry 2021 Jun;8(6):535-550. [doi: 10.1016/S2215-0366(21)00025-0] [Medline: 33639109]
40. González-García M, Álvarez JC, Pérez EZ, Fernandez-Carriba S, López JG. Feasibility of a brief online mindfulness and compassion-based intervention to promote mental health among university students during the COVID-19 pandemic. Mindfulness (N Y) 2021 May 17;12(7):1-11 [FREE Full text] [doi: 10.1007/s12671-021-01632-6] [Medline: 34025814]
41. Bäuerle A, Jahre L, Teufel M, Jansen C, Musche V, Schweda A, et al. Evaluation of the e-mental health mindfulness-based and skills-based "CoPE It" intervention to reduce psychological distress in times of COVID-19: results of a bicentre longitudinal study. Front Psychiatry 2021 Nov 4;12:768132. [doi: 10.3389/fpsyt.2021.768132] [Medline: 34803775]
42. Sommerlad A, Marston L, Huntley J, Livingston G, Lewis G, Steptoe A, et al. Social relationships and depression during the COVID-19 lockdown: longitudinal analysis of the COVID-19 Social Study. Psychol Med 2021 Jan 13:1-10 [FREE Full text] [doi: 10.1017/S0033291721000039] [Medline: 33436126]
43. Daly M, Sutin AR, Robinson E. Longitudinal changes in mental health and the COVID-19 pandemic: evidence from the UK Household Longitudinal Study. Psychol Med 2020 Nov 13:1-10 [FREE Full text] [doi: 10.1017/S0033291720004432] [Medline: 33183370]
44. Stroud I, Gutman LM. Longitudinal changes in the mental health of UK young male and female adults during the COVID-19 pandemic. Psychiatry Res 2021 Sep;303:114074 [FREE Full text] [doi: 10.1016/j.psychres.2021.114074] [Medline: 34271372]
45. Daly M, Robinson E. Longitudinal changes in psychological distress in the UK from 2019 to September 2020 during the COVID-19 pandemic: Evidence from a large nationally representative study. Psychiatry Res 2021 Jun;300:113920 [FREE Full text] [doi: 10.1016/j.psychres.2021.113920] [Medline: 33882397]
46. Fancourt D, Steptoe A, Bu F. Trajectories of anxiety and depressive symptoms during enforced isolation due to COVID-19 in England: a longitudinal observational study. Lancet Psychiatry 2021 Feb;8(2):141-149. [doi: 10.1016/s2215-0366(20)30482-x]
47. Shapira S, Yeshua-Katz D, Cohn-Schwartz E, Aharonson-Daniel L, Sarid O, Clarfield A. A pilot randomized controlled trial of a group intervention via Zoom to relieve loneliness and depressive symptoms among older persons during the COVID-19 outbreak. Internet Interv 2021 Apr;24:100368 [FREE Full text] [doi: 10.1016/j.invent.2021.100368] [Medline: 33527072]
48. Bantjes J, Kazdin AE, Cuijpers P, Breet E, Dunn-Coetzee M, Davids C, et al. A web-based group cognitive behavioral therapy intervention for symptoms of anxiety and depression among university students: open-label, pragmatic trial. JMIR Ment Health 2021 May 27;8(5):e27400 [FREE Full text] [doi: 10.2196/27400] [Medline: 34042598]
49. El Morr C, Ritvo P, Ahmad F, Moineddin R, MVC Team. Correction: Effectiveness of an 8-week web-based mindfulness virtual community intervention for university students on symptoms of stress, anxiety, and depression: randomized controlled trial. JMIR Ment Health 2020 Sep 30;7(9):e24131 [FREE Full text] [doi: 10.2196/24131] [Medline: 32997643]
50. Ferrari M, Hunt C, Harrysunker A, Abbott MJ, Beath AP, Einstein DA. Self-compassion interventions and psychosocial outcomes: a meta-analysis of RCTs. Mindfulness 2019 Apr 8;10(8):1455-1473. [doi: 10.1007/s12671-019-01134-6]

51. COVID-19: $48.1 million for National Mental Health and Wellbeing Pandemic Response Plan. Government of Australia Department of Health. URL: https://www.health.gov.au/ministers/the-hon-greg-hunt-mp/media/covid-19-481-million-for-national-mental-health-and-wellbeing-pandemic-response-plan [accessed 2022-02-14]
52. Feldstein AC, Glasgow RE. A practical, robust implementation and sustainability model (PRISM) for integrating research findings into practice. Jt Comm J Qual Patient Saf 2008 Apr;34(4):228-243. [doi: 10.1016/s1553-7250(08)34030-6] [Medline: 18468362]
53. Bolea-Alamanac BM. Promotion of mental health: a gender perspective. In: Bährer-Kohler S, Carod-Artal FJ, editors. Global mental health: prevention and promotion. Cham: Springer International Publishing; 2017:141-155.

Abbreviations

DASS-21: Short Form Depression, Anxiety and Stress Scale
MSPSS: Multidimensional Scale of Perceived Social Support
SCS: Social Connectedness and the Social Assurance Scales
SCS-SF: Self-Compassion Short Form
VU: Victoria University

Edited by A Mavragani; submitted 19.12.21; peer-reviewed by J Bantjes, A Kazdin; comments to author 08.01.22; revised version received 18.01.22; accepted 18.01.22; published 25.02.22.

Please cite as:
Parker A, Dash S, Bourke M, Patten R, Craike M, Baldwin P, Hosking W, Levinger I, Apostolopoulos V, de Courten M, Sharples J, Naslund M, Stavropoulos V, Woessner M, Sonn C, Stansen C, Pascoe M
A Brief, Daily, Online Mental Health and Well-being Intervention for University Staff During the COVID-19 Pandemic: Program Description and Outcomes Using a Mixed Methods Design
JMIR Form Res 2022;6(2):e35776
URL: https://formative.jmir.org/2022/2/e35776
doi:10.2196/35776
PMID:35044304

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org