JMIR Formative Research

Impact Factor (2023): 2.0
Volume 5 (2021), Issue 1 ISSN 2561-326X Editor in Chief: Amaryllis Mavragani, PhDc

Contents

Original Papers

Reviews

Corrigenda and Addenda

Original Paper

The Use of Small Electronic Devices and Health: Feasibility of Interventions for a Forthcoming Crossover Design

Lisbeth Hoekjaer Larsen[1], MSc, PhD; Maja Hedegaard Lauritzen[1], BA; Sirin Wilhelmsen Gangstad[2,3], MSc, PhD; Troels Wesenberg Kjaer[1,4], MD, PhD

[1]Department of Neurology, Zealand University Hospital, Roskilde, Denmark

[2]Department of Applied Mathematics and Computer Science, Technical University of Denmark, Lyngby, Denmark

[3]Uneeg Medical, Lynge, Denmark

[4]Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

Corresponding Author:
Lisbeth Hoekjaer Larsen, MSc, PhD
Department of Neurology
Zealand University Hospital
Sygehusvej 10
Roskilde,
Denmark
Phone: 45 41 55 85 92
Email: lisbla@regionsjaelland.dk

Abstract

Background: Modern lifestyle is heavily affected by technology such as smartphones, tablets, and other small computers; yet it remains unclear how our health and well-being are affected by the heavy use of these devices.

Objective: This feasibility study aims to test two different interventions of an experimental protocol for a forthcoming large-scale community-based study and get estimates of parameters for sample size calculation. The aim of the large-scale study is to investigate the effect of (1) a wearable tracking device on aerobic capacity (VO_2max/kg) and the effect of (2) restricting media use on total sleep time.

Methods: Twenty healthy participants were included and equipped with a wrist-worn device tracking physical activity and sleep. Participants were allocated to either a physical activity group, which was instructed to use the wrist-worn device to support exercise, or a sleep silent group, which was instructed to remove or switch off all electronic devices in the bedroom (except the wrist-worn tracking device). The intervention lasted approximately 4 weeks. Data collected included blood pressure, submaximal cycle ergometer test, self-reported technology use, and compliance of using the wearable tracking device.

Results: All participants wore the wearable tracking device 95.8% (SD 4.4%) of the time. Participants in the physical activity group increased aerobic capacity from 30.38 (SD 8.98) to 32.1 (SD 8.71) mL/kg/min (t=–2.31, P=.046) and decreased their systolic blood pressure from 126.5 (SD 15.8) mm Hg to 121.8 (SD 11.7) mm Hg (t=2.72, P=.02). The sleep silent group prolonged their time offline before bedtime from 18.1 (SD 19.4) minutes to 27.2 (SD 17.3) minutes (t=–2.94, P=.02).

Conclusions: The two interventions are feasible to conduct. Participants were willing to wear the tracking device on their wrist and restrict all media use in their bedroom and thereby reduce bedtime technology use. Our results also suggest that tracking physical activity using a wearable device is accompanied by noteworthy health benefits. We outline necessary adjustments for a forthcoming large-scale study.

(JMIR Form Res 2021;5(1):e20410) doi:10.2196/20410

KEYWORDS

accelerometer; activity trackers; aerobic capacity; insufficient sleep; media use; screen time; sleep problems; smartphones; wearable tracking devices

Introduction

Progress in technology has revolutionized the way we live in modern society. Small and convenient electronic devices are with us everywhere and play a central role in our lives and the way we work, communicate, interact, search for information, do chores, and pass time. Yet it remains unclear how our health and well-being are affected by the use of these devices. In this feasibility study, we test an experimental protocol designed to investigate how the use of a wearable tracking device (WTD) and bedtime technology use affect physical activity and sleep, respectively. More knowledge of the effect of technology is needed as inactivity and insufficient sleep pose serious public health implications in modern society.

In western culture, physical inactivity and sedentary lifestyle are increasing and, as a consequence, so are health-related problems and health care costs [1]. Global Health Observatory data estimates that 37% of the adult population in high-income countries is insufficiently physically active [2]. It has been suggested that WTDs may encourage physically active behavior [3]. WTDs are wearable computers able to monitor different health-related parameters such as steps, distance covered, and pulse continuously under real-life conditions, and they are already widely used by consumers. The self-monitoring is made possible by different sensors and algorithms and is often accompanied by mobile apps. Modern WTDs have the opportunity to incorporate principles for behavior change in the promotion of physical activity including feedback, tailored information, gamification, rewards, goal setting, prompts, social comparison, and connectivity [3,4]. Despite the promising features embedded in WTDs, results are mixed from previous studies investigating the effect of increasing physical activity with WTDs on different health parameters [5-7]. Part of the discrepancy between studies may relate to study populations, interventions, comparators, and outcomes. The effect of using WTD on VO_2max seems to be less studied, although this health parameter is known to be an important indicator of health-risk status. Epidemiologic studies have reported that a low VO_2max is a more powerful predictor of risk for adverse outcomes than traditional risk factors, including hypertension, lipid abnormalities, smoking, physical inactivity, obesity, and diabetes mellitus [1,8,9].

Insufficient sleep constitutes another health risk in modern society. Recent evidence demonstrates the proportion of people getting less than the recommended hours of sleep is rising [10]. A survey conducted by the National Sleep Foundation found that the proportions of people sleeping fewer than 7 hours are 40% in Japan, 27% in the United States, and 21% in Germany [11]. Insufficient sleep can have multiple negative consequences, such as cognitive impairment, obesity, hypertension and insulin resistance (diabetes), and substantial economic losses [10,12]. It has been proposed that the increased use of media via smartphones, tablets, and other handheld devices before bedtime is worsening the challenge because the screen light significantly suppresses the secretion of melatonin and consequently disrupts sleep [13]. Furthermore, the contents received from these handheld devices may induce arousal and stress reactions, making it difficult to fall asleep [14]. Studies on smartphone use and sleep have quite consistently shown an association between bedtime technology use and sleep descriptors [14-16]. According to a study from Denmark, 40% of 815 young Danish students gave likes or sent messages during the night [17]. It is, however, unclear from this study and many similar studies whether smartphone activity is causing an increase in sleep onset latency and sleep interruption or if smartphone activity is used as an entertainment device among those with sleep impairment due to other causes [14,17]. Of note, a study with 942 Canadian students demonstrated that sleep problems predicted media use and not the opposite [18]. Most studies today are based on cross-sectional design, meaning that the causality is difficult to ascertain [14]. Recruiting participants for an experimental protocol may pose a challenge due to a lack of motivation to negotiate changes in bedtime smartphone use [19]. Nevertheless, more experimental research is needed on how bedtime uses of smartphones affect sleep measured over a longer period of time.

The purpose of this study was to test the feasibility of two different interventions of an experimental protocol and to get estimates of parameters for sample size calculation in order to refine the protocol for a forthcoming large-scale study. The aim of the forthcoming study is to investigate the effect of (1) using a WTD on aerobic capacity and (2) removing electronic devices from the bedroom on total sleep time (TST). The forthcoming study will contain both a baseline and an intervention period, but this study aims at investigating the feasibility of the interventions only.

Methods

Participants

Twenty able-bodied participants (4 males, with a mean age of 48 [SD 9] years) were recruited to participate in the study through local advertisement in the municipality of Naestved, Denmark. Participants were required to be aged 18 to 75 years, to own a smartphone or tablet, and to be able to exercise on their own. People already exercising for more than 15 hours weekly were not eligible. The sample size for this study was set to 20 participants, which we estimated to be adequate to test the experimental protocol and get estimates of parameters for sample size calculation to the necessary degree of precision [20]. All participants gave informed consent to the experimental procedure, which was approved by the local ethics committee (SJ-743). The study was performed in accordance with the Declaration of Helsinki.

Experimental Protocol

Participants attended 2 test days (T1 and T2) with 33 (SD 8) days in between (Figure 1). BMI, blood pressure (BP, mean of 3 repeated measures), and a submaximal cycle ergometer test to estimate VO_2max [21] were conducted at both test days. Furthermore, participants answered questions on a tablet regarding their level of moderate-to-vigorous physical activity (MVPA) in minutes per week with the Nordic Physical Activity Questionnaire-short (NPAQ-short) [22], their current level of sleep problems with the Insomnia Severity Index (ISI) [23], and

their time offline (TO) before and after sleep. The NPAQ-short is a 2-item questionnaire to monitor physical activity (time and intensity) and compliance with the World Health Organization (WHO) recommendations. The ISI is a 7-item questionnaire where participants rate symptoms of their sleep problems using a Likert-type scale. Each item is rated on a 0 to 4 scale, and the total score ranges from 0 to 28. A higher score suggests more severe insomnia.

Figure 1. Experimental setup.

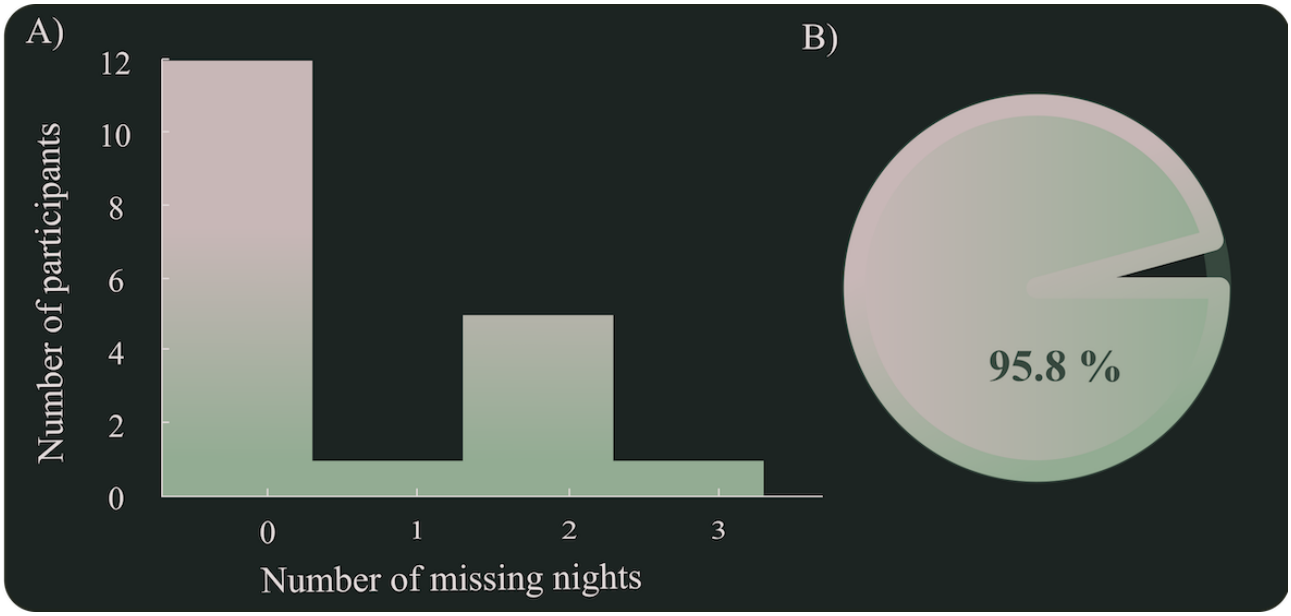
On T1, participants were allocated to either a physical activity (PA) group or a sleep silent (SS) group by minimization in order to ensure a balanced age distribution in both groups. This was done to examine the feasibility of both interventions among different age groups. In the PA group, participants were encouraged to challenge themselves with a realistic self-chosen fitness goal for the intervention period based on their resources and prior training level. The goal could be specific, such as accomplishing running 5 kilometers without stopping, or more general, such as meeting WHO's minimum recommendation of 150 minutes MVPA weekly [24]. Participants in the PA group were introduced to the WTD and instructed to follow their progress on the accompanying mobile app Garmin Connect. All exercise was performed independently by the participants. Participants were asked after the intervention about their use of WTD and whether they wished to continue using a wrist-worn tracking device in the future. The SS group was instructed to remove or switch off all electronic devices in the bedroom (except the WTD). Several different technologies (such as computers, tablets, and other handheld devices) are used for the same activities as a smartphone and, therefore, use of all electronic devices was restricted. Analog alarm clocks were distributed, and participants were asked not to check their smartphone and other digital screens and devices from bedtime until they get up in the morning.

On T1, all participants were equipped with a WTD (Vivosmart 4, Garmin Ltd), and instructed to download the mobile app Garmin Connect and set up a user account. One participant already used a WTD (Fenix 5X, Garmin Ltd), which the participant continued to use instead of Vivosmart 4. All participants were instructed to wear their WTD on their wrist for the entire period of approximately 4 weeks. The small device detects physical activity, heart rate, and sleep via an embedded triaxial accelerometer, optical photoplethysmography signals, and associated algorithms. It automatically records intensity and type and duration of different activity patterns such as walking, running, and biking for at least 10 minutes and attempts to detect sleep onset, sleep end, sleep stages (light, deep, rapid eye movement, and wake), and level of movement during sleep. Based on the time stamps of the WTD measurements, the compliance of wearing the WTD was investigated for each participant. The amount of time a pulse measurement was available in the recorded data relative to the length of the intervention period was computed. The pulse was chosen as it is sampled relatively frequently (1 sample per 2 minutes). The percentage of available pulse data was used as a proxy for the percentage of time the participant wore the WTD. Furthermore, the amount of nights with missing TST estimates in the WTD recordings was investigated, as the TST is an important parameter in the upcoming study.

Results

Twenty participants were recruited for this study. One participant from the SS group lost the WTD after 3 weeks and was excluded. Another participant from the SS group got an injured little finger while walking a dog (not related to study activities) and therefore did not perform the cycle ergometer test at T2. Throughout the intervention period, the participants wore the WTD device 95.8% (SD 4.4%) of the time. Seven participants missed 1 to 3 nights of data due to not charging the battery (Figure 2).

Figure 2. Compliance with wearing the wearable tracking device. A) shows number of missing nights per participant, and B) shows compliance in percentages for all participants during the entire intervention period.



From T1 to T2, the PA group increased their estimated VO_2max from 30.38 (SD 8.98) to 32.1 (SD 8.71) mL/kg/min (t=–2.31, P=.046) and the systolic BP decreased from 126.5 (SD 15.8) mm Hg at T1 to 121.8 (SD 11.7) mm Hg at T2 (t=2.72, P=.02) while no difference was observed in the diastolic BP (from 84.3 [SD 10.1] mm Hg to 80.4 [SD 8.8] mm Hg; t=1.96, P=.08) or in the BMI (see Table 1). According to self-reported exercise behavior, 2 participants in the PA group did not meet the WHO's minimum recommendation of 150 MVPA minutes per week at T1, while 2 participants had filled out the questionnaire incorrectly. At T2, 1 participant in the PA group reported an activity level below 150 MVPA minutes per week and no difference was reported in MVPA between T1 and T2 in the PA group (from 310 [SD 216] to 375 [SD 172] minutes per week, t=–1.5, P=.18; Table 1). Three participants in the PA group reported sleep problems (ISI value 8.3 [SD 2], n=3) at T1 and 2 participants (ISI value 4.7 [SD 4], n=3) at T2. No change was observed in time offline (TO) in the PA group (Table 1). All participants in the PA group reported to use the WTD to track their activity level in the intervention period. Eight participants wished to continue using a WTD after T2, while 2 participants were reluctant due to a lack of interest in the information collected and stress associated with self-monitoring respectively.

In the SS group, 6 participants reported sleep problems (ISI value 9 [SD 2], n=6) at T1 and 3 participants (ISI value 4.7 [SD 6], n=6) at T2. The SS group prolonged the TO before bedtime (from 18.1 [SD 19.4] to 27.2 [SD 17.3] minutes; t=–2.94, P=.02) while no change was observed in TO in the morning (Table 2). In the SS group, no change was observed in the estimated VO_2max, BP, or BMI between test days (Table 2). According to self-reported exercise behavior, 3 participants in SS group did not meet the WHO's minimum recommendation of 150 MVPA minutes per week at both T1 and T2, and no change in the MVPA was observed in the group (from 274 [SD 172] to 283 [SD 203] minutes per week; t=–0.18, P=.86). One participant had filled out the questionnaire incorrectly.

Table 1. Pre-post measurements in the physical activity group.

Characteristic	T1, mean (SD)	T2, mean (SD)	t	P value
BMI	26.73 (4.79)	26.69 (4.77)	0.61	.56
Systolic BP[a] (mm Hg)	126.5 (15.77)	121.8 (11.71)	2.72	.02
Diastolic BP (mm Hg)	84.3 (10.05)	80.37 (8.75)	1.96	.08
VO_2max[b]/kg	30.38 (8.98)	32.1 (8.71)	–2.31	.046
SR MVPA[c] (min/week)	310 (216)	375 (172)	–1.50	.18
SR TO[d], evening (min)	11.4 (15.55)	15.1 (16.67)	–1.13	.29
SR TO, morning (min)	21.9 (28.5)	22.8 (29.78)	–0.14	.89

[a]BP: blood pressure.

[b]VO_2max: maximal oxygen uptake.

[c]SR MVPA: self-reported moderate-to-vigorous physical activity.

[d]SR TO: self-reported time offline.

Table 2. Pre-post measurements in the sleep silent group.

Characteristic	T1, mean (SD)	T2, mean (SD)	*t*	*P* value
BMI	27.35 (5.59)	27.30 (5.6)	0.26	.80
Systolic BP[a] (mm Hg)	127.15 (16.9)	125.78 (17.78)	0.30	.77
Diastolic BP (mm Hg)	84.67 (10.41)	87.41 (11.58)	–1.16	.28
VO_2max[b]/kg	30.42 (8.57)	32.18 (9.25)	–1.31	.23
SR MVPA[c] (min/week)	274 (172)	283 (203)	–0.18	.86
SR TO[d], evening (min)	18.1 (19.36)	27.22 (17.34)	–2.94	.02
SR TO, morning (min)	28 (28.08)	32.22 (19.70)	–1.81	.11

[a]BP: blood pressure.

[b]VO_2max: maximal oxygen uptake.

[c]SR MVPA: self-reported moderate-to-vigorous physical activity.

[d]SR TO: self-reported time offline.

Discussion

Principal Findings

Results from this study suggest that the experimental protocol is feasible to conduct: participants were willing to wear the wrist-worn tracking device and keep track of their exercise or remove their smartphone from the bedroom. The participants wore the WTD nearly 96% of the time they were enrolled in the study, demonstrating an extremely high compliance considering the participants wear the WTD around-the-clock and it includes an inevitable loss of data points due to necessary charging of battery one or twice a week.

The majority (60%) of participants in this feasibility study had a low or somewhat low VO_2max at T1 in both groups according to Astrands classification of aerobic capacity by age and gender [25]. It has been demonstrated that a low VO_2max is associated with a 2- to 5-fold increase in cardiovascular disease or all-cause mortality, independent of other cardiovascular disease risk factors [26]. Importantly, relatively small improvements in aerobic capacity such as 1 metabolic equivalent (3.5 mL/kg/minute) have been associated with 8% to 35% reductions in mortality [26]. From this perspective, an average VO_2max increase of 1.71 mL/kg/minute in the PA group could suggest a noteworthy health benefit if the participants maintain the level of exercise from the intervention period in future.

In the PA group we also observed an average decrease of systolic BP of 4.7 mm Hg. Hypertension significantly increases the risks of heart, brain, and other diseases. In a meta-analysis by Lewington et al [27], the age-specific relevance of usual BP to vascular mortality was assessed from one million adults in 61 prospective studies. The authors found that a reduction in systolic BP of just 2 mm Hg reduces apoplexy mortality by 10% and death of ischemic heart disease by 7% among middle-aged people. In light of this, our observed decrease in average systolic BP of 4.7 mm Hg is also highly relevant. A recent review evaluated the effect of using WTDs on metabolic outcomes such as BP, blood glucose level, and cholesterol levels in patients [5]. Based on the 6 included studies, the authors conclude that WTDs play a role as a facilitator in motivating and accelerating physical activity, but current data do not suggest other consistent health benefits for patients. Two other recent reviews conclude that people using wearable devices improved their daily step counts regardless of age, sex and health status [6,7]. Of note, Brickwood et al [7] also found a significant increase in MVPA, while Lynch et al [6] did not find this positive effect. A great challenge in this field is that the literature remains limited primarily to short-term studies, and many of these are underpowered feasibility or pilot studies [5,28]. Personal preferences and adverse effects related to self-monitoring may also play a role in the disagreement between studies. For instance, modern WTDs allow individuals to gain insight into their own activity level 24 hours a day, and studies have demonstrated that for some individuals self-monitoring is valued and can prompt further goal-directed behavior while for other individuals the inability to meet goals can trigger negative experiences [3]. Two participants in the PA group reported skepticism to continue using a WTD due to a lack of interest in measurements and stress associated with self-monitoring, respectively. Thus, the effect of self-monitoring of PA behavior may be affected by personality. Large studies that can accommodate the fast pace of advances in technology are needed to examine if WTDs can enhance important health outcomes and determine which populations are most receptive to WTDs.

Participants in the SS group prolonged the TO before bedtime, demonstrating a willingness to incorporate restrictions on bedtime technology use. Previous studies suggest that bedtime technology use is negatively related to sleep outcomes, but few longitudinal studies have been conducted with an experimental setup. A strength of this study is therefore the interventional and feasible study design. The few existing experimental studies that have been conducted report contradictory results on sleep measures [19,29-31]. For instance, restricting mobile phone use before bedtime for 4 weeks had no effect on sleep measures in a study conducted by Harris el al [29] in Norwegian high school athletes, while He et al [30] found several improvements in both sleep measures and working memory in Japanese university students. Of note, the inclusion criteria differed in the 2 studies: He et al [30] only included participants with poor sleep and a

habit of using a mobile phone during bedtime, while Harris et al [29] did not have such inclusion criteria. In our feasibility study, we did not have any inclusion criteria regarding smartphone use meaning that we included both light and heavy smartphone users. A recent telephone-based survey showed that 42% of participants reported using electronic devices in bed after lights out, and 27% of adults who reported always using electronic devices in bed were spending over an hour per night using them [32]. The survey demonstrates a large variance in the habits of bedtime technology among adults, which is important to consider in order to illuminate how the use of smartphones affect sleep and sleep quality.

Current research on the associations between sleep measures and smartphone use has mainly focused on children, adolescents, or university students, which compromises generalizability of the results to the population above 25 to 30 years [14,32]. In our feasibility study, the average age of participants was 48 years ranging from age 24 to 60 years. This may also explain why we did not encounter similar challenges in the recruiting process as Bartel et al [19], who only included adolescents. Nevertheless, a study including all age groups can contribute to cover a gap in the literature.

The effect of bedtime technology use has mainly been investigated with self-reported outcome measures and may thereby be prone to misclassification, recall difficulty, recall bias, and response-style bias [14]. Only a few studies have applied objective sleep measures based on actigraphy and examined the association between sleep and self-reported media use [33,34]. These studies report that self-reported bedtime technology use is negatively related to objective sleep measures in adolescents. Although the literature shows that actigraphy reliably detects sleep-wake patterns in normal individuals [35], we are currently investigating the validity of the sleep detection provided by the Vivosmart 4 in a separate study. One study has used a screen time detecting app to examine the relationship between self-reported sleep and screen time measured objectively. Increased screen time was associated with poor self-reported sleep outcomes (sleep quality, sleep duration, sleep efficiency, and longer sleep onset latency) [36]. Future research should ideally combine a large-scale intervention with objective measures of both sleep and screen time in an adult population in order to draw valid conclusions about cause and effect of the association between bedtime technology use and sleep measures.

Limitations

The feasibility study design had some limitations, which preferably should be adjusted in the forthcoming large-scale study. First, we did not have any inclusion criteria regarding usual smartphone use and physical activity level, meaning that for some participants the intervention made little change to their established pattern. An advantage of such broad inclusion criteria is that it enables a generalization to a broad population group. However, a disadvantage is that the result may be contaminated. Hence, in the large-scale study inclusion criteria should be added in order to ensure examination of relevant participants and the content of the interventions needs to be specified. Second, this study design did not include actual control observations, which is necessary in order to determine an effect of an intervention. Finally, the intervention period of 4 weeks is short and should be expanded in order to investigate long-term effects. Meeting these limitations in a forthcoming large-scale study can contribute with experimental evidence of the effect of using WTDs on aerobic capacity and restricting bedtime technology use on sleep length.

Conclusions

The experimental protocol in this study was feasible to conduct. Participants were willing to wear the WTD around-the-clock and use the wrist-worn device to support exercise or remove their smartphone from the bedroom. We observed that tracking PA using a wearable device is accompanied by noteworthy health benefits and that restricting technology use in the bedroom reduce participant use of bedtime technology. In a forthcoming large-scale study, sample size calculations will be based on collected estimates of VO_2max and TST. Furthermore, in order to obtain experimental evidence of the effect of using WTDs on aerobic capacity and illuminate causal claims of restricting bedtime technology use on TST, adjustment highlighted in the previous section should be prioritized.

Acknowledgments

The study was supported by the European Union via Interreg, which is funded by the European Regional Development Fund. The study was designed by LHL and TWK; LHL and MHL performed the data collection. All authors were involved in the analysis and interpretation of data and all authors revised and finally approved the version to be published.

Conflicts of Interest

None declared.

References

1. Blair SN. Physical inactivity: the biggest public health problem of the 21st century. Br J Sports Med 2009 Jan;43(1):1-2. [Medline: 19136507]
2. Guthold R, Stevens G, Riley L, Bull F. Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1·9 million participants. The Lancet Global Health 2018 Oct;6(10):e1077-e1086 [FREE Full text] [doi: 10.1016/S2214-109X(18)30357-7]

3. Carter DD, Robinson K, Forbes J, Hayes S. Experiences of mobile health in promoting physical activity: a qualitative systematic review and meta-ethnography. PLoS One 2018 Dec;13(12):e0208759 [FREE Full text] [doi: 10.1371/journal.pone.0208759] [Medline: 30557396]
4. Olander EK, Fletcher H, Williams S, Atkinson L, Turner A, French DP. What are the most effective techniques in changing obese individuals' physical activity self-efficacy and behaviour: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 2013;10:1-15 [FREE Full text] [doi: 10.1186/1479-5868-10-29] [Medline: 23452345]
5. Jo A, Coronel BD, Coakes CE, Mainous AG. Is there a benefit to patients using wearable devices such as Fitbit or health apps on mobiles? A systematic review. Am J Med 2019 Dec;132(12):1394-1400. [doi: 10.1016/j.amjmed.2019.06.018] [Medline: 31302077]
6. Lynch C, Bird S, Lythgo N, Selva-Raj I. Changing the physical activity behavior of adults with fitness trackers: a systematic review and meta-analysis. Am J Health Promot 2019 Dec 20;34(4):418-430. [doi: 10.1177/0890117119895204] [Medline: 31858812]
7. Brickwood K, Watson G, O'Brien J, Williams AD. Consumer-based wearable activity trackers increase physical activity participation: systematic review and meta-analysis. JMIR Mhealth Uhealth 2019 Apr 12;7(4):e11819 [FREE Full text] [doi: 10.2196/11819] [Medline: 30977740]
8. Myers J, Prakash M, Froelicher V, Do D, Partington S, Atwood JE. Exercise capacity and mortality among men referred for exercise testing. N Engl J Med 2002 Mar 14;346(11):793-801. [doi: 10.1056/nejmoa011858] [Medline: 11893790]
9. Kodama S, Saito K, Tanaka S, Maki M, Yachi Y, Asumi M, et al. Cardiorespiratory fitness as a quantitative predictor of all-cause mortality and cardiovascular events in healthy men and women: a meta-analysis. JAMA 2009 May 20;301(19):2024-2035. [doi: 10.1001/jama.2009.681] [Medline: 19454641]
10. Chattu V, Manzar M, Kumary S, Burman D, Spence D, Pandi-Perumal S. The global problem of insufficient sleep and its serious public health implications. Healthcare 2018 Dec 20;7(1):1. [doi: 10.3390/healthcare7010001] [Medline: 30577441]
11. International Bedroom Poll. National Sleep Foundation. 2013. URL: https://www.sleepfoundation.org/sites/default/files/inline-files/RPT495a.pdf [accessed 2020-12-02]
12. Chattu VK, Sakhamuri SM, Kumar R, Spence DW, BaHammam AS, Pandi-Perumal SR. Insufficient sleep syndrome: is it time to classify it as a major noncommunicable disease? Sleep Sci 2018;11(2):56-64 [FREE Full text] [doi: 10.5935/1984-0063.20180013] [Medline: 30083291]
13. Cajochen C, Frey S, Anders D, Späti J, Bues M, Pross A, et al. Evening exposure to a light-emitting diodes (LED)-backlit computer screen affects circadian physiology and cognitive performance. J Appl Physiol (1985) 2011 May;110(5):1432-1438 [FREE Full text] [doi: 10.1152/japplphysiol.00165.2011] [Medline: 21415172]
14. Thomée S. Mobile phone use and mental health. A review of the research that takes a psychological perspective on exposure. Int J Environ Res Public Health 2018 Nov 29;15(12):2692. [doi: 10.3390/ijerph15122692] [Medline: 30501032]
15. Hale L, Guan S. Screen time and sleep among school-aged children and adolescents: a systematic literature review. Sleep Med Rev 2015 Jun;21:50-58. [doi: 10.1016/j.smrv.2014.07.007] [Medline: 25193149]
16. Exelmans L, Van den Bulck J. Bedtime mobile phone use and sleep in adults. Soc Sci Med 2016 Jan;148:93-101. [doi: 10.1016/j.socscimed.2015.11.037] [Medline: 26688552]
17. Rod NH, Dissing AS, Clark A, Gerds TA, Lund R. Overnight smartphone use: a new public health challenge? A novel study design based on high-resolution smartphone data. PLoS ONE 2018 Oct 16;13(10):1-12. [doi: 10.1371/journal.pone.0204811] [Medline: 30325929]
18. Tavernier R, Willoughby T. Sleep problems: predictor or outcome of media use among emerging adults at university? J Sleep Res 2014 Feb 19;23(4):389-396. [doi: 10.1111/jsr.12132] [Medline: 24552437]
19. Bartel K, Scheeren R, Gradisar M. Altering adolescents' pre-bedtime phone use to achieve better sleep health. Health Communication 2018 Jan 09;34(4):456-462. [doi: 10.1080/10410236.2017.1422099] [Medline: 29313721]
20. 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;10:161 [FREE Full text] [doi: 10.1186/1471-2288-10-67] [Medline: 20637084]
21. Beekley MD, Brechue WF, deHoyos DV, Garzarella L, Werber-Zion G, Pollock ML. Cross-validation of the YMCA submaximal cycle ergometer test to predict VO2max. Res Q Exerc Sport 2004 Sep;75(3):337-342. [doi: 10.1080/02701367.2004.10609165] [Medline: 15487296]
22. Danquah IH, Petersen CB, Skov SS, Tolstrup JS. Validation of the NPAQ-short—a brief questionnaire to monitor physical activity and compliance with the WHO recommendations. BMC Public Health 2018 May 08;18(1):601 [FREE Full text] [doi: 10.1186/s12889-018-5538-y] [Medline: 29739383]
23. Bastien CH, Vallières A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. Sleep Med 2001 Jul;2(4):297-307. [Medline: 11438246]
24. Global recommendations on physical activity for health. Geneva: World Health Organization; 2010. URL: https://apps.who.int/iris/rest/bitstreams/52834/retrieve [accessed 2020-12-02]
25. Astrand I. Aerobic work capacity in men and women with special reference to age. Acta Physiol Scand Suppl 1960;49(169):1-92. [doi: 10.1016/b978-1-4831-9819-4.50022-4] [Medline: 13794892]

26. Ross R, Blair SN, Arena R, Church TS, Després J, Franklin BA, American Heart Association Physical Activity Committee of the Council on LifestyleCardiometabolic Health, Council on Clinical Cardiology, Council on Epidemiology and Prevention, Council on Cardiovascular and Stroke Nursing, Council on Functional Genomics and Translational Biology, Stroke Council. Importance of assessing cardiorespiratory fitness in clinical practice: a case for fitness as a clinical vital sign: a scientific statement from the American Heart Association. Circulation 2016 Dec 13;134(24):e653-e699. [doi: 10.1161/CIR.0000000000000461] [Medline: 27881567]
27. Lewington S, Clarke R, Qizilbash N, Peto R, Collins R, Prospective SC. Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. Lancet 2002 Dec 14;360(9349):1903-1913. [Medline: 12493255]
28. Stuckey MI, Carter SW, Knight E. The role of smartphones in encouraging physical activity in adults. Int J Gen Med 2017;10:293-303 [FREE Full text] [doi: 10.2147/IJGM.S134095] [Medline: 28979157]
29. Harris A, Gundersen H, Mørk-Andreassen P, Thun E, Bjorvatn B, Pallesen S. Restricted use of electronic media, sleep, performance, and mood in high school athletes—a randomized trial. Sleep Health 2015 Dec;1(4):314-321. [doi: 10.1016/j.sleh.2015.09.011] [Medline: 29073407]
30. He J, Tu Z, Xiao L, Su T, Tang Y. Effect of restricting bedtime mobile phone use on sleep, arousal, mood, and working memory: a randomized pilot trial. PLoS ONE 2020 Feb 10;15(2):e0228756. [doi: 10.1371/journal.pone.0228756] [Medline: 32040492]
31. Hughes N, Burke J. Sleeping with the frenemy: how restricting ‘bedroom use’ of smartphones impacts happiness and wellbeing. Comput Human Behav 2018 Aug;85:236-244. [doi: 10.1016/j.chb.2018.03.047]
32. Lastella M, Rigney G, Browne M, Sargent C. Electronic device use in bed reduces sleep duration and quality in adults. Sleep Biol Rhythms 2020 Jan 3;18(2):121-129. [doi: 10.1007/s41105-019-00251-y]
33. Fobian AD, Avis K, Schwebel DC. Impact of media use on adolescent sleep efficiency. J Dev Behav Pediatr 2016 Jan;37(1):9-14 [FREE Full text] [doi: 10.1097/DBP.0000000000000239] [Medline: 26651090]
34. Tavernier R, Heissel JA, Sladek MR, Grant KE, Adam EK. Adolescents' technology and face-to-face time use predict objective sleep outcomes. Sleep Health 2017 Aug;3(4):276-283. [doi: 10.1016/j.sleh.2017.04.005] [Medline: 28709515]
35. Sadeh A. The role and validity of actigraphy in sleep medicine: an update. Sleep Med Rev 2011 Aug;15(4):259-267. [doi: 10.1016/j.smrv.2010.10.001] [Medline: 21237680]
36. Christensen MA, Bettencourt L, Kaye L, Moturu ST, Nguyen KT, Olgin JE, et al. Direct measurements of smartphone screen-time: relationships with demographics and sleep. PLoS One 2016;11(11):e0165331 [FREE Full text] [doi: 10.1371/journal.pone.0165331] [Medline: 27829040]

Abbreviations

BP: blood pressure
ISI: Insomnia Severity Index
MVPA: moderate-to-vigorous physical activity
NPAQ-short: Nordic Physical Activity Questionnaire-short
PA: physical activity
SS: sleep silent
T1: day 1
T2: day 2
TO: time offline
TST: total sleep time
VO_2max: maximal oxygen uptake
WHO: World Health Organization
WTD: wearable tracking device

Edited by G Eysenbach; submitted 18.05.20; peer-reviewed by A Jo, S Hayes, S Blair; comments to author 21.08.20; revised version received 15.10.20; accepted 28.10.20; published 04.01.21.

Please cite as:
Larsen LH, Lauritzen MH, Gangstad SW, Kjaer TW
The Use of Small Electronic Devices and Health: Feasibility of Interventions for a Forthcoming Crossover Design
JMIR Form Res 2021;5(1):e20410
URL: https://formative.jmir.org/2021/1/e20410
doi:10.2196/20410
PMID:33393907

Original Paper

Cancer Patients' Perceived Value of a Smartphone App to Enhance the Safety of Home-Based Chemotherapy: Feasibility Study

Nina Kongshaug[1,2], MD; John-Arne Skolbekken[3], PhD; Arild Faxvaag[4], MD, PhD; Eva Hofsli[2], MD, PhD

[1]Cancer Clinic, St. Olav's University Hospital, Trondheim, Norway

[2]Department of Clinical and Molecular Medicine, Faculty of Medicine and Health Scienes, Norwegian University of Science and Technology, Trondheim, Norway

[3]Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, Trondheim, Norway

[4]Department of Neuromedicine and Movement Science, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, Trondheim, Norway

Corresponding Author:
Nina Kongshaug, MD
Cancer Clinic
St. Olav's University Hospital
Postbox 8905
Trondheim, 7491
Norway
Phone: 47 99703704
Email: nina.kongshaug@stolav.no

Abstract

Background: Oral anticancer therapies can be self-administered by patients outside the hospital setting, which poses challenges of adherence to a drug plan and monitoring of side effects. Modern information technology may be developed and implemented to address these pertinent issues.

Objective: The aim of this study was to explore how a smartphone app developed through a stepwise, iterative process can help patients using oral chemotherapy to take their drug, and to report adherence and side effects in a reliable and verifiable manner.

Methods: Fourteen patients starting capecitabine treatment were included in this study and used the smartphone app in addition to regular follow up of capecitabine treatment. Nine of these patients fulfilled the treatment plan and were interviewed based on a semistructured interview guide and the System Usability Scale (SUS). In addition, two focus groups were completed with 7 oncologists and 7 oncology nurses, respectively. Interview data were analyzed in accordance with the principles of systematic text condensation. Features of the app were also assessed.

Results: The smartphone app provided the patients with a feeling of reassurance regarding correct adherence of their oral chemotherapy treatment. They used the app as a memory tool about their treatment and possible serious side effects, as well as for treatment education. Patients expressed concerns about using the app to report side effects that were not considered to be obviously serious, fearing overreporting. The health personnel expressed an overall positive attitude to integrate this new tool in their everyday work.

Conclusions: Patients on oral chemotherapy treatment at home felt safe and found the app to be helpful. The app promoted learning about their treatment and made the patients more independent of the cancer clinic, reducing the need for the clinic's limited resources for follow up of patients on oral anticancer medications.

(JMIR Form Res 2021;5(1):e20636) doi:10.2196/20636

KEYWORDS

mhealth; mobile app; smartphone app; oral chemotherapy; patient safety; home-based cancer treatment

Introduction

Medical cancer treatment has changed from traditional intravenous chemotherapy given at hospitals to home-based oral anticancer therapies [1]. Oral chemotherapy can potentially be administered without the patient having to visit a cancer clinic given the advantages of avoiding the need for intravenous administration and its associated complications. Nevertheless, adherence is a widely known challenge for cancer patients on oral chemotherapy [2]. Given that patients will adhere to the drug plan and that side effects can be handled properly, such home-based therapy will make cancer patients more independent of the cancer clinic and also have potential to save costs [1,3]. Traditionally, the follow up of oral treatment adherence and side effects when patients are at home is done by a phone call to the patient and appointments at the clinic.

Vincent [4] defines patient safety as the management of risk over time to maximize benefit and minimize harm to patients in the health care system. Patient safety is an aspect of diagnostic services (eg, diagnostic safety [5]), therapy (eg, medication safety [6]), and health care coordination (eg, errors of commission and errors of omission [7]). Because patient safety is such an essential property of a health care system, health institutions are obliged to assess, monitor, and continuously work to improve the patient safety aspects of their services.

The context of this study is the safety of home-based therapy with capecitabine, an oral chemotherapeutic drug that is used in the treatment of gastrointestinal and breast cancers. Capecitabine has a wide dosage range and potential serious side effects such as diarrhea, hematologic toxicity, and hand-foot syndrome, although fatigue, nausea, and stomatitis are also frequently reported [8]. In recent years, the cancer clinic at St. Olav's University Hospital in central Norway experienced hospital admissions due to diarrhea followed by acute renal failure as severe adverse events of capecitabine use, two of which were fatal. To improve the safety of home-based capecitabine treatment, the clinic immediately changed to a stricter dispensing and monitoring regimen, including the use of pill dispensers and follow-up phone calls. The clinic also developed a smartphone app to be used for reminding the patients to take the drug and to report side effects. This idea is also supported by a recent article describing that educating patients with timely medical information through their smartphones improves patient knowledge, treatment adherence, and clinical outcomes [9].

Approximately two-thirds of all people worldwide own a mobile phone [10]. This has created an ecosystem for mobile health (mHealth), the practice of medicine and public health supported by mobile devices [11]. Among their many prospects, mHealth apps offer the possibility for health care institutions to reach out to and interact with patients staying at home [10,12]. An mHealth taxonomy developed in 2015 described eight different use cases: point-of-care diagnostics, patient monitoring, wellness, compliance, education and reference, behavior modification, efficiency and productivity, and environmental monitoring [13]. There are examples of mHealth apps within oncology, including tools for point-of-care diagnostics (eg, melanoma diagnostic services) and tools for assessing patient-reported outcomes [14,15]. mHealth tools have been shown to increase medication adherence in patients with diabetes [16], but literature of their effects on adherence to oral anticancer therapies is lacking [10].

For an mHealth app to have an impact as a patient safety tool, it must be taken into use and perceived as useful by a majority of the patients in the target group [17]. We therefore sought to explore patients' use of an app from an institutional perspective (ie, patient safety) as well as from the perspective of the patient (eg, perceived usefulness). The objective of this study was to explore how a smartphone app can assist patients in adhering to the capecitabine medication plan and for reporting side effects, and to also characterize the main features that the patients find to be most useful within the app.

Methods

Study Design

We performed a feasibility study with 14 cancer patients and 14 health care providers. Patients, physicians, and nurses were recruited at the cancer clinic of St. Olav's University Hospital in central Norway in the period of March to October 2017. Nine of the 14 patients completed the test period and subsequently underwent a semistructured interview. The reasons for the 5 patients not completing the test period were as follows: capecitabine discontinued due to side effects (n=2), follow up by an oncologist outside St. Olav's University Hospital (n=1), technical problems with downloading the app to the patient's smartphone (n=1), and insufficient smartphone competence (n=1).

The oncology nurses assisted patients in downloading the app and setting up the treatment plan on their smartphones. In addition, two focus groups were completed with 7 oncologists and 7 oncology nurses, respectively. The main characteristics of the participants are summarized in Table 1.

Table 1. Demographic and treatment characteristics of the patients in this study (N=9).

Characteristic	Value
Gender, n (%)	
Men	6 (67)
Women	3 (33)
Age (years), n (%)	
40-49	2 (22)
50-59	4 (44)
60-69	2 (22)
70-79	1 (11)
Oncology treatment plan, n (%)	
Chemoradiotherapy (radiation plus concomitant capecitabine)	6 (67)
Intravenous chemotherapy every 3rd week plus capecitabine	3 (33)
Smartphone system, n (%)	
IOS (iPhone)	6 (67)
Android (Samsung)	3 (33)
Level of education	
College/university	7 (78)
High school	2 (22)

All patients were chemotherapy-naive, had a gastrointestinal cancer, and were indicated for capecitabine treatment. Inclusion criteria were patients >18 years of age with the ability to independently manage their medication and having good knowledge of how to use a smartphone (ie, used their phone for more than text messages and phone calls). The patients installed the app on their personal smartphone when they started the capecitabine treatment. All patients used the smartphone app in addition to regular follow up.

Intervention

The smartphone app prototype was developed by a stepwise, iterative process in a multiprofessional group from St. Olav's University Hospital and Norwegian University of Science and Technology (NTNU) in cooperation with information and technology communication system developers and designers, facilitated by the Technology Transfer Office of NTNU. The app is based on knowledge about capecitabine treatment and the current procedures for monitoring of these patients at the cancer clinic of St. Olav's University Hospital. The source for the side effect component was Common Toxicity Criteria, version 4.03 [18]. Before starting the feasibility study, the prototype version of the app was tested on 10 colleagues to ensure acceptable usability.

The app has two main features: (1) supporting adherence to the medication (Figure 1) and (2) management and reporting of side effects (Figure 2). The app alerts and reminds the patient to take the drug at the right time and offers a calendar visualization of the medication plan. The patients can register side effects, and the app provides a patient decision support system to call the nurse at the cancer clinic if needed. The prototype also provides a summary of all side effects registered in each capecitabine treatment cycle.

Figure 1. Features of the app for supporting medication adherence.

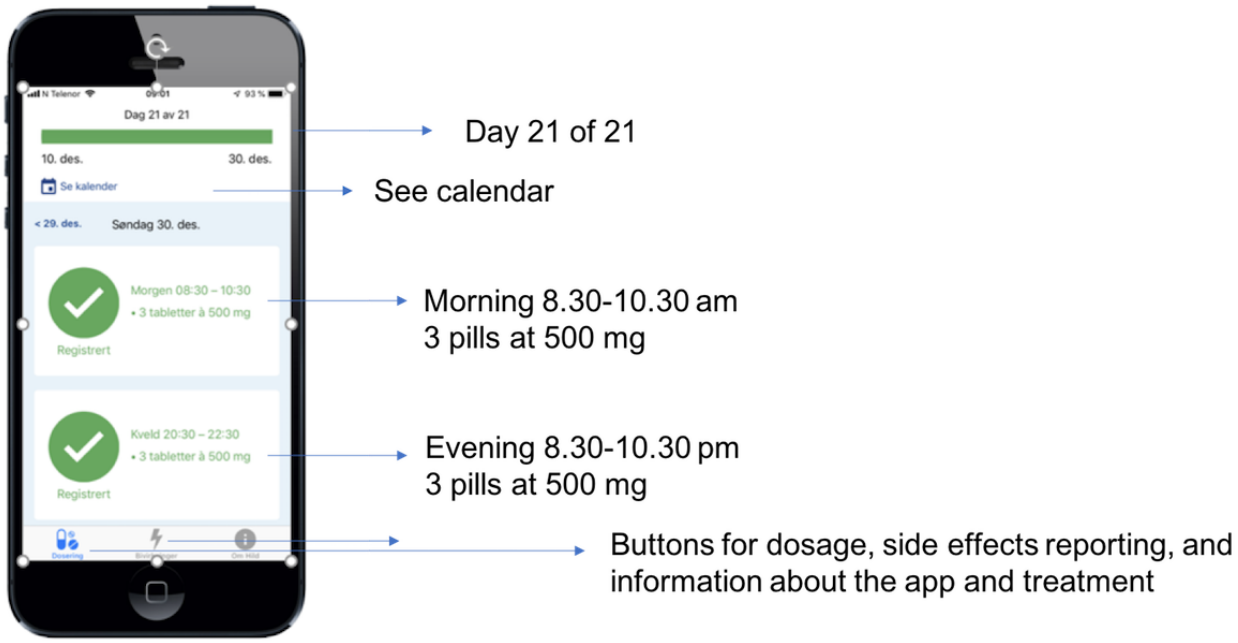
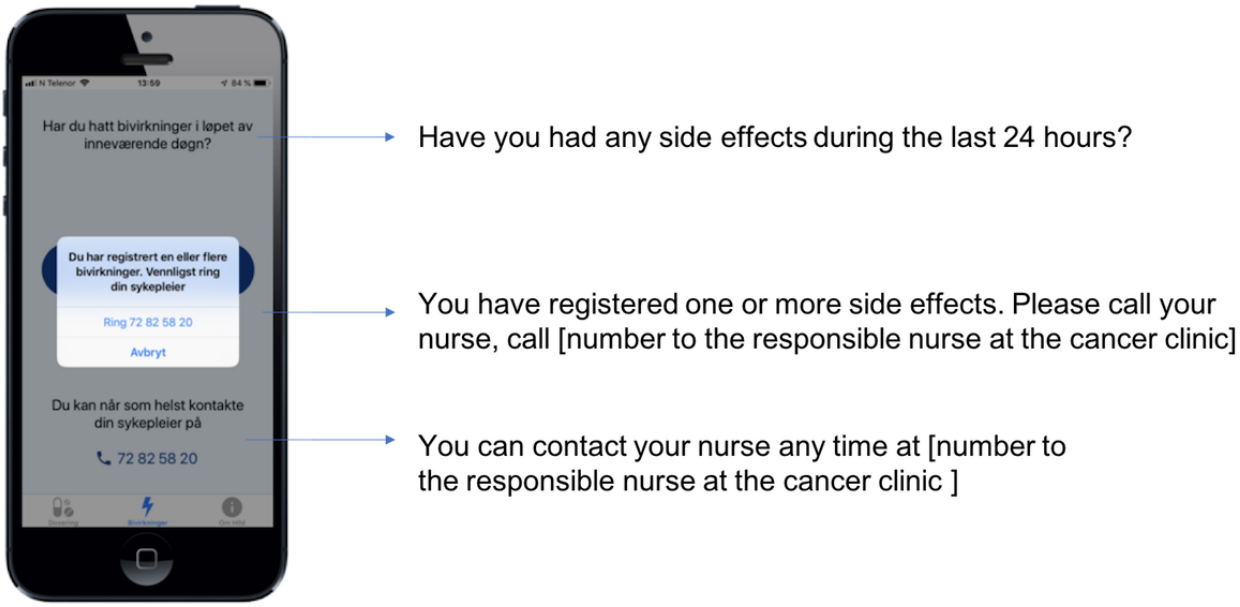


Figure 2. Features of the app for supporting management and reporting of side effects.



Measures

Data were collected through semistructured interviews and with the System Usability Scale (SUS) as well as through two focus groups [19-21]. The SUS is considered to be an easy, quick, and reliable test of usability that is technology-agnostic, which is also available in a validated Norwegian translated version [20]. The questionnaire was not designed for statistical use in this study but was rather used as a starting point for the semistructured interviews.

Data Collection

After 2 to 3 weeks on capecitabine treatment, a semistructured interview was performed with the patient. All patients were interviewed once. The interview guide focused on the patients' experiences with the use of the app regarding correct capecitabine adherence and reporting of important side effects, experiences of safety of home-based chemotherapy treatment, and the possibility to obtain adequate help from health personnel when needed. The patients also provided information on a Norwegian validated version of the SUS questionnaire. This information was used as a starting point for the interviews. Patient interviews lasted from 12 to 25 minutes. The patient interviews took place at the hospital in an undisturbed room. The focus groups (with physicians and nurses) took place in a meeting room at the cancer clinic and lasted 60 minutes each. One of the authors acted as group moderator. An external

researcher was co-moderator and took field notes. The focus group conversations covered the same topics as the patient interviews, but from a health personnel perspective. All interviews were digitally recorded and subsequently transcribed. Finally, the transcriptions were controlled against the recording.

Data Analysis

Data were analyzed according to the principles of systematic text condensation [22]. This procedure consisted of four steps: (1) getting a total impression by reading all of the text materials and identifying preliminary themes; (2) identifying meaning units from both the technical aspects of the smartphone app and its use by patients, oncology physicians, and nurses; (3) abstracting condensates from each group and subgroup; and (4) creating synthesized descriptions of the patients', oncology physicians', and nurses' experiences and opinions about the use of the smartphone app in the follow up of patients on capecitabine treatment. To some extent, we performed a stepwise analysis before completing data collection.

Ethical Considerations and Approval

Participants provided informed consent based on oral and written information about the study and its purpose. The patients used the smartphone app as a supplement to regular follow up with pill dispensers and phone calls from the cancer clinic. All interviews were audiotaped and transcribed without any identifiable information so as to preserve the participants' confidentiality. The soundtracks were deleted after transcription. Data stored on the patients' smartphones were secured by a pin code. In this version of the app, the treatment plan was set up by a nurse on the patient's personal phone, protected with a pin code unknown to the patient.

The Regional Committee for Medical and Health Research Ethics, South East, Norway confirmed that their approval was not required for this study (REK 2015/1581). The study was approved by the Norwegian Centre for Research Data and the Data Protection Officer for both the NTNU and St. Olav's University Hospital.

Results

Overall Perspectives

The patients reported that the app aided them in adhering to the drug plan through reminders to take the drug and self-reporting of drug usage and side effects. The app also served as a memory aid, enabling them to learn more about the drug they were taking, and provided reassurance. Health personnel at the cancer clinic were concerned about the balance between making themselves more available to the patients and being able to handle the anticipated increase in the number of requests. When the study participants visited the outpatient clinic, nurses and physicians (with some exceptions) did not check the medication history via the app. Patients reported far fewer side effects than anticipated.

Perceived Safety

One of the main findings of this study was that patients who used the app felt safer. Both the alerts on when to take the drug and which dose to take contributed to this feeling:

> *..you get an alarm on your mobile. You always have your mobile with you, it is a safety net. This was the greatest benefit with it.*

Despite being instructed otherwise, some patients came to believe that the information they recorded on the app was shared with the clinic without delay. Unsurprisingly, the thought of having health personnel continuously monitoring their treatment and eventual side effects increased their sense of safety:

> *...so you feel that you are better followed-up [by the cancer clinic]. You know that if you register [the data], that someone will see it. It probably gives a better feeling of safety.*

However, there was no such feature on the app version that the patients were testing. Hence, the use of the app made the patients believe they were being followed up more closely by the clinic than they actually were.

Improved Memory and Interaction With the Clinic

The patients appreciated having access to a correct and always updated phone number to a nurse in the clinic:

> *Even if I have good control of my [information] sheets, it [the mobile app] is easier and more available. I had something to report about side effects, and they were there when you touched the screen, and then you get a phone number, and I got in touch with a nurse immediately… That helps a lot.*

According to the nurse informants, this contrasted with previous patient reporting, where they spent whole days waiting for a nurse to call at day 3, 10, and 17 in the treatment cycle. Even if they did not have bothersome symptoms, they focused on the call at those specific days:

> *Many patients are at home all day, waiting for that phone call.*

Patients that recorded their medication history on the app reported that they used these recordings to recount the details of their experiences with taking the drug:

> *For instance, when you are on chemotherapy, your memory is not as good as before you became ill, so it's a benefit that you record if the side effects started on Tuesday or Thursday. It's a nice aid… Because, there is a relation between the [treatment] doses, and then it's easier to understand.*

Hence, the app gave the patients an overview and a deeper insight into their side effect profile, which seemed to support a richer and more purposeful interaction with the clinic.

Learning Promotion and Independence From the Clinic

According to the patients, the overview of serious side effects was always readily available on their mobile device, and it was quicker to open the app than having to find the information sheet provided by the cancer clinic. Some patients regularly used the app's side effect component to assess their own side effects and decide whether to report the side effects to the clinic.

> *I have been into this side effect part [of the app] several times and assessed whether these are symptoms I have or not. So, I haven't had side effects which should be reported [to the cancer clinic].*

This finding was supported by the nurses who focused on the patients' opportunity to act more independently while treated at home as the app assisted them in managing their own treatment and conceivable side effects. The clinicians also perceived the app as strengthening patients' adherence to the right dosage and helping them to become more responsible for their own treatment.

Suggestions for Improvement

The patients wanted to share information about adherence and side effects with their nurse and doctor:

> *I don't know if this [app] will be connected to the electronic health record at the hospital or something like that? … Then I think it really can be useful, when the physicians can follow the adherence as well as the side effects. I think everything about surveillance and follow up is a good thing.*

They also wanted an overview of all their hospital appointments integrated in the app, including receiving short messages if any of their appointments were changed:

> *But, there is something about the administration of the letters that we receive about appointments. They could have been dropped. Could have used the app instead. I think there are many opportunities here.*

The patients expressed that they were ready for more digital communication than was available through this app and welcomed use of internet and smartphone tools for cancer treatment follow up.

Reporting Side Effects

The physicians focused on the risk of information overload and how to filter what they needed to know and act on versus what not to engage in, given that the patients with the app could report on side effects whenever they wanted.

> *We walk around with [smart] watches and measurements of blood pressure… Why do we need all this information?… We need to have the information which impacts on the cancer treatment.*

The nurses also emphasized the need for a good system for monitoring the patient registrations at the hospital, including when to act on them. However, they also focused on how the app could help patients take more responsibility for their cancer-related symptoms and treatment.

In its present design, the app did not allow for direct transfer of side effect reports. Instead, the app encouraged the patients to call the hospital whenever they experienced side effects that the clinic should be made aware of. However, patients were reluctant to use this function.

> *I had skin symptoms, then I saved and kept going to the next one, but then I got the message that said I should call my nurse, and... God, maybe I shouldn't have done that?*

As a result, patients reported far fewer side effects than anticipated, in contrast to the clinicians' fear of information overload.

Discussion

Principal Findings

In this feasibility study, we have shown that cancer patients can use a smartphone app to be reminded to take a drug and report on their adherence to a cytostatic drug regimen in a reliable manner. Despite the fact that the app enabled reporting of side effects and offered side effect–specific advice, the app obviously failed to make patients comply with the hospitals' guidelines for immediate reporting of serious side effects and adaptive adjustment of the therapeutic regimen.

Adherence is a known challenge for cancer patients on oral chemotherapy [2]. Patients regarded the drug-take reminding function useful and believed that it improved adherence. This observation is in line with those of previous studies that have explored the effects of drug-reminder apps in other clinical domains [23]. The drug-take reporting function of the app points toward a more comprehensive documentation of pharmaceutical interventions in oncology. However, whether the app actually increases adherence to the drug needs to be tested in a randomized clinical trial.

Patients using the app reported that they learned more about their treatment and that this made them less dependent on the cancer clinic. This might imply that an app can be an important supplement to the follow up by health care providers of cancer patients on oral anticancer treatment. This is in line with the results of Kessel et al [24] who showed that health-related quality of life reporting from oncological patients through a mobile app was accepted by patients.

An overall effect of the app was that it made the patients feel safer by working as a proxy for the clinic. The app offered the patients reassurance, assuming that they were very closely monitored by the cancer clinic despite being informed that the study version of the app did not have any feature allowing for automatic communication with the clinic. This effect was not intended and is an example of an unintended positive effect of health information technology [25]. In our study, all of the patients received standard follow up in addition to the app, and therefore there were no related ethical or patient safety issues. The next version of the app will be connected to the hospital network, enabling clinicians to follow up on patient-generated reports in a population health manner [26,27]. The ability of an app to provide reassurance to patients that suffer from a chronic, potentially life-threatening disease could increase patients' adherence to the app and hence limit the well-known problem of user attrition [28,29]. This line of thought will be explored in future designs of the app.

In addition to objective parameters such as blood tests, correct reporting of side effects is a key for optimizing chemotherapy dosage [30-32]. Despite potential benefits, there are both technological and administrative challenges with integrating side effect reporting into practice [33]. We found that the side effect reporting function in the app served as a source of

knowledge about side effects, but that the coupling between registering side effects and the following of rapid advice directly from the clinic often made the patients refrain from reporting. Taken together, the side effect reports failed to give a complete picture of what the patients were experiencing. This fear of reporting side effects could also be due to a fear of cessation of medication, possibly affecting their treatment negatively [34]. Motivated by the possibility of the side effect reporting, further work on the design of the side effect component of the app is needed, focusing on balancing the patients' needs and understanding of reporting, as well as the health personnel's needs to avoid information overload. With these aspects in mind, the new design of the side effect component of the app should allow the patients to register nuanced grading of side effects in which only specified severe side effects triggers an alert to the cancer clinic. Further work also includes a design change toward the clinicians' need to find all patient data as a part of the electronic health record.

Strengths and Limitations

Despite consistent findings in the patient interviews in this study, the small number of patients is a limitation to be overcome with future research. Another weakness is that the physicians, with a few exceptions, did not use the app in their daily work. This may be due to the fact that all of the data were stored on the patients' private smartphones and that many of the physicians in a busy workday did not know who was included in the study and subsequently omitted to ask the patients.

The context of the use of this app differs from most mHealth apps that are oriented toward achieving wellness, as this is about illness and all potential dangers associated with having cancer and being exposed to risky therapies [35]. To our knowledge, this is the first study that indicates an app's impact on the feeling of reassurance while using potentially toxic cancer medication. These results also provide a more complete picture of the adherence and side effects than we recently obtained with phone calls to the patients on specific days during the treatment schedule.

Conclusion

The growing number of new oral anticancer therapies encourages new thinking of the follow-up routines for this specific patient group. In conclusion, this app can be a helpful tool for supporting patients in the home-based part of their cancer treatment. The app must meet both patients' and clinicians' needs, but the patients' and clinicians' requirements for usefulness are not necessarily identical.

Acknowledgments

This work was supported by the Technology Transfer Office, NTNU, Norway. The authors would like to thank the patients and the health care professionals at the cancer clinic of St. Olav's University Hospital, central Norway for making this work possible. In addition, we extend appreciation to Rickard-Johan Skjong for his services as a co-moderator during the focus group discussions.

Conflicts of Interest

None declared.

References

1. Burhenn PS, Smudde J. Using tools and technology to promote education and adherence to oral agents for cancer. Clin J Oncol Nurs 2015 Jun;19(3 Suppl):53-59. [doi: 10.1188/15.S1.CJON.53-59] [Medline: 26030395]
2. Kawakami K, Yokokawa T, Kobayashi K, Sugisaki T, Suzuki K, Suenaga M, et al. Self-Reported Adherence to Capecitabine on XELOX Treatment as Adjuvant Therapy for Colorectal Cancer. Oncol Res 2017 Nov 02;25(9):1625-1631. [doi: 10.3727/096504017X15012905098071] [Medline: 28766482]
3. Knottnerus JA, Tugwell P. Prioritization of health care and research given limited and too limited resources. J Clin Epidemiol 2017 Jun;86:1-2. [doi: 10.1016/j.jclinepi.2017.06.018] [Medline: 28705426]
4. Vincent R, Amalberti R. Safer Healthcare: Strategies for the Real World. Cham: Springer; 2016.
5. Singh H, Graber ML, Kissam SM, Sorensen AV, Lenfestey NF, Tant EM, et al. System-related interventions to reduce diagnostic errors: a narrative review. BMJ Qual Saf 2012 Feb;21(2):160-170 [FREE Full text] [doi: 10.1136/bmjqs-2011-000150] [Medline: 22129930]
6. Allan EL, Barker KN. Fundamentals of medication error research. Am J Hosp Pharm 1990 Mar;47(3):555-571. [Medline: 2180287]
7. Shalom E, Shahar Y, Parmet Y, Lunenfeld E. A multiple-scenario assessment of the effect of a continuous-care, guideline-based decision support system on clinicians' compliance to clinical guidelines. Int J Med Inform 2015 Apr;84(4):248-262. [doi: 10.1016/j.ijmedinf.2015.01.004] [Medline: 25649843]
8. Masuda N, Lee S, Ohtani S, Im Y, Lee E, Yokota I, et al. Adjuvant Capecitabine for Breast Cancer after Preoperative Chemotherapy. N Engl J Med 2017 Jun 01;376(22):2147-2159. [doi: 10.1056/NEJMoa1612645] [Medline: 28564564]
9. Timmers T, Janssen L, Kool RB, Kremer JA. Educating Patients by Providing Timely Information Using Smartphone and Tablet Apps: Systematic Review. J Med Internet Res 2020 Apr 13;22(4):e17342 [FREE Full text] [doi: 10.2196/17342] [Medline: 32281936]

10. Skrabal Ross X, Gunn KM, Patterson P, Olver I. Mobile-Based Oral Chemotherapy Adherence-Enhancing Interventions: Scoping Review. JMIR Mhealth Uhealth 2018 Dec 21;6(12):e11724 [FREE Full text] [doi: 10.2196/11724] [Medline: 30578182]
11. Adibi S. Mobile Health: A Technology Road Map. Cham: Springer; 2015.
12. Kauw D, Koole MAC, van Dorth JR, Tulevski II, Somsen GA, Schijven MP, et al. eHealth in patients with congenital heart disease: a review. Expert Rev Cardiovasc Ther 2018 Sep;16(9):627-634. [doi: 10.1080/14779072.2018.1508343] [Medline: 30079780]
13. Olla P, Shimskey C. mHealth taxonomy: a literature survey of mobile health applications. Health Technol 2015 Jan 30;4(4):299-308. [doi: 10.1007/s12553-014-0093-8]
14. Chao E, Meenan CK, Ferris LK. Smartphone-Based Applications for Skin Monitoring and Melanoma Detection. Dermatol Clin 2017 Oct;35(4):551-557. [doi: 10.1016/j.det.2017.06.014] [Medline: 28886812]
15. Benze G, Nauck F, Alt-Epping B, Gianni G, Bauknecht T, Ettl J, et al. PROutine: a feasibility study assessing surveillance of electronic patient reported outcomes and adherence via smartphone app in advanced cancer. Ann Palliat Med 2019 Apr;8(2):104-111. [doi: 10.21037/apm.2017.07.05] [Medline: 29156896]
16. Ryan EA, Holland J, Stroulia E, Bazelli B, Babwik SA, Li H, et al. Improved A1C Levels in Type 1 Diabetes with Smartphone App Use. Can J Diabetes 2017 Feb;41(1):33-40. [doi: 10.1016/j.jcjd.2016.06.001] [Medline: 27570203]
17. Han A, Min S, Ahn S, Min S, Hong H, Han N, et al. Mobile medication manager application to improve adherence with immunosuppressive therapy in renal transplant recipients: A randomized controlled trial. PLoS One 2019;14(11):e0224595 [FREE Full text] [doi: 10.1371/journal.pone.0224595] [Medline: 31689320]
18. Common Terminology Criteria for Adverse Events (CTCAE) Version 4.0. US Department of Health and Human Services, National Institutes of Health, National Cancer Institute. 2009 May 28. URL: https://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_4.03/CTCAE_4.03_2010-06-14_QuickReference_8.5x11.pdf [accessed 2020-12-20]
19. Brinkmann S, Kvale S. Interviews: Learning the craft of qualitative research interviewing 3rd edition. Newbury Park, CA: Sage Publishing; 2015.
20. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usab Stud 2009;4(3).
21. Krueger R, Casey M. Focus groups: A practical guide for applied research 5th edition. Newbury, CA: Sage Publishing; 2014.
22. Malterud K. Systematic text condensation: a strategy for qualitative analysis. Scand J Public Health 2012 Dec;40(8):795-805. [doi: 10.1177/1403494812465030] [Medline: 23221918]
23. Park JYE, Li J, Howren A, Tsao NW, De Vera M. Mobile Phone Apps Targeting Medication Adherence: Quality Assessment and Content Analysis of User Reviews. JMIR Mhealth Uhealth 2019 Jan 31;7(1):e11919 [FREE Full text] [doi: 10.2196/11919] [Medline: 30702435]
24. Kessel KA, Vogel MM, Alles A, Dobiasch S, Fischer H, Combs SE. Mobile App Delivery of the EORTC QLQ-C30 Questionnaire to Assess Health-Related Quality of Life in Oncological Patients: Usability Study. JMIR Mhealth Uhealth 2018 Feb 20;6(2):e45 [FREE Full text] [doi: 10.2196/mhealth.9486] [Medline: 29463489]
25. Kuziemsky CE, Borycki E, Nøhr C, Cummings E. The nature of unintended benefits in health information systems. Stud Health Technol Inform 2012;180:896-900. [Medline: 22874322]
26. Dyck LA, Snelling S, Morrison V, Haworth-Brockman M, Atkinson D. Equity reporting: a framework for putting knowledge mobilization and health equity at the core of population health status reporting. Health Promot Chronic Dis Prev Can 2018 Mar;38(3):116-124. [doi: 10.24095/hpcdp.38.3.02] [Medline: 29537769]
27. Kruse CS, Stein A, Thomas H, Kaur H. The use of Electronic Health Records to Support Population Health: A Systematic Review of the Literature. J Med Syst 2018 Sep 29;42(11):214 [FREE Full text] [doi: 10.1007/s10916-018-1075-6] [Medline: 30269237]
28. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar 31;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
29. Van der Mispel C, Poppe L, Crombez G, Verloigne M, De Bourdeaudhuij I. A Self-Regulation-Based eHealth Intervention to Promote a Healthy Lifestyle: Investigating User and Website Characteristics Related to Attrition. J Med Internet Res 2017 Jul 11;19(7):e241 [FREE Full text] [doi: 10.2196/jmir.7277] [Medline: 28698168]
30. Basch E. Patient-reported outcomes: an essential component of oncology drug development and regulatory review. Lancet Oncol 2018 May;19(5):595-597. [doi: 10.1016/S1470-2045(18)30141-4] [Medline: 29726373]
31. Basch E. The missing voice of patients in drug-safety reporting. N Engl J Med 2010 Mar 11;362(10):865-869 [FREE Full text] [doi: 10.1056/NEJMp0911494] [Medline: 20220181]
32. Laugsand EA, Sprangers MAG, Bjordal K, Skorpen F, Kaasa S, Klepstad P. Health care providers underestimate symptom intensities of cancer patients: a multicenter European study. Health Qual Life Outcomes 2010 Sep 21;8:104 [FREE Full text] [doi: 10.1186/1477-7525-8-104] [Medline: 20858248]
33. Basch E, Barbera L, Kerrigan CL, Velikova G. Implementation of Patient-Reported Outcomes in Routine Medical Care. Am Soc Clin Oncol Educ Book 2018 May 23;38:122-134 [FREE Full text] [doi: 10.1200/EDBK_200383] [Medline: 30231381]

34. Odeh B, Kayyali R, Nabhani-Gebara S, Philip N. Optimizing cancer care through mobile health. Support Care Cancer 2015 Jul;23(7):2183-2188. [doi: 10.1007/s00520-015-2627-7] [Medline: 25649121]
35. 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]

Abbreviations

mHealth: mobile health
NTNU: Norwegian University of Science and Technology
SUS: System Usability Scale

Edited by G Eysenbach; submitted 07.06.20; peer-reviewed by A Miguel-Cruz, C Reis; comments to author 10.08.20; revised version received 05.10.20; accepted 09.11.20; published 06.01.21.

Please cite as:
Kongshaug N, Skolbekken JA, Faxvaag A, Hofsli E
Cancer Patients' Perceived Value of a Smartphone App to Enhance the Safety of Home-Based Chemotherapy: Feasibility Study
JMIR Form Res 2021;5(1):e20636
URL: https://formative.jmir.org/2021/1/e20636
doi:10.2196/20636
PMID:33404505

Original Paper

Improving Primary Care Medication Processes by Using Shared Electronic Medication Plans in Switzerland: Lessons Learned From a Participatory Action Research Study

Benjamin Bugnon[1,2], MSc, PharmD; Antoine Geissbuhler[3], MD, PhD; Thomas Bischoff[2], MD; Pascal Bonnabry[1], PharmD, PhD; Christian von Plessen[2,4,5], MD, PhD

[1]Institute of Pharmaceutical Sciences of Western Switzerland, University of Geneva, Geneva, Switzerland

[2]Direction Générale de la Santé, État de Vaud, Lausanne, Switzerland

[3]Department of Radiology and Medical Informatics, University of Geneva, Geneva, Switzerland

[4]Department of Clinical Research, University of Southern Denmark, Odense, Denmark

[5]Center for Primary Care and Public Health, Unisanté, Lausanne, Switzerland

Corresponding Author:
Benjamin Bugnon, MSc, PharmD
Institute of Pharmaceutical Sciences of Western Switzerland
University of Geneva
Rue Michel-Servet 1, CMU Bâtiment B
Geneva, 1211
Switzerland
Phone: 41 796250180
Email: benjamin.bugnon@gmail.com

Abstract

Background: Several countries have launched health information technology (HIT) systems for shared electronic medication plans. These systems enable patients and health care professionals to use and manage a common list of current medications across sectors and settings. Shared electronic medication plans have great potential to improve medication management and patient safety, but their integration into complex medication-related processes has proven difficult, and there is little scientific evidence to guide their implementation.

Objective: The objective of this paper is to summarize lessons learned from primary care professionals involved in a pioneering pilot project in Switzerland for the systemwide implementation of shared electronic medication plans. We collected experiences, assessed the influences of the local context, and analyzed underlying mechanisms influencing the implementation.

Methods: In this formative action research study, we followed 5 clusters of health care professionals during 6 months. The clusters represented rural and urban primary care settings. A total of 18 health care professionals (primary care physicians, pharmacists, and nurses) used the pilot version of a shared electronic medication plan on a secure web platform, the precursor of Switzerland's electronic patient record infrastructure. We undertook 3 group interviews with each of the 5 clusters, analyzed the content longitudinally and across clusters, and summarized it into lessons learned.

Results: Participants considered medication plan management, digitalized or not, a core element of good clinical practice. Requirements for the successful implementation of a shared electronic medication plan were the integration into and simplification of clinical routines. Participants underlined the importance of an enabling setting with designated reference professionals and regular high-quality interactions with patients. Such a setting should foster trusting relationships and nurture a culture of safety and data privacy. For participants, the HIT was a necessary but insufficient building block toward better interprofessional communication, especially in transitions. Despite oral and written information, the availability of shared electronic medication plans did not generate spontaneous demand from patients or foster more engagement in their medication management. The variable settings illustrated the diversity of medication management and the need for local adaptations.

Conclusions: The results of our study present a unique and comprehensive description of the sociotechnical challenges of implementing shared electronic medication plans in primary care. The shared ownership among multiple stakeholders is a core challenge for implementers. No single stakeholder can build and maintain a safe, usable HIT system with up-to-date medication information. Buy-in from all involved health care professionals is necessary for consistent medication reconciliation along the

entire care pathway. Implementers must balance the need to change clinical processes to achieve improvements with the need to integrate the shared electronic medication plan into existing routines to facilitate adoption. The lack of patient involvement warrants further study.

(JMIR Form Res 2021;5(1):e22319) doi:10.2196/22319

KEYWORDS

shared electronic medication plan; medication list; medication reconciliation; electronic health records; primary care; national eHealth strategy; Switzerland; participatory action research; complex adaptive system; eHealth; medication; health information technology; implementation

Introduction

Medication processes are crucial for improving patient outcomes, and at the same time, medication-related errors are one of the main causes of the overall burden associated with adverse events [1]. Only 4% to 21% of patients receive the optimum benefits from their medication use [2]. Avoidable adverse drug events account for approximately 5% of hospital admissions [3]. According to the World Health Organization, a more responsible use of medicines could save up to US $42 billion annually worldwide by reducing medication-related harm [4,5].

Medication errors occur in all health care settings, but more commonly in ambulatory care [6] and during care transitions because of the loss or incomplete transfer of information about patients' medications [7-9]. About 55% of patients risk having one or more unexplained differences in their documented treatment plans across different health care services [10]. The problem is ubiquitous and has an impact on patients and health care systems globally [7], including in Switzerland [11-13].

Medication reconciliation (MedRec), the process of creating and managing the most accurate list of the medications that a patient is taking [14], can prevent such events at interfaces of care. However, it is difficult because medication regimens are increasingly complex [15] and multiple disparate actors are involved [16,17]. It is perhaps unsurprising that despite significant efforts to implement it, MedRec is still only progressing slowly in many countries [17]. For example, in Switzerland, systematic MedRec has only been tested in a few pilot projects and has yet to be implemented across the whole country [11].

Health care organizations have invested in health information technology (HIT) systems to address these difficulties [18]. Such systems should help overcome insufficient access to up-to-date information, low efficiency, and organizational issues [17]. Moreover, they should help reduce stress among patients and workloads among staff caused by lack of information while avoiding risky workarounds and improving the quality of care [19]. Although the great potential for HIT investment is acknowledged internationally, approaches and strategies vary [20].

Several countries have launched HIT systems for shared electronic medication plans, which allow multiple health care professionals to use and manage their common patient's current list of medications [20-22]. The core information in a shared electronic medication plan system is made up of the clinical decisions related to the treatment plan, such as adding, adapting, or stopping medications. The architecture used for a shared electronic medication plan system (eg, in Denmark [23]) contrasts with that of other systems that automatically calculate a patient's current medication list from dispensing and e-prescribing databases (eg, France [24], Ireland [25], Netherlands [26]), but not all clinical or self-care decisions necessarily end up on paper, in an electronic prescription, or in dispensing notes. The latter automatic systems, therefore, appear limited in terms of information accuracy, whereas a digital shared medication plan fundamentally relies on the system's joint and regular use in clinical practice to ensure consistently reconciled medication information along the patient's entire care pathway.

Implementing HIT systems for shared medication plans is challenging. System usability and its integration into clinical workflows is essential for medication list accuracy [23,27,28]. Attention should be paid to clinical and administrative workflows and system design [22,29,30] as well as to easily accessible information technology and clinical support [31,32]. The need to clarify professionals' responsibilities has often been raised [28,33,34]. Similarly, introducing a predefined process for using and managing patients' shared medication plans has been claimed as a solution [30,31,33]. In addition, trust must be built into the system by making the shared information reliable [27,28,34] and ensuring the privacy and security of data [28,32,35]. Unfortunately, evidence-based strategies for implementing such a system cannot be derived from these often heterogeneous and highly contextual studies.

However, these studies have illustrated the sociotechnical nature and complexity of implementing a digital shared medication plan. Systemwide HIT implementation projects should embrace this complexity and consider strategic, managerial, and social aspects in addition to technological challenges [36-38]. One approach is using formative research to create collaborative learning opportunities in these complex situations [39-41]. Such research aims to interpret and understand the potential effects of an HIT implementation project rather than predict them. Insights into the key mechanisms affecting the success or failure of complex programs of change, such as the implementation of shared electronic medication plans, can support stakeholders as they seek to build on local experiences.

With this in mind, we designed a formative action research study of a pioneering Swiss pilot project using shared electronic medication plans on an eHealth platform. We aimed to produce practical knowledge for use in the implementation of shared electronic medication plans on a larger scale. The study

objectives were to learn from the local experience of 5 clusters of primary care professionals, assess the influences of context, and describe related mechanisms in order to achieve the efficient use of digital shared medication plans for safer, more effective patient care.

Methods

Design

This formative participatory action research (PAR) study followed 5 clusters of health care professionals over 6 months using 3 interviews per cluster and a model to guide an iterative inductive thematic analysis.

PAR is a collective, self-reflective investigation undertaken by researchers and participants together [42]. It connects actions influenced by context, culture, and history and is embedded in social dynamics. The strengths of PAR are responsiveness to context, the engagement of frontline health care professionals, and a focus on the mechanisms of implementation that can help bring about real-world service improvements [43].

Throughout the successive meetings with participants, we followed the iterative process proposed by Loewenson et al [44], using the steps of systematizing experience, collectively analyzing and problematizing, reflecting on and choosing an action, taking and evaluating action, and systematizing learning. We invited each group of participants to define their collective commitments at the first meeting. The reflective process was stimulated by asking questions such as “What is going on?” “How do we continue?” and “What are our main lessons learned?” We also ensured that all the lessons learned that were documented by researchers were proposed for further discussion or refinement.

Context

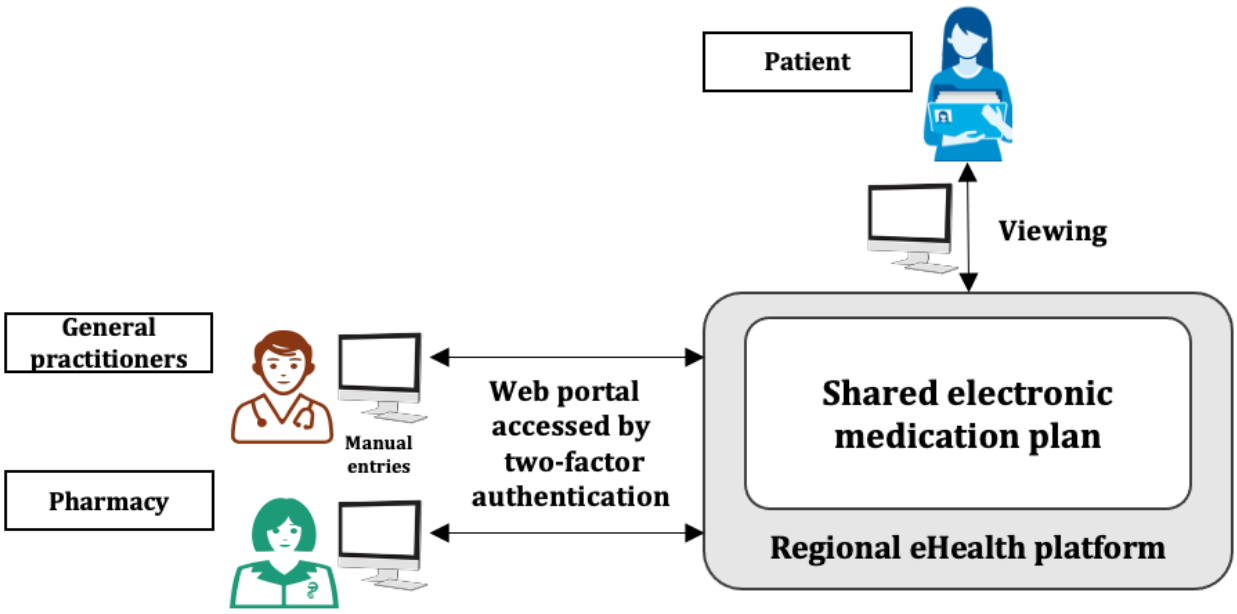
The study was embedded in the pilot project for the implementation of shared electronic medication plans on the regional eHealth platform for the Nord Broye region in the canton of Vaud [45]. We recruited health care professionals into local study clusters from among the 36 general practitioners (GPs) and 36 pharmacies who had cared for the 193 patients participating in the pilot project from 2013 to 2018 (Figure 1). Primary care professionals were free to enroll in the pilot project led by the regional network for care coordination, which was sponsored by their respective corporation and public authority. Patients using at least three medications regularly were invited to join the pilot project’s medication management program. Care professionals communicated to patients directly, while the pilot project team provided leaflets and information online. They nominated a GP and a pharmacy as reference points to manage their medications, and they committed to consulting and procuring their medication only from them while sharing all necessary information completely.

The digital solution chosen for the shared medication plans was an online platform for creating, using, and managing a list of all the medications a patient was taking and had taken in the past [22,46]. A shared electronic medication plan must be accessible, complete, and updated at every contact between the patient and an intervening health care professional. Technologically, this solution was envisioned as an interoperable system based on the pharmacy profiles defined by the Integrating the Healthcare Enterprise (IHE) consortium [47]. During the pilot project, the definition of a national e-medication interoperability standard based on IHE pharmacy HL7 CDA was under development (interprofessional working group from 2015, recommendation published in 2017) [22,48]. The users accessed the shared medication plan through a secure web portal with two-factor authentication. Patients could access their medication plan online or receive a printed one. Professionals had to enter all data manually in addition to filling out the usual paper documentation because their clinical software applications were not yet integrated.

The solution is a module of the web platform developed for the cantons of Geneva and Vaud in anticipation of Switzerland’s electronic patient record (EPR) system, a national digital inventory of all the relevant health data concerning the country’s patients [22]. The EPR is based on decentralized information exchange infrastructure. Several regional platforms have been implemented that are run by private or public entities but overseen nationally by the federal law of 2017 [49]. Patients own their data and share them with health care professionals of their own free will. Primary care physicians are free to choose whether they want to join the EPR (opt in), whereas all hospitals are obliged to be connected. Swiss national policy acknowledges the importance of e-medication [50,51], but an overall strategy has not yet been defined.

Switzerland’s political culture is liberal, and the 26 cantons of the federal state have far-reaching autonomy regarding the organization of health care [52]. Patients can access the health care professionals or specialist physicians of their choice. GPs only have a gatekeeper function in some optional insurance plans. As of 2020, there are no shared patient registers. Among the member states of the Organisation for Economic Co-operation and Development (OECD), Switzerland has below-average digital maturity [53], and 30% of its GPs still use paper-based patient records, far behind their colleagues from the European Union, of whom only 4% rely on paper. Finally, there are no regulatory or other specific incentives for the vendors of medical or pharmaceutical record software.

Figure 1. Description of the shared electronic medication plan system used in the regional pilot project. The web portal was not integrated with usual systems used by care professionals, as the national standard for e-medication based on IHE Pharmacy HL7 CDA was a work in progress during the pilot.



Theoretical and Conceptual Framework

Organizational innovations in health care often fail because the complexity and adaptability of the health system are underestimated [54-56]. Recognizing health care as a complex adaptive system (CAS) means focusing on the dynamic interactions between individuals and organizations across the entire system. When seeking to initiate change in a CAS, sensemaking and learning about it are more important than planning and controlling the change itself. By definition, a CAS is unpredictable, but some simple rules (ie, guiding principles), can help foster transformation [55,57-59].

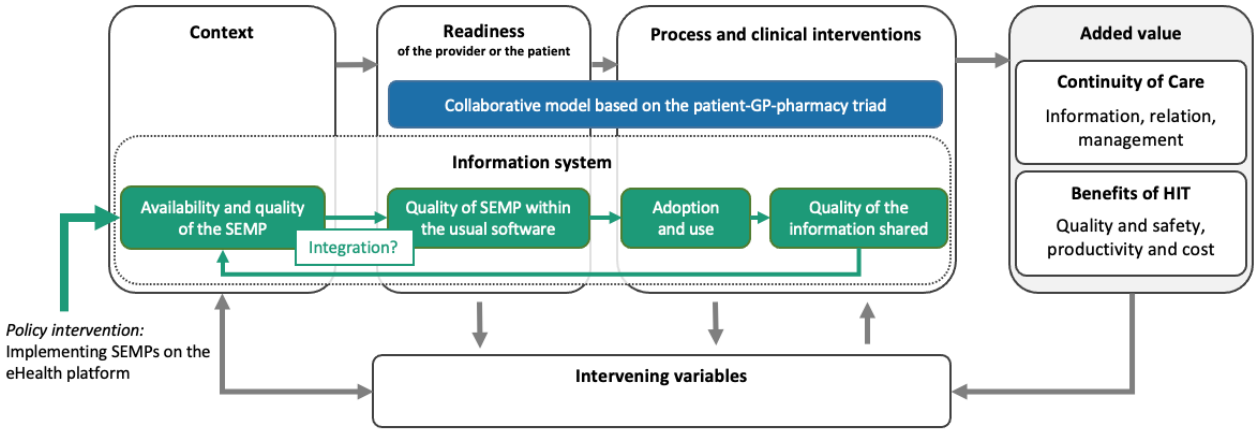
Creating a common electronic patient health record incorporating a shared medication plan is in itself a complex sociotechnical intervention; the technological component of the intervention is influenced by and influences every user's behavior, as well as the organization and context [60]. We developed a model (Figure 2) of how different elements of the implementation of the shared electronic medication plan might be linked to expected results.

We based our model on Lilford et al's [61] approach to mapping policy and service interventions with regard to structures, processes, outcomes, and intervening variables. In the pilot study, the introduction of the shared electronic medication plan system affected the participating health care organization structures and required adaptions of their work processes. Some participants also combined the shared medication plan use with other clinical interventions, such as medication review. All these elements as a whole system led to health care outcomes. Although we did not want to predefine the intended outcomes of safer and better patient care, we did specify that the continuity of care, claimed as a main policy ambition for the pilot project, should be not only at the informational level but also at the relational and management levels, as per Haggerty et al's [62] definition. It is also essential to consider the intervening variables, as they are interrelated with the structural and process factors mediating the outcomes. For instance, a patient's trust in their pharmacy and its staff (intervening variable) is influenced by the availability of a space in the pharmacy where they can talk in confidence (structure), whether a dedicated pharmacist follows up with a chronic patient (generic process), how information is given when dispensing a pillbox (clinical process), and the consequent safe use of medicines (outcomes). In the present study, the main intervention is at the policy level: implementing shared electronic medication plans on the eHealth platform in the region. Our study sought to leverage health care professionals' experiences to assess contextual influences from a systemic perspective. For this reason, the model specifies both the context and the readiness of the provider or the patient, as structural factors can be respectively external or internal of the health care providers.

The information system itself was added to the model as a transversal dimension, based on the eHealth Clinical Adoption framework defined by Lau et al [63]. Those authors described how the successful adoption and benefits of HIT depend on its quality. The overall quality of HIT is made up of the qualities of the system, the service, and the information available. For a shared record system, because the quality of information is made up of shared content, it is strongly dependent on the quality of usage. This is why our model illustrates the interrelation of the perceived quality of the shared electronic medication plan system, the quality of its usage, and the quality of the shared content as distinct dimensions. Finally, the model describes the shared electronic medication plan's overall added value in terms of the improved elements in the continuity of care and the benefits of HIT.

All the dimensions in our model helped us to break down and make sense of the implementation of the HIT, potential interventions, and the points requiring study. The two essential new elements brought in by the pilot project were the addition of a shared electronic medication plan system onto an existing eHealth platform and a new collaborative model based on the patient-GP-pharmacy triad to make primary care medication management safer in cases involving polypharmacy.

Figure 2. Proposed model for the implementation of a shared electronic medication plan system. GP: general practitioner; HIT: health information technology; SEMP: shared electronic medication plan.



Recruitment in the Study, Sampling, and Ethics

Invitations were sent to the group of 36 GPs and 36 pharmacies who had been enrolled in the pilot project. From those who volunteered, we created clusters consisting of at least one pharmacist from an enrolled pharmacy and one GP who were responsible for at least one common patient enrolled in the pilot project. Each cluster could invite other primary health care professionals involved in their local settings, such as a home care nurse. We characterized each cluster by urbanization density classification [64]. All participants consented according to the canton of Vaud's legal, privacy, and ethical requirements. No patient data were collected.

Data Collection and Analysis

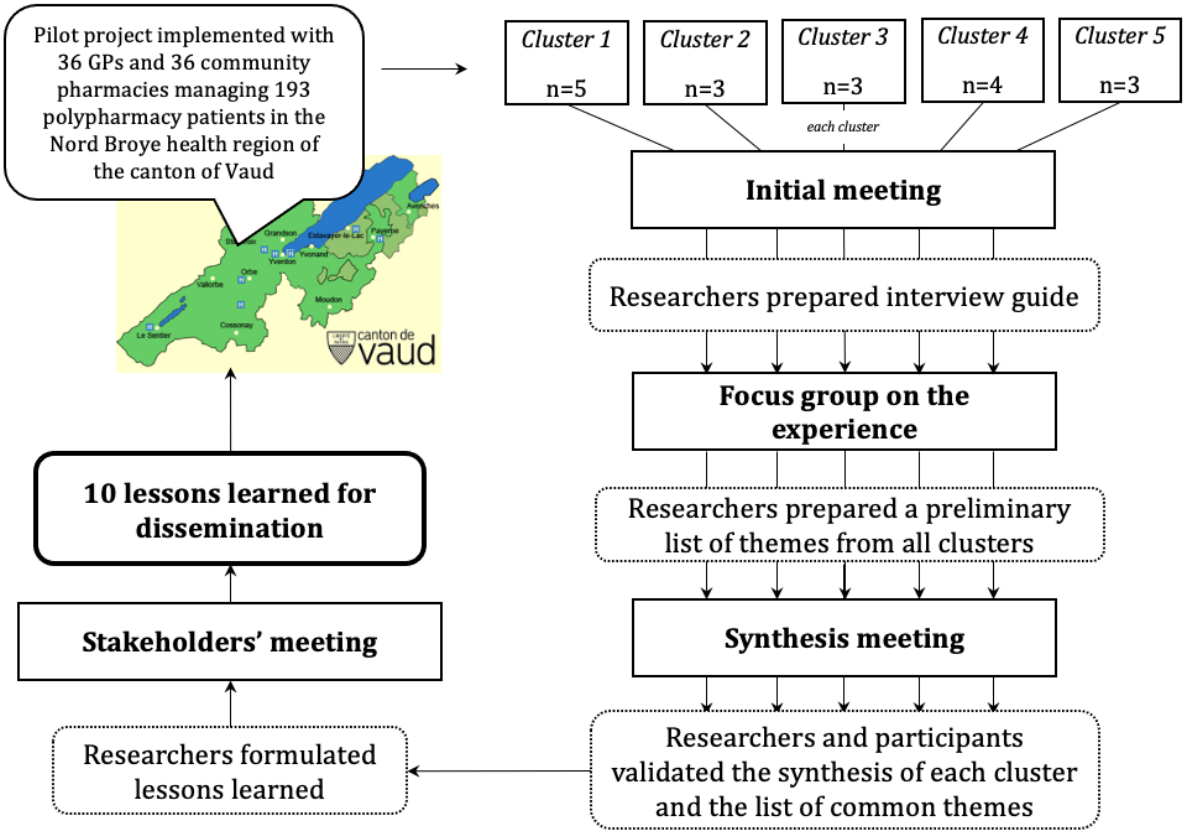
Two researchers, a pharmacist and GP respectively, both with research training and experience, collected and analyzed the data (Figure 3). They were not enrolled in the pilot project (ie, did not count in clusters) but were familiar with the settings, and they knew some of the participants professionally.

Between May 2018 and January 2019, each cluster participated in 3 group interviews whose main topics of investigation were, respectively, (1) motivations and commitment, (2) experience and refinement, and (3) synthesis and learnings. Interview guides (Multimedia Appendix 1) for each round were prepared using the conceptual model as a basis. The investigators guided participants toward thinking about the added value of the eHealth platform and the collaborative model of care as 2 interdependent components associated with the implementation of patients' shared medication plans. Participants were first encouraged to share and reflect on their experiences of initiating and managing the medication plan, using the platform, interacting with patients about medication lists, and collaborating with other professionals. These experiences then nurtured discussions on the contextual or organizational factors influencing implementation and on the role of shared electronic medication plans in achieving safer, more effective care. The investigators facilitated exploration of the different themes that emerged from each group in order to increase diversity across clusters. During the last cluster meeting, participants also discussed and summarized the most important practical knowledge that should be disseminated to stakeholders involved in the future development and scale-up of shared electronic medication plan systems.

Data collection and analysis were iterative so that each group's experiences could be collected and synthesized longitudinally and data across clusters could be analyzed horizontally to condense them into themes. Finally, we conducted a secondary thematic content analysis to condense the themes into lessons learned, or lessons intended to describe the simple rules underlying the mechanisms related to the implementation in a CAS.

Following the principles of PAR, we proposed refinements to and requested validation from participants at each step of the study. Furthermore, we presented the lessons learned at a stakeholders meeting, which included participants and representatives of other stakeholders in the regional pilot project. This gave time for discussions and dialogue on setting prioritized next steps. We also attempted to enhance the reliability of our research by having data analyzed by the 2 main researchers and then by a researcher outside the pilot project [65].

Figure 3. Data collection and analysis. GP: general practitioner.



Results

Participants

Among the 36 GPs and 36 pharmacies who were enrolled in the pilot project, 31 volunteered for this action research study. A total of 5 clusters were identified, including 13 care professionals. Consequently, 18 volunteers had no patients in common with any other volunteer professionals and thus were excluded in this study. The 13 participants invited 5 extra primary health care professionals into their clusters, for a total of 18 participants (Table 1). We conducted 15 group interviews that lasted 60 to 105 minutes.

Table 1. Characteristics of the 5 clusters.

Cluster	Cluster location	Participants
1	Town (semidense)	GP[a], 2 pharmacists from different pharmacies, a medical secretary specialized in care coordination, and a home care nurse
2	Town (semidense)	GP also working in local hospital emergency unit and 2 pharmacists from different pharmacies
3	Rural area (dispersed)	GP, pharmacist, and independent nurse in GP practice
4	University center for primary care in a city	GP, 2 pharmacists, and scientific collaborator
5	City	GP, pharmacist, and home care nurse, all responsible for a nursing home

[a]GP: general practitioner.

Lessons Learned

An overview of the lessons learned is presented in Textbox 1.

Textbox 1. Lessons learned.

Lessons learned, to be used in the strategy for the systemwide implementation of shared electronic medication plans improving primary care medication processes

1. Information sharing during clinical routines must be simplified and secured by integrating shared electronic medication plans into existing processes and health information technology systems.
2. A medication plan, whether digital or not, is a matter of good clinical practice.
3. Designating reference professionals ensures the exhaustivity and continuity of the medication information communicated.
4. Regular high-quality interactions between patients and professionals strengthen the continuity of medication plan management.
5. Implementing a new tool, ensuring good clinical practice, and increasing interactions for coordination require more resources and an adapted organizational model.
6. The availability of the shared electronic medication plan did not generate spontaneous demand from patients or foster more engagement in their medication management.
7. Adopting a shared electronic medication plan is triggered by a culture of patient safety and data privacy.
8. Fostering trusting relationships at all levels is essential.
9. Legal, financial, and governance framework conditions influence the uptake and impact of shared electronic medication plans.
10. A shared electronic medication plan is a necessary building block of communication about medication, especially at transitions, but it is not a sufficient one.

Lesson No. 1: Information Sharing During Clinical Routines Must Be Simplified and Secured by Integrating Shared Electronic Medication Plans Into Existing Processes and HIT Systems

Participants consistently emphasized the need to integrate the shared electronic medication plan system into their usual electronic medical records systems and pharmacy management systems:

> *Its integration into my usual software is crucial to simplifying my work.* [GP, cluster 3]
>
> *The workflow is sometimes intense… and we are a team… only integration can enable reliable information sharing on any contact with the patient.* [Pharmacist, cluster 5]

During the pilot project, participants had to document the medication-related decisions in both the shared system and their usual patient record system. They feared this double documentation could cause errors, and they expressed frustration about redundant work:

> *For the small number of patients we are following [about 10], it's okay, but we couldn't do it properly for every patient without a certain degree of automatization and integration with our usual system.* [Pharmacist 1, cluster 2]

The shared electronic medication plan system's overall good usability and integration with current clinical software was considered a sine qua non for meaningful implementation.

Participants highlighted integration issues as crucial, and they deplored their dependence on their software vendors to better integrate the shared electronic medication plan in their own system. They were critical of the national strategy, which foresees standards of interoperability but leaves system integration to market forces:

> *As clients, we are captives of our medical software vendor. What can you [the public administration] do to leverage integration?* [GP, cluster 3]

Poor current levels of competition in the market for medical records or pharmacy systems was also mentioned as a barrier to integration.

Lesson No. 2: A Medication Plan, Whether Digital or Not, Is a Matter of Good Clinical Practice

Participants proposed that professional attitudes and clinical work processes were even more important than HIT systems for improving medication management:

> *You [the investigators] are working to set up a great, relevant system…but we could likely do better with a less sophisticated tool.…Working with a medication plan should be a matter of good practice!* [GP, cluster 1]

During group discussions, participants mentioned that prescribing drugs without a holistic view of all the medications a patient is taking and communication of the current medication plan to other health care professionals involved were both not uncommon.

Notwithstanding that health care professionals are legally responsible for the safe use of medications, the responsibilities for creating, maintaining, and communicating medication plans were not always clear. Multiple physicians write prescriptions, but they do not always maintain an overview of the patient's entire list of medications, which risks causing the patient serious problems. For example, participants revealed that some older patients accumulated numerous medications from various prescribers with no awareness of the potential for drug-drug interactions. It was argued that procedures, standards, or even regulatory actions were needed to clarify responsibilities, regardless of the implementation of any new HIT systems.

Lesson No. 3: Designating Reference Professionals Ensures the Exhaustivity and Continuity of the Medication Information Communicated

Participants recognized that formalized roles and relationships between patients and their GPs and pharmacists improved the exchange of information about medications. During the pilot project, patients registered with one GP and one pharmacy as health care professional reference points and committed to sharing all their medication-related information with them. This was a major change from the usual practice in Swiss health care:

> *When she enrolled, one of our patients informed us that she received a neuroleptic drug from a specialist by post. None of us knew! She is a typical polypharmacy patient who regularly comes to us for her medicines.…We did not expect that from her at all.* [Pharmacist, cluster 3]

Formalizing these associations led to more accurate medication lists through improved relational continuity and clear channels of communication with other health care providers and professionals, such as hospitals or specialist physicians. These formal reference persons were also seen as key facilitators during the scaling-up transition period from multiple sources of medication information to one systematically used shared electronic medication plan system.

Lesson No. 4: Regular High-Quality Interactions Between Patients and Professionals Strengthen the Continuity of Medication Plan Management

Whereas the shared electronic medication plan improves documentation and information exchange, the validity and relevance of information about medications depend on the quality and regularity of the interactions between patients and professionals:

> *After some time…after doing regular reviews and interacting with the patient,…that's how you get to know—when the trust is built—the things that matter to them, their worries…and they may even confess how they really manage their medication! From there, you can really care for them and support them on their pathway.* [Nurse, cluster 3]

To illustrate this point, participants mentioned common activities, such as medication reviews, the identification of side effects, and the evaluation of adherence or support for administration. These interventions could also serve as important checkpoints for the accuracy of the medication plan. Furthermore, participants suggested that associating the implementation of shared electronic medication plans with these other important activities could accelerate their adoption.

Lesson No. 5: Implementing a New Tool, Ensuring Good Clinical Practice, and Increasing Interactions for Coordination Require More Resources and an Adapted Organizational Model

Not all primary health care professionals are equally ready to adapt their daily clinical practice for better patient follow-up and coordination activities. Although a shared electronic medication plan system has the potential to increase efficiency, the adoption capacity of providers depends on the availability of competent staff, flexibility, adequate facilities, and an effective organizational model:

> *With the pharmacy team, we have participated in several pilot projects on new services.…We hired an extra pharmacist…but the ones [ie, other pharmacies and their staff] that do not invest will likely not manage to evolve and will struggle more with the regular follow-up of patients who do have a [shared electronic medication plan].…* [Pharmacist, cluster 5]

> *We have now agreed on how we proceed with patients who are followed by the practice [from the cluster] and come to the pharmacy after hospital discharge…and that the nurse provides communication if there is a change.* [Pharmacist, cluster 3]

They highlighted that the introduction of new roles and competencies, such as the medical secretary specialized in care coordination, the independent nurses in GP practices, or the clinical pharmacist for pharmaceutical care, was still at the early stage of development in the region and that the financing model was not yet well established.

Lesson No. 6: The Availability of the Shared Electronic Medication Plan Did Not Generate Spontaneous Demand From Patients or Foster More Engagement in Their Medication Management

Participants reported that very few patients showed interest in exploring or using the shared electronic medication plan. Oral and written information given out at project inclusion and through promotional flyers in the waiting areas of GPs' practices, in pharmacies, or online had not seemed to make a difference. Some speculated about explanations for this apparent lack of interest:

> *Some young and some elderly [declined access to the web portal]. It did not seem to be a matter of age, even if there were some technological barriers in some cases.* [Nurse, cluster 3]

> *They accepted [participating in the project] because I stated that it would be good for them.* [GP, cluster 2]

Patients seemed to have a limited understanding of the processes of medication management and had difficulties viewing its potential in terms of improvements to quality and safety. Accordingly, the rationale for the shared electronic medication plan and how it functioned remained obscure to them:

> *When we came to this patient, with all these forms, to ask him if he'd sign to agree that his regular GP and pharmacy—who he'd known for a long time—could communicate about his medication…he was like, "How come? You do that usually, don't you?" He was very surprised and kind of worried!* [Pharmacist 2, cluster 4]

Apparently, this patient had taken it for granted that reasonable communication processes existed between his GP and his

pharmacy. He had not been aware of the regulatory and practical barriers to sharing health-related information.

Study participants further argued that the intention behind the design of the shared electronic medication plan system had not been to engage patients:

> *The medication plan could also be a tool for extra interventions with the patient, like patient education, but it can also just be simply printed from our software.…At the moment,…the [shared electronic medication plan] isn't designed as a specific tool to foster patient engagement.* [Pharmacist 1, cluster 4]

Thus, to date, patients have not been considered active participants in their medication management, and the HIT system was not designed to foster patient empowerment.

Lesson No. 7: Adopting a Shared Electronic Medication Plan Is Triggered by a Culture of Patient Safety and Data Privacy

Participants noted the ambiguity between sharing health-related information to improve medication management and safety and the need for data privacy and confidentiality:

> *It's a question of balancing benefits and risks. Chronic patients with polypharmacy are more likely to benefit and realize its importance.* [GP, cluster 5]

The fear of privacy related to digital technology, often fueled in the media, can hinder adoption. Participants pointed to the need to address habits and culture during the shared electronic medication plan system's implementation:

> *The use of shared records is essential for medication safety, but this challenges habits and perceived responsibilities, especially among older generations of doctors. This cultural shift should be supported, and it should start with new doctors during their education.* [GP, cluster 4]

Transparent evaluation was identified as a means of demonstrating the clinical benefits and nurturing a dialogue on privacy and patient safety.

Lesson No. 8: Fostering Trusting Relationships at All Levels Is Essential

Participants repeatedly highlighted the importance of trust in the implementation of the shared electronic medication plan:

> *It [the shared electronic medication plan's use by the patient-GP-pharmacy triad] should be based on trust.* [Pharmacist, cluster 1]

Trust between patients and professionals is required for medication plans to have any value; trust between professionals fosters information exchange; and trust between HIT providers, health care professionals, and the state facilitates implementation. Trust in the HIT can be diminished by breaches of confidentiality and the misguided implementation of eHealth systems. Conversely, participants appreciated the present study's collaborative design because it fostered trusting relationships among them:

> *It [participating in the study] brought us around the table, gave us time to get to know each other and discuss.…Although we regularly interact, it is always brief.* [Pharmacist, cluster 1]

> *It [participating in the study] helped to reach a better mutual understanding and create a climate of collaboration.* [GP, cluster 2]

Lesson No. 9: Legal, Financial, and Governance Framework Conditions Influence the Uptake and Impact of Shared Electronic Medication Plans

Group discussions repeatedly mentioned the crucial importance of the legal, governance, and financial conditions surrounding medication management. Questions were raised about the mandatory or facultative use of the shared electronic medication plan system, its legal status, and different users' legal responsibilities in the case of adverse events, discrepancies, and incompleteness:

> *If the [shared electronic medication plan's] use were mandatory by law [for all health care professionals], at least then I'd think that we could rely on it more.…If not, you will always wonder if it is complete or not.…You're supposed to trust the list, not just consider if it's the truth or not when you are making decisions.…* [GP, cluster 2]

In the pilot project launching phase in particular, concerns were raised that a lack of professional adherence would impede scale-up:

> *I need to be sure the plan is complete and updated.…If not, I won't use it. But if everyone avoids using it for the same reason,…no one will ever update it.* [GP, cluster 5]

Indeed, the participants were divided about whether to make the shared electronic medication plan mandatory. Some emphasized the legitimacy of an official status, arguing for mandatory participation for all health care professionals. Others advocated for a more specific strategy to enhance the involvement of health care professionals, for example, via financial incentives for both patients and health care professionals when they signed up for a collaborative model of care.

Participants were concerned about the shared electronic medication plan system's governance and how their active involvement to manage it would be financed. Here, they perceived the liberal approach to organizing Switzerland's health care to be a major challenge:

> *It is important to clarify the roles and responsibilities.…But who should decide?* [Nurse, cluster 1]

The current model of reimbursement for health care professionals' activities was also considered a barrier because of its poor financial incentives for collaborative care management activities and the lack of consistency among reimbursement models:

> *Updating the plan, making sure it is complete; explaining; answering questions the patient may have—it all takes time! But to date, we are not directly paid for this.…The negotiations with the health assurance companies [ie, the payers in the system] are going to be complicated.* [Pharmacist, cluster 3]

One example of the inconsistency in health care professionals' payments is the support for medication management:

> *Patients we have known for a long time…suddenly disappear because the GP calls for homecare services to follow-up. They prepare the pillbox at the dining table while chatting with the patient or the family—there is a much higher risk of errors than in a secure double-checked process in the pharmacy. We often know the patient's preferences, their habits, history, story,…but we are not involved anymore.* [Pharmacist 1, cluster 1]

Even when there is a local consensus, the financial reality is that:

> *For homecare services, it's the way they are financed for entering the home to better assess and follow-up a situation that is getting more complicated.…The psychosocial support is not really reimbursed.…"* [Nurse, cluster 1]

The current reimbursement system for coordinating activities (especially in complex cases), reviewing medication, and supporting patients with their medication use and adherence was perceived to be a hindrance to regular, in-depth updating of the shared electronic medication plan. Switzerland's general governance and financing systems for health care services may themselves pose a challenge to safe and meaningful scale-up of shared electronic medication plans.

Lesson No. 10: A Shared Electronic Medication Plan System Is a Necessary Building Block of Communication About Medication, Especially at Transitions, but It Is Not a Sufficient One

While participants appreciated the shared electronic medication plan as a useful building block in a system for medication management, they cautioned that communication problems, especially during transitions, were much broader:

> *All the issues related to care transitions go beyond the scope of medication information.…You need to take into account many factors to adapt care, starting from the patients' pathways and their specific medical conditions.* [GP, cluster 2]

For example, participants mentioned that few hospital units had properly implemented MedRec and that the introduction of shared electronic medication plans alone would not directly change that.

Participants lamented the lack of standardized communication, especially between the GPs or the pharmacy and home care services or hospitals:

> *Actually,…communication between the pharmacy and the doctor works pretty well. We work with the prescriptions, and sometimes we call each other if needed,…but the main issues are with the multiple homecare services organizations operating for our patients.…Even public organizations work in different ways [eg, medication management, communication of lists].* [GP, cluster 3]

> *Most of the troubles come when the patient's hospitalized.…* [GP, cluster 1]

They highlighted the risk of losing or misunderstanding information due to heterogeneous communication habits and multiple channels of exchange. They hoped that the shared electronic medication plan system would contribute to the standardization of medication information and encourage better communication among professionals.

Overall, participants highlighted that the shared electronic medication plan system had the potential to trigger improvements beyond its original specific scope:

> *It's like a big, complex ball of wool, with many knots.…You have to start somewhere, to pinch one strand to start untangling it.…You cannot pull it in all directions at once.* [GP, cluster 2]

Discussion

Principal Findings

Health care professionals and patients alike need an accessible, common, complete, and accurate list of all the medications the patient is taking. However, introducing shared medication plans has proven difficult in several countries, and guidance for their implementation seems needed. We have presented 10 lessons learned from the first pilot project in Switzerland attempting to implement shared electronic medication plans, and we discuss this in light of studies from other contexts.

Clearly, no single organization can create and implement a comprehensive, robust, and user-friendly shared electronic medication plan system alone; HIT companies, policy makers, project teams, and the system's users—both professionals and patients—must also collaborate. Given the systemic and safety implications of implementing eHealth projects, public health authorities are taking significant steps to improve the usability of HIT systems [66]. The pilot project suffered from a lack of cooperation among HIT, clinical, and policy stakeholders and from weak enabling framework conditions, especially at the federal level. These external issues prolonged the project phase, contributed to the lack of evolution of the eHealth platform and the absence of integration with other HIT applications, and ultimately led to disengagement by health care professionals. Usability "does not heal by itself" [67] through market competition. Federating the stakeholders in an appropriate, adaptable framework involving collaboration and policy coordination is a sine qua non for the successful implementation of ambitious eHealth projects. Building a shared electronic medication plan system implies a shared ownership.

Implementing a shared electronic medication plan system and improving clinical practice is a complex process. Stakeholders face a dilemma. On the one hand, better clinical practice requires change, which technology can support. On the other hand, the

new technology needs to be fitted to an existing process to increase its acceptance. Study participants emphasized that improvements required good clinical practice, trust, and collaboration. Technology alone, therefore, is clearly neither a prerequisite nor a guarantee for safer work processes; rather, it acts as a catalyst [60] for the simultaneous innovation of the technology, processes, and relationships [68,69]. eHealth platforms could be better implemented by using approaches from quality improvement [70] and service design [71]. Combining HIT system design and clinical practice improvement within a shared electronic medication plan's implementation strategy could likely prevent the chicken-or-the-egg dilemma and better leverage synergies.

A shared electronic medication plan should improve coordination in variable and changing contexts by using the same common regional HIT system. We accepted any proposition from the participants and found variability in how the shared electronic medication plan was initiated and updated across the 5 clusters of health care professionals. In the clusters, the main professional who regularly reviewed and updated the shared electronic medication plan was different; in 1 cluster it was a pharmacist, in 2 it was a GP, and in 2 it was a nurse or the medical secretary with a care coordination role in the GP practice. The basic rule was that they needed to define how they would manage the shared medication plan together in routine practice to ensure its accuracy. The model was easily adopted in every case because it was based on a consensus and the professionals' preferences on how they wanted to manage it. Participants acknowledged their inherent shared ownership of the shared electronic medication plan. At the same time, clear processes and responsibilities are called for, both in HIT design and among health care professionals [28,32]. Our findings suggested that there was no one-size-fits-all solution; thus, strictly enforcing the implementation of a rigid solution could be difficult and could cause unintended consequences, and it would be unlikely to be achieved through policy making in Switzerland's context. A strategy enabling all health care professionals to be involved in a patient's care via a shared system and promoting basic principles of use seems more appropriate. Such an approach facilitates regular updates directly when interventions are made or discrepancies are identified. It may also increase the sense of shared ownership and favor self-organization at the local or the patient level. Knowing the issues related to the complex workflows that hinder the implementation of MedRec [17], an eHealth platform will not likely solve every problem. Standardization, automation, user constraints, and clear roles and processes [28-31] all need to be carefully balanced, with room for adaptations to local variables, in order to support a mutual commitment to using patients' common medication plans along the continuum of care.

The process of managing a shared electronic medication plan also raises questions about patients' roles and responsibilities. The pilot project HIT system implemented in the present study was not designed to empower patients and facilitate their engagement in their own medication management. The only function available to patients via the patient portal was a view of their medication plan. They could not make adjustments. The limitations of such an approach are obvious: safe, efficient medication management requires contributions from all stakeholders, including patients and their relatives. Despite increasing evidence of the benefits of comanaging digital medication systems with patients [72-75], they are still mostly treated as the passive recipients of medication lists produced by and for health care professionals. The German experience is insightful. Despite a clear policy for the systematic production of medication plans by professionals, expectations have fallen short. Few of the eligible patients ended up with the accurate list they were supposed to have [76], and when they did, only about half of them understood its content [77]. We plan to further study patients' perspectives in our ensuing work. Today, any service improvements or innovation should acknowledge the coproduction of value by patients and health care professionals together as partners [78].

Guiding stakeholders' actions towards the meaningful use of shared electronic medication plans should start by acknowledging the shared ownership and complexity of the process. From a CAS perspective, a strategy for driving major changes relies on the power of an attractor [79], a vision shared among stakeholders, that can inspire independent people and organizations to self-organize and evolve in a coherent, synergistic manner within the broader health care system. Advocating for a shared electronic medication plan comanaged by patients and health care professionals as a shared vision is even more important in settings where stakeholder fragmentation and autonomy are high. Our study stimulated collaborative actions by raising awareness of the value and shared ownership of a shared electronic medication plan, which encourages leadership at every level and supports collective learning. Indeed, these are some of the key ingredients for successfully enabling transformation in a CAS [57,80,81].

Strengths and Limitations of Our Participatory Action Research Study

Mobilizing stakeholders through formative action research is a promising approach to dealing with the complex sociotechnical challenges related to shared electronic medication plans. This type of research can nurture the implementation dynamic; policymakers cannot mandate the required motivation and trust. Local networks and cultures can vary and have a significant influence on whether a new shared HIT system gains acceptance. Disregarding them has contributed to ineffective communication with the public or failure to engage with health care professionals [38,82]. Health care professionals want to be considered long-term partners in major HIT projects, not simply clients [83]. The series of cluster meetings during our study helped to enhance mutual understanding, collective learning, and trust. Similar benefits have been reported from facilitating an interprofessional dynamic [33,84], especially when it was a core focus of the implementation strategy [29,34]. Our study participants rarely have opportunities for dialogue and reflection at the local level, and this was appreciated and even triggered some further collaborations. We also realized that the mixed status of our 2 main investigators, who were clinicians, researchers, and employees of the public health authorities, strengthened our participants' motivation to get involved in the study. They considered involvement to be a meaningful way to facilitate communication and mutual understanding between

the people in the field and decision makers. We argue that further formative action research could be a key facilitator in the implementation of new shared electronic medication plan systems.

Our study has some limitations. First, we only included GPs, pharmacists, and nurses involved in primary care. Thus, we could only investigate issues related to care transitions from the primary care perspective. Second, we report experiences from a relatively small region of French-speaking Switzerland, which might limit the study's transferability to other contexts. The lessons learned could, nevertheless, support learning in other settings. Third, the participants who volunteered for the study were likely early adopters and highly motivated. Additionally, because implementation intensity was low, some specific use and implementation issues that are likely to be encountered in the future require more assessment.

However, our novel approach, which used 5 clusters and iterative participatory analysis, is a strength of our study. We maximized diversity by including rural and urban settings, whereas earlier studies were mostly limited to university medical centers. Early adopters are not the majority, of course, but they are often the determinants for the diffusion of any innovation. Leveraging their experience can benefit other individuals less keen to explore that innovation. Moreover, we sought to embrace complexity by using a systems perspective to support sensemaking and awareness. These can help guide stakeholders and likely support further learnings.

Conclusion

The 10 lessons learned from this study give an overview of the mechanisms and dimensions related to the implementation of shared electronic medication plans in primary care settings. This paper gives practical guidance on implementation and describes some of the key sociotechnical challenges that will face implementors aiming to instill the regular, meaningful use of shared electronic medication plans—plans that should be consistently reconciled along the patient's entire care pathway—in clinical practice.

We consider the poor spontaneous patient involvement with their shared electronic medication plan to be a significant shortcoming and a point that has clearly not met the policy ambitions of fostering patient empowerment and medication adherence. Nevertheless, the local adaptability of the participating clusters was striking, as was their ability to reach consensus around useful solutions. This suggests that implementation strategies should facilitate the emergence of local engagement rather than implementing rigid top-down processes. HIT systems should be able to support various configurations of use in practice while maintaining predefined basic principles agreed among stakeholders. Last but not least, collective leadership is essential to handle the inherent shared ownership of a medication plan and to make change happen at every level, from direct patient care to the policy framework.

Future research should explore experiences in different countries in order to determine how system characteristics, stakeholder cooperation, health care policy, patients' and professionals' responsibilities, and implementation strategies affect the uptake of such shared systems by health care professionals and the benefits these shared medication plans bring to patients and health care services overall. Integrating patients so that they begin to comanage their medication plans also raises important questions. Finally, we suggest that formative participatory action research, including qualitative and quantitative methodologies, should play a key facilitating role in achieving a safe and meaningful use of shared electronic medication plans to create an efficient learning health system.

Acknowledgments

We are very grateful to all the participants in the clusters. No specific funding was used for this study.

Authors' Contributions

BB and TB designed the study and collected the data. BB analyzed the data, with iterative reviews by TB and CP. BB and CP wrote the manuscript. PB and AG were involved in the design. All the authors participated in the interpretation of the results and critical revision of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide and meetings goals.
[PDF File (Adobe PDF File), 80 KB - formative_v5i1e22319_app1.pdf]

References

1. Slawomirski L, Auraaen A, Klazinga N. The economics of patient safety: Strengthening a value-based approach to reducing patient harm at national level. Organisation for Economic Co-operation and Development. 2017. URL: https://www.oecd-ilibrary.org/social-issues-migration-health/the-economics-of-patient-safety_5a9858cd-en [accessed 2020-12-16]

2. Garfield S, Barber N, Walley P, Willson A, Eliasson L. Quality of medication use in primary care--mapping the problem, working to a solution: a systematic review of the literature. BMC Med 2009 Sep 21;7:50 [FREE Full text] [doi: 10.1186/1741-7015-7-50] [Medline: 19772551]
3. Winterstein AG, Sauer BC, Hepler CD, Poole C. Preventable drug-related hospital admissions. Ann Pharmacother 2002;36(7-8):1238-1248. [doi: 10.1345/aph.1A225] [Medline: 12086559]
4. Donaldson LJ, Kelley ET, Dhingra-Kumar N, Kieny M, Sheikh A. Medication Without Harm: WHO's Third Global Patient Safety Challenge. The Lancet 2017 Apr;389(10080):1680-1681. [doi: 10.1016/s0140-6736(17)31047-4]
5. Aitken M, Gorokhovich L. Advancing the Responsible Use of Medicines: Applying Levers for Change. SSRN J 2012 Sep 17 [FREE Full text] [doi: 10.2139/ssrn.2222541]
6. Elliott RA, Camacho E, Jankovic D, Sculpher MJ, Faria R. Economic analysis of the prevalence and clinical and economic burden of medication error in England. BMJ Qual Saf 2020 Jun 11. [doi: 10.1136/bmjqs-2019-010206] [Medline: 32527980]
7. Medication Safety in Transitions of Care. World Health Organization. 2019. URL: https://apps.who.int/iris/bitstream/handle/10665/325453/WHO-UHC-SDS-2019.9-eng.pdf [accessed 2020-12-16]
8. Tam VC, Knowles SR, Cornish PL, Fine N, Marchesano R, Etchells EE. Frequency, type and clinical importance of medication history errors at admission to hospital: a systematic review. CMAJ 2005 Aug 30;173(5):510-515 [FREE Full text] [doi: 10.1503/cmaj.045311] [Medline: 16129874]
9. Lehnbom EC, Stewart MJ, Manias E, Westbrook JI. Impact of medication reconciliation and review on clinical outcomes. Ann Pharmacother 2014 Oct;48(10):1298-1312. [doi: 10.1177/1060028014543485] [Medline: 25048794]
10. Redmond P, Grimes T, McDonnell R, Boland F, Hughes C, Fahey T. Impact of medication reconciliation for improving transitions of care. Cochrane Database Syst Rev 2018 Aug 23;8:CD010791 [FREE Full text] [doi: 10.1002/14651858.CD010791.pub2] [Medline: 30136718]
11. Executive summary - Programme pilote progress?! La sécurité de la médication aux interfaces. Sécurité des patients Suisse. 2017. URL: https://www.securitedespatients.ch/fileadmin/user_upload/1_Projekte_und_Programme/progress_sichere_med._schnittstellen/progress_sms_f/Executive_summary_progress__Se__curite___de_la_me__dication.pdf [accessed 2020-12-16]
12. Anguish I, Wick HD, Fonjallaz M, Luchez PC, Jotterand S, Locca J, et al. [Hospital's discharge prescription: a challenge for continuity of care and the interprofessional collaboration]. Rev Med Suisse 2013 May 15;9(386):1021-1025. [Medline: 23745235]
13. Giannini O, Rizza N, Pironi M, Parlato S, Waldispühl Suter B, Borella P, et al. Prevalence, clinical relevance and predictive factors of medication discrepancies revealed by medication reconciliation at hospital admission: prospective study in a Swiss internal medicine ward. BMJ Open 2019 May 27;9(5):e026259 [FREE Full text] [doi: 10.1136/bmjopen-2018-026259] [Medline: 31133583]
14. How-to Guide: Prevent Adverse Drug Events by Implementing Medication Reconciliation. Institute for Healthcare Improvement. 2011. URL: http://www.ihi.org/resources/Pages/Tools/HowtoGuidePreventAdverseDrugEvents.aspx [accessed 2020-12-16]
15. Medication Safety in Polypharmacy. World Health Organization. 2019. URL: https://apps.who.int/iris/bitstream/handle/10665/325454/WHO-UHC-SDS-2019.11-eng.pdf?sequence=1&isAllowed=y [accessed 2020-02-12]
16. Meguerditchian AN, Krotneva S, Reidel K, Huang A, Tamblyn R. Medication reconciliation at admission and discharge: a time and motion study. BMC Health Serv Res 2013 Nov 21;13:485 [FREE Full text] [doi: 10.1186/1472-6963-13-485] [Medline: 24261516]
17. Pevnick JM, Shane R, Schnipper JL. The problem with medication reconciliation. BMJ Qual Saf 2016 Sep;25(9):726-730 [FREE Full text] [doi: 10.1136/bmjqs-2015-004734] [Medline: 26795914]
18. Bassi J, Lau F, Bardal S. Use of information technology in medication reconciliation: a scoping review. Ann Pharmacother 2010 May;44(5):885-897. [doi: 10.1345/aph.1M699] [Medline: 20371752]
19. Remen VM, Grimsmo A. Closing information gaps with shared electronic patient summaries: how much will it matter? Int J Med Inform 2011 Nov;80(11):775-781. [doi: 10.1016/j.ijmedinf.2011.08.008] [Medline: 21956001]
20. Claeys C, Foulon V, de Winter S, Spinewine A. Initiatives promoting seamless care in medication management: an international review of the grey literature. Int J Clin Pharm 2013 Dec;35(6):1040-1052. [doi: 10.1007/s11096-013-9844-2] [Medline: 24022724]
21. Manskow U, Lind K, Bergmo T. Digital solutions for a shared medication list A narrative literature review. 2019 Presented at: 17th Scandinavian Conference on Health Informatics; Nov 12-13, 2019; Oslo, Norway.
22. Gall W, Aly A, Sojer R, Spahni S, Ammenwerth E. The national e-medication approaches in Germany, Switzerland and Austria: A structured comparison. Int J Med Inform 2016 Sep;93:14-25. [doi: 10.1016/j.ijmedinf.2016.05.009] [Medline: 27435943]
23. Munck LK, Hansen KR, Mølbak AG, Balle H, Kongsgren S. The use of shared medication record as part of medication reconciliation at hospital admission is feasible. Dan Med J 2014 May;61(5):A4817 [FREE Full text] [Medline: 24814735]
24. Jurado C, Calmels V, Lobinet E, Divol E, Hanaire H, Metsu D, et al. The Electronic Pharmaceutical Record: A new method for medication reconciliation. J Eval Clin Pract 2018 Aug;24(4):681-687. [doi: 10.1111/jep.12942] [Medline: 29761596]

25. Grimes T, Fitzsimons M, Galvin M, Delaney T. Relative accuracy and availability of an Irish National Database of dispensed medication as a source of medication history information: observational study and retrospective record analysis. J Clin Pharm Ther 2013 Jun;38(3):219-224. [doi: 10.1111/jcpt.12036] [Medline: 23350784]
26. Uitvlugt EB, van den Bemt BJF, Chung WL, Dik J, van den Bemt PMLA, Karapinar-Çarkit F. Validity of a nationwide medication record system in the Netherlands. Int J Clin Pharm 2019 Jun;41(3):687-690. [doi: 10.1007/s11096-019-00839-x] [Medline: 31028600]
27. Jøsendal A, Strand BT. How Discrepancies in Medication Records Affect the Creation and Trust in a Shared Electronic Medication List in Norway. 2019 Presented at: 17th Scandinavian Conference on Health Informatics; Nov 12-13, 2019; Oslo, Norway.
28. Hammar T, Ekedahl A, Petersson G. Implementation of a shared medication list: physicians' views on availability, accuracy and confidentiality. Int J Clin Pharm 2014 Oct;36(5):933-942. [doi: 10.1007/s11096-014-0012-0] [Medline: 25193264]
29. Rangachari P, Dellsperger KC, Fallaw D, Davis I, Sumner M, Ray W, et al. A Mixed-Method Study of Practitioners' Perspectives on Issues Related to EHR Medication Reconciliation at a Health System. Qual Manag Health Care 2019;28(2):84-95 [FREE Full text] [doi: 10.1097/QMH.0000000000000208] [Medline: 30801417]
30. Børgaard H, Jensen L, Jepsen N, Clausen M, Madsen N, Marquart H, et al. Complex Medication Reconciliation in the Danish Medication System: Shared Medication Record for patients in transition of care across sectors. 2017 Presented at: 15th Scandinavian Conference on Health Informatics; Aug 29-30, 2017; Kristiansand, Norway.
31. Stock R, Mahoney ER, Gauthier D, Center L, Minniti M, Scott J, et al. Developing a Community-Wide Electronic Shared Medication List. Advances in Patient Safety: New Directions and Alternative Approaches 2008;4 [FREE Full text] [Medline: 21249957]
32. Ammenwerth E, Duftschmid G, Gall W, Hackl WO, Hoerbst A, Janzek-Hawlat S, et al. A nationwide computerized patient medication history: evaluation of the Austrian pilot project "e-Medikation". Int J Med Inform 2014 Sep;83(9):655-669. [doi: 10.1016/j.ijmedinf.2014.06.004] [Medline: 24986321]
33. Kivekäs E, Luukkonen I, Mykkänen J, Saranto K. Improving the coordination of patients' medication management: a regional Finnish development project. Stud Health Technol Inform 2014;201:175-180. [Medline: 24943541]
34. Rangachari P, Dellsperger KC, Rethemeyer RK. A qualitative study of interprofessional learning related to electronic health record (EHR) medication reconciliation within a social knowledge networking (SKN) system. J Healthc Leadersh 2019;11:23-41 [FREE Full text] [doi: 10.2147/JHL.S198951] [Medline: 31114416]
35. Hackl WO, Hoerbst A, Duftschmid G, Gall W, Janzek-Hawlat S, Jung M, et al. Crucial factors for the acceptance of a computerized national medication list: insights into findings from the evaluation of the Austrian e-Medikation pilot. Appl Clin Inform 2014;5(2):527-537 [FREE Full text] [doi: 10.4338/ACI-2014-04-RA-0032] [Medline: 25024766]
36. Deutsch E, Duftschmid G, Dorda W. Critical areas of national electronic health record programs-is our focus correct? Int J Med Inform 2010 Mar;79(3):211-222. [doi: 10.1016/j.ijmedinf.2009.12.002] [Medline: 20079685]
37. Deetjen U. European E-Prescriptions: Benefits and Success Factors. University of Oxford. 2016 Jun. URL: https://www.politics.ox.ac.uk/materials/publications/15224/workingpaperno5ulrikedeetjen.pdf [accessed 2020-12-16]
38. Greenhalgh T, Morris L, Wyatt JC, Thomas G, Gunning K. Introducing a nationally shared electronic patient record: case study comparison of Scotland, England, Wales and Northern Ireland. Int J Med Inform 2013 May;82(5):e125-e138. [doi: 10.1016/j.ijmedinf.2013.01.002] [Medline: 23434362]
39. van Gemert-Pijnen JEWC, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011 Dec 05;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]
40. Greenhalgh T, Stramer K, Bratan T, Byrne E, Mohammad Y, Russell J. Introduction of shared electronic records: multi-site case study using diffusion of innovation theory. BMJ 2008 Oct 23;337:a1786 [FREE Full text] [doi: 10.1136/bmj.a1786] [Medline: 18948344]
41. Stetler CB, Legro MW, Wallace CM, Bowman C, Guihan M, Hagedorn H, et al. The role of formative evaluation in implementation research and the QUERI experience. J Gen Intern Med 2006 Feb;21(Suppl 2):S1-S8 [FREE Full text] [doi: 10.1111/j.1525-1497.2006.00355.x] [Medline: 16637954]
42. McIntyre A. Participatory action research. In: Qualitative Research Methods. Los Angeles, CA: Sage Publications; 2008.
43. Singer SJ. The time is now for action research. Isr J Health Policy Res 2013 Oct 23;2(1):41 [FREE Full text] [doi: 10.1186/2045-4015-2-41] [Medline: 24153050]
44. Loewenson R, Laurell A, Hogstedt C, Shroff Z. Participatory action research in health systems: a method reader. EQUINET Africa. Harare, Zimbabwe: TARSC, AHPSR, WHO, IDRC Canada, Equinet; 2014. URL: https://equinetafrica.org/sites/default/files/uploads/documents/PAR_Methods_Reader2014_for_web.pdf [accessed 2020-12-16]
45. Perez F. Projet Sécurisation de la médication tout au long de la chaîne de soins: Rapport final. Réseau Santé Nord Broye. 2016. URL: https://www.reseau-sante-nord-broye.ch/system/files/2016/05/rapport_final_du_projet-plan_de_medication.pdf [accessed 2020-12-16]
46. Spahni S, Guardia A, Boggini T, Geissbuhler A. Design and implementation of a shared treatment plan in a federated health information exchange. Stud Health Technol Inform 2013;192:1090. [Medline: 23920864]

47. Spahni S. Implementation Case Study: Geneva, Switzerland - Shared Medication Treatment Plan. Integrating the Healthcare Enterprise. 2017. URL: https://www.ihe.net/wp-content/uploads/uploadedFiles/Documents/CASE_STUDIES/IHE-based-Implementation-Case-Study-Shared-Medication-Treatment-Plan.pdf [accessed 2020-06-20]
48. Existing interoperability standard. eHealth Suisse. URL: https://www.e-health-suisse.ch/fr/technique-semantique/interoperabilite-semantique/formats-dechange/formats-dechange-existants.html [accessed 2020-09-06]
49. De Pietro C, Francetic I. E-health in Switzerland: The laborious adoption of the federal law on electronic health records (EHR) and health information exchange (HIE) networks. Health Policy 2018 Feb;122(2):69-74 [FREE Full text] [doi: 10.1016/j.healthpol.2017.11.005] [Medline: 29153922]
50. 3512 | Droit à un plan de médication en vue de renforcer la sécurité des patients. Swiss Parliament. URL: https://www.parlament.ch/fr/ratsbetrieb/amtliches-bulletin/amtliches-bulletin-die-verhandlungen?SubjectId=44066 [accessed 2019-07-11]
51. Politique de la santé: les priorités du Conseil Fédéral - Stratégie Santé2020. Confédération Suisse. 2013. URL: https://www.bag.admin.ch/bag/fr/home/strategie-und-politik/gesundheit-2020/prioritaeten-gesundheit2020.html [accessed 2019-07-11]
52. Organisation for Economic Cooperation and Development, World Health Organization. OECD Reviews of Health Systems: Switzerland 2011. Organisation for Economic Cooperation and Development. 2011. URL: https://www.oecd-ilibrary.org/social-issues-migration-health/oecd-reviews-of-health-systems-switzerland-2011_9789264120914-en [accessed 2020-12-16]
53. Thiel R, Deimel L, Schmidtmann D, Piesche K, Hüsing T, Rennoch T, et al. #SmartHealthSystems: International comparison of digital strategies. Bartelsmann Stiftung. 2019. URL: https://www.bertelsmann-stiftung.de/en/publications/publication/did/smarthealthsystems-1 [accessed 2020-12-16]
54. Evans J, Baker G, Berta W, Barnsley J. The evolution of integrated health care strategies. Adv Health Care Manag 2013;15:125-161. [doi: 10.1108/s1474-8231(2013)0000015011] [Medline: 24749215]
55. Tsasis P, Evans JM, Owen S. Reframing the challenges to integrated care: a complex-adaptive systems perspective. Int J Integr Care 2012;12:e190 [FREE Full text] [doi: 10.5334/ijic.843] [Medline: 23593051]
56. Paina L, Peters DH. Understanding pathways for scaling up health services through the lens of complex adaptive systems. Health Policy Plan 2012 Aug;27(5):365-373. [doi: 10.1093/heapol/czr054] [Medline: 21821667]
57. Edgren L, Barnard K. Complex adaptive systems for management of integrated care. Leadership Health Serv 2012 Jan 27;25(1):39-51. [doi: 10.1108/17511871211198061]
58. The Health Foundation. Complex adaptive systems. Evidence Scan. 2010. URL: https://www.health.org.uk/publications/complex-adaptive-systems [accessed 2019-06-26]
59. Institute of Medicine Committee on Quality of Health Care in America. Appendix B: Redesigning Health Care with Insights from the Science of Complex Adaptive Systems. In: Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: National Academies Press; 2001.
60. Meyer I, Muller S, Kubitschke L. Achieving effective integrated e-care beyond the silos. Int J Care Pathways 2015;2015(1):31. [doi: 10.4018/978-1-4666-6138-7]
61. Lilford RJ, Chilton PJ, Hemming K, Girling AJ, Taylor CA, Barach P. Evaluating policy and service interventions: framework to guide selection and interpretation of study end points. BMJ 2010 Aug 27;341:c4413. [doi: 10.1136/bmj.c4413] [Medline: 20802000]
62. Haggerty JL, Reid RJ, Freeman GK, Starfield BH, Adair CE, McKendry R. Continuity of care: a multidisciplinary review. BMJ 2003 Nov 22;327(7425):1219-1221 [FREE Full text] [doi: 10.1136/bmj.327.7425.1219] [Medline: 14630762]
63. Lau F, Price M, Keshavjee K. From benefits evaluation to clinical adoption: making sense of health information system success in Canada. Healthc Q 2011;14(1):39-45. [doi: 10.12927/hcq.2011.22157] [Medline: 21301238]
64. UN Statistical Commission. A recommendation on the method to delineate cities, urban and rural areas for international statistical comparisons. European Commission. 2020. URL: https://ec.europa.eu/eurostat/cros/system/files/bg-item3j-recommendation-e.pdf [accessed 2020-09-25]
65. Coghlan D, Brannick T. Doing action research in your own organization. London, United Kingdom: SAGE Publications; 2010.
66. Kushniruk AW, Bates DW, Bainbridge M, Househ MS, Borycki EM. National efforts to improve health information system safety in Canada, the United States of America and England. Int J Med Inform 2013 May;82(5):e149-e160. [doi: 10.1016/j.ijmedinf.2012.12.006] [Medline: 23313431]
67. Kaipio J, Lääveri T, Hyppönen H, Vainiomäki S, Reponen J, Kushniruk A, et al. Usability problems do not heal by themselves: National survey on physicians' experiences with EHRs in Finland. Int J Med Inform 2017 Jan;97:266-281 [FREE Full text] [doi: 10.1016/j.ijmedinf.2016.10.010] [Medline: 27919385]
68. Kubitschke L, Meyer I, Müller S. Do All Roads Lead to Rome? Models for Integrated e-Care Services in Europe. In: Achieving Effective Integrated E-Care Beyond the Silos. Los Angeles, CA: IGI Global; 2014:22-40.
69. Billings J, Carretero S, Kagialaris G, Mastroyiannakis T, Meriläinen-Porras S. The Role of Information Technology in Long-Term Care for Older People. In: Long-Term Care in Europe. London, United Kingdom: Palgrave Macmillan; 2013:252-277.
70. Nicholson C, Jackson C, Marley J. A governance model for integrated primary/secondary care for the health-reforming first world - results of a systematic review. BMC Health Serv Res 2013 Dec 20;13:528 [FREE Full text] [doi: 10.1186/1472-6963-13-528] [Medline: 24359610]

71. Shaw J, Agarwal P, Desveaux L, Palma DC, Stamenova V, Jamieson T, et al. Beyond "implementation": digital health innovation and service design. NPJ Digit Med 2018;1:48 [FREE Full text] [doi: 10.1038/s41746-018-0059-8] [Medline: 31304327]
72. Prey J. Patient engagement to improve medication safety in the hospital [thesis]. Columbia University Academic Commons. 2016. URL: https://academiccommons.columbia.edu/doi/10.7916/D8F18ZX4 [accessed 2020-12-16]
73. Leonhardt KK, Pagel P, Bonin D, Moberg DP, Dvorak ML, Hatlie MJ. Creating an Accurate Medication List in the Outpatient Setting Through a Patient-Centered Approach. In: Advances in Patient Safety: New Directions and Alternative Approaches. Vol. 3: Performance and Tools. Rockville, MD: Agency for Healthcare Research and Quality; 2008.
74. Dullabh PM, Sondheimer NK, Katsh E, Evans MA. How patients can improve the accuracy of their medical records. EGEMS (Wash DC) 2014;2(3):1080 [FREE Full text] [doi: 10.13063/2327-9214.1080] [Medline: 25848614]
75. Fylan B, Armitage G, Naylor D, Blenkinsopp A. A qualitative study of patient involvement in medicines management after hospital discharge: an under-recognised source of systems resilience. BMJ Qual Saf 2018 Jul;27(7):539-546. [doi: 10.1136/bmjqs-2017-006813] [Medline: 29146681]
76. Mueller MA, Opitz R, Grandt D, Lehr T. The federal standard medication plan in practice: An observational cross-sectional study on prevalence and quality. Res Social Adm Pharm 2020 Oct;16(10):1370-1378. [doi: 10.1016/j.sapharm.2020.01.013] [Medline: 32061549]
77. Botermann L, Monzel K, Krueger K, Eickhoff C, Wachter A, Kloft C, et al. Evaluating patients' comprehensibility of a standardized medication plan. Eur J Clin Pharmacol 2016 Oct;72(10):1229-1237. [doi: 10.1007/s00228-016-2082-5] [Medline: 27305905]
78. Batalden M, Batalden P, Margolis P, Seid M, Armstrong G, Opipari-Arrigan L, et al. Coproduction of healthcare service. BMJ Qual Saf 2016 Jul;25(7):509-517 [FREE Full text] [doi: 10.1136/bmjqs-2015-004315] [Medline: 26376674]
79. Sturmberg J, O'Halloran DM, Martin C. Understanding health system reform - a complex adaptive systems perspective. J Eval Clin Pract 2012 Feb;18(1):202-208. [doi: 10.1111/j.1365-2753.2011.01792.x] [Medline: 22221420]
80. Swanson R, Cattaneo A, Bradley E, Chunharas S, Atun R, Abbas K, et al. Rethinking health systems strengthening: key systems thinking tools and strategies for transformational change. Health Policy Plan 2012 Oct;27(Suppl 4):iv54-iv61 [FREE Full text] [doi: 10.1093/heapol/czs090] [Medline: 23014154]
81. Best A, Greenhalgh T, Lewis S, Saul JE, Carroll S, Bitz J. Large-system transformation in health care: a realist review. Milbank Q 2012 Sep;90(3):421-456 [FREE Full text] [doi: 10.1111/j.1468-0009.2012.00670.x] [Medline: 22985277]
82. Bratan T, Stramer K, Greenhalgh T. 'Never heard of it'- understanding the public's lack of awareness of a new electronic patient record. Health Expect 2010 Dec;13(4):379-391 [FREE Full text] [doi: 10.1111/j.1369-7625.2010.00608.x] [Medline: 20579117]
83. Beger B, Wilks M. Physicians' Perspectives on eHealth. In: Managing eHealth. London, United Kingdom: Palgrave Macmillan UK; 2014:77-94.
84. Stock R, Mahoney ER, Gauthier D, Center L, Minniti M, Scott J, et al. Developing a Community-Wide Electronic Shared Medication List. In: Henriksen K, Battles JB, Keyes MA, Grady ML, editors. Advances in Patient Safety: New Directions and Alternative Approaches. Vol 4: Technology and Medication Safety. Rockville, MD: Agency for Healthcare Research and Quality; 2008.

Abbreviations

CAS: complex adaptive system
EPR: electronic patient record
GP: general practitioner
HIT: health information technology
IHE: Integrating the Healthcare Enterprise
MedRec: medication reconciliation
PAR: participatory action research

Edited by G Eysenbach; submitted 08.07.20; peer-reviewed by Y Chu, T Nguyen, YH Kwan, J Appalasamy; comments to author 29.07.20; revised version received 23.09.20; accepted 07.12.20; published 07.01.21.

Please cite as:
Bugnon B, Geissbuhler A, Bischoff T, Bonnabry P, von Plessen C
Improving Primary Care Medication Processes by Using Shared Electronic Medication Plans in Switzerland: Lessons Learned From a Participatory Action Research Study
JMIR Form Res 2021;5(1):e22319
URL: http://formative.jmir.org/2021/1/e22319/
doi:10.2196/22319
PMID:33410753

Review

Mobile App–Based Self-Report Questionnaires for the Assessment and Monitoring of Bipolar Disorder: Systematic Review

Eric C Chan[1], BSc, MSc, MD, FRCPC; Yuting Sun[1]; Katherine J Aitchison[1], BA, BM BCh, MA, PhD, FRCPsych; Sudhakar Sivapalan[1], MSc, MD, FRCPC

Department of Psychiatry, University of Alberta, Edmonton, AB, Canada

Corresponding Author:
Eric C Chan, BSc, MSc, MD, FRCPC
Department of Psychiatry
University of Alberta
1E1 Walter Mackenzie Health Sciences Center
8440 112 St NW
Edmonton, AB, T6G 2B7
Canada
Phone: 1 7802332709
Email: echantai@ualberta.ca

Abstract

Background: Bipolar disorder is a chronic, progressive illness characterized by recurrent episodes of mania and depression. Self-report scales have historically played a significant role in the monitoring of bipolar symptoms. However, these tools rely on episodic memory, which can be unreliable and do not allow the clinician to monitor brief episodic symptoms or the course of symptoms over shorter periods of time. Mobile app–based questionnaires have been suggested as a tool to improve monitoring of patients with bipolar disorder.

Objective: This paper aims to determine the feasibility and validity of mobile app–based self-report questionnaires.

Methods: We performed a systematic review of the literature according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. The PubMed, PsycInfo, Web of Science, Ovid MEDLINE, and EMBASE databases were searched for papers published in English that assessed adherence to and the validity of mobile app–based self-report questionnaires. Relevant studies published from database creation to May 22, 2020, were identified, and results examining the validity of and rates of adherence to app-based self-report questionnaires are reported.

Results: A total of 13 records were identified for inclusion in this review. Of these studies, 4 assessed the concurrent validity of mobile app–based self-report tools, with the majority of findings indicating significant associations between data collected using these tools and the Young Mania Rating Scale, Hamilton Depression Rating Scale-17, or Montgomery-Åsberg Depression Rating Scale (P<.001 to P=.24). Three studies comparing the variability or range of symptoms between patients with bipolar disorder and healthy controls suggested that these data are capable of differentiating between known groups. Two studies demonstrated statistically significant associations between data collected via mobile app–based self-report tools and instruments assessing other clinically important factors. Adherence rates varied across the studies examined. However, good adherence rates (>70%) were observed in all but 1 study using a once-daily assessment. There was a wide range of adherence rates observed in studies using twice-daily assessments (42%-95%).

Conclusions: These findings suggest that mobile app–based self-report tools are valid in the assessment of symptoms of mania and depression in euthymic patients with bipolar disorder. Data collected using these tools appear to differ between patients with bipolar disorder and healthy controls and are significantly associated with other clinically important measures. It is unclear at this time whether these tools can be used to detect acute episodes of mania or depression in patients with bipolar disorder. Adherence data indicate that patients with bipolar disorder show good adherence to self-report assessments administered daily for the duration of the study periods evaluated.

(JMIR Form Res 2021;5(1):e13770) doi:10.2196/13770

KEYWORDS

mobile apps; mental health; bipolar disorder; smartphone; cell phone

Introduction

Bipolar disorder is a chronic, progressive illness characterized by recurrent episodes of mania and depression. The international 12-month prevalence of bipolar I disorder is 0.0% to 0.6%, and the international 12-month prevalence of bipolar II disorder is 0.3% [1]. Both manic and depressive episodes are associated with impairments in social and occupational functioning, and the World Health Organization's World Mental Health Surveys identified the disorder as having the second-strongest effect on days out of role compared with other common physical and mental illnesses [2-5]. In addition, bipolar disorder is associated with a high risk of suicide, with one-third to one-half of patients attempting suicide at least once in their lifetime and 15% to 20% of suicide attempts completed [6]. Given such adverse consequences of mania and depression, timely detection of relapse is an important aspect in the psychiatric care of the disease.

No biomarker has been approved for the diagnosis or assessment of bipolar disorder, so medical practitioners must rely on clinical assessment and reports from the patient and collateral sources in order to monitor the disease. However, detection of mood episodes can be delayed, with previous data indicating that the interval between illness onset and hospitalization is often 3 weeks or more [7]. One challenge for the detection of mood episodes is the lack of insight that can occur in patients with bipolar disorder, especially during episodes of pure mania [8]. Previous data suggest, however, that some patients in acute mania may retain awareness of their diagnosis and its potential consequences despite having impaired insight into their current symptoms [9]. Given patients' preserved awareness of their diagnosis even in the context of active symptoms, the use of self-report questionnaires has the potential to facilitate symptom monitoring, including changes over time.

Self-report scales, such as the Mood Disorder Questionnaire (MDQ) and the Altman Self-Rating Mania Scale (ASRM), have previously been developed for use in the monitoring of bipolar symptoms. These scales have been validated in inpatient populations with bipolar disorder, with respective sensitivities and specificities of 86% and 71% for the MDQ and 93% and 33% for the ASRM [10-12].

Traditionally, self-report scales have been administered via pen and paper; however, some limitations exist with this form of data collection. When administered in the context of visits with a health care provider, these tools rely on retrospective reporting of symptoms, which can be unreliable and do not allow the clinician to monitor symptoms associated with brief mood episodes or the course of symptoms over shorter periods of time [12-14]. In a study asking participants to complete paper diaries on a daily basis, participants were found to record entries outside of the requested time frame and inaccurately report the date of these entries, reducing the accuracy of the data collected [15]. In addition, the frequency with which the clinician is able to review responses obtained via pen and paper is limited by the frequency in which the responses are forwarded to the provider. This often occurs on clinic visits, which limits the ability of the health care provider to respond in a timely fashion if the patient deteriorates between scheduled appointments.

The administration of self-report scales using mobile apps has the potential to circumvent some of these issues. Automatic transmission of data using a mobile device could allow clinicians to monitor symptoms in real time, improving their ability to proactively detect and engage the patient when symptoms relapse. In addition, scale administration using a mobile app may be less disruptive for the patient, increasing the frequency that the patient is willing to complete the scale. For example, one study described a mobile app for monitoring nonaffective psychosis that yielded more data points and took less time compared with the text messaging–only equivalent [16]. The increased data collection afforded by the use of mobile apps may also have uses in research settings. Frequent administration of scales may allow researchers to better characterize the course of illness over time and to identify warning signs that mark early deterioration.

Given the variability in the course of symptoms in bipolar disorder, the use of mobile apps in this population has been of considerable recent interest, with 35 apps identified using the Google Play and iOS stores in a previous systematic review [17]. Studies have shown that 60% to 70% of patients with mental illness would be interested in using a mobile app to monitor their mental health condition, and a study examining publicly available consumer reviews of 48 apps for bipolar disorder, the majority of which were symptom-monitoring apps (1911/2173, 87.9%), found that 1608 of 2173 (74.0%) reviews included positive appraisals of the app discussed [13,18-20]. Additionally, a recent study evaluating 2 smartphone-based self-monitoring systems for bipolar disorder showed acceptable usefulness, usability, feasibility, and technical stability for both systems evaluated [21]. However, a 2015 review showed that 60% of symptom-monitoring apps available did not use validated screening measures [17]. Furthermore, it is possible that for a given validated screening tool, data collected via a mobile app may differ from those collected via a pen-and-paper version.

The validity of a scale is defined as "the extent to which an instrument indeed measures the latent dimension or construct it was developed to evaluate" [22]. The major forms of validity are content validity, criterion validity, and construct validity. Content validity refers to whether the measure adequately assesses the domain of interest, and it is primarily assessed through evaluation by experts and the target population. Criterion validity refers to whether the results of a measure relate to another measure of relevance. It includes predictive validity (the ability of the measure to predict a future result or answer a future question) and concurrent validity (the strength of the relationship between the new measure and a gold standard measurement made at a similar time). Construct validity refers to the degree to which the measure assesses the construct of concern. Construct validity can be evaluated through convergent validity, discriminant or divergent validity, differentiation or comparison between known groups, or correlational analysis [22].

The aim of this systematic review was to assess the feasibility and validity of self-report questionnaire-based mobile apps as tools for bipolar symptom monitoring through a systematic review of the literature. We identified studies in which patients with bipolar disorder were monitored using self-report scales administered by a mobile app with or without comparison to a traditional form of symptom monitoring, such as pen-and-paper rating scales or standardized clinician interviews. The outcomes of interest in this review were adherence rates and the criterion or construct validity of self-report scales administered by mobile app.

Methods

In order to identify data describing the feasibility and validity of mobile apps in the assessment of bipolar disorder, we conducted searches of the PubMed, PsycInfo, Web of Science, Ovid MEDLINE, and EMBASE databases. One researcher (YS) searched these databases using the following keywords: "mental disorders," "psychiatry," or "mental health" AND "mobile application," "cell phone," or "smartphone," excluding the term "substance-related disorders." All records published in English listed from database creation to May 22, 2020, were identified. In addition, the references on the full paper of the records assessed were reviewed in order to identify other potential candidates for inclusion.

YS and ECC independently screened the records to identify papers suitable for inclusion in this review. In the case of disagreement between the 2 authors, records were evaluated by a third author (SS), who determined whether the paper would be forwarded to the next step of screening. There was no disagreement between authors following the review of the full papers.

Titles and abstracts of records were screened using the following exclusion criteria: (1) the study did not refer to the use of mobile apps, smartphones, or mobile phone or technology as the primary intervention of interest, or the intervention of interest was solely text message based; (2) bipolar disorder was not the primary condition of interest; (3) the interventions studied did not include self-report symptom monitoring as a component; and (4) the study did not present data from an applied intervention (such as a protocol paper, review paper, or response or correction to another paper).

The full text of the remaining studies were evaluated, and studies were excluded if they met one of the following criteria: (1) the study did not present data on adherence or validity; (2) the study did not present data from an applied intervention (such as a protocol or review paper); (3) the study did not refer to symptom assessment via self-report by mobile app, smartphone, or mobile phone or technology as a primary intervention of interest; (4) the intervention of interest was solely text message based; and (5) bipolar disorder was not the primary condition of interest.

Studies identified for inclusion in this review were then evaluated for data on the adherence rates and validity of mobile app–based symptom monitoring tools with or without comparison to standardized pen-and-paper or clinical interview–based measures. ECC and YS assessed each of the identified studies for bias using the Cochrane Risk of Bias 2 tool or the Cochrane Risk of Bias in Non-Randomized Studies of Interventions assessment tool. These tools were developed for the assessment of bias in randomized and nonrandomized studies, respectively [23,24]. These assessments were reviewed by another author (SS) and are available in Multimedia Appendix 1.

Results

Identified Records

The flow diagram of the search method is depicted in Figure 1. Initial searches produced 2827 unique records following the removal of duplicates. A total of 50 records were identified following screening of the abstracts, and their references were also searched for further relevant studies. Following the search procedure described above, 13 records were identified for inclusion in this review; study characteristics are listed in Table 1. Findings of each study are listed separately (Table 2). The assessments of the risk of bias are described in Multimedia Appendix 1.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram.

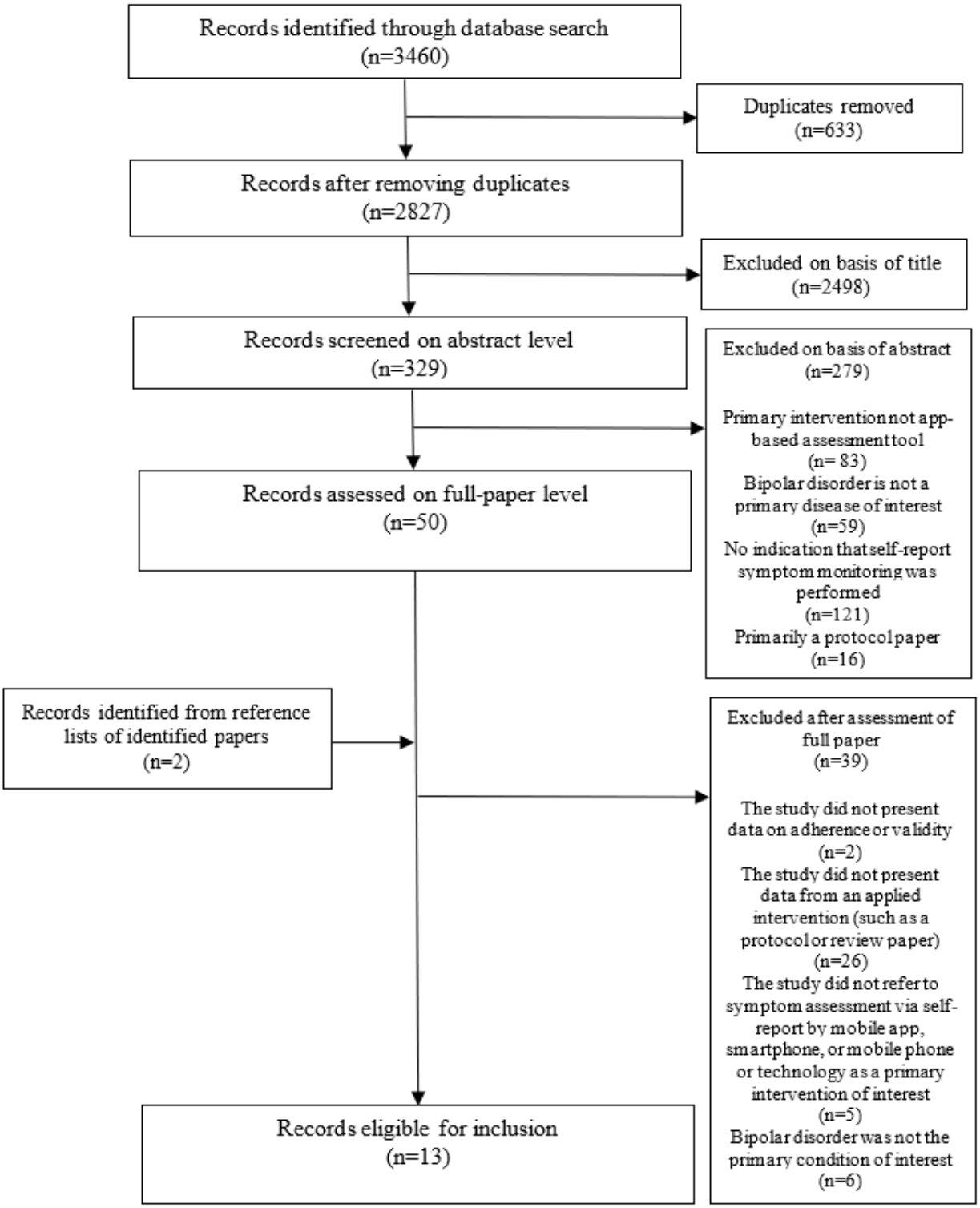


Table 1. Characteristics of included studies.

Reference	Location	Participants, n	Mobile app–based intervention	Comparison (if applicable)	Duration
Busk et al (2020) [25]	Copenhagen, Denmark	84	Monsenso system: self-monitoring of 10 symptoms completed daily.	YMRS[a] and HDRS[b] at baseline and after 4 weeks, 3 months, 6 months, and 9 months.	9 months
Carr et al (2018) [26]	Oxford, United Kingdom	43 (bipolar disorder); 26 (borderline personality disorder); 44 (healthy controls)	Mood Zoom smartphone app: 6-item assessment of mood and related items administered 10 times daily.	N/A[c]	3 months
Depp et al (2012) [27]	San Diego, CA	18 (intervention) 22 (comparison)	9-point bipolar anchored scale completed twice per day. Could not be completed after 2 hours.	Daily paper-and-pencil mood charts. MADRS[d] and YMRS completed at baseline and 6 weeks and 12 weeks after baseline.	12 weeks
Depp et al (2015) [28]	San Diego, CA	51 (intervention), (41 analyzed); 53 (comparison), (41 analyzed)	PRISM: 10 questions followed by rating of current mood state on a 9-point bipolar anchored scale completed twice per day.	Daily pencil-and-paper mood charts.	10 weeks
Faurholt-Jepsen et al (2015) [29]	Copenhangen, Denmark	30	MONARCA: self-monitoring of 11 symptoms completed daily. Allowed for retrospective data entry up to 2 days later.	Monthly clinical assessment via HDRS-17 and YMRS. Scores compared to those obtained via app from day of assessment and 3 previous days.	6 months
Faurholt-Jepsen et al (2015) [30]	Copenhagen, Denmark	39 (intervention); 39 (comparison)	MONARCA: self-monitoring of 11 symptoms completed daily. Allowed for retrospective data entry up to 2 days later.	Participants provided with a smartphone without the MONARCA system.	6 months
Faurholt-Jepsen et al (2019) [31]	Copenhagen, Denmark	84 patients (participants in MONARCA II trial)	Monsenso app for daily self-monitoring of mood, mixed mood, and irritability level.	HDRS, YMRS, FAST[e], PSS[f], and WHOQoL[g] carried out at 4 weeks, 3 months, 6 months, and 9 months.	9 months
Faurholt-Jepsen et al (2019) [32]	Copenhagen, Denmark	84 patients with bipolar disorder (participants in MONARCA II trial)	Monsenso app for daily self-monitoring of mood and related symptoms.	No comparison used for outcomes of interest.	9 months
Hidalgo-Mazzei et al (2016) [33]	Barcelona, Spain	51	SIMPLe app: short 5-item screening tests completed daily. Weekly yes/no questions for DSM-5[h] criteria of manic and depressive episodes.	N/A	3 months
Hidalgo-Mazzei et al (2018) [34]	Barcelona, Spain	201	SIMPLe 1.5 (improved version of SIMPLe 1.0): short 5-item screening tests completed daily. Weekly Yes/No questions for DSM-5 criteria of manic and depressive episodes. Additional features included medication reminders, personalized prodromal symptoms, gamification module, mood chart sharing, and psychoeducational messages.	N/A	6 months
Li et al (2019) [35]	Hershey, PA	10 (bipolar disorder); 10 (healthy controls)	Twice-daily mood and stress self-report, once daily sleep measures.	N/A	14 days

Reference	Location	Participants, n	Mobile app–based intervention	Comparison (if applicable)	Duration
Saunders et al (2017) [36]	Oxford, United Kingdom	21	Mood Zoom app: daily mood monitoring. True Colours system: weekly mood measures.	N/A	12 weeks
Schwartz et al (2016) [37]	Pennsylvania	10 (bipolar I or II); 10 (healthy controls)	4 items on visual analog scale and 1 item on Likert scale completed twice per day.	N/A	2 weeks

[a]YMRS: Young Mania Rating Scale.

[b]HDRS: Hamilton Depression Rating Scale.

[c]N/A: not applicable.

[d]MADRS: Montgomery-Åsberg Depression Rating Scale.

[e]FAST: Functional Assessment Short Test.

[f]PSS: Perceived Stress Scale.

[g]WHOQoL: World Health Organization Quality of Life (abbreviated).

[h]DSM-5: Diagnostic and Statistical Manual of Mental Disorders, fifth edition.

Table 2. Summary of findings on mobile app use in bipolar symptom monitoring.

Reference	Completion rates	Correlation between data obtained via mobile app and comparator
Busk et al (2020) [25]	Average self-assessment adherence: 82.8%	Mood scores: • HDRS[a]: r=–0.40; P<.001; • YMRS[b]: r=0.22; P<.001
Carr et al (2018) [26]	20/43 (47%) of patients with bipolar disorder; 14/26 (54%) of patients with borderline personality disorder; 20/44 (45%) of healthy controls had satisfactory data; 14/26 (54%) of patients with borderline personality disorder	Variability of negative mood: • BD[c] median: –0.99 (IQR 0.85); • BPD[d] median: 1.71 (IQR 1.11); • Healthy control median: 0.35 (IQR 0.47) • BD vs BPD (FDR[e]): 1.57×10^{-3}; • BD vs HC (FDR): 2.31×10^{-2} Variability of positive mood: • BD median: –0.91 (IQR 0.70); • BPD median: 1.42 (IQR 0.56); • Healthy control median: 0.62 (IQR 0.52) • BP vs BPD (FDR): 1.21×10^{-3}; • BP vs HC (FDR): 6.13×10^{-1} (nonsignificant) Variability of irritability: • BD median: –0.56 (IQR 0.43); • BPD median: 1.01 (IQR 0.49); • Healthy control median: 0.33 (IQR 0.46) • BP vs BPD (FDR): 1.87×10^{-3}; • BP vs HC (FDR): 2.39×10^{-2}
Depp et al (2012) [27]	Intervention: 42.1%; Comparison: 82.9%; t_{35}=5.8; P<.001	**Mood ratings:** Intervention: • MADRS[f]: r=–0.567; P=.01 • YMRS: r=0.294; P=.24 Comparison: • r=–0.243; P=.35 • r=0.452; P=.07
Depp et al (2015) [28]	Intervention: 65%; Comparison: 83%	—[g]
Faurholt-Jepsen et al (2015) [29]	—	Mood: • HDRS-17: β=–0.058; P<.001 • YMRS: β=0.039; P<.001 Sleep: • HDRS-17: β=0.02; P=.21 • YMRS: β=–0.047; P=.03 Activity: • HDRS-17: β=–0.042; P<.001 • YMRS: β=0.048; P<.001 Stress: • HDRS-17: β=0.046; P<.001 • YMRS: β=0.012; P=.35
Faurholt-Jepsen et al (2015) [30]	Intervention: 93.03% (7.15% done retrospectively)	—

Reference	Completion rates	Correlation between data obtained via mobile app and comparator
Faurholt-Jepsen et al (2019) [31]	Adherence rate: 72.6%	Mood ratings: • HDRS: β=–0.033; *P*<.001 • YMRS: β=0.044; *P*<.001 Self-reported mixed symptoms: • Clinically rated mixed symptoms: β=3.40; *P*=.02 • PSS[h]: β=14.08; *P*<.001 • WHOQoL[i]: β=–7.80; *P*=.15 • FAST[j]: β=–2.02; *P*=.72 Irritability: • YMRS: β=0.023; *P*<.001 • PSS: β=11.32; *P*<.001 • WHOQoL: β=–11.59; *P*<.001 • FAST: β=–9.90; *P*<.001
Faurholt-Jepsen et al (2019) [32]	Reported in previous study [29]	Mood instability factor (number of mood changes over period evaluated by scale): • FAST: β=–12.04; *P*<.001 • PSS: β=10.52; *P*<.001 • WHOQoL: β=–12.17; *P*<.001
Hidalgo-Mazzei et al (2016) [33]	88% completion rate; 74% of users actively using app after 3 months	—
Hidalgo-Mazzei et al (2018) [34]	70/201 (35%) users dropped out during the first month; 30% of participants using the app regularly after 6 months	—
Li et al (2019) [35]	70% completion rate in bipolar patients and healthy controls	**Variability of symptoms:** Mood: • Bipolar ICC[k]: 0.55; healthy control ICC: 0.72; *P*<.001 Energy: • Bipolar ICC: 0.49; healthy control ICC: 0.61; *P*<.001 Speed of thoughts: • Bipolar ICC: 0.40; healthy control ICC: 0.67; *P*<.001 Impulsivity: • Bipolar ICC: 0.16; healthy control ICC: 0.68; *P*<.001 Sleep: • Bipolar ICC: 0.46; healthy control ICC: 0.30; *P*<.001
Saunders et al (2017) [36]	Daily questionnaire: median 86.67%; Weekly questionnaire: median 100%	—

Reference	Completion rates	Correlation between data obtained via mobile app and comparator
Schwartz et al (2016) [37]	Bipolar: 95%; Controls: 88%; *P*=.68	14-day mean of mood: • Bipolar median: 48.6; control median: 53.2; *P*=.04 14-day mean of energy: • Bipolar median: 44.7; control median: 52.1; *P*=.007 14-day range of mood: • Bipolar median: 48.0; control median: 32.5; *P*=.04 14-day range of thoughts: • Bipolar median: 59.5; control median: 26.5; *P*=.002 14-day range of impulsivity: • Bipolar median: 76; control median: 28.5; *P*=.005

[a]HDRS: Hamilton Depression Rating Scale.
[b]YMRS: Young Mania Rating Scale.
[c]BD: bipolar disorder.
[d]BPD: borderline personality disorder.
[e]FDR: false discovery rate.
[f]MADRS: Montgomery-Åsberg Depression Rating Scale.
[g]Not available.
[h]PSS: Perceived Stress Scale.
[i]WHOQoL: World Health Organization Quality of Life (abbreviated).
[j]FAST: Functional Assessment Short Test.
[k]ICC: intraclass correlation coefficient.

Data on Validity

A total of 4 papers identified for inclusion assessed the concurrent validity of mobile app–based self-report tools, all compared against the Young Mania Rating Scale (YMRS) and either the Hamilton Depression Rating Scale (HDRS) or the Montgomery-Åsberg Depression Rating Scale (MADRS) [25,27,29,31]. All 4 studies found a statistically significant association between mood ratings collected via self-report using a mobile app and clinical assessment using the HDRS or MADRS. In addition, 3 studies found a statistically significant association between mood ratings collected via self-report using a mobile app and clinical assessment using the YMRS [25,29,31]. The fourth study, however, did not observe a statistically significant relationship [27]. One study also found a statistically significant relationship between self-reported mixed symptoms and clinically rated mixed symptoms, as well as a statistically significant relationship between self-reported irritability and YMRS scores [31]. One study examined mood ratings that were reported using a paper-and-pencil tool as well [27]. They did not find a statistically significant correlation between mood ratings reported using a paper-and-pencil tool and either the MADRS or YMRS [27].

A total of 3 studies examined the ability of self-report scales administered via a mobile app to differentiate between known groups, a form of construct validity [26,35,37]. Of these, 2 studies evaluated the differences in the variability of symptoms (mood, irritability, energy, speed of thoughts, impulsivity, or sleep) between patients with bipolar disorder and healthy controls [26,35]. These studies found statistically significant differences in the variability of symptoms between the 2 groups, with the exception of variability of positive mood [26]. One study also compared the variability of negative mood, positive mood, and irritability between patients with bipolar disorder and patients with borderline personality disorder; this study observed a statistically significant difference between the 2 groups for all 3 variables studied [26]. One study examined the difference in the 14-day mean of participants' mood and energy, as well as the 14-day range of mood, thoughts, and impulsivity between patients with bipolar disorder and healthy controls [37]. Statistically significant differences were observed between the 2 groups for all 5 of these variables [37].

Additionally, 2 studies examined the convergent validity of self-report symptom assessments administered via a mobile app with instruments assessing related factors: the Functional Assessment Short Test (FAST), the Cohen Perceived Stress Scale (PSS), and the abbreviated World Health Organization Quality of Life scale (WHOQoL-BREF) [31,32]. A statistically significant relationship was observed between self-reported mixed symptoms and PSS scores, but not with WHOQoL-BREF or FAST scores [31]. A statistically significant association was observed for both irritability and mood instability determined using self-report compared with the FAST, PSS, and WHOQoL-BREF [31,32].

Data on Adherence

Varying levels of adherence to the reporting protocol, ranging from 42% to 95%, were reported among studies in which measures were administered once or twice daily, with all but 1 study that used once-daily administration having adherence rates >70% [25,27,28,30,31,33,35-37]. Two studies reported high dropout rates [26,34]. In 1 study, participants were asked to complete a 6-item assessment 10 times daily, with 59 out of 113 (52.2%) of participants dropping out across all 3 study groups [26]. The other study reported that 70 out of 201 (34.8%) participants dropped out during the first month, which was higher than the percentage of participants dropping out in another study using a similar mobile app [33,34]. Compliance rates were substantially higher for the paper-and-pencil conditions in the 2 studies reported by Depp et al [27,28]. However, the frequency of measure completion was not the same between the 2 groups, and the paper-and-pencil condition could complete the measure at any time, whereas the phone condition was time limited [27,28]. These differences may have contributed substantially to the differences in completion rates between conditions.

Discussion

Principal Findings

The overall results of this review suggest that mobile app–based self-report questionnaires demonstrate concurrent validity when compared with established measures of depression and mania and convergent validity when compared with other related assessment tools. Furthermore, current evidence indicates that mobile app–based self-report questionnaires are able to differentiate between patients with bipolar disorder and patients with borderline personality disorder or healthy controls. In terms of protocol adherence, variability was observed in completion rates, with higher overall adherence rates in participants completing questionnaires daily compared with twice daily. High dropout rates were observed when participants were asked to complete the measure 10 times per day.

In this review, 4 studies analyzed the association between the self-reporting of symptoms via a mobile app and clinical assessment tools. While all 4 studies found a statistically significant association between mood ratings collected via self-report and clinical assessment tools for depression, only 3 out of 4 studies found a statistically significant association between mood ratings collected via self-report and the YMRS. Of note, the study in which no statistically significant correlation was found compared YMRS scores to data collected over the entire study duration and to those collected during the first 6 weeks of the study [27]. As the YMRS assesses symptoms over the preceding 48 hours, the poor correlation may be at least partly attributable to the difference in time periods observed. Only 1 other study reported the period of data used in the comparison, comparing YMRS scores to data collected over the preceding 3 days [29]. This may be a more appropriate comparison, especially as one goal of app-based self-report scales is the detection of acute mood states and changes in symptoms over time.

Furthermore, data collected via the paper-and-pencil condition did not have a statistically significant correlation with either the MADRS or YMRS [27]. This suggests that app-based self-report scales may more accurately collect data on depressive symptoms compared with their paper-based counterparts. While there are few data comparing mobile assessments with rating scales administered via paper and pencil, it has been suggested elsewhere that participants may be more forthcoming when reporting symptoms through mobile assessments [38]. In addition, it has been shown that participants completing measures via paper and pencil may complete the entries retrospectively and hence, outside the specified time frame being assessed [17]. This may explain the seemingly increased accuracy of symptoms reported via app-based measures compared with paper and pencil.

A manic or depressive episode at study onset was an exclusion criterion for many of the studies identified [27-33]. In addition, 3 other studies indicated that patients were euthymic for the duration of the study [25,26,36]. The remaining studies did not state whether any participants experienced acute episodes of mania or depression. As such, it is unclear whether mobile app–based self-report tools can detect acute mood episodes in patients with bipolar disorder.

Some studies assessed the ability of mobile app–based self-report tools to differentiate between known groups [26,35,37]. These studies found statistically significant differences between patients diagnosed with bipolar disorder and healthy controls. While differences in mean mood and mean energy were observed between the 2 groups in 1 study, the magnitude of the difference in range of thoughts and range of impulsivity between the 2 groups was higher [37]. The 2 other studies comparing 2 known groups also observed differences in the variability of symptoms associated with bipolar disorder [26,35]. These findings suggest that the range and course of symptoms measured using mobile app–based self-report tools may allow us to distinguish patients with bipolar disorder from healthy controls.

Studies comparing data collected via self-report assessments administered via a mobile app to the FAST, PSS, and WHOQoL-BREF observed statistically significant associations between some data collected and these measures. As the FAST, PSS, and WHOQoL-BREF assess functional impairment, psychological distress, and quality of life, these findings suggest that data collected via self-report using a mobile app may also reflect other factors of clinical importance [39-41].

Lower rates of adherence to the protocol were observed in most studies in which assessments were administered twice daily compared with studies in which assessments were administered once daily. Furthermore, 1 study in which assessments were administered 10 times per day observed high dropout rates during its 3-month course [26]. These findings suggest that users may have difficulty completing multiple assessments per day but are able to manage assessments occurring once daily. Different proportions of participants dropped out in 2 studies administering similar mobile apps [33,34]. The reason for this is unclear. Previous data indicate that users value apps that are simple and intuitive to use [42]. The study in which higher

dropout rates were observed used a version of the app containing numerous additional features, so it is possible that users found the app more complicated and were less willing to continue regular use as a result [34].

Limitations

In this review, only English studies from peer-reviewed journals were considered. As very few (n=49) non-English papers were identified prior to screening, this was felt to have minimal impact on overall results. As there were large numbers of protocol papers identified, for which it is not possible to exclude unpublished data, it is also possible that publication bias may have resulted in missed negative findings. While 13 papers were identified for inclusion in this review, only 5 different research groups seem to be represented, based on the names and affiliations of authors. One group is represented in 5 studies, which is over one-third of those identified for inclusion [25,29-32]. This may contribute to bias; however, it is reassuring that the reported findings appear to be fairly consistent across the different groups included. As noted above, no study reported on the ability of mobile app–based self-report tools to detect acute mood episodes. As such, it is unclear whether these tools are suitable for this purpose.

Future Research

Further studies on the validity of mobile app–based assessment tools, especially studies evaluating the ability of these tools to detect acute mood states, will better inform us about the potential utility of these tools in clinical settings. Future research into the course of symptoms measured using these tools may also provide insights into the differences between patients with bipolar disorder and healthy controls. Furthermore, the use of repeated self-report questionnaires combined with physiological and behavioral monitoring, which have been examined elsewhere [43], and with other biomarkers also bears further investigation and may further our understanding of bipolar disorder.

Conclusions

These findings suggest that mobile app–based self-report tools are valid in the assessment of symptoms of mania and depression in euthymic patients with bipolar disorder. These findings also suggest that data on the range and variability of symptoms collected using a mobile app differ between patients with bipolar disorder and healthy controls and are significantly associated with other clinically important measures. It is unclear at this time whether these tools can be used to detect acute episodes of mania or depression in patients with bipolar disorder. Adherence data indicate that patients with bipolar disorder show good adherence to self-report assessments administered daily for the duration of the study periods evaluated.

Acknowledgments

This research was supported by the University of Alberta Department of Psychiatry Residency Project Fund. We thank Dr Nicholas Mitchell, assistant professor at the University of Alberta Department of Psychiatry, for his comments, which improved the manuscript.

Conflicts of Interest

ECC was the named fellow on a grant from Janssen Inc Canada, and the funds were used to contribute to support for research assistants involved in another research project. ECC is also the author of a mobile version of a rating scale for the assessment of suicidality. KJA has received two research grants in the last two years from Janssen Inc Canada (fellowship grants for trainees) and provided consultancy services (unpaid) for HLS Therapeutics. YS and SS have no relevant conflicts of interest to declare. The funder had no role in the design and conduct of the study or in the data collection, analysis, interpretation, or writing of this publication.

Multimedia Appendix 1
Assessment of risk of bias in included studies.
[PDF File (Adobe PDF File), 179 KB - formative_v5i1e13770_app1.pdf]

References

1. American Psychiatric Association. Bipolar and Related Disorders. In: Diagnostic and Statistical Manual of Mental Disorders. 5th ed. Arlington, VA: American Psychiatric Association; 2013:123-154.
2. O'Donnell LA, Deldin PJ, Grogan-Kaylor A, McInnis MG, Weintraub J, Ryan KA, et al. Depression and executive functioning deficits predict poor occupational functioning in a large longitudinal sample with bipolar disorder. J Affect Disord 2017 Jun;215:135-142. [doi: 10.1016/j.jad.2017.03.015] [Medline: 28319690]
3. Parker G, McCraw S, Hadzi-Pavlovic D. Unipolar and bipolar patient responses to a new scale measuring the consequences of depression. Psychiatry Res 2015 Dec 15;230(2):676-681. [doi: 10.1016/j.psychres.2015.10.024] [Medline: 26545613]
4. Rosa AR, Reinares M, Michalak EE, Bonnin CM, Sole B, Franco C, et al. Functional impairment and disability across mood states in bipolar disorder. Value Health 2010 Dec;13(8):984-988 [FREE Full text] [doi: 10.1111/j.1524-4733.2010.00768.x] [Medline: 20667057]

5. Alonso J, Petukhova M, Vilagut G, Chatterji S, Heeringa S, Üstün TB, et al. Days out of role due to common physical and mental conditions: results from the WHO World Mental Health surveys. Mol Psychiatry 2011 Dec;16(12):1234-1246 [FREE Full text] [doi: 10.1038/mp.2010.101] [Medline: 20938433]
6. Grande I, Berk M, Birmaher B, Vieta E. Bipolar disorder. Lancet 2016 Apr 09;387(10027):1561-1572. [doi: 10.1016/S0140-6736(15)00241-X] [Medline: 26388529]
7. Francis A, Gasparo P. Interval between symptom onset and hospitalization in mania. J Affect Disord 1994 Jul;31(3):179-185. [doi: 10.1016/0165-0327(94)90027-2] [Medline: 7963070]
8. Cassidy F. Insight in bipolar disorder: relationship to episode subtypes and symptom dimensions. Neuropsychiatr Dis Treat 2010 Oct 05;6:627-631 [FREE Full text] [doi: 10.2147/NDT.S12663] [Medline: 20957122]
9. Silva RDAD, Mograbi DC, Bifano J, Santana CMT, Cheniaux E. Insight in bipolar mania: evaluation of its heterogeneity and correlation with clinical symptoms. J Affect Disord 2016 Jul 15;199:95-98. [doi: 10.1016/j.jad.2016.04.019] [Medline: 27093493]
10. Hirschfeld RMA. The Mood Disorder Questionnaire: A Simple, Patient-Rated Screening Instrument for Bipolar Disorder. Prim Care Companion J Clin Psychiatry 2002 Feb;4(1):9-11 [FREE Full text] [doi: 10.4088/pcc.v04n0104] [Medline: 15014728]
11. Altman E, Hedeker D, Peterson JL, Davis JM. A comparative evaluation of three self-rating scales for acute mania. Biol Psychiatry 2001 Sep 15;50(6):468-471. [doi: 10.1016/s0006-3223(01)01065-4] [Medline: 11566165]
12. Kung S, Palmer BA, Lapid MI, Poppe KA, Alarcon RD, Frye MA. Screening for bipolar disorders: Clinical utilization of the Mood Disorders Questionnaire on an inpatient mood disorders unit. J Affect Disord 2015 Dec 01;188:97-100. [doi: 10.1016/j.jad.2015.08.060] [Medline: 26344754]
13. Areàn PA, Hoa Ly K, Andersson G. Mobile technology for mental health assessment. Dialogues Clin Neurosci 2016 Jun;18(2):163-169 [FREE Full text] [Medline: 27489456]
14. Torous J, Friedman R, Keshavan M. Smartphone ownership and interest in mobile applications to monitor symptoms of mental health conditions. JMIR Mhealth Uhealth 2014;2(1):e2 [FREE Full text] [doi: 10.2196/mhealth.2994] [Medline: 25098314]
15. Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR. Patient compliance with paper and electronic diaries. Control Clin Trials 2003 Apr;24(2):182-199. [doi: 10.1016/s0197-2456(02)00320-3] [Medline: 12689739]
16. Ainsworth J, Palmier-Claus JE, Machin M, Barrowclough C, Dunn G, Rogers A, et al. A comparison of two delivery modalities of a mobile phone-based assessment for serious mental illness: native smartphone application vs text-messaging only implementations. J Med Internet Res 2013;15(4):e60 [FREE Full text] [doi: 10.2196/jmir.2328] [Medline: 23563184]
17. Nicholas J, Larsen ME, Proudfoot J, Christensen H. Mobile Apps for Bipolar Disorder: A Systematic Review of Features and Content Quality. J Med Internet Res 2015;17(8):e198 [FREE Full text] [doi: 10.2196/jmir.4581] [Medline: 26283290]
18. Firth J, Cotter J, Torous J, Bucci S, Firth JA, Yung AR. Mobile Phone Ownership and Endorsement of "mHealth" Among People With Psychosis: A Meta-analysis of Cross-sectional Studies. Schizophr Bull 2016 Mar;42(2):448-455 [FREE Full text] [doi: 10.1093/schbul/sbv132] [Medline: 26400871]
19. Torous J, Chan SR, Yee-Marie TS, Behrens J, Mathew I, Conrad EJ, et al. Patient Smartphone Ownership and Interest in Mobile Apps to Monitor Symptoms of Mental Health Conditions: A Survey in Four Geographically Distinct Psychiatric Clinics. JMIR Ment Health 2014;1(1):e5 [FREE Full text] [doi: 10.2196/mental.4004] [Medline: 26543905]
20. Nicholas J, Fogarty AS, Boydell K, Christensen H. The Reviews Are in: A Qualitative Content Analysis of Consumer Perspectives on Apps for Bipolar Disorder. J Med Internet Res 2017 Apr 07;19(4):e105 [FREE Full text] [doi: 10.2196/jmir.7273] [Medline: 28389420]
21. Faurholt-Jepsen M, Torri E, Cobo J, Yazdanyar D, Palao D, Cardoner N, et al. Smartphone-based self-monitoring in bipolar disorder: evaluation of usability and feasibility of two systems. Int J Bipolar Disord 2019 Jan 04;7(1):1 [FREE Full text] [doi: 10.1186/s40345-018-0134-8] [Medline: 30610400]
22. Boateng GO, Neilands TB, Frongillo EA, Melgar-Quiñonez HR, Young SL. Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. Front Public Health 2018;6:149 [FREE Full text] [doi: 10.3389/fpubh.2018.00149] [Medline: 29942800]
23. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019 Aug 28;366:l4898. [doi: 10.1136/bmj.l4898] [Medline: 31462531]
24. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016 Oct 12;355:i4919 [FREE Full text] [doi: 10.1136/bmj.i4919] [Medline: 27733354]
25. Busk J, Faurholt-Jepsen M, Frost M, Bardram JE, Vedel Kessing L, Winther O. Forecasting Mood in Bipolar Disorder From Smartphone Self-assessments: Hierarchical Bayesian Approach. JMIR Mhealth Uhealth 2020 Apr 01;8(4):e15028 [FREE Full text] [doi: 10.2196/15028] [Medline: 32234702]
26. Carr O, Saunders KEA, Tsanas A, Bilderbeck AC, Palmius N, Geddes JR, et al. Variability in phase and amplitude of diurnal rhythms is related to variation of mood in bipolar and borderline personality disorder. Sci Rep 2018 Dec 26;8(1):1649 [FREE Full text] [doi: 10.1038/s41598-018-19888-9] [Medline: 29374207]

27. Depp CA, Kim DH, de Dios LV, Wang V, Ceglowski J. A Pilot Study of Mood Ratings Captured by Mobile Phone Versus Paper-and-Pencil Mood Charts in Bipolar Disorder. J Dual Diagn 2012 Jan 01;8(4):326-332 [FREE Full text] [doi: 10.1080/15504263.2012.723318] [Medline: 23646035]
28. Depp CA, Ceglowski J, Wang VC, Yaghouti F, Mausbach BT, Thompson WK, et al. Augmenting psychoeducation with a mobile intervention for bipolar disorder: a randomized controlled trial. J Affect Disord 2015 Mar 15;174:23-30. [doi: 10.1016/j.jad.2014.10.053] [Medline: 25479050]
29. Faurholt-Jepsen M, Vinberg M, Frost M, Christensen EM, Bardram JE, Kessing LV. Smartphone data as an electronic biomarker of illness activity in bipolar disorder. Bipolar Disord 2015 Nov;17(7):715-728. [doi: 10.1111/bdi.12332] [Medline: 26395972]
30. Faurholt-Jepsen M, Frost M, Ritz C, Christensen EM, Jacoby AS, Mikkelsen RL, et al. Daily electronic self-monitoring in bipolar disorder using smartphones - the MONARCA I trial: a randomized, placebo-controlled, single-blind, parallel group trial. Psychol Med 2015 Oct;45(13):2691-2704. [doi: 10.1017/S0033291715000410] [Medline: 26220802]
31. Faurholt-Jepsen M, Frost M, Christensen EM, Bardram JE, Vinberg M, Kessing LV. The association between mixed symptoms, irritability and functioning measured using smartphones in bipolar disorder. Acta Psychiatr Scand 2019 May;139(5):443-453. [doi: 10.1111/acps.13021] [Medline: 30865288]
32. Faurholt-Jepsen M, Frost M, Busk J, Christensen EM, Bardram JE, Vinberg M, et al. Is smartphone-based mood instability associated with stress, quality of life, and functioning in bipolar disorder? Bipolar Disord 2019 Nov;21(7):611-620. [doi: 10.1111/bdi.12796] [Medline: 31081991]
33. Hidalgo-Mazzei D, Mateu A, Reinares M, Murru A, Del Mar Bonnín C, Varo C, et al. Psychoeducation in bipolar disorder with a SIMPLe smartphone application: Feasibility, acceptability and satisfaction. J Affect Disord 2016 Aug;200:58-66. [doi: 10.1016/j.jad.2016.04.042] [Medline: 27128358]
34. Hidalgo-Mazzei D, Reinares M, Mateu A, Nikolova VL, Del Mar Bonnín C, Samalin L, et al. OpenSIMPLe: A real-world implementation feasibility study of a smartphone-based psychoeducation programme for bipolar disorder. J Affect Disord 2018 Dec 01;241:436-445. [doi: 10.1016/j.jad.2018.08.048] [Medline: 30145515]
35. Li H, Mukherjee D, Krishnamurthy VB, Millett C, Ryan KA, Zhang L, et al. Use of ecological momentary assessment to detect variability in mood, sleep and stress in bipolar disorder. BMC Res Notes 2019 Dec 04;12(1):791 [FREE Full text] [doi: 10.1186/s13104-019-4834-7] [Medline: 31801608]
36. Saunders KEA, Bilderbeck AC, Panchal P, Atkinson LZ, Geddes JR, Goodwin GM. Experiences of remote mood and activity monitoring in bipolar disorder: A qualitative study. Eur Psychiatry 2017 Mar;41:115-121. [doi: 10.1016/j.eurpsy.2016.11.005] [Medline: 28135594]
37. Schwartz S, Schultz S, Reider A, Saunders EFH. Daily mood monitoring of symptoms using smartphones in bipolar disorder: A pilot study assessing the feasibility of ecological momentary assessment. J Affect Disord 2016 Feb;191:88-93. [doi: 10.1016/j.jad.2015.11.013] [Medline: 26655117]
38. Torous J, Staples P, Shanahan M, Lin C, Peck P, Keshavan M, et al. Utilizing a Personal Smartphone Custom App to Assess the Patient Health Questionnaire-9 (PHQ-9) Depressive Symptoms in Patients With Major Depressive Disorder. JMIR Ment Health 2015;2(1):e8 [FREE Full text] [doi: 10.2196/mental.3889] [Medline: 26543914]
39. Bonnín CM, Martínez-Arán A, Reinares M, Valentí M, Solé B, Jiménez E, et al. Thresholds for severity, remission and recovery using the functioning assessment short test (FAST) in bipolar disorder. J Affect Disord 2018 Nov;240:57-62. [doi: 10.1016/j.jad.2018.07.045] [Medline: 30053684]
40. Nielsen MG, Ørnbøl E, Vestergaard M, Bech P, Larsen FB, Lasgaard M, et al. The construct validity of the Perceived Stress Scale. J Psychosom Res 2016 May;84:22-30. [doi: 10.1016/j.jpsychores.2016.03.009] [Medline: 27095155]
41. World Health Organization. The World Health Organization Quality of Life assessment (WHOQOL): position paper from the World Health Organization. Soc Sci Med 1995 Nov;41(10):1403-1409. [doi: 10.1016/0277-9536(95)00112-k] [Medline: 8560308]
42. Mendiola MF, Kalnicki M, Lindenauer S. Valuable features in mobile health apps for patients and consumers: content analysis of apps and user ratings. JMIR Mhealth Uhealth 2015 May 13;3(2):e40 [FREE Full text] [doi: 10.2196/mhealth.4283] [Medline: 25972309]
43. Seppälä J, De Vita I, Jämsä T, Miettunen J, Isohanni M, Rubinstein K, M-RESIST Group, et al. Mobile Phone and Wearable Sensor-Based mHealth Approaches for Psychiatric Disorders and Symptoms: Systematic Review. JMIR Ment Health 2019 Feb 20;6(2):e9819 [FREE Full text] [doi: 10.2196/mental.9819] [Medline: 30785404]

Abbreviations

ASRM: Altman Self-Rating Mania Scale
FAST: Functional Assessment Short Test
HDRS: Hamilton Depression Rating Scale
MADRS: Montgomery-Åsberg Depression Rating Scale
MDQ: Mood Disorder Questionnaire
PSS: Cohen Perceived Stress Scale

WHOQoL-BREF: abbreviated World Health Organization Quality of Life
YMRS: Young Mania Rating Scale

Edited by G Eysenbach; submitted 21.02.19; peer-reviewed by D Hidalgo-Mazzei, M Faurholt-Jepsen, D Fulford, A Paglialonga; comments to author 19.08.19; revised version received 13.01.20; accepted 02.10.20; published 08.01.21.

Please cite as:
Chan EC, Sun Y, Aitchison KJ, Sivapalan S
Mobile App–Based Self-Report Questionnaires for the Assessment and Monitoring of Bipolar Disorder: Systematic Review
JMIR Form Res 2021;5(1):e13770
URL: https://formative.jmir.org/2021/1/e13770
doi:10.2196/13770
PMID:33416510

Original Paper

Attitudes and Perceptions Toward COVID-19 Digital Surveillance: Survey of Young Adults in the United States

Lauren Maytin[1]; Jason Maytin[1]; Priya Agarwal[1]; Anna Krenitsky[1]; JoAnn Krenitsky[1], MPH; Robert S Epstein[1], MD, MS

Epstein Health LLC, Woodcliff Lake, NJ, United States

Corresponding Author:
Robert S Epstein, MD, MS
Epstein Health LLC
50 Tice Blvd
Suite 340
Woodcliff Lake, NJ, 07677
United States
Phone: 1 201 285 5800
Email: repstein@epsteinhealth.com

Abstract

Background: COVID-19 is an international health crisis of particular concern in the United States, which saw surges of infections with the lifting of lockdowns and relaxed social distancing. Young adults have proven to be a critical factor for COVID-19 transmission and are an important target of the efforts to contain the pandemic. Scalable digital public health technologies could be deployed to reduce COVID-19 transmission, but their use depends on the willingness of young adults to participate in surveillance.

Objective: The aim of this study is to determine the attitudes of young adults regarding COVID-19 digital surveillance, including which aspects they would accept and which they would not, as well as to determine factors that may be associated with their willingness to participate in digital surveillance.

Methods: We conducted an anonymous online survey of young adults aged 18-24 years throughout the United States in June 2020. The questionnaire contained predominantly closed-ended response options with one open-ended question. Descriptive statistics were applied to the data.

Results: Of 513 young adult respondents, 383 (74.7%) agreed that COVID-19 represents a public health crisis. However, only 231 (45.1%) agreed to actively share their COVID-19 status or symptoms for monitoring and only 171 (33.4%) reported a willingness to allow access to their cell phone for passive location tracking or contact tracing.

Conclusions: Despite largely agreeing that COVID-19 represents a serious public health risk, the majority of young adults sampled were reluctant to participate in digital monitoring to manage the pandemic. This was true for both commonly used methods of public health surveillance (such as contact tracing) and novel methods designed to facilitate a return to normal (such as frequent symptom checking through digital apps). This is a potential obstacle to ongoing containment measures (many of which rely on widespread surveillance) and may reflect a need for greater education on the benefits of public health digital surveillance for young adults.

(JMIR Form Res 2021;5(1):e23000) doi:10.2196/23000

KEYWORDS

attitude; perception; young adult; COVID-19; digital surveillance; population health technologies; surveillance; population; survey; adolescent

Introduction

The COVID-19 pandemic reached a disturbing milestone in the United States on November 21, 2020, as the number of confirmed cases surpassed 12 million, with the virus now spreading more rapidly and more broadly than ever before [1]. Since the summer of 2020, when states reopened businesses and public spaces, there has been a resurgence of cases as lockdowns were lifted and community spread intensified.

Young adults are believed to have played a major role in the increased number of cases and the heightened transmission of COVID-19 as social gatherings resumed and colleges and universities returned to campus [2]. Though hospitalization rates and mortality are currently lower for young adults than older adults [3], growing evidence suggests that younger generations are a major vector of COVID-19 transmission, comprising a relatively large proportion of the total confirmed cases [4]. Furthermore, COVID-19 incidence has increased among those aged 0-39 years as the pandemic has progressed, shifting the age distribution of cases over time from older to younger demographics [5].

While young adults play a significant role in the spread of COVID-19, they also tend to display a greater indifference toward the health risk posed by the pandemic and may be more resistant to policies aimed at reducing transmission. Young adults are significantly more likely to refuse a COVID-19 vaccine [6], less likely to closely follow COVID-19 news [7], and tend to see the pandemic as a greater risk to their finances than their health [8]. The relatively low mortality and severity of COVID-19 symptoms in this age group may contribute to this mindset. In addition, asymptomatic and mild cases comprise the majority of young COVID-19 cases and are believed to contribute significantly to community spread [9,10]. All of these factors (lower perceived vulnerability, reduced disease severity, higher engagement in social activities, and relatively high infection rate) made young adults a driving force of resurgences of COVID-19 [11].

As part of reopening plans, most states hired thousands of contact tracers to conduct public health surveillance and outreach to control and contain the spread of COVID-19 [12]. However, a recent study demonstrated that controlling the epidemic by manual contact tracing is not feasible given the infectiousness of COVID-19 and the high incidence of transmissions from presymptomatic or asymptomatic individuals [13]. These researchers proposed that an app providing instant digital contact tracing is needed for epidemic control. Beyond this kind of passive digital proximity and contact tracing, active symptom monitoring using mobile technology is also viewed as a key component for public health entities to better assess the community burden of COVID-19 [14]. Taken together, the active and passive surveillance of populations with digital public health tools has the potential to enable monitoring of COVID-19 status in real time and can be deployed rapidly and at scale, allowing targeted interventions to control spread.

To be successful, any COVID-19 precision public health control efforts that include digital surveillance must have a significant acceptance by the community. In particular, it is important to know how young adults would use these population health technologies for COVID-19 monitoring and whether they believe them to be necessary or beneficial. Therefore, we sought to understand the views of young adults regarding digital surveillance, including which aspects they would accept and which they would not, and define the factors that may influence their willingness to participate in digital monitoring of their movements or health status to help control the spread of COVID-19.

Methods

Study Design

The study was designed to be a national cross-sectional survey of young adults aged 18-24 years. Participants were recruited to answer an online questionnaire in which most questions were closed-ended by design, with 2-5 response options.

Target Sample

To be representative of the target US young adult population (those aged 18-24 years), the sample size was determined to be 500, assuming 95% confidence intervals, a 5% margin of error, and a completion rate of approximately 80%. Age and gender balancing were prespecified to ensure these strata were generalizable to US census data. All US regions were targeted.

Questionnaire Development

The survey instrument was designed to meet the specified study objectives. The survey scope and questionnaire inputs were based on a review of the limited COVID-19–related published literature available at that time [15,16]. In addition, 2 experts in the design and development of survey instruments for research reviewed the survey and provided additional comments and are noted in the acknowledgment section.

The online survey was designed so that a respondent was required to answer each question before they were directed to the next question. Respondents were not able to go back and change answers already entered. No identifying questions were asked and all survey responses were deidentified. Unanswered questions were not permitted, with the exception of the single open-ended question, which was optional.

Prior to the start of the survey, participants were provided introductory information that described COVID-19 and how it is transmitted. In addition, the concept of digital monitoring was defined and examples given. The language provided to the participants and the survey questions are included as a supplemental file (Multimedia Appendix 1).

Study Population (Inclusion and Exclusion Criteria)

The inclusion criteria for the study population were participants aged 18-24 years, of all genders, and residing in any of the census regions of the United States. Participants who did not meet these inclusion criteria were excluded from completing the online survey.

Survey Platform and Participant Recruitment

This online survey was fielded and conducted using the SurveyMonkey platform [17]. SurveyMonkey panels are recruited from a database of over 2.5 million people in the United States. These panels are representative of a current, diverse online population that voluntarily joins the SurveyMonkey platform for survey research. All panelists share demographic information about themselves such as gender, age, and region, and other targeting attributes such as job type or technology usage.

SurveyMonkey balances its panels according to census data of age and gender. Panelist profiles are regularly refreshed to

ensure respondent profiles are always current, and email and location verification is used to detect fraud and identify exclusions to prevent duplicate responses to the same survey. Ongoing panel calibration studies ensure response quality is on par with national benchmarks [18].

SurveyMonkey reaches panelists through technological means such as computer or mobile devices and offers a charitable incentive model. Panelists take surveys for charity and a chance to play a sweepstakes instant-win game. Panelists earn credits for completing surveys that they can redeem for gift cards or donate to charity [19].

Data Management and Analysis

Deidentified survey responses were collected on the SurveyMonkey platform and exported for analysis. The data were aggregated to ensure anonymity and key findings were summarized using descriptive statistics. Survey respondents who disagreed with participating in any form of digital surveillance for COVID-19 were categorized and compared to all other respondents. Chi-square tests were used to calculate *P* values for categorical variables and *t* tests were used for comparing continuous variables.

Ethical Considerations

Prior to patient recruitment and to comply with human subjects research requirements, we submitted our protocol and questionnaire to the Western Institutional Review Board. They determined the study was exempt under 45 CFR § 46.104(d)(2) because the research involved no more than minimal risk to subjects and only included interactions involving educational tests, survey procedures, interview procedures, or observations of public behavior.

Individuals aged 18-24 years were invited to participate and consented to participate via acceptance of a SurveyMonkey survey invitation. A brief introduction to the survey content was provided before participants opted to stop or continue to the question and answer portion of the survey. Each question of the survey included a “no response” option should the respondent prefer not to share that information. Respondents were allowed to withdraw from the survey at any time.

Results

SurveyMonkey audience sampling identified 809 prescreened panelists who were invited to participate. A total of 548 respondents initiated the survey. Of these, 35 participants abandoned the survey and 513 completed it, for a survey completion rate of 93.6%. Nearly all (99.8%) respondents used a mobile phone or tablet to complete the questionnaire. Table 1 summarizes the sociodemographics of participants. With a mean age of 20.6 years, nearly two-thirds of participants had partially completed or completed college. When asked whether they knew someone who had contracted COVID-19 or if they had contracted it themselves, 192 (37.4%) answered in the affirmative.

Table 1. Sociodemographic characteristics of survey participants (N=513).

Variables	Values
Mean age, years (SD)	20.6 (2)
Age range, years	18-24
Gender, n (%)	
Male	261 (51)
Female	246 (48)
Other	6 (1)
Highest educational level, n (%)	
Some high school	27 (5.3)
High school	134 (26)
Some college	198 (38.6)
College	121 (23.6)
Graduate/professional degree	33 (6.4)
Race/ethnicity, n (%)	
American Indian/Alaskan	6 (1.2)
Asian	79 (15.4)
Black/African American	60 (11.7)
Hispanic/Latino	96 (18.7)
Native Hawaiian	5 (1.0)
White	228 (44.4)
Multiethnic	39 (7)
Geographic region of residence, n (%)	
Northeast	124 (24.2)
Midwest	100 (19.5)
Southeast	100 (19.5)
South	80 (15.6)
West	39 (7.6)
Pacific	63 (12.3)
Alaska or Hawaii	7 (1.4)
Participant, close friend, and/or family had COVID-19, n (%)	
Yes	192 (37.4)
No	321 (62.6)

Most (n=383, 74.7%) young adults agreed that the COVID-19 pandemic is a public health crisis that poses significant risk to the health and safety of the US population (Table 2). However, only 56.9% (n=292) agreed that digital monitoring would be effective in helping to stop COVID-19 transmission. Even fewer young adults (n=236; 46.0%) agreed that digital monitoring would be necessary for a return to normal. Over half of young adult participants expressed privacy concerns about personal information being used in digital surveillance systems for COVID-19 monitoring.

Table 2. Responses of young adults in the United States (N=513) to questions about their beliefs and concerns regarding COVID-19 and digital surveillance.

Beliefs/concerns and responses	Participants, n (%)
Believes the COVID-19 pandemic is a public health crisis	
Strongly disagree	30 (5.8)
Disagree	44 (8.6)
Neutral	56 (10.9)
Agree	186 (36.3)
Strongly agree	197 (38.4)
Believes digital tracking/monitoring would help stop the spread of COVID-19	
Strongly disagree	43 (8.9)
Disagree	53 (10.3)
Neutral	125 (24.4)
Agree	188 (36.7)
Strongly agree	104 (20.3)
Believes digital tracking/monitoring is necessary to return to normal	
Strongly disagree	58 (11.3)
Disagree	76 (14.8
Neutral	143 (27.9)
Agree	145 (28.3)
Strongly agree	91 (17.8)
Concerns about the privacy of my information used for tracking/monitoring	
Strongly disagree	42 (8.2)
Disagree	85 (16.6)
Neutral	126 (24.6)
Agree	139 (27.1)
Strongly agree	121 (23.6)

Young adults expressed differences in opinion regarding their willingness to participate in the two types of potential COVID-19 tracking. Nearly half (n=231; 45.1%) stated their willingness for active monitoring (eg, they would manually input data via their phone/tablet), while only 33.3% (n=171) stated a willingness for passive monitoring (eg, monitoring of location and contacts by cell phone; Table 3). Approximately 25% of responses to each question were neutral. Young adults also appeared more willing to share personal health information than location or contact data.

Table 3. Responses of young adults in the United States (N=513) to questions about their willingness to participate in aspects of COVID-19 digital surveillance.

Questions and responses	Participants, n (%)
Willing to allow cell phone to passively monitor	
Strongly disagree	99 (19.3)
Disagree	113 (22)
Neutral	130 (25.3)
Agree	121 (23.6)
Strongly agree	50 (9.8)
Willing to actively input specific health data via phone/tablet	
Strongly disagree	58 (11.9)
Disagree	96 (18.7)
Neutral	25 (24.4)
Agree	167 (32.6)
Strongly agree	64 (12.5)
Willing to share results of any COVID-19 virus or antibody tests	
Strongly disagree	60 (11.7)
Disagree	52 (10.1)
Neutral	112 (21.8)
Agree	200 (39)
Strongly agree	89 (17.4)
Willing to share symptom information such as coughing, tiredness, or temperature	
Strongly disagree	68 (13.3)
Disagree	66 (12.9)
Neutral	106 (20.7)
Agree	193 (37.6)
Strongly agree	80 (15.6)
Willing to share my location and where I have been, tracked by my phone	
Strongly disagree	146 (28.5)
Disagree	105 (20.5)
Neutral	120 (23.4)
Agree	103 (20.1)
Strongly agree	39 (7.6)
Willing to share personal contact data such as who I was with, tracked by my phone	
Strongly disagree	138 (26.9)
Disagree	117 (22.8)
Neutral	114 (22.2)
Agree	99 (19.3)
Strongly agree	45 (8.8)

When asked to select the entities with whom they would be willing to share their personal information (Table 4), 62.6% (n=321) of young adults endorsed sharing information with their doctor/health care provider. Government agencies (local and federal) were the next most trusted option, with a similar number willing to share their data with local or federal government agencies and researchers. Only about one-third of the respondents were willing to share their information with schools or employers. By far, the group that respondents trusted least with their information was private companies.

Respondents were asked whether they would agree to participate in monitoring prior to engaging in select activities to help control

viral transmission (Table 5). None of the activities generated agreement from more than half of participants, and nearly one-fourth of respondents stated that they would not agree to any monitoring to participate in the listed activities.

There were no significant differences in mean age, gender, or education between those willing or neutral regarding being tracked for COVID-19 risk behaviors/symptoms versus those who were unwilling (Table 6). Significant differences were noted in race/ethnicity in univariate analyses.

There was a stepwise relationship between disbelief that COVID-19 is a public health crisis and lack of willingness to participate in digital tracking (Figure 1). At the extremes, only 15.2% (30/197) of respondents who strongly agreed that COVID-19 is a public health crisis were unwilling to be tracked, whereas 53.3% (16/30) of respondents who strongly disagreed that COVID-19 is a public health crisis were unwilling to be tracked.

Table 4. Responses of young adults in the United States (N=513) to a question regarding their willingness to share personal health data with individuals or organizations for COVID-19 digital surveillance.

Potential data recipient	Willing to share information, n (%)
Your doctor or other health care provider responsible for your care	321 (62.6)
Local, county, or state health department	199 (38.8)
Federal agencies or researchers (such as the Centers for Disease Control and Prevention or the National Institutes of Health)	189 (36.8)
Your school	174 (33.9)
Your employer	156 (30.4)
Companies such as Google, Microsoft, Apple, and Facebook	78 (15.2)
None of these	76 (14.8)

Table 5. Responses of young adults in the United States (N=513) to a question regarding activities they would be willing to agree to COVID-19 digital surveillance prior to participating.

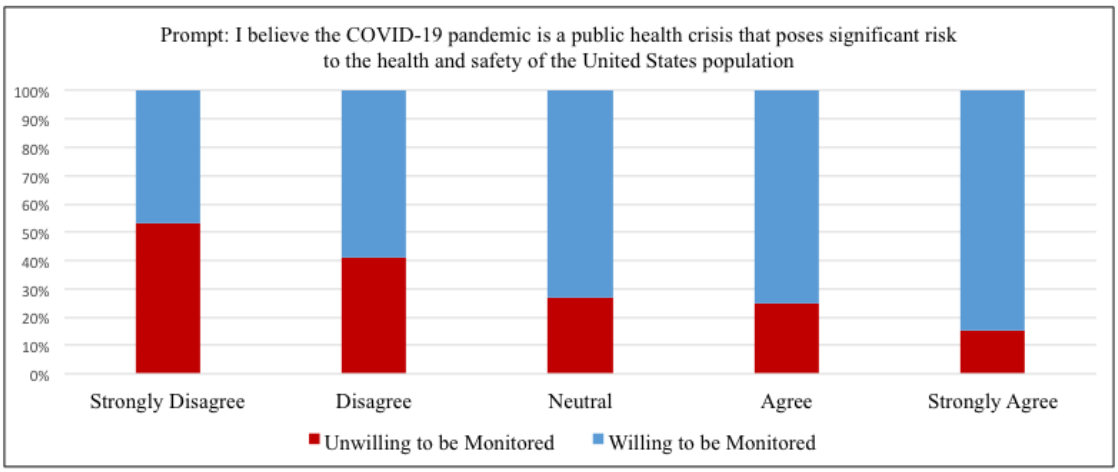
Activity	Willing participants, n (%)
Travel by airplane	222 (43.3)
Visit elderly or sick family members in a nursing home or hospital	189 (36.8)
Return to in-person attendance at school or work	183 (35.7)
Attend large social gatherings (eg, wedding, graduation, birthday party)	182 (35.5)
Shop at an indoor mall	176 (34.3)
Travel by public transportation (eg, bus, train, subway)	173 (33.7)
Dine indoors at a restaurant	166 (32.4)
Be a spectator at a sporting event, concert, or movie theatre	164 (32.0)
Be in public without a mask or face covering	158 (30.8)
Participate in organized team sporting events	150 (29.2)
Return to indoor places of worship	149 (29.0)
Gather with family and friends that do not live with you	142 (27.7)
None of these	118 (23.0)

Table 6. Characteristics of young adults in the United States (N=513) unwilling to be passively or actively digitally monitored for COVID-19.

Factors	All others (n=388)	Unwilling (n=125)	*P* value
Mean age, years (SD)	20.6 (4)	20.6 (4)	.69
Gender (female), n (%)	182 (47)	64 (51)	.59
Location, n (%)			.09
Alaska/Hawaii	4 (1)	3 (2.4)	N/A[a]
Midwest	77 (19.8)	23 (18.4)	N/A
Northeast	96 (24.7)	28 (22.4)	N/A
Pacific	48 (12.4)	15 (12)	N/A
South	65 (16.7)	15 (12)	N/A
Southeast	76 (19.6)	24 (19.2)	N/A
West	22 (5.6)	17 (13.6)	N/A
Highest education, n (%)			.45
Some high school	17 (4.3)	10 (0.1)	N/A
High school	100 (25.8)	34 (27.2)	N/A
Some college	149 (38.4)	49 (39.2)	N/A
College	95 (24.4)	26 (20.8)	N/A
Graduate school	27 (6.9)	6 (4.8)	N/A
Race/ethnicity, n (%)			.002
American Indian/Alaska Native	3 (0.1)	3 (2.4)	N/A
Asian	68 (17.5)	11 (8.8)	N/A
Black/African American	53 (13.7)	7 (5.6)	N/A
Hispanic/Latino or Spanish	72 (18.6)	24 (19.2)	N/A
Multiethnic	23 (5.9)	16 (12.8)	N/A
Native Hawaiian/Islander	5 (1.2)	5 (4)	N/A
White	164 (42.3)	64 (51.2)	N/A
Close friend or family member with COVID-19, n (%)	140 (36.1)	52 (41.6)	.32

[a]N/A: not applicable.

Figure 1. Relationship between belief in COVID-19 as a significant public health crisis and willingness to be digitally monitored from survey of young adults in the United States (N=513).



Discussion

Although population health technologies such as digital contact tracing and symptom surveillance have the potential to be instrumental in controlling the spread of COVID-19, the young adults in our survey were not in agreement about the necessity and effectiveness of these public health practices. Although the vast majority agreed that COVID-19 is a public health threat, they were nearly evenly divided in their opinions about digital monitoring being a viable solution.

This is in stark contrast to public health recommendations to rely on surveillance systems to safely participate in daily activities such as work or school during the pandemic. A report released by the Duke-Margolis Center for Health Policy recommends that a national surveillance system be established to control the ongoing COVID-19 pandemic and suppress future resurgences [20]. The authors assert that widespread serologic and diagnostic testing capabilities, routine data sharing, syndromic surveillance, and the use of digital tools and resources (particularly digital apps already in use by some health systems) will be essential for rapidly tracing and quarantining new cases and for ongoing COVID-19 surveillance.

However, our research indicates that a somewhat large and consistent percentage (20%-25%) of young adults feel neutral regarding questions of surveillance effectiveness or necessity and expressed privacy concerns related to digital surveillance. Young adults may be especially resistant to or ambivalent about digital monitoring efforts due to the increased importance of peer-group social interactions during this stage of development as well as underestimation of the risk involved in such behavior. Adolescent self-esteem and sense of identity is greatly influenced by socialization with peers, in ways that may make young adults reluctant to abide by strict distancing and monitoring guidelines [21]. There may be a need for education directed at this demographic on the usefulness of these precision public health tools for infection control.

We also sought to understand if there were differences between those who might endorse active digital surveillance (eg, inputting daily symptoms in a digital app) versus passive digital surveillance (eg, location tracking using mobile phone information). Respondents showed a clear preference for active monitoring, with almost 12% more respondents agreeing to active monitoring than passive monitoring. This was surprising given the popularity among young adults of various social media apps (such as Snapchat) that use passive location-tracking services [22]. It may be that active monitoring provides young adults with a sense of control over choosing the information they share. Passive monitoring (such as location tracking) may feel less transparent and more invasive.

This finding is similar to other research that has shown that many Americans believe that passive monitoring through cell phones would be ineffective against COVID-19 and such surveillance is unacceptable [23]. Distrust toward the government and concerns about security and privacy are the main barriers to adoption of digital surveillance tools to control COVID-19 [24]. Young Americans in particular have been shown to have lower levels of trust compared to older Americans. This lower interpersonal level of trust extends to institutions such as elected officials, police officers, the military, and other civic leaders [25], and perhaps provides context for why young adults may perceive contact-tracing efforts as an invasion of their privacy.

Our survey results underscore the trust and privacy concerns young adults have toward entities conducting digital surveillance and about the potential misuse of their personal data. Respondents frequently cited potential abuse of information as a major concern created by COVID-19 digital monitoring, including such open-ended remarks as “I do not trust anyone in power with this information” and “I worry about who would have access to the data and the potential for federal overreach.” Others reported privacy concerns with statements such as “Strong invasion of privacy,” “My information would be used for something else and sold to companies,” and similar sentiments.

An unexpected finding of our research was that respondents felt more comfortable sharing medical data such as symptoms and COVID-19 test results rather than information such as location or personal contacts. Moreover, they were most comfortable sharing this information with their physician or other health care provider responsible for their care and not public health authorities, schools, or other entities.

Most importantly, young adult respondents showed overwhelming distrust toward sharing these data with private companies listed in the questionnaire as Google, Apple, Microsoft, or Facebook, with only 15.2% agreeing to share their information with such entities. Pandemic surveillance programs developed by these private companies may be severely hindered by this distrust, regardless of their functionality. This is similar to research that suggests that private technology entities are the least likely source for which individuals would be willing to use surveillance apps and that there is no single, authoritative provider to which everyone would be willing to share the data necessary for digital tracking apps to be effective [26].

It may be that these trust, privacy, and personal data concerns will pose a significant challenge in convincing individuals to use digital tracking applications [27]. However, for young adults frustrated with social distancing policies, digital monitoring may provide something of a middle ground, allowing them to continue social activities while also granting some control over COVID-19 spread. However, our research indicated that even for social activities that may involve the risk of possible infection, young adults were not supportive of surveillance measures that could keep them or others safe. There was a lack of consensus regarding the types of activities for which it would be acceptable for young adults to be monitored prior to participating. Agreement to participate in digital monitoring was fairly low for all activities, even those such as visiting elderly family members in a hospital or nursing home, where the risk to themselves or others would presumably be greatest. There was no majority agreement on any one activity that would benefit from surveillance and 23% (n=118) of respondents were not willing to participate in surveillance for any activity.

This may support the idea that some individuals are unaware of the risks of their behaviors or underestimate their personal

risk of infection relative to others [28]. Young adults have been shown to exhibit a higher willingness to accept risk in situations where the likelihood of positive or negative outcomes is unknown [29]. As demonstrated during the COVID-19 pandemic, this tendency manifested itself in continued socialization with peers despite social distancing guidelines, and a willingness to neglect public health safety measures such as mask wearing and handwashing when compared to older adults [30].

It is for this very reason (the underestimation of health risk) that the use of digital public health tools for COVID-19 is needed. The utility of population health technologies for disease surveillance has been shown in the past in studies of the transmission of influenza-like illness [31]. Radin et al [31] found that wearable device data significantly improved predictions of influenza-like illness transmission and concluded that such data collection systems could be crucial in guiding responses to suppress future outbreaks. For young adults, digital monitoring could be essential for a safe “return to normalcy,” but digital public health technologies’ effectiveness against COVID-19 will be dependent on widespread trust and uptake [32].

In addition, agencies cannot implement surveillance programs without a better understanding of obstacles to their success. Although many digital health tools have been rapidly deployed in response to the COVID-19 crisis, continuous improvements, modifications, and customizations will have to be made to these tactics so they may be personalized for the various populations they are meant to protect. Such tactics must take into account that young adults have been more likely to experience job or wage loss because of the COVID-19 outbreak and are more likely to report high levels of emotional distress during the COVID-19 pandemic compared to older Americans [8]. These impacts may be motivating the behaviors of young adults during this pandemic and influencing their attitudes toward surveillance efforts. Further research is needed on how young adults perceive their behavior and risks of COVID-19 to inform the future use of digital health technologies to monitor and control this disease in this population. For young adults, it may be that more education about the benefits of such precision public health efforts and the involvement of their trusted health care providers would be a path worth exploring to achieve digital surveillance goals.

One limitation of this study is that the sample population is not a random sample of the United States. The survey was conducted using the SurveyMonkey platform, where all of the respondents previously agreed to survey participation and should thus be viewed as a convenience sample. However, the age, gender, and ethnic and racial distribution of our survey participants is representative of the general young adult population in the United States [33]. The regional distribution of responses is also relatively representative of the US population, with the highest proportion residing in the Northeast (n=124; 24.2%). An additional limitation is that due to the rapidly changing nature of the pandemic, new information becoming available, and case numbers and personal circumstances changing, the findings of this research may not be reflective of shifting opinions.

In conclusion, despite largely agreeing that COVID-19 represents a serious public health risk, a large proportion of young adults are reluctant to participate in digital monitoring to manage the pandemic. This is true for both commonly used methods of public health surveillance (such as contact tracing) and novel methods designed to facilitate a return to normal (such as frequent symptom checking through digital apps). As a major vector of COVID-19 transmission, the participation of young adults in digital COVID-19 monitoring is important to its success. Ultimately, the hesitancy of young adults to participate in digital monitoring must be addressed for these public health surveillance systems to be effective.

Acknowledgments

The authors would like to acknowledge Cynthia Girman, DrPH, and Molly Aldridge, MPH, for their expertise and reviewing and providing feedback on the study protocol and Institutional Review Board review submissions.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Questionnaire used to survey young adults in the United States regarding COVID-19 digital surveillance.
[PDF File (Adobe PDF File), 98 KB - formative_v5i1e23000_app1.pdf]

References

1. Maxouris C. US coronavirus cases top 12 million. An expert says spread is now 'faster' and 'broader' than ever. CNN. 2020 Nov 21. URL: https://www.cnn.com/2020/11/21/health/us-coronavirus-saturday/index.html [accessed 2020-11-21]
2. Cai W, Ivory D, Smith M, Lemonides A, Higgins L. What we know about coronavirus cases on campus. The New York Times. 2020 Jul 28. URL: https://www.nytimes.com/interactive/2020/07/28/us/covid-19-colleges-universities.html [accessed 2020-11-23]
3. Garg S, Kim L, Whitaker M, O'Halloran A, Cummings C, Holstein R, et al. Hospitalization Rates and Characteristics of Patients Hospitalized with Laboratory-Confirmed Coronavirus Disease 2019 - COVID-NET, 14 States, March 1-30, 2020.

MMWR Morb Mortal Wkly Rep 2020 Apr 17;69(15):458-464 [FREE Full text] [doi: 10.15585/mmwr.mm6915e3] [Medline: 32298251]
4. Boehmer TK, DeVies J, Caruso E, van Santen KL, Tang S, Black CL, et al. Changing Age Distribution of the COVID-19 Pandemic - United States, May-August 2020. MMWR Morb Mortal Wkly Rep 2020 Oct 02;69(39):1404-1409 [FREE Full text] [doi: 10.15585/mmwr.mm6939e1] [Medline: 33001872]
5. Malmgren J, Guo B, Kaplan H. COVID-19 confirmed case incidence age shift to young persons age 0-19 and 20-39 years over time: Washington State March - April 2020. medRxiv. Preprint published online on May 28, 2020 [FREE Full text] [doi: 10.1101/2020.05.21.20109389]
6. Expectations for a COVID-19 vaccine. AP-NORC at the University of Chicago. URL: https://apnorc.org/projects/expectations-for-a-covid-19-vaccine/ [accessed 2020-11-25]
7. Jurkowitz M, Mitchell A. Older Americans continue to follow COVID-19 news more closely than younger adults. Pew Research Center. 2020. URL: https://www.journalism.org/2020/04/22/older-americans-continue-to-follow-covid-19-news-more-closely-than-younger-adults/ [accessed 2020-11-25]
8. Schaeffer K, Raine L. Experiences with the COVID-19 outbreak can vary for Americans of different ages. Pew Research Center. 2020 Jun 16. URL: https://www.pewresearch.org/fact-tank/2020/06/16/experiences-with-the-covid-19-outbreak-can-vary-for-americans-of-different-ages/ [accessed 2020-11-23]
9. Epidemiology Working Group for NCIP Epidemic Response, Chinese Center for Disease Control and Prevention. [The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in China]. Zhonghua Liu Xing Bing Xue Za Zhi 2020 Feb 10;41(2):145-151. [doi: 10.3760/cma.j.issn.0254-6450.2020.02.003] [Medline: 32064853]
10. Yang R, Gui X, Xiong Y. Comparison of Clinical Characteristics of Patients with Asymptomatic vs Symptomatic Coronavirus Disease 2019 in Wuhan, China. JAMA Netw Open 2020 May 01;3(5):e2010182 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.10182] [Medline: 32459353]
11. Bosman J, Mervosh S. As virus surges, younger people account for 'disturbing' number of cases. The New York Times. 2020 Jun 25. URL: https://www.nytimes.com/2020/06/25/us/coronavirus-cases-young-people.html [accessed 2020-11-23]
12. Simmons-Duffin S. States nearly doubled plans for contact tracers since NPR surveyed them 10 days ago. National Public Radio. 2020 May 07. URL: https://www.npr.org/sections/health-shots/2020/04/28/846736937/we-asked-all-50-states-about-their-contact-tracing-capacity-heres-what-we-learned [accessed 2020-11-23]
13. Ferretti L, Wymant C, Kendall M, Zhao L, Nurtay A, Abeler-Dörner L, et al. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. Science 2020 May 08;368(6491):eabb6936 [FREE Full text] [doi: 10.1126/science.abb6936] [Medline: 32234805]
14. Chan AT, Brownstein JS. Putting the Public Back in Public Health - Surveying Symptoms of Covid-19. N Engl J Med 2020 Aug 13;383(7):e45. [doi: 10.1056/NEJMp2016259] [Medline: 32501663]
15. Zhong B, Luo W, Li H, Zhang Q, Liu X, Li W, et al. Knowledge, attitudes, and practices towards COVID-19 among Chinese residents during the rapid rise period of the COVID-19 outbreak: a quick online cross-sectional survey. Int J Biol Sci 2020;16(10):1745-1752 [FREE Full text] [doi: 10.7150/ijbs.45221] [Medline: 32226294]
16. McFadden SM, Malik AA, Aguolu OG, Willebrand KS, Omer SB. Perceptions of the adult US population regarding the novel coronavirus outbreak. PLoS One 2020;15(4):e0231808 [FREE Full text] [doi: 10.1371/journal.pone.0231808] [Medline: 32302370]
17. SurveyMonkey. URL: https://www.surveymonkey.com/ [accessed 2020-11-15]
18. Market Research Solutions. SurveyMonkey. URL: https://www.surveymonkey.com/market-research/data-quality/ [accessed 2020-11-15]
19. SurveyMonkey Contribute. SurveyMonkey. URL: https://help.surveymonkey.com/articles/en%20_US/kb/SurveyMonkey-Contribute?bc=Buying_Responses#Charities [accessed 2020-11-15]
20. McClellan M, Gottlieb S, Mostashari F, Rivers C, Silvis L. A national COVID-19 surveillance system: achieving containment. Margolis Center for Health Policy. 2020. URL: https://healthpolicy.duke.edu/sites/default/files/2020-06/a_national_covid_surveillance_system.pdf [accessed 2020-06-15]
21. Brown B, Lohr M. Peer-group affiliation and adolescent self-esteem: An integration of ego-identity and symbolic-interaction theories. Journal of Personality and Social Psychology 1987;52(1):47-55 [FREE Full text] [doi: 10.1037/0022-3514.52.1.47]
22. Constine J. Snapchat launches location-sharing feature Snap Map. TechCrunch. 2017 Jun 21. URL: https://techcrunch.com/2017/06/21/snap-map/ [accessed 2020-11-15]
23. Anderson M, Auxier B. Most Americans don't think cellphone tracking will help limit COVID-19, are divided on whether it's acceptable. Pew Research Center. 2020 Apr 16. URL: https://www.pewresearch.org/fact-tank/2020/04/16/most-americans-dont-think-cellphone-tracking-will-help-limit-covid-19-are-divided-on-whether-its-acceptable/ [accessed 2020-11-20]
24. Altmann S, Milsom L, Zillessen H, Blasone R, Gerdon F, Bach R, et al. Acceptability of App-Based Contact Tracing for COVID-19: Cross-Country Survey Study. JMIR mHealth uHealth 2020 Aug 28;8(8):e19857 [FREE Full text] [doi: 10.2196/19857] [Medline: 32759102]

25. Gramlich J. Young Americans are less trusting of other people and key institutions than their elders. Pew Research Center. 2019. URL: https://www.pewresearch.org/fact-tank/2019/08/06/young-americans-are-less-trusting-of-other-people-and-key-institutions-than-their-elders/ [accessed 2020-11-15]
26. Hargittai E, Redmiles E. Will Americans be willing to install COVID-19 tracking apps? Scientific American. URL: https://blogs.scientificamerican.com/observations/will-americans-be-willing-to-install-covid-19-tracking-apps/ [accessed 2020-11-10]
27. Fahey RA, Hino A. COVID-19, digital privacy, and the social limits on data-focused public health responses. Int J Inf Manage 2020 Dec;55:102181 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2020.102181] [Medline: 32836638]
28. Wise T, Zbozinek T, Michelini G, Hagan CC, Mobbs D. Changes in risk perception and protective behavior during the first week of the COVID-19 pandemic in the United States. PsyArXiv. Preprint posted online on March 19, 2020. [doi: 10.31234/osf.io/dz428]
29. Tymula A, Rosenberg Belmaker LA, Roy AK, Ruderman L, Manson K, Glimcher PW, et al. Adolescents' risk-taking behavior is driven by tolerance to ambiguity. Proc Natl Acad Sci U S A 2012 Oct 16;109(42):17135-17140 [FREE Full text] [doi: 10.1073/pnas.1207144109] [Medline: 23027965]
30. Hutchins HJ, Wolff B, Leeb R, Ko JY, Odom E, Willey J, et al. COVID-19 Mitigation Behaviors by Age Group - United States, April-June 2020. MMWR Morb Mortal Wkly Rep 2020 Oct 30;69(43):1584-1590 [FREE Full text] [doi: 10.15585/mmwr.mm6943e4] [Medline: 33119562]
31. Radin J, Wineinger N, Topol E, Steinhubl S. Harnessing wearable device data to improve state-level real-time surveillance of influenza-like illness in the USA: a population-based study. The Lancet Digital Health 2020 Feb;2(2):e85-e93 [FREE Full text] [doi: 10.1016/s2589-7500(19)30222-5]
32. Gasser U, Ienca M, Scheibner J, Sleigh J, Vayena E. Digital tools against COVID-19: taxonomy, ethical challenges, and navigation aid. The Lancet Digital Health 2020 Aug;2(8):e425-e434 [FREE Full text] [doi: 10.1016/S2589-7500(20)30137-0] [Medline: 32835200]
33. Status trends in the education of racial and ethnic groups. National Center for Education Statistics. URL: https://nces.ed.gov/programs/raceindicators/indicator_RAA.asp [accessed 2020-11-15]

Edited by G Eysenbach; submitted 30.07.20; peer-reviewed by D Levine, C García; comments to author 22.09.20; revised version received 25.11.20; accepted 19.12.20; published 08.01.21.

Please cite as:
Maytin L, Maytin J, Agarwal P, Krenitsky A, Krenitsky J, Epstein RS
Attitudes and Perceptions Toward COVID-19 Digital Surveillance: Survey of Young Adults in the United States
JMIR Form Res 2021;5(1):e23000
URL: http://formative.jmir.org/2021/1/e23000/
doi:10.2196/23000
PMID:33347420

Original Paper

Telemedicine for Remote Surgical Guidance in Endoscopic Retrograde Cholangiopancreatography: Mixed Methods Study of Practitioner Attitudes

Hedvig Aminoff[1], MA; Sebastiaan Meijer[1], PhD; Urban Arnelo[2,3], PhD, MD; Susanne Frennert[4], PhD

[1]Biomedical Engineering and Health Systems, KTH Royal Institute of Technology, Stockholm, Sweden

[2]Department of Surgical and Perioperative Sciences, Surgery, Umeå University, Umeå, Sweden

[3]Division of Surgery, CLINTEC, Karolinska Institutet, Stockholm, Sweden

[4]Department of Computer Science and Media Technology, Internet of Things and People Research Center, Malmö University, Malmö, Sweden

Corresponding Author:
Hedvig Aminoff, MA
Biomedical Engineering and Health Systems
KTH Royal Institute of Technology
Hälsovägen 11
Stockholm, 141 57 Huddinge
Sweden
Phone: 46 8 790 80 71
Email: hedvigam@kth.se

Abstract

Background: Telemedicine innovations are rarely adopted into routine health care, the reasons for which are not well understood. Teleguidance, a promising service for remote surgical guidance during endoscopic retrograde cholangiopancreatography (ERCP) was due to be scaled up, but there were concerns that user attitudes might influence adoption.

Objective: Our objective was to gain a deeper understanding of ERCP practitioners' attitudes toward teleguidance. These findings could inform the implementation process and future evaluations.

Methods: We conducted semistructured interviews with ERCP staff about challenges during work and beliefs about teleguidance. Theoretical constructs from the technology acceptance model (TAM) guided the thematic analysis. Our findings became input to a 16-item questionnaire, investigating surgeons' beliefs about teleguidance's contribution to performance and factors that might interact with implementation.

Results: Results from 20 interviews with ERCP staff from 5 hospitals were used to adapt a TAM questionnaire, exchanging the standard "Ease of Use" items for "Compatibility and Implementation Climate." In total, 23 ERCP specialists from 15 ERCP clinics responded to the questionnaire: 9 novices (<500 ERCP procedures) and 14 experts (>500 ERCP procedures). The average agreement ratings for usefulness items were 64% (~9/14) among experts and 75% (~7/9) among novices. The average agreement ratings for compatibility items were somewhat lower (experts 64% [~9/14], novices 69% [~6/9]). The averages have been calculated from the sum of several items and therefore, they only approximate the actual values. While 11 of the 14 experts (79%) and 8 of the 9 novices (89%) agreed that teleguidance could improve overall quality and patient safety during ERCP procedures, only 8 of the 14 experts (57%) and 6 of the 9 novices (67%) agreed that teleguidance would not create new patient safety risks. Only 5 of the 14 experts (36%) and 3 of the 9 novices (33%) were convinced that video and image transmission would function well. Similarly, only 6 of the 14 experts (43%) and 6 of the 9 novices (67%) agreed that administration would work smoothly. There were no statistically significant differences between the experts and novices on any of the 16 items (P<.05).

Conclusions: Both novices and experts in ERCP procedures had concerns that teleguidance might disrupt existing work practices. However, novices were generally more positive toward teleguidance than experts, especially with regard to the possibility of developing technical skills and work practices. While newly trained specialists were the main target for teleguidance, the experts were also intended users. As experts are more likely to be key decision makers, their attitudes may have a greater relative impact on adoption. We present suggestions to address these concerns. We conclude that using the TAM as a conceptual framework can support user-centered inquiry into telemedicine design and implementation by connecting qualitative findings to well-known analytical themes.

(JMIR Form Res 2021;5(1):e20692) doi:10.2196/20692

KEYWORDS

telemedicine; telementoring; teleconsulting; technology acceptance model; professional users; specialties, surgical/education; attitude; clinical staff; surgery; framework; surgeon; user-centered; interview; survey

Introduction

Background

Rapid development of surgical techniques and medical technology creates a continual need for retraining among surgeons [1,2]. Remote surgical guidance through telementoring and teleconsulting [3] can be a cost-effective way to facilitate teaching and training for less experienced surgeons [4] and support safe adoption of new clinical methods among experienced practitioners [5-9]. However, telemedicine innovations rarely move from the pilot stage to routine delivery [10,11]. As of yet, the factors contributing to acceptance and adoption of telemedicine are not well understood [12-15].

This study focuses on a promising telemedicine service for remote surgical guidance called teleguidance. The innovation was based on videoconferencing combined with transmission of high-quality endoscopic video and fluoroscopy. In this way, a high-volume clinical center could provide intraoperative consultation to a low-volume center during endoscopic retrograde cholangiopancreatography (ERCP), which is a highly specialized procedure for the diagnosis and treatment of biliary and pancreatic disease. A feasibility study demonstrated the impact on the clinical outcomes [16]. However, when teleguidance was to be scaled up to include more hospitals, some practitioners appeared less interested than anticipated. This raised concerns about implementation and about whether teleguidance could become an accepted way of working. We therefore conducted a theory-driven, user-centered study to gain a deeper understanding of practitioners' attitudes toward teleguidance.

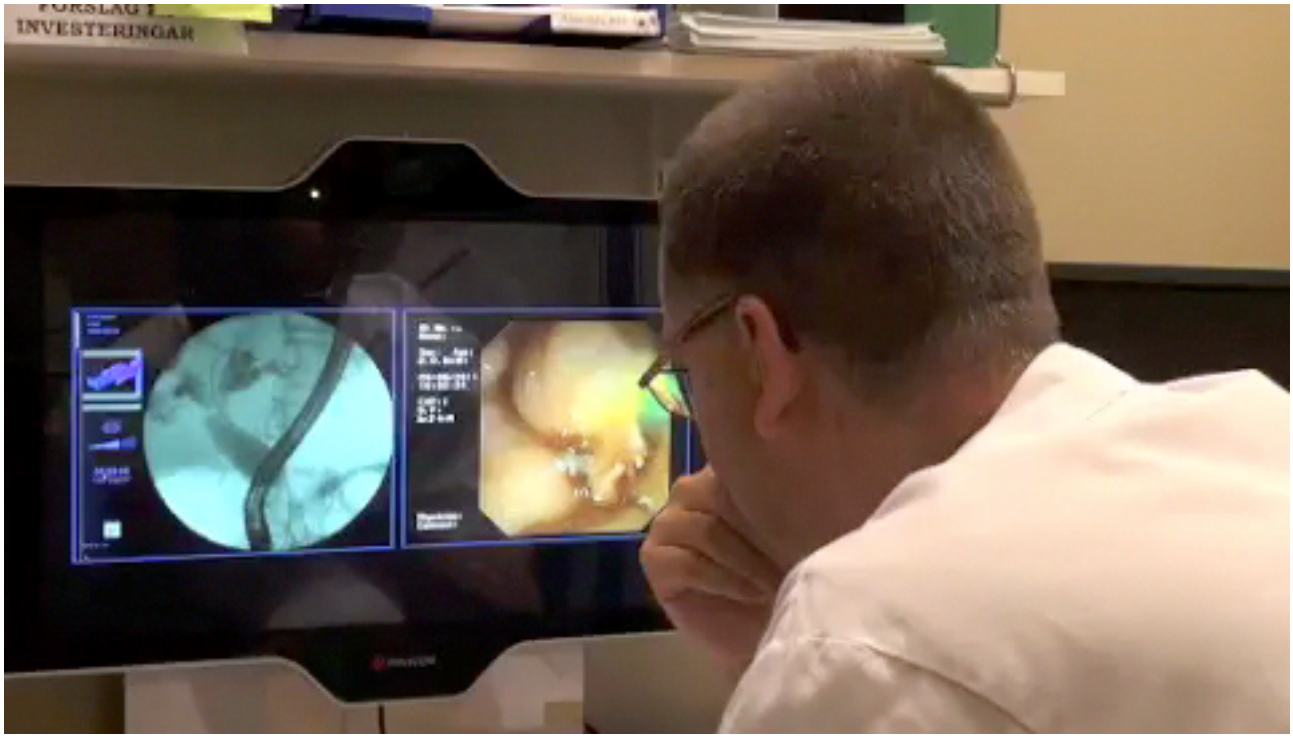
In the following sections, the clinical procedure and the telemedicine innovation are described. These sections also provide a general background to studying attitudes toward new technologies and a description of our methodological approach.

Teleguidance in ERCP

ERCP is a technically advanced procedure for the diagnosis and treatment of biliary and pancreatic disease. ERCP has a long learning curve in both technical skills and decision making. After the initial specialty training, it is necessary to perform a certain number of cases per year to sustain the acquired skills, which may be difficult in low-volume clinics. Continual retraining, necessary for keeping up with new surgical advances [17], is also sometimes difficult at hospitals with fewer resources for education and research. This can have consequences in the case of unusual conditions or if complications arise during the procedure. Difficulties during ERCP can lead to delays in diagnosis and treatment or painful or even life-threatening complications for patients who already have serious underlying health issues [18].

Practitioners in need of advice during an ERCP procedure, but without the option to ask an experienced colleague on site, can opt for alternative procedures or refer the patient to another hospital. Another option is to get in touch with fellow specialists by telephone. Teleguidance was developed to enhance this practice through videoconferencing coupled with high-quality video transmission of videos and radiographic imagery (Figure 1).

Figure 1. Teleguidance in endoscopic retrograde cholangiopancreatography.



A feasibility study has shown that teleguidance between a tertiary referral center and a low-volume hospital can result in improved quality of care [16]. This raised the prospect of extending this telemedicine service. A series of user-centered design efforts were initiated [19,20] as well as health-economic modeling [4]. Prior to expanding the service, an exploratory survey (Multimedia Appendix 1) showed variations in how clinicians rated their need of support. Subsequently, we wished to investigate practitioners' attitudes toward teleguidance more thoroughly.

Related Research

Traditionally, advanced surgical skills are learned by working together with experienced surgeons as mentors, progressing from shadowing to increasingly independent work. Necessary on-site training and retraining is sometimes difficult due to practical reasons and costs [9]. Numerous case studies—the earliest dating from the 1960s—have presented videoconferencing as a safe and efficacious way of providing surgical mentoring, enabling safe adoption of new techniques through remote expert guidance [2,6,7,21]. However, surgical telementoring is not widely used [5] and its impact over time is not well understood [6]. Implementation outcomes in health care are strongly affected by organizational context and how well an innovation answers to user needs [22]. This complexity makes it challenging to identify measurable determinants that provide an adequate image of implementation [23], in particular, regarding the quality and safety of products and services that function across multiple organizations [24].

Technology acceptance, a concept that relates beliefs and attitudes to use, is often considered an important determinant for technology implementation [22]. The technology acceptance model (TAM) [25] hypothesizes that people are more likely to use a technology if they believe it will be useful and easy to use. TAM was developed to provide validated measures for efficient early acceptance testing during development of office information systems [25,26]. The model defines two fundamental constructs: Perceived Usefulness signifying "the degree to which a person believes that using a particular system would enhance his or her job performance" and Perceived Ease of Use, representing "the degree to which a person believes that using a particular system would be free from effort" [27]. These central constructs were to be measured by a parsimonious questionnaire and were also intended to be transferable across different technologies and users [27]. TAM has been applied in many domains and TAM research has provided validation, extension, and elaboration of the central constructs [28]. However, surprisingly, few studies have investigated whether the model actually is a reliable predictor of use [28-31] or what makes a system useful [32]. Despite these weaknesses, the model is so frequently used that it has been described as a paradigm [30,33,34].

Various versions of TAM are commonly used in health care, and TAM has been extensively applied in studies of telemedicine [14,35-39]. While physician acceptance is commonly considered an important success factor [40], there does not appear to be any "optimal" version of TAM for telemedicine [14]. Despite its frequent use, TAM has shown shortcomings in health care [14,33,35]. Some of these have been attributed to the model's narrow focus on individual users' needs [39].

Another critique is that TAM invites quantitative treatment of narrowly defined theoretical constructs: the constructs themselves are treated as "black boxes," which in the end has led to a state of theoretical confusion and chaos "around the TAM's contribution to the understanding of technology acceptance and use" [32]. TAM was developed for prototype usability testing or system selection for office information technology systems, and the original definitions are grounded

in research about cognitive and affective factors affecting the use of single-user computer software [25]. Transferring TAM to health care raises fundamental issues about how the carefully designed and validated TAM scale items (Multimedia Appendix 2) match the concepts being studied [41]. Holden et al [41] posit that acceptance studies in health care could benefit from a broad set of perceptions about usefulness and adapting the variables to the context in question. Many studies have added extensions to include a wider range of situational and social influences than were originally defined [14], and alternative conceptualizations of "Ease of Use" for health care broaden the focus from individual users' beliefs about usability issues to include an organizational context [39].

The many adaptations and variants of TAM and the discussions about its relevance in health care highlight the importance of carefully considering what "Usefulness" and "Ease of Use" mean in each specific case. However, the model's construct definitions can support qualitative data collection, analyses, and the interpretation of findings [42,43].

The combination of ambivalent attitudes toward teleguidance and research showing that telemedicine adoption appears problematic motivated us to investigate ERCP practitioners' attitudes toward teleguidance. Guided by TAM, we studied the ways in which teleguidance might be perceived as "useful" and "easy to use" and how these perceptions vary across ERCP practitioners. These results are intended to inform the design and implementation process and to be valuable for understanding if and how teleguidance will be used at different clinical sites over time.

Methods

Questionnaire Design

Behavioral questionnaires should be grounded in the understandings of what is to be measured through contact with domain experts, and research in the relevant behavior domain should guide the construction of the items [44]. We conducted interviews to understand ERCP work and stakeholder beliefs about teleguidance's contribution to procedures or other interactions with ERCP work. This was followed by thematic analysis [45], where TAM served as a theoretical framework. The interview findings served as a basis for adapting the standard TAM questionnaire. The interviews and questionnaire are described in more detail in the following sections.

Interview Procedure

Initial key contacts with clinical staff at the different locations were set up by senior physicians at the tertiary referral center providing teleguidance, and a snowballing technique [46] gave us access to additional respondents. A total of 20 semistructured interviews with 10 ERCP specialists, 5 ERCP assistants, 3 technical staff, and 2 administrative staff from 5 hospitals were conducted.

Prior to each interview, the background, design, and purpose of the study, as well as the implications of participation were explained and also presented in printed from in order to gain informed consent [47] (Multimedia Appendix 3). Each participant was given a verbal presentation of the telemedicine service and a printed presentation with text and images describing teleguidance. Interviews were audio recorded and transcribed verbatim and treated as realist accounts. The interview length ranged from 30 minutes to 2 hours.

Coding and analysis proceeded through several iterations of reading the interview transcripts and refining the coding and themes. After coding, the data extracts were collated to help review patterns and relationships. The thematic analysis focused on identifying issues that might affect the use of teleguidance. At the outset, a number of themes were defined from the TAM model: statements related to performance, productivity, and effectiveness were to be coded as "Usefulness" issues, and issues related to expected usability or design issues as "Ease of Use" issues.

Questionnaire Procedure

The 16-item questionnaire used a 7-point Likert-type scale, with a midpoint alternative to respond as "neutral." It was made accessible as a closed, web-based 1-page survey provided through a web-based survey service during a 6-week period. Email invitations were sent to 25 physicians regularly performing ERCP at 15 different ERCP clinics, providing a link, information about estimated time to complete the survey (5 minutes), and information about data management and analysis (Multimedia Appendix 4). The number of practicing ERCP specialists in Sweden is small, and we made an effort to reach as many specialists in the field as possible that we had not yet interviewed. Analytical themes were operationalized as questionnaire items (Multimedia Appendix 5). Questionnaire items and phrasing were reworked a number of times to provide a succinct format and secure a high response rate. The order of the questions was randomized to avoid order effects. A few questions were also negatively phrased. Subjects' age, gender, and professional experience, and an option to add comments was included. The questionnaire was pilot tested [46] by 2 ERCP specialists at the University hospital. The results were treated with exploratory data analysis methods, and we created visual representations of the score distributions in the form of stacked columns (Multimedia Appendix 6 and Multimedia Appendix 7). To gain interpretability and improve the stability of the ratings, we dichotomized the Likert scale ratings [48] with a cut between disagreement and neutral (1,2,3,4) and agreement (5,6,7). We also ran a Mann-Whitney *U* test in SPSS (IBM Corp), a rank-based nonparametric test, to investigate differences in the attitude scores between experts and novices for each questionnaire item. For all tests, a *P* value less than .05 was considered statistically significant.

Results

Interview Results

The interviews gave us insight into practitioner beliefs about teleguidance's possible contributions to performance and factors that might interact with implementation. Four analytical themes (Multimedia Appendix 8) were defined through an iterative process of reviewing the transcripts, codes, and themes [49]. As the interview study progressed, it became clear that the respondents were not mentioning standard "Ease of Use" factors related to usability issues or design. What we found instead was

mention about how teleguidance might interact with the work system, workflow issues, patients' and management's attitudes, and whether the telemedicine service would cause practical/technical or administrative issues. Risk was an additional theme that emerged inductively from the data sets. On this basis, we replaced the concept "Perceived ease of use" with 2 concepts defined in the Consolidated Framework for Implementation Research (CFIR) [50]: compatibility and implementation climate. Compatibility refers to the fit between the innovation and the current work systems. Implementation climate is intended to reflect users' beliefs about whether the use of teleguidance would be expected and supported among important stakeholders. The themes were used as the basis for the questionnaire design.

Questionnaire Results

In this study, 25 ERCP specialists—14 experts (>500 procedures) and 9 novices (<500 procedures)—provided complete responses (100% completion rate); 2 respondents were removed as they had previously participated in the interviews. The perceived usefulness items and average dichotomized agreement scores are shown in Table 1.

Table 1. Dichotomized agreement scores of perceived usefulness by experts and novices in endoscopic retrograde cholangiopancreatography.

Perceived usefulness items	Experts (n=14), n (%)	Novices (n=9), n (%)
Performance		
Teleguidance can be a way for me to improve my technical skills in ERCP[a]	9 (64)	7 (78)
The ERCP that we perform are challenging enough for teleguidance to be of value	7 (50)	6 (67)
Overall, teleguidance would be beneficial for the quality and patient safety of the ERCP that we perform	11 (79)	8 (89)
Teleguidance would help to further develop the ERCP activities at this clinic	10 (71)	8 (89)
Effectiveness and productivity		
Teleguidance would allow my patients to get the appropriate treatment faster	10 (71)	6 (67)
Teleguidance can allow my patients to receive a better ERCP treatment	11 (79)	8 (89)
Teleguidance would allow us to perform a greater number of ERCP procedures	8 (57)	4 (44)
Teleguidance can help me get the most out of the time I set aside for ERCP	8 (57)	7 (78)

[a]ERCP: endoscopic retrograde cholangiopancreatography.

The average agreement ratings for usefulness items were 64% (~9/14) among experts and 75% (~7/9) among novices. The average agreement ratings for compatibility items were somewhat lower (experts 64% [~9/14], novices 69% [~6/9]). The averages were calculated from the sum of several items and therefore, they only approximate the actual values. Experts and novices tended to agree that teleguidance could contribute to better overall ERCP treatment for patients (11/14, 79% and 8/9, 89%; respectively) and improve quality and patient safety during ERCP procedures (11/14, 79% and 8/9, 89%; respectively). However, only 7 of the 14 experts (50%) thought that the procedures they performed were challenging enough for teleguidance to be of value, while 6 of the 9 novices (67%) agreed. The novices also agreed to a higher extent (7/9, 78%) than experts (8/14, 57%) that teleguidance could contribute to their effectiveness. Fewer experts (10/14, 71%) than novices (8/9, 89%) believed teleguidance could help develop ERCP activities at the clinic. Multimedia Appendix 6 shows the score distributions of the Usefulness items.

Experts gave high agreement scores (>75%) on both implementation climate items (Table 2), while relatively few novices agreed that management would be positive toward teleguidance (experts 11/14, 79%; novices 7/9, 56%). There were also concerns about the quality of video and image transmission and administration between hospitals, with relatively low agreements on "The quality of video and image transmission between hospitals will function well" (experts 5/14, 36%; novices 3/9, 33%) and "Administration between hospitals will function well" (experts 6/14, 43%; novices 6/9, 67%). Similarly, only 8 of the 14 experts (57%) and 6 of the 9 novices (67%) agreed that teleguidance would not create new patient safety risks. This contrasts with the scores for "teleguidance would be beneficial for the quality and patient safety of the ERCP that we perform," where both groups expressed positive expectations about the service's contribution to quality and patient safety.

Table 2. Dichotomized agreement scores on implementation climate and compatibility by experts and novices in endoscopic retrograde cholangiopancreatography.

Perceived ease of use items	Experts (n=14), n (%)	Novices (n=9), n (%)
Implementation climate		
Management would be positive that I use teleguidance	11 (79)	5 (56)
My patients would be positive that I use teleguidance	11 (79)	7 (78)
Compatibility		
Teleguidance is a way of working that could suit me and my workplace	10 (71)	9 (100)
The quality of video and image transmission between hospitals will function well	5 (36)	3 (33)
Administration between hospitals will function well	6 (43)	6 (67)
Teleguidance is unlikely to create risks for patients' confidentiality and integrity	10 (71)	7 (78)
Teleguidance is unlikely to create risks for staff integrity	11 (79)	7 (78)
Teleguidance is unlikely to create new patient safety risks	8 (57)	6 (67)

The graphs showing score distributions (Multimedia Appendix 7) illustrate that agreement was generally high; however, there was a large portion of neutral ratings on items, namely, “Management would be positive that I use teleguidance,” “The quality of video and image transmission between hospitals will function well,” and “Administration between hospitals will function well.”

Mann-Whitney *U* tests showed that there were no statistically significant differences between the novices and the experts on any of the 16 items ($P<.05$) (Multimedia Appendix 9).

Discussion

Principal Results

The interviews provided insight into the types of benefits teleguidance could provide and into the everyday ERCP work practices that might be affected by teleguidance. This served as important input for our questionnaire, where we defined the construct “Usefulness” in terms of how teleguidance might contribute to performance, productivity, and effectiveness. We exchanged the standard TAM construct “Ease of Use” for “Implementation climate and Compatibility” to better reflect concerns about how teleguidance might fit with the existing work system and if teleguidance might introduce new risks. Our main focus was to develop an understanding of the complexity of the domain and of the diversity among respondents, grounded in qualitative data.

The questionnaire allowed us to expand our inquiry to include a larger number of specialists in a domain where access to practitioners can be very difficult [51]. In the interviews, many expressed positive expectations about teleguidance, particularly that it could answer to challenges that novices were facing. However, many staff members also expressed concerns about how teleguidance would fit in with existing work system. The questionnaire results similarly showed that most respondents believed that teleguidance could contribute to the quality and safety of procedures and to the many anticipated technical and administrative issues—possibly even new patient safety issues. This indicates that practitioners had concerns that teleguidance might disrupt work.

While novices were the primary target group of the telemedicine service, teleguidance was designed with both novices and more experienced ERCP specialists in mind. We found that some experts were consistently skeptical toward teleguidance. As senior clinicians are more likely to be key decision makers [52], the attitudes among this group can have a greater relative importance for implementation than novices’ attitudes. Below, we discuss some possible reasons for and consequences of the differences between novices and experts and comment on the methodological concerns. We conclude with some practical suggestions for the implementation of teleguidance.

Differences Between Novices and Experts

Our interviews indicated that novices could be under considerable pressure during key phases of the procedure and they often saw room for improvement in current work processes. This was reflected in the questionnaire results, where novices had higher agreement scores on all the performance items and on the items for individual and organizational effectiveness. Novices may also have a lower threshold to work with videoconferencing than their more experienced colleagues, who also were older; previous use of information and communication technology in everyday life has been seen as a significant predictor for physicians’ telemedicine use [53]. The score distributions show that there were some items with many neutral responses, especially the “Implementation climate and Compatibility” items. This not only draws down the dichotomized score but it also indicates a challenge in asking potential users to form an opinion of a complex intervention, which might have complex outcomes, eg, patient safety issues. Experienced practitioners displayed a lower level of agreement that teleguidance could improve their individual performance, which may be explained by a less imperative need for support. However, developing integrative competence and taking part in a surgical innovation is an important aspect of sustaining acquired surgical competence [54], which is one of the aims of teleguidance.

While more experts than novices believed that management would be positive toward teleguidance, they also expected more administrative challenges. The differences in how experts and novices weighed these organizational aspects of teleguidance

may be explained both by differences in roles and in experience: senior practitioners were more likely to have managerial functions and hence might have a different perspective of management priorities and the changes that teleguidance might entail [2]. We can only speculate about the experts who gave negative ratings consistently: senior ERCP specialists with established practices and status may perceive teleguidance more as a challenge to traditional routines [55] or professional autonomy [56], and as a consequence, tend to prefer existing work practices [57]. However, it is likely that negative attitudes are a common source of bias in implementation studies, as these practitioners may very well decline to participate at all. Research has shown that differences in power and politics among professional groups influence the use of new technologies in health care [52]. Our findings underline the importance of including a variety of experience and roles in this type of study: some experts thought that novices would prefer "hands-on" help to remote guidance, while many novices themselves said the opposite. In addition, many of the interviewed nurses mentioned concerns about staff and patient integrity issues, but nurses were not included in the questionnaire nor was this view reflected in the doctors' questionnaire responses.

Practical Implications for Implementation

Teleguidance was initially developed to meet a wide range of challenges regarding the quality of ERCP procedures; it was not exclusively intended to serve the practical training needs of novices. Negative attitudes, even among a smaller group of experienced practitioners, may offset implementation efforts. However, our findings provide some guidance for design and implementation. The interviews provided insight into ERCP as a time-sensitive, collaborative team effort, highly dependent on medical technologies. Staff concerns that teleguidance might be an extra burden is based on daily experiences of ERCP work. Viewing teleguidance as a service rather than a new technology can widen the design perspective to include considerations what happens when two work systems are bridged by telemedicine. Experts' concerns about administrative issues and compatibility of work practices should be taken seriously; implementation efforts could benefit from identifying workflow issues, defining staff roles and tasks, and designating scheduling allowances for teleguidance-related tasks. If teleguidance is to be widely used, it is important to define and communicate the value of teleguidance even for experienced practitioners and to investigate incentives for experts' participation, since teleguidance also aims to support learning among experts. The SAGES telementoring initiative [2] differentiates between telementoring and teleconsultation, answering to different needs among experts and novices. Their definition of telementoring emphasizes a learning relationship between a mentor and a mentee and that telementoring occurs within an educational framework. Framing teleguidance in a similar way could benefit all parties; by clarifying relationships and objectives, teleguidance may be implemented as an explicit training effort for novices. Well-defined educational objectives might serve as incentives for novices to participate as well as increase the management support of telementoring. This could be a way to avoid inadvertently challenging the power and autonomy of incumbent experts.

As a contrast, teleguidance between two qualified experts might be defined as teleconsultation. This would signify different content and implications of the practice, with an emphasis on an exchange between peers, which may be experienced as less of a threat or intrusion by the more experienced ERCP specialists.

Limitations

This study has limitations due to the lack of lack of internal validity tests, which were beyond the scope of the study. This study does not attempt to exhaustively identify themes that may affect attitudes toward teleguidance, as the TAM guided toward predefined factors of interest. The number of respondents may be questioned, but as the total population of practicing ERCP specialists in Sweden is small and our respondents are highly representative, we claim to adequately cover variations among the groups. This study was exploratory, focused on developing an understanding of the complexity of the domain and of diversity among respondents. The quality of our findings is grounded in the qualitative data, rather than in statistical inference [58]. This study represents "the scientific discovery phase" [42], where empirical findings from a complex setting can ground hypotheses about behavior and design. In this sense, items with low agreement or ambiguous findings such as the seemingly contradictory beliefs about patient safety are valuable indications about how similar studies can be refined.

Conclusion

In our interviews, practitioners' descriptions of ERCP work and their beliefs about teleguidance did not resonate with the classical TAM questionnaire; they had no need to "work more quickly" or "make the job easier" nor did the interviews provide any statements about "Ease of Use" issues such as design features or usability. Instead, staff spoke of organizational demands deeply infused with clinical work, intense team collaboration, and constant organizational pressure for effectiveness and efficiency.

This means that teleguidance does not just have to answer to individual users' needs but also to organizational demands and priorities. Our findings show that introducing teleguidance is not "just" introducing new technology; teleguidance will change collaborative practices, linking locations that have their own sets of practices and priorities, which also can cause disruptions. We believe that the main cause for negative attitudes toward teleguidance is based on these concerns, which can be addressed during design and implementation. This study is an example of how TAM can support theory-guided user-centered design approaches to telemedicine development [31]. This may be a way to tackle the complexity of introducing technology in health care [59]. Using TAM in this way is also a return to the original intentions of the TAM, namely, to provide early user feedback to the system development processes, so as to gain better understanding of "how to improve user acceptance through design" [26].

Future Work

We suggest that using theories to guide the investigation of relevant user needs and expectations in a specific context is a way to inform the development and implementation of

telemedicine. By connecting findings to well-known analytical themes such as usefulness and terminology and concepts in frameworks such as CFIR [50], this type of qualitative approach can contribute to understanding the forces that shape the adoption of telemedicine and contribute to its effects. The complexity of introducing teleguidance across multiple clinical sites and ERCP teams will make evaluation particularly challenging [60,61]. Theories that accommodate complexity in studies of technological change are increasingly emphasized [59,62]. Building on our insights from this study, we hope to apply sociotechnical methods that are developed for understanding changes in complex and adaptive settings [59] and follow the introduction of teleguidance over time in a real-world context to study the ways in which teleguidance affects user behaviors.

Acknowledgments

The authors wish to thank Dr Erik Haraldsson at Skaraborgs Sjukhus, Skövde for the exploratory survey, which raised interest about practitioner attitudes toward teleguidance. This study would not have been possible without the support of Innovationsplatsen, Karolinska University Hospital and the staff at the participating ERCP clinics.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Results of the exploratory needs survey.
[DOCX File , 18 KB - formative_v5i1e20692_app1.docx]

Multimedia Appendix 2
Original technology acceptance model items (Davis, 1989).
[PNG File , 20 KB - formative_v5i1e20692_app2.png]

Multimedia Appendix 3
Consent form for the interviews.
[DOCX File , 16 KB - formative_v5i1e20692_app3.docx]

Multimedia Appendix 4
Invitation and information for participation in the questionnaire.
[DOCX File , 17 KB - formative_v5i1e20692_app4.docx]

Multimedia Appendix 5
Operationalizations and questionnaire items.
[PNG File , 96 KB - formative_v5i1e20692_app5.png]

Multimedia Appendix 6
Distribution of agreement scores on usefulness.
[PNG File , 51 KB - formative_v5i1e20692_app6.png]

Multimedia Appendix 7
Distribution of agreement scores on implementation climate and compatibility.
[PNG File , 57 KB - formative_v5i1e20692_app7.png]

Multimedia Appendix 8
Analytical themes and examples of coded content.
[PNG File , 68 KB - formative_v5i1e20692_app8.png]

Multimedia Appendix 9
Mann-Whitney U test for comparison between novices and experts: scale 1 (strongly disagree) to 7 (strongly agree).
[PNG File , 74 KB - formative_v5i1e20692_app9.png]

References

1. No authors listed. Statements on emerging surgical technologies and the evaluation of credentials. American College of Surgeons. Surg Endosc 1995 Feb;9(2):207-208. [doi: 10.1007/BF00191968] [Medline: 7597595]
2. Schlachta CM, Nguyen NT, Ponsky T, Dunkin B. Project 6 Summit: SAGES telementoring initiative. Surg Endosc 2016 Jun 6;30(9):3665-3672. [doi: 10.1007/s00464-016-4988-5]
3. Huang EY, Knight S, Guetter CR, Davis CH, Moller M, Slama E, et al. Telemedicine and telementoring in the surgical specialties: A narrative review. Am J Surg 2019 Oct;218(4):760-766. [doi: 10.1016/j.amjsurg.2019.07.018] [Medline: 31350010]
4. Brinne Roos J, Bergenzaun P, Groth K, Lundell L, Arnelo U. Telepresence-teleguidance to facilitate training and quality assurance in ERCP: a health economic modeling approach. Endosc Int Open 2020 Mar;8(3):E326-E337 [FREE Full text] [doi: 10.1055/a-1068-9153] [Medline: 32118106]
5. Augestad KM, Han H, Paige J, Ponsky T, Schlachta CM, Dunkin B, et al. Educational implications for surgical telementoring: a current review with recommendations for future practice, policy, and research. Surg Endosc 2017 Oct;31(10):3836-3846. [doi: 10.1007/s00464-017-5690-y] [Medline: 28656341]
6. Erridge S, Yeung DKT, Patel HRH, Purkayastha S. Telementoring of Surgeons: A Systematic Review. Surg Innov 2019 Feb;26(1):95-111. [doi: 10.1177/1553350618813250] [Medline: 30465477]
7. Shimizu S, Itaba S, Yada S, Takahata S, Nakashima N, Okamura K, et al. Significance of telemedicine for video image transmission of endoscopic retrograde cholangiopancreatography and endoscopic ultrasonography procedures. J Hepatobiliary Pancreat Sci 2011 May;18(3):366-374. [doi: 10.1007/s00534-010-0351-8] [Medline: 21127912]
8. El-Sabawi B, Magee W. The evolution of surgical telementoring: current applications and future directions. Ann Transl Med 2016 Oct;4(20):391 [FREE Full text] [doi: 10.21037/atm.2016.10.04] [Medline: 27867943]
9. Antoniou SA, Antoniou GA, Franzen J, Bollmann S, Koch OO, Pointner R, et al. A comprehensive review of telementoring applications in laparoscopic general surgery. Surg Endosc 2012 Aug;26(8):2111-2116. [doi: 10.1007/s00464-012-2175-x] [Medline: 22350150]
10. Zanaboni P, Wootton R. Adoption of telemedicine: from pilot stage to routine delivery. BMC Med Inform Decis Mak 2012 Jan 04;12:1 [FREE Full text] [doi: 10.1186/1472-6947-12-1] [Medline: 22217121]
11. Standing C, Standing S, McDermott M, Gururajan R, Kiani Mavi R. The Paradoxes of Telehealth: a Review of the Literature 2000-2015. Syst Res 2016 Dec 01;35(1):90-101. [doi: 10.1002/sres.2442]
12. Flodgren G, Rachas A, Farmer A, Inzitari M, Shepperd S. Interactive telemedicine: effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2015 Sep 07(9):CD002098 [FREE Full text] [doi: 10.1002/14651858.CD002098.pub2] [Medline: 26343551]
13. Ryu S. Telemedicine: Opportunities and Developments in Member States: Report on the Second Global Survey on eHealth 2009 (Global Observatory for eHealth Series, Volume 2). Healthc Inform Res 2012;18(2):153. [doi: 10.4258/hir.2012.18.2.153]
14. 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]
15. Augestad KM, Bellika JG, Budrionis A, Chomutare T, Lindsetmo R, Patel H, Mobile Medical Mentor (M3) Project. Surgical telementoring in knowledge translation--clinical outcomes and educational benefits: a comprehensive review. Surg Innov 2013 Jun;20(3):273-281. [doi: 10.1177/1553350612465793] [Medline: 23117447]
16. Påhlsson HI, Groth K, Permert J, Swahn F, Löhr M, Enochsson L, et al. Telemedicine: an important aid to perform high-quality endoscopic retrograde cholangiopancreatography in low-volume centers. Endoscopy 2013;45(5):357-361. [doi: 10.1055/s-0032-1326269] [Medline: 23468194]
17. Society of American GastrointestinalEndoscopic Surgeons (SAGES) Guidelines Committee. Guidelines for training in diagnostic and therapeutic endoscopic retrograde cholangiopancreatography (ERCP). Surg Endosc 2007 Jun;21(6):1010-1011. [doi: 10.1007/s00464-007-9341-6] [Medline: 17410399]
18. Olsson G, Arnelo U, Swahn F, Törnqvist B, Lundell L, Enochsson L. The H.O.U.S.E. classification: a novel endoscopic retrograde cholangiopancreatography (ERCP) complexity grading scale. BMC Gastroenterol 2017 Mar 09;17(1):38 [FREE Full text] [doi: 10.1186/s12876-017-0583-z] [Medline: 28274206]
19. Groth K, Arnelo U, Bergenzaun P, Lundell L, Aminoff H, Frykholm O, et al. Designing for interactivity in a tele-guidance setting. Int J Integr Care 2016 Nov 09;16(5):19. [doi: 10.5334/ijic.2568]
20. Frykholm O, Aminoff H, Groth K, Arnelo U. User-centered design of ERCP teleguidance. Göteborg: Gothenburg University Publications Electronic Archive; 2015. URL: http://hdl.handle.net/2077/38625 [accessed 2020-12-15]
21. Lee SP, Lee HL, Hahm JS, Choi HS, Joe I, Shimizu S. International live endoscopic multichannel demonstration using superfast broadband internet connections. Clin Endosc 2012 Mar;45(1):73-77 [FREE Full text] [doi: 10.5946/ce.2012.45.1.73] [Medline: 22741135]
22. Greenhalgh T, Swinglehurst D, Stones R. Rethinking resistance to 'big IT': a sociological study of why and when healthcare staff do not use nationally mandated information and communication technologies. HEALTH SERVICES AND DELIVERY RESEARCH 2014 Nov;39(2) [FREE Full text] [doi: 10.3310/hsdr02390] [Medline: 27466649]
23. Plsek PE, Greenhalgh T. Complexity science: The challenge of complexity in health care. BMJ 2001 Sep 15;323(7313):625-628 [FREE Full text] [doi: 10.1136/bmj.323.7313.625] [Medline: 11557716]

24. Carayon P. Human factors of complex sociotechnical systems. Appl Ergon 2006 Jul;37(4):525-535. [doi: 10.1016/j.apergo.2006.04.011] [Medline: 16756937]
25. Davis FD. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly 1989 Sep 21;13(3):319-340. [doi: 10.2307/249008]
26. Davis FD. User acceptance of information technology: system characteristics, user perceptions and behavioral impacts. International Journal of Man-Machine Studies 1993 Mar;38(3):475-487. [doi: 10.1006/imms.1993.1022]
27. Davis FD, Bagozzi RP, Warshaw PR. User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. Management Science 1989 Aug;35(8):982-1003. [doi: 10.1287/mnsc.35.8.982]
28. Lee Y, Kozar KA, Larsen KR. The Technology Acceptance Model: Past, Present, and Future. CAIS 2003;12. [doi: 10.17705/1cais.01250]
29. 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]
30. Turner M, Kitchenham B, Brereton P, Charters S, Budgen D. Does the technology acceptance model predict actual use? A systematic literature review. Information and Software Technology 2010 May;52(5):463-479. [doi: 10.1016/j.infsof.2009.11.005]
31. Harst L, Lantzsch H, Scheibe M. Theories Predicting End-User Acceptance of Telemedicine Use: Systematic Review. J Med Internet Res 2019 May 21;21(5):e13117 [FREE Full text] [doi: 10.2196/13117] [Medline: 31115340]
32. Benbasat I, Barki H. Quo vadis TAM? JAIS 2007 Apr;8(4):211-218. [doi: 10.17705/1jais.00126]
33. Holden RJ, Karsh B. The technology acceptance model: its past and its future in health care. J Biomed Inform 2010 Feb;43(1):159-172 [FREE Full text] [doi: 10.1016/j.jbi.2009.07.002] [Medline: 19615467]
34. Bagozzi R. The Legacy of the Technology Acceptance Model and a Proposal for a Paradigm Shift. JAIS 2007 Apr;8(4):244-254. [doi: 10.17705/1jais.00122]
35. Ammenwerth E. Technology Acceptance Models in Health Informatics: TAM and UTAUT. Stud Health Technol Inform 2019 Jul 30;263:64-71. [doi: 10.3233/SHTI190111] [Medline: 31411153]
36. Hayotte M, Thérouanne P, Gray L, Corrion K, d'Arripe-Longueville F. The French eHealth Acceptability Scale Using the Unified Theory of Acceptance and Use of Technology 2 Model: Instrument Validation Study. J Med Internet Res 2020 Apr 15;22(4):e16520 [FREE Full text] [doi: 10.2196/16520] [Medline: 32293569]
37. Safi S, Danzer G, Schmailzl KJ. Empirical Research on Acceptance of Digital Technologies in Medicine Among Patients and Healthy Users: Questionnaire Study. JMIR Hum Factors 2019 Nov 29;6(4):e13472 [FREE Full text] [doi: 10.2196/13472] [Medline: 31782741]
38. Schweitzer M, Huber L, Gorfer T, Hörbst A. Experiences With Developing and Using Vital Sign Telemonitoring to Support Mobile Nursing in Rural Regions: Feasibility and Usability Study. JMIR Nursing 2020 Apr 23;3(1):e17113. [doi: 10.2196/17113]
39. Jacob C, Sanchez-Vazquez A, Ivory C. Social, Organizational, and Technological Factors Impacting Clinicians' Adoption of Mobile Health Tools: Systematic Literature Review. JMIR Mhealth Uhealth 2020 Feb 20;8(2):e15935 [FREE Full text] [doi: 10.2196/15935] [Medline: 32130167]
40. Wade VA, Eliott JA, Hiller JE. Clinician acceptance is the key factor for sustainable telehealth services. Qual Health Res 2014 May;24(5):682-694. [doi: 10.1177/1049732314528809] [Medline: 24685708]
41. Holden RJ, Brown RL, Scanlon MC, Karsh B. Modeling nurses' acceptance of bar coded medication administration technology at a pediatric hospital. J Am Med Inform Assoc 2012;19(6):1050-1058 [FREE Full text] [doi: 10.1136/amiajnl-2011-000754] [Medline: 22661559]
42. Roth E, Patterson E. Using observational study as a tool for discovery: Uncovering cognitive and collaborative demands and adaptive strategies. In: Montgomery H, Lipshitz R, Brehmer B, editors. How Professionals Make Decisions. Boca Raton: CRC Press; 2005:379-393.
43. Roth EM. Uncovering the Requirements of Cognitive Work. Hum Factors 2008 Jun;50(3):475-480. [doi: 10.1518/001872008x288556]
44. Anastasi A. Evolving Concepts of Test Validation. Annu Rev Psychol 1986 Jan 21;37(1):1-16. [doi: 10.1146/annurev.ps.37.020186.000245]
45. Pettigrew AM. Longitudinal Field Research on Change: Theory and Practice. Organization Science 1990 Aug;1(3):267-292. [doi: 10.1287/orsc.1.3.267]
46. Breakwell G, Hammond S, Fife-Schaw C, editors. Surveys and sampling issues. In: Research methods in psychology. Thousand Oaks, USA: Sage Publications; 2002:89-104.
47. Arksey H, Knight P. Protecting Rights and Welfare. In: Arksey H, Knight P, editors. Interviewing for social scientists: An introductory resource with examples. London: Sage Publications; Jan 1, 2012:126-140.
48. Merkys G, Bubeliene D. Optimization of Data Processing and Presentation in Social Surveys: From Likert-Means to “Yes Percentage”. In: Agarwal N, Sakalauskas L, Weber GW, editors. Modeling and Simulation of Social-Behavioral Phenomena in Creative Societies. Cham, Switzerland: Springer International Publishing; Aug 22, 2019:12-25.
49. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]

50. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci 2009 Aug 07;4(1):50 [FREE Full text] [doi: 10.1186/1748-5908-4-50] [Medline: 19664226]
51. Furniss D, Randell R, O'Kane A, Taneva S, Mentis H, Blandford A. Fieldwork for Healthcare: Guidance for Investigating Human Factors in Computing Systems. In: Furniss D, Randell R, O'Kane A, Taneva S, editors. Synthesis Lectures on Assistive, Rehabilitative, and Health-Preserving Technologies. San Rafael: Morgan & Claypool Publishers; Dec 22, 2014:1-146.
52. Robert G, Greenhalgh T, MacFarlane F, Peacock R. Adopting and assimilating new non-pharmaceutical technologies into health care: a systematic review. J Health Serv Res Policy 2010 Oct;15(4):243-250. [doi: 10.1258/jhsrp.2010.009137] [Medline: 20592046]
53. Lacasta Tintorer D, Flayeh Beneyto S, Manresa JM, Torán-Monserrat P, Jiménez-Zarco A, Torrent-Sellens J, et al. Understanding the discriminant factors that influence the adoption and use of clinical communities of practice: the ECOPIH case. BMC Health Serv Res 2015 Sep 10;15:373 [FREE Full text] [doi: 10.1186/s12913-015-1036-4] [Medline: 26358037]
54. Wani S, Keswani RN, Petersen B, Edmundowicz SA, Walsh CM, Huang C, et al. Training in EUS and ERCP: standardizing methods to assess competence. Gastrointest Endosc 2018 Jun;87(6):1371-1382. [doi: 10.1016/j.gie.2018.02.009] [Medline: 29709305]
55. Chau PY, Hu PJ. Investigating healthcare professionals' decisions to accept telemedicine technology: an empirical test of competing theories. Information & Management 2002 Jan;39(4):297-311. [doi: 10.1016/S0378-7206(01)00098-2]
56. Walter Z, Lopez MS. Physician acceptance of information technologies: Role of perceived threat to professional autonomy. Decision Support Systems 2008 Dec;46(1):206-215. [doi: 10.1016/j.dss.2008.06.004]
57. Kim HW, Kankanhalli A. Investigating User Resistance to Information Systems Implementation: A Status Quo Bias Perspective. MIS Quarterly 2009;33(3):567. [doi: 10.2307/20650309]
58. Jansen H. The Logic of Qualitative Survey Research and its Position in the Field of Social Research Methods. Forum: Qualitative Social Research 2010 Mar 2;11(2) [FREE Full text] [doi: 10.17169/fqs-11.2.1450]
59. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Analysing the role of complexity in explaining the fortunes of technology programmes: empirical application of the NASSS framework. BMC Med 2018 May 14;16(1):66 [FREE Full text] [doi: 10.1186/s12916-018-1050-6] [Medline: 29754584]
60. Woods D, Dekker S. Anticipating the effects of technological change: A new era of dynamics for human factors. Theoretical Issues in Ergonomics Science 2000 Jan;1(3):272-282. [doi: 10.1080/14639220110037452]
61. Greenhalgh T, Papoutsi C. Studying complexity in health services research: desperately seeking an overdue paradigm shift. BMC Med 2018 Jun 20;16(1):95 [FREE Full text] [doi: 10.1186/s12916-018-1089-4] [Medline: 29921272]
62. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. BMJ 2015 Mar 19;350:h1258 [FREE Full text] [doi: 10.1136/bmj.h1258] [Medline: 25791983]

Abbreviations

CFIR: consolidated framework for implementation research
ERCP: endoscopic retrograde cholangiopancreatography
TAM: technology acceptance model

Edited by G Eysenbach; submitted 26.05.20; peer-reviewed by IC Lin, À Salvador Verges; comments to author 25.07.20; revised version received 02.10.20; accepted 26.10.20; published 11.01.21.

Please cite as:
Aminoff H, Meijer S, Arnelo U, Frennert S
Telemedicine for Remote Surgical Guidance in Endoscopic Retrograde Cholangiopancreatography: Mixed Methods Study of Practitioner Attitudes
JMIR Form Res 2021;5(1):e20692
URL: https://formative.jmir.org/2021/1/e20692
doi:10.2196/20692
PMID:33427670

Original Paper

Online Pelvic Floor Group Education Program for Women With Persistent Genital Arousal Disorder/Genito-Pelvic Dysesthesia: Descriptive Feasibility Study

Robyn A Jackowich[1], MSc; Kayla M Mooney[1], MSc; Evelyn Hecht[2], PT, ACT; Caroline F Pukall[1], PhD

[1]Department of Psychology, Queen's University, Kingston, ON, Canada

[2]EMH Physical Therapy, New York, NY, United States

Corresponding Author:
Caroline F Pukall, PhD
Department of Psychology
Queen's University
62 Arch Street, Humphrey Hall
Kingston, ON, K7L 3N6
Canada
Phone: 1 613 533 3200
Fax: 1 613 533 2499
Email: caroline.pukall@queensu.ca

Abstract

Background: Persistent genital arousal disorder/genito-pelvic dysesthesia (PGAD/GPD) is a highly distressing yet poorly understood condition characterized by persistent genito-pelvic sensations, often described as “genital arousal,” which occur in the absence of sexual desire. PGAD/GPD is associated with significant impairment in psychosocial and daily functioning; however, there are currently no empirically validated treatment algorithms for PGAD/GPD. Pelvic floor physical therapy exercises have been found to be effective at reducing other forms of genito-pelvic discomfort, such as vulvodynia, and may also be beneficial to those experiencing PGAD/GPD. Many individuals with PGAD/GPD report difficulty finding a health care provider who is knowledgeable about PGAD/GPD; therefore, pelvic floor education and exercises in an online format may have the potential to reach more individuals in need.

Objective: This study examined the feasibility of an online pelvic floor group education program; descriptively assessed outcomes related to distress, discomfort, catastrophizing, and mood; and obtained feedback from participants in order to inform the development of improved online group programs.

Methods: Fourteen women with current symptoms of PGAD/GPD attended an online, 8-session pelvic floor group education program. Participants completed questionnaires of symptoms (ie, symptom distress, discomfort) and psychosocial well-being (ie, depression, anxiety, symptom catastrophizing) prior to the group sessions (Time 1), immediately after the final group session (Time 2), and 6 months following the final group session (Time 3). Participants also completed an anonymous feedback questionnaire immediately following the group program.

Results: Overall, participants who attended a larger number of the group sessions (>5 sessions, n=7) appeared to report lower baseline (Time 1) symptoms and psychosocial impairment than those who attended fewer sessions (<5 sessions, n=7). A pattern of small improvements was seen following the group sessions on symptom and psychosocial outcomes. In the feedback questionnaire, breathing and relaxation exercises were described to be the most helpful home practice exercises, and participants rated sessions on (1) the relationship between emotions and PGAD/GPD symptoms and (2) relaxation exercises to be the most helpful. A number of barriers to participation in the group program were also identified, including comorbid health concerns and lack of personal time to complete the program/exercises.

Conclusions: Online interventions provide an opportunity to reach international participants who may otherwise struggle to access a knowledgeable provider for their PGAD/GPD symptoms. Addressing barriers may help to increase participants’ abilities to engage in the program. Future programs may seek to integrate a greater focus on relaxation strategies and cognitive-affective strategies for managing PGAD/GPD symptoms.

(JMIR Form Res 2021;5(1):e22450) doi:10.2196/22450

KEYWORDS

persistent genital arousal disorder; genitopelvic dysesthesia; online program; pelvic floor; pilot

Introduction

Persistent genital arousal disorder/genito-pelvic dysesthesia (PGAD/GPD) is a highly distressing yet misunderstood condition characterized by distressing genito-pelvic sensations (ie, dysesthesias) often described as "genital arousal" [1]. Although many people assume that sensations of genital arousal are pleasant, wanted, and aligned with one's internal sense of feeling "turned on", PGAD/GPD represents a clear example of disagreement between the physical sensations of what would commonly occur in response to effective sexual stimulation and the subjective experience of those sensations. Part of the negative experience of these sensations is due to their extreme nature: the sensations are most often prolonged, persistent, and difficult—if not almost impossible—to stop [2]. In addition, they are commonly described as painful [1,3]. These sensations can occur in response to a variety of sexual and nonsexual triggers, or they may occur suddenly and unexpectedly [2]. Studies have indicated that levels of distress in response to these symptoms are predominantly moderate to high, defined as a mean of around 7 on a scale from 0 (no distress) to 10 (extremely high distress) [2,4].

Despite the recent emergence of clinical and research attention to this pattern of symptoms, the most commonly used diagnostic manuals do not yet include a formal diagnosis of PGAD/GPD, with the exception of the most recent version of the International Classification of Diseases (ICD-11 [5]). Additional efforts have been made, and the most recent classification system for sexual dysfunctions published by the International Society for the Study of Women's Sexual Health (ISSWSH) includes criteria for PGAD/GPD [6]. The ISSWSH criteria are based on expert opinion and consist of the following: persistent or recurrent, unwanted or intrusive, distressing feelings of genital arousal, or being on the verge of orgasm, not associated with concurrent sexual interest, thoughts, or fantasies with a duration of 6 months or more. These feelings can be associated with: (a) limited resolution, no resolution, or aggravation of symptoms by sexual activity with or without aversive or compromised orgasm; (b) aggravation of symptoms by certain circumstances; (c) despair, emotional lability, catastrophizing, or suicidality; and (d) inconsistent evidence of genital arousal (eg, vaginal lubrication) during symptoms.

Although the prevalence of PGAD/GPD is unknown because of a lack of large-scale epidemiological studies, estimates from other sources exist. Based on their sample of women who presented at a sexual health clinic in the United Kingdom, Garvey and colleagues (2009) estimated that PGAD/GPD may affect approximately 1% of women [7]. More recently, three community samples from Canada, the United States, and Italy have found a similar prevalence rate, with 0.6% to 2.7% of women endorsing all of the characteristic features of PGAD/GPD at a moderate or higher frequency [8,9]. It is important to note that most of the clinical and research literature focuses on women with symptoms of PGAD/GPD; however, case studies describing similar symptoms in men have also emerged [10]. Given the significant representation of women in the current research literature, the research cited in this paper focuses on PGAD/GPD in women.

No empirically validated treatment algorithms for PGAD/GPD exist. Management commonly consists of pharmacological approaches, psychological interventions, and pelvic floor physical therapy [10], with some health care providers offering surgical interventions [11]. However, none of these treatment options have been formally tested or validated. A conservative approach to symptom management is often recommended, with the options being, in part, guided by the patients' preferences and level of distress and the health care providers' expertise and referral base [12]. Although there may be variations in the specific options and the timing of these options offered by various health care providers, many agree that the symptom-related distress must be specifically addressed in those with PGAD/GPD given the high frequency of self-reported suicidal ideation in this group [13]. However, access to treatment remains a barrier for many of those with PGAD/GPD due to the nature of the symptoms; many affected patients report shame and embarrassment surrounding the communication of their symptoms to others, including health care providers [14]. Even those who approach their health care providers may leave the situation feeling misunderstood and stigmatized because PGAD/GPD and its possible treatment options are not well known or understood [14]. Furthermore, the practical aspects of travelling to see a health care provider who is knowledgeable about PGAD/GPD may present major obstacles, ranging from financial constraints, to geographic barriers, to the challenges posed by travelling, which may significantly aggravate symptoms (eg, vibrations from car, prolonged sitting [2,14]).

In an effort to examine the feasibility of an accessible therapeutic option geared toward alleviating distress for those with PGAD/GPD, we piloted an online group program focusing on pelvic floor education and exercises and distress reduction for women with PGAD/GPD. Our focus on the pelvic floor in this group is based on a case study of a woman with PGAD/GPD who was successfully treated via pelvic floor rehabilitation [15] as well as on empirically based recommendations for other conditions characterized by genito-pelvic discomfort/dysesthesia (eg, vulvodynia [16]). Our aims were to (1) examine the feasibility of an online group program; (2) descriptively assess outcomes related to distress, discomfort, catastrophizing, and mood; and (3) obtain feedback from participants in order to inform the development of improved online group programs.

Methods

Participants

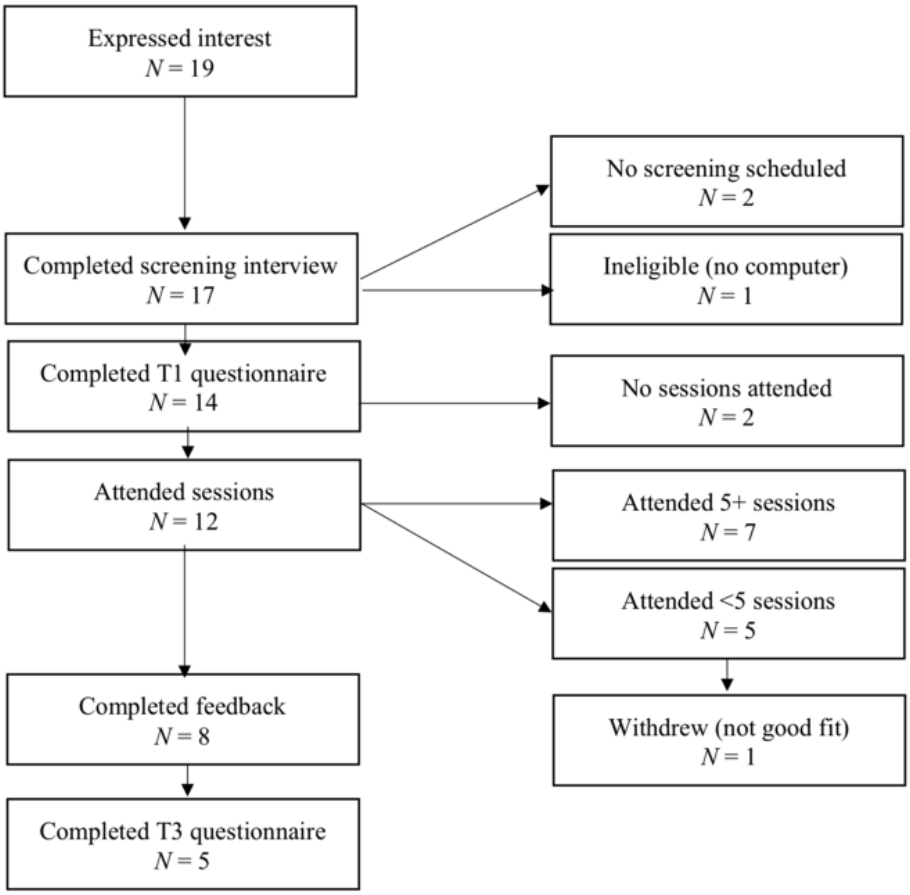
Participants were women who were experiencing symptoms of PGAD/GPD for a minimum of 3 months. The inclusion criteria for PGAD/GPD were based on its clinical descriptors [6,17]. PGAD/GPD includes experiencing feelings of persistent, involuntary genital arousal sensations that (1) are not fully relieved by one or more orgasms; (2) occur in the absence of

subjective feelings of sexual arousal; (3) persist (ie, last longer than 30 minutes); (4) are experienced as intrusive and unwanted; and (5) are experienced as subjectively distressing.

In order to be eligible to participate, participants were also required to be 18 years of age or older, fluent in English, and not experiencing any other serious mental health concerns that would interfere with their ability to participate in the group sessions (examples included, but were not limited to, substance use disorder, borderline personality disorder, and psychosis). In order to determine the effectiveness of the program, participants were asked to not make any changes to their PGAD/GPD treatments or medications during the course of the 8-session weekly program.

Participants were recruited online via social media advertisements (ie, Facebook, Twitter, blog posts) and postings on relevant websites and listservs to patients and health care professionals. Our laboratory also has a database of participants from past research studies who consented to be notified about additional research opportunities. Initially, 19 women contacted the study team to express interest in participating, and 17 underwent a phone screening with a member of the study team to confirm eligibility. See Figure 1 for a depiction of participant flow.

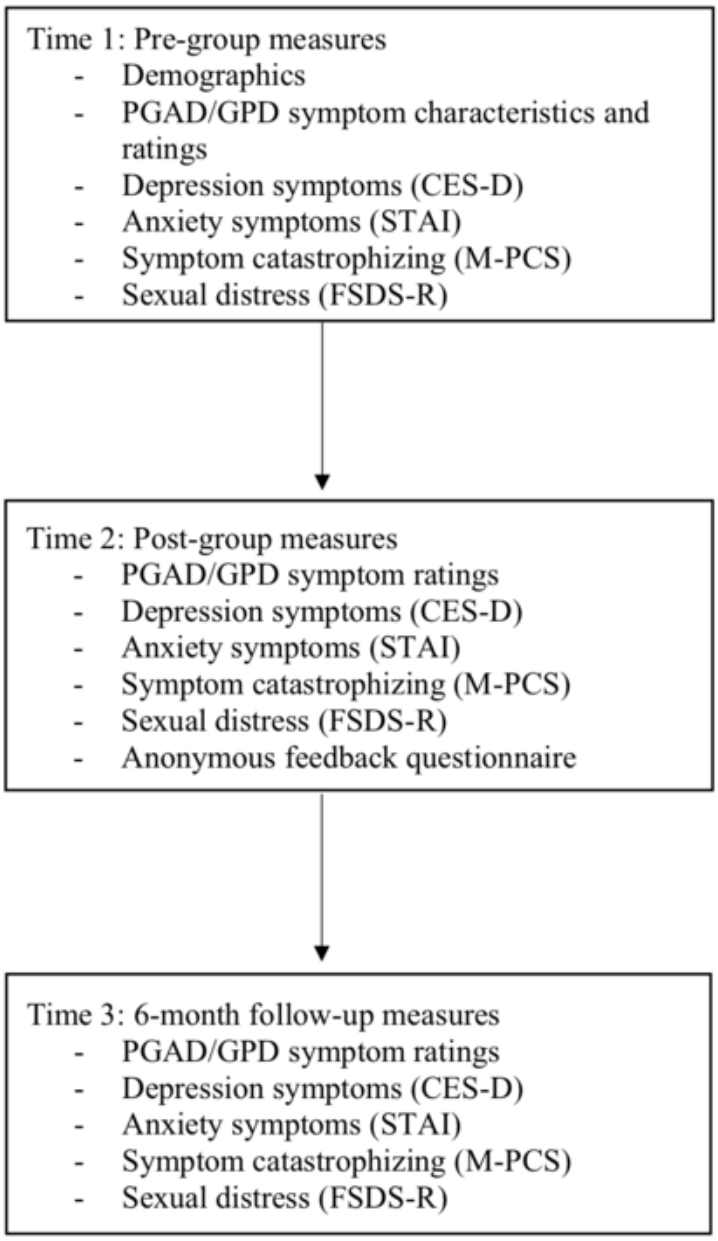
Figure 1. Flow of participation.



Measures

Questionnaires were completed before the start of the group (Time 1), immediately after the final group session (Time 2), and 6 months following the final group session (Time 3). See Figure 2 for an outline of the measures included at the different time points of the program.

Figure 2. Self-report measures included at each time point of the study. PGAD/GPD: persistent genital arousal disorder/genito-pelvic dysesthesia.



Demographic Information

Participants were asked to provide sociodemographic information, including age, ethnicity, education, occupation, relationship status, and sexual orientation.

Symptom Characteristics

Participants were asked to report on a number of PGAD/GPD symptom characteristics: the approximate date of their PGAD/GPD symptom onset, the proportion of time PGAD/GPD symptoms are present (*from 0% to 100%*), the distress associated with their PGAD/GPD symptoms (*0=none to 10=most distress ever*), and the discomfort associated with their PGAD/GPD symptoms (*0=none to 10=most discomfort ever*). Participants were also asked about other gynecological concerns that they experience and the number of health care providers that they had approached regarding their PGAD/GPD symptoms.

Depression Symptoms (Center for Epidemiologic Studies Depression Scale)

Symptoms of depression were assessed using the Center for Epidemiologic Studies Depression Scale (CES-D) [18]. The CES-D is a 20-item scale designed to assess the frequency of symptoms of depression over the past week. The measure is scored on a 4-point scale, with response options ranging from 0 (*Rarely, or none of the time; less than 1 day*) to 3 (*Most or all of the time; 5-7 days*). Higher scores on the CES-D represent more depressive symptoms.

Anxiety Symptoms (State-Trait Anxiety Inventory)

Symptoms of anxiety were assessed using the trait subscale of the State-Trait Anxiety Inventory (STAI) [19]. The STAI trait subscale is comprised of 20 statements. Participants were asked to rate how well each statement describes them on a 4-point scale. Response options range from 1 (*Almost never*) to 4

(*Almost always*). Higher scores on the STAI trait subscale indicate greater trait anxiety.

Symptom Catastrophizing (Modified Pain Catastrophizing Scale)

Symptom catastrophizing was measured using a modified version of the Pain Catastrophizing Scale (M-PCS) [20]. The PCS was originally designed to assess catastrophizing related to pain experiences but was modified for the purposes of this study by replacing the word "pain" with "vulvar sensations" throughout. The measure includes 13 statements that represent thoughts or feelings that may occur during PGAD/GPD experiences (eg, "I worry all the time about whether the vulvar sensations will end," "I feel I can't go on," etc), and participants were asked to report on the frequency at which they experience these thoughts or feelings from 0 (*Not at all*) to 4 (*All the time*). Total M-PCS scores range from 0 to 52, with higher scores indicating higher levels of symptom catastrophizing. Scores of 30 and above on the original scale suggest clinical levels of catastrophizing [20].

Sexual Distress (Female Sexual Distress Scale – Revised)

The Female Sexual Distress Scale – Revised (FSDS-R) [21] was used to assess sexually-related personal distress. The measure contains a list of 13 feelings and problems that some people have regarding their sexuality (eg, distressed about your sex life, unhappy about your sexual relationship). Participants were asked to rate how often each problem had bothered them or caused them distress over the past 30 days, on a 5-point scale ranging from 0 (*Never*) to 4 (*Always*). Higher scores on the FSDS-R indicate greater levels of sexual distress.

Global Perception of Improvement

The two post-group questionnaires (Times 2 and 3) included a single-item question estimating percent overall improvement of PGAD/GPD symptoms (including emotional well-being, pain, sexual functioning, relationship, etc from 0% to 100%) as a result of attending the program.

Anonymous Feedback Questionnaire

A feedback questionnaire was administered to all participants who attended at least one of the group sessions following the final session of the program to gather information about the acceptability of the session format and content. Participants were asked which sessions they found the most and least helpful. Feedback was also solicited on the acceptability of the length and frequency of the sessions as well as on homework exercises, including the degree to which participants were able to complete the exercises (*from 0=not at all to 4=a high degree*), the most and least helpful exercises, and factors that would have facilitated homework completion. Participants also commented on the potential benefit of including partners in the group.

Procedure

The study received approval from the Queen's University (Kingston, ON) Health Sciences and Affiliated Hospitals Research Ethics Board. Following the telephone screening, eligible participants were directed to an online survey hosted by Qualtrics survey software to complete before the group began (Time 1). After reading a Letter of Information and providing their consent to participate, participants completed the survey, which took approximately 30 to 45 minutes.

The online group education program ran from January to March of 2018. The program included 8 weekly sessions, each lasting 120 minutes. The sessions were hosted on Zoom videoconferencing software. The Zoom software transmits only encrypted information, with multilayer security and end-to-end encryption ("Encryption for Meetings", 2019). As such, Zoom is in accordance with the Health Insurance Portability and Accountability Act (HIPAA, 1996). Participants were provided with detailed instructions on how to use Zoom, and a team member (KMM) was present during all sessions to assist with any technical difficulties. Moreover, participants in the group were instructed to respect and maintain the confidentiality of other members of the group. The topics presented in each of the weekly education sessions are presented in Table 1.

Table 1. Topics covered in each of the online pelvic floor group education sessions.

Session number	Session topic
Session 1	Introduction to the science of pain (ie, processes that occur within the nervous system when one is in pain or discomfort [such as PGAD/GPD symptoms])
Session 2	Diaphragmatic breathing and its connection to the pelvic floor
Session 3	The benefits of movement and exercise on pelvic floor function
Session 4	Mindfulness
Session 5	Strategies for communicating one's needs while experiencing discomfort, including tips for educating clinicians and sharing information with loved ones
Session 6	The role of nutrition in inflammation and experiences of discomfort
Session 7	Sleep hygiene and strength training (general body conditioning exercises to help a person with chronic symptoms become more functional in their activities of daily living)
Session 8	Emotion regulation and self-compassion

Each of the 8 sessions followed a set structure. First, participants were welcomed, and the educational topics for the session were introduced. Then, participants were guided to perform a breath technique to center/focus on the present. Following the breathing exercise, the educational topic of the session was presented by a registered physical therapist (EH), who then demonstrated the exercise introduced that session. Following the demonstration, the physical therapist then provided verbal guidance for

participants to perform the exercise. Participants then engaged in a discussion about the topic, and the session ended with a question and answer period. There was no cost to participants for the group educational sessions, and no compensation was provided. When participants missed a session, they were provided with handouts summarizing the material presented in that session. Immediately following the completion of the online program, participants were sent a link to complete the online questionnaires for a second time (Time 2). Participants were also sent the same online questionnaires 6 months following completion of the program (Time 3).

Results

Data Considerations

The final sample contained 14 participants at the start of the group (Time 1), 6 participants at the end of the group (Time 2), and 5 participants at the 6-month follow-up (Time 3; see Figure 1).

Quantitative Results

Prior to conducting analyses, the data were examined for missing values and outliers where appropriate. No outliers were identified, defined as values more than 3 times the interquartile range [22]. No missing data were imputed for sociodemographic or symptom questions. On validated questionnaires with more than 10 items (CES-D, STAI, M-PCS, FSDS-R), if fewer than 15% of the items were missing for each individual, missing values were replaced with the individual's mean response on that questionnaire. If more than 15% were missing, that individual's questionnaire was excluded from the analyses. Quantitative results are presented as means and standard deviations. Analyses were conducted using IBM SPSS Version 26.

Feedback Questionnaire

Responses to open-ended questions in the feedback questionnaires are presented.

Sample Demographics and Symptom Characteristics

Demographic and symptom characteristics for participants who completed the Time 1 questionnaires (n=14) are presented in Table 2. Participants were, on average, 43.71 years old (SD 17.65; range: 18 to 71). With respect to PGAD/GPD symptoms, participants reported a long average duration of symptoms (mean 7.43 years, SD 10.25), and a moderate to high level of associated distress and discomfort (Table 2).

Table 2. Sociodemographic information and symptom characteristics for participants who completed the pre-program questionnaires (Time 1; n=14).

Characteristic	Values
Ethnicity, n (%)	
American	8 (57)
Northern European (except British Isles)	3 (21)
French	1 (7)
Latin American	1 (7)
American and Eastern European	1 (7)
Occupational status, n (%)	
Full-time	2 (14)
Part-time	5 (36)
Unemployed	2 (14)
Retired	1 (7)
On disability	4 (29)
Education, n (%)	
All/some high school	2 (14)
All/some college/undergraduate degree	7 (50)
All/some graduate school/professional training	5 (36)
Relationship status, n (%)	
Single	1 (7)
Dating	3 (21)
Married/cohabitating	8 (57)
Divorced	2 (14)
Sexual orientation, n (%)	
Mixed-sex oriented	9 (64)
Same-sex oriented	2 (14)
Bisexual	2 (14)
Asexual	1 (7)
PGAD/GPD[a] symptoms, mean (SD)	
Time since PGAD/GPD symptom onset, years	7.43 (10.25)
Time PGAD/GPD present, %	63.31 (28.83)
Distress (0=none, 10=most distress ever)	7.64 (2.27)
Discomfort (0=none, 10=most discomfort ever)	6.43 (3.63)
HCPs[b] seen for PGAD/GPD, n	5.50 (5.40)
Other gynecological concerns[c], n	2.14 (1.99)

[a]PGAD/GPD: persistent genital arousal disorder/genito-pelvic dysesthesia.

[b]HCP: health care provider.

[c]Examples of other gynecological concerns include interstitial cystitis, pelvic inflammatory disease, endometriosis, and sexually transmitted infections.

Who Attends the Majority of the Group Program?

Of those who completed the Time 1 questionnaire, 7 attended 5 or more of the sessions, and 7 attended less than 5 sessions. For most absences, no reason was provided; when reasons were provided, the most common ones were a scheduling conflict (9 absences) or being too sick/tired (5 absences). To better understand who attended the majority of the program sessions, health history, symptoms, and psychosocial well-being are presented for those who attended 5 or more versus less than 5 sessions (Table 3). Visual inspection of the data suggested that those who attended 5 or more sessions were younger, reported more gynecological comorbidities, and reported less severe PGAD/GPD symptoms (lower associated discomfort, lower

associated distress, and symptoms were present for shorter amount of time). They also reported lower baseline depressive and anxiety symptoms and lower symptom catastrophizing; however, they reported greater sexual distress. Overall, it appears that those with less severe PGAD/GPD symptoms and associated psychosocial concerns attended the majority of the program sessions. Graphs representing all individual responses for each individual participant are presented in Multimedia Appendix 1.

Table 3. Average scores (and SD) of participants who attended 5 or more of the 8 online educational group sessions (n=7) and those who attended less than 5 sessions (n=7).

Characteristic	Attended 5+ sessions, mean (SD)	Attended <5 sessions, mean (SD)
Age, years	41.7 (16.3)	45.7 (20.0)
Other gynecological concerns, n	2.4 (1.5)	1.9 (2.5)
Time PGAD/GPD[a] present, %	59.8 (33.7)	66.3 (26.4)
Distress score	7.6 (2.5)	7.7 (2.2)
Discomfort score	5.4 (3.9)	7.4 (3.4)
CES-D[b]	26.9 (11.9)	32.3 (12.6)
STAI[c]	54.3 (14.1)	58.7 (15.3)
M-PCS[d]	32.5 (13.7)	34.7 (13.3)
FSDS-R[e]	33.4 (14.0)	29.3 (8.3)

[a]PGAD/GPD: persistent genital arousal disorder/genito-pelvic dysesthesia.

[b]CES-D: Center for Epidemiologic Studies Depression Scale.

[c]STAI: State-Trait Anxiety Inventory.

[d]M-PCS: Modified Pain Catastrophizing Scale.

[e]FSDS-R: Female Sexual Distress Scale – Revised.

Symptoms and Psychosocial Well-being Before and After the Program

At Time 2, participants rated their overall perceived improvement to be 13.5% (SD 20.29; range: 0% to 50%; n=6). At Time 3, participants rated their overall perceived improvement to be slightly higher (mean 15.0%, SD 11.18; range: 0% to 30%; n=5). Descriptive information about symptom characteristics, psychosocial adjustment, and sexual well-being at all 3 time points is presented in Table 4. An overall pattern emerged, such that PGAD/GPD symptoms and psychosocial well-being improved across time, with the exception of discomfort associated with PGAD/GPD symptoms, which increased at Time 2 but decreased at Time 3. These results are consistent with the small improvements reported on the global improvement measure. Depression symptoms, anxiety symptoms, and sexual distress decreased following the group; however, the average scores remained within the range indicating clinically significant levels. The average score of symptom catastrophizing (M-PCS) fell in the range indicating clinically significant catastrophizing at Time 1 but decreased at Times 2 and 3 to a score that no longer fell in the range indicating clinical significance. Individual responses on each of the outcome variables, plotted across the 3 study time points, are presented in Multimedia Appendix 2.

Table 4. Average scores prior to attending the online educational group program (Time 1), at the end of the 8-week group program (Time 2), and 6 months following the program (Time 3).

Characteristic	Time 1, mean (SD)	Time 2, mean (SD)	Time 3, mean (SD)
	n	n	n
Time PGAD/GPD[a] present, %	63.31 (28.83)	54.83 (34.96)	51.25 (33.26)
	13	6	4
Distress score	7.64 (2.27)	7.67 (3.14)	6.80 (2.39)
	14	6	5
Discomfort score	6.43 (3.63)	7.33 (2.58)	5.40 (2.97)
	14	6	5
CES-D[b]	29.57* (12.12)	28.67* (12.04)	23.0* (8.28)
	14	6	5
STAI[c]	56.5* (14.34)	52.17* (13.41)	46.80* (12.85)
	14	6	5
M-PCS[d]	34.21* (12.60)	29.67 (13.71)	25.40 (11.61)
	14	6	5
FSDS-R[e]	31.36* (11.27)	30.33* (17.34)	19.00* (15.75)
	14	6	5

[a]PGAD/GPD: persistent genital arousal disorder/genito-pelvic dysesthesia.

[b]CES-D: Center for Epidemiologic Studies Depression Scale.

[c]STAI: State-Trait Anxiety Inventory.

[d]M-PCS: Modified Pain Catastrophizing Scale.

[e]FSDS-R: Female Sexual Distress Scale – Revised.

*Scores that fall above established cutoffs, suggesting clinically significant symptoms.

Feedback Questionnaire: What Did Participants Think About the Group Program?

At Time 2, 8 participants completed an anonymous feedback questionnaire about their experience attending the program. Overall, breathing and relaxation exercises were described to be the most helpful home practice exercises by the majority of participants (n=7). On average, participants reported a moderate ability to complete the home practice exercises (mean 2.0, SD 0.9). Participants reported that the most helpful topics were sessions that (1) discussed the relationship between emotions, discomfort, and PGAD/GPD symptoms (n=3) and (2) included relaxation exercises (ie, breathing, visualization; n=2). The sessions that were rated as least helpful were those on (1) nutrition (n=2) and (2) sleep and strength training (n=1). Ratings of mindfulness and the science of pain/discomfort received a mixed response (1 positive and 1 negative rating for each). All participants reported that they were happy with the number and length of the sessions, although one additionally specified that they would prefer biweekly sessions. Only 1 participant indicated that the inclusion of partners in the group would be helpful (no: n=5; not sure: n=2).

Open-ended responses on the feedback questionnaire are presented in Textbox 1. When asked what was helpful about the sessions, themes emerged of *normalization* (eg, “not as alone as I feel”), *support to complete ongoing interventions* for the PGAD/GPD symptoms (“reminders of some of what I was already doing,” “having a focus each week”), and *hopefulness* (“made me once again think about what I can do to handle this”). When asked what would help participants to complete the home practice exercises regularly, 2 participants identified barriers to completing the exercises: *timing of the sessions* and *health concerns*. Two participants offered concrete changes to the structure of home practice exercises: a brief note summarizing the home practice exercises at the end of each session (n=1) and more personal interaction during the program (n=1).

Textbox 1. Open-ended responses from participants (n=8) who attended the 8-session group educational program.

Question 1: In what ways did you find the sessions helpful? What did you find most useful?

- Knowing about how the brain works
- Reminder that emotional suffering, anxiety etc. can worsen physical conditions. Support that I am not as 'alone' as I feel trying to ENDURE this horrid condition!
- I appreciate having a focus each week and the science and research presented behind the methods.
- I learned something new in every session. The handouts after the sessions were over. I found that I was able to underline certain points and it was easier to look back over. I also liked the videos as I am a very visual learner.
- Made me once again think about what I can do to handle this while I still hope to get help from the doctors (that they will find a reason for PGAD/GPD).
- I knew most of the information already...
- Being able to see how others responded and what input they had was very valuable to me.
- Reminders of some of what I was already doing: meditations, yoga, physical exercise, correct nutrition.

Question 2: What would have made it easier for you to complete the home exercises?

- More personal interaction.
- More personal time. The sessions came at a very busy time in my life […] I didn't have the capacity to slow down.
- I think timing had a lot to do with issues for me. Having had two hospitalizations within a month […] has left me severely restricted at this time.
- A separate note with just the exercise after each session.
- Better health.

Discussion

Principal Findings

To our knowledge, this study is the first to examine the feasibility of an online pelvic floor group education program for PGAD/GPD. PGAD/GPD is associated with significant negative psychosocial impact [13], and there is great need for empirically based treatment interventions that address both PGAD/GPD symptoms and their consequences [10]. Given the limited information on and treatments available for PGAD/GPD [10,14], online interventions have the opportunity to reach many individuals who may not otherwise have access to treatment for PGAD/GPD. There is growing evidence that web-based health interventions can be effective in promoting knowledge and behavioral change in the management of other chronic illnesses [23].

Descriptive results regarding participants who attended the majority of the group sessions suggested that they were younger and reported less severe PGAD/GPD symptoms and associated psychosocial concerns (eg, symptom catastrophizing, depression, and anxiety) than those who attended fewer sessions. In addition, those who attended the majority of the sessions and completed the Time 2 and 3 outcome measures reported small improvements in PGAD/GPD symptoms (proportion of time present, distress, and discomfort) and psychosocial well-being (depression symptoms, anxiety symptoms, catastrophizing of PGAD/GPD symptoms, and sexual distress). These results indicate that regular participation in an online group program may be beneficial in terms of outcome. However, these findings also suggest that, even when intervention is presented via an online format, participants with more severe PGAD/GPD symptoms and associated psychosocial consequences may still face barriers to engaging in such an intervention.

Indeed, previous studies have found that PGAD/GPD symptoms interfere with daily activities, such as sitting for prolonged periods of time and the ability to concentrate [2]. This interference may have prevented those with more severe symptoms from more fully attending the intervention-based sessions, even remotely. Future intervention programs for PGAD/GPD should consider modifications and accommodations that would help to address barriers to participation. For example, the integration of asynchronous content to allow participants more flexibility in the timing of some of the more educational aspects of the program, in combination with shorter synchronous sessions, with frequent rest breaks for the more experiential and interactive components of the program, may be helpful. More significant psychosocial correlates of PGAD/GPD symptoms (ie, greater depression symptoms, anxiety symptoms, and symptom catastrophizing) could interfere with participation by reducing motivation to attend sessions and complete home exercises or by increasing avoidance of discussions or exercises that may be perceived to increase symptoms. Individuals experiencing significant symptoms of depression, anxiety, or both may benefit from strategies to reduce depression/anxiety prior to, or concurrently with, pelvic health education.

We also collected feedback on the most and least helpful aspects of the program, with a view to use the knowledge gained to aid in the redevelopment of an online group intervention for those with PGAD/GPD. Based on the feedback, sessions that focused on stress management and the role of cognitions and emotions in the management of genito-pelvic discomfort and unwanted arousal seemed most helpful and should be included in future programs. Treatment approaches that focus on these aspects

have been found to be effective for reducing pain and associated psychosocial difficulties in women with other forms of genito-pelvic discomfort (eg, vulvodynia [24-26]). Integration of these components early in the program may also help to address some of the barriers (ie, reducing distress and improving psychosocial well-being) to attending the program. Overall, in the feedback questionnaires, participants highlighted the value of normalization, hopefulness, and motivation/support to continue seeking treatments. The group format may be particularly valuable for individuals experiencing a condition, such as PGAD/GPD, that is surrounded by high levels of shame and hopelessness [3,13], and is often unknown within the health care community [14].

Limitations

The results of this study must be considered in the context of several limitations. This study was descriptive in nature and relied on a small sample size. Future studies may seek to conduct a similar program in a larger sample, while addressing the barriers identified in this study. In addition, participants did not undergo a clinical exam to confirm their diagnosis of PGAD/GPD. However, an extensive phone screening interview reviewing the diagnostic criteria for PGAD/GPD was conducted, and previous research has found high agreement between self-reported symptoms and clinical diagnosis for samples of women with other forms of genito-pelvic discomfort (ie, vulvodynia [27,28]). Finally, there is no information about how PGAD/GPD symptoms change over time without intervention to use as a comparison to these results. While this sample reports a long duration of symptoms (7.43 years on average), the chronicity of PGAD/GPD is unknown. Research on other forms of genito-pelvic discomfort (ie, vulvodynia) has found that chronicity is heterogenous [29-31]. These studies also indicate that individuals with a longer duration of symptoms are more likely to report greater pain intensity, anxiety, comorbid chronic pain conditions, and a primary symptom onset [29-31]. More information about the chronicity of PGAD/GPD symptoms will aid in interpreting future treatment outcome research. Finally, the sample was limited to individuals with access to the Internet and a computer. Online interventions may have the ability to reach individuals who cannot travel for in-person interventions; however, a limitation is access to, knowledge of, and comfort with telehealth technology (such as videoconferencing software).

Conclusions

Online interventions provide an opportunity to reach international participants who may otherwise struggle to access a knowledgeable provider for their PGAD/GPD symptoms. The group format may also help to encourage hopefulness and normalize the experience of symptoms that are often surrounded by feelings of shame. This study is the first online intervention study for PGAD/GPD, and it describes the feasibility of an 8-session pelvic floor education program. Overall, participants were satisfied with the length and frequency of the sessions. Deep breathing and relaxation exercises were reported to be beneficial by almost all participants. A number of barriers to participating in the program were identified (greater symptom and psychosocial impairment, timing of the sessions, concurrent health concerns) that could be addressed to help improve the efficacy of future interventions and increase the ability of participants to fully engage in the program. Future programs for PGAD/GPD may increase their focus on stress management strategies and working with thoughts and emotions related to PGAD/GPD symptoms. Finally, the results reinforce that PGAD/GPD is a highly distressing condition associated with significant burden. More research is needed to identify treatments and interventions to support individuals with PGAD/GPD.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Individual pre- and post-program self-report measures.
[DOCX File , 379 KB - formative_v5i1e22450_app1.docx]

Multimedia Appendix 2
Individual changes in outcome variables across the three study time points.
[DOCX File , 49 KB - formative_v5i1e22450_app2.docx]

References

1. Pukall CF, Jackowich R, Mooney K, Chamberlain SM. Genital Sensations in Persistent Genital Arousal Disorder: A Case for an Overarching Nosology of Genitopelvic Dysesthesias? Sex Med Rev 2019 Jan;7(1):2-12. [doi: 10.1016/j.sxmr.2018.08.001] [Medline: 30301706]
2. Jackowich R, Pink L, Gordon A, Poirier É, Pukall CF. Symptom Characteristics and Medical History of an Online Sample of Women Who Experience Symptoms of Persistent Genital Arousal. J Sex Marital Ther 2018 Feb 17;44(2):111-126. [doi: 10.1080/0092623X.2017.1321598] [Medline: 28459348]
3. Jackowich RA, Pink L, Gordon A, Poirier É, Pukall CF. An Online Cross-Sectional Comparison of Women With Symptoms of Persistent Genital Arousal, Painful Persistent Genital Arousal, and Chronic Vulvar Pain. J Sex Med 2018 Apr;15(4):558-567. [doi: 10.1016/j.jsxm.2018.02.007] [Medline: 29609914]

4. Leiblum S, Brown C, Wan J, Rawlinson L. Persistent sexual arousal syndrome: a descriptive study. J Sex Med 2005 May;2(3):331-337. [doi: 10.1111/j.1743-6109.2005.20357.x] [Medline: 16422863]
5. World Health Organization. International classification of diseases for mortality and morbidity statistics (11th Revision). 2018. URL: https://icd.who.int/browse11/l-m/en
6. Parish SJ, Goldstein AT, Goldstein SW, Goldstein I, Pfaus J, Clayton AH, et al. Toward a More Evidence-Based Nosology and Nomenclature for Female Sexual Dysfunctions-Part II. J Sex Med 2016 Dec;13(12):1888-1906. [doi: 10.1016/j.jsxm.2016.09.020] [Medline: 27843072]
7. Garvey LJ, West C, Latch N, Leiblum S, Goldmeier D. Report of spontaneous and persistent genital arousal in women attending a sexual health clinic. Int J STD AIDS 2009 Aug;20(8):519-521. [doi: 10.1258/ijsa.2008.008492] [Medline: 19625580]
8. Jackowich R, Pukall C. Prevalence of Persistent Genital Arousal Disorder in 2 North American Samples. J Sex Med 2020 Dec;17(12):2408-2416 [FREE Full text] [doi: 10.1016/j.jsxm.2020.09.004] [Medline: 33067159]
9. Dèttore D, Pagnini G. Persistent Genital Arousal Disorder: A Study on an Italian Group of Female University Students. J Sex Marital Ther 2021;47(1):60-79. [doi: 10.1080/0092623X.2020.1804022] [Medline: 32762421]
10. Jackowich RA, Pink L, Gordon A, Pukall CF. Persistent Genital Arousal Disorder: A Review of Its Conceptualizations, Potential Origins, Impact, and Treatment. Sex Med Rev 2016 Oct;4(4):329-342. [doi: 10.1016/j.sxmr.2016.06.003] [Medline: 27461894]
11. Komisaruk B, Goldstein I. Persistent genital arousal disorder: Current conceptualizations and etiologic mechanisms. Curr Sex Heal Reports 2017;9(4):177-182 [FREE Full text] [doi: 10.1007/s11930-017-0122-5]
12. Pukall C, Goldmeier D. Persistent genital arousal disorder. In: Binik YM, Hall K. eds. Principles and Practice of Sex Therapy (6th Ed.). New York, NY: Guilford Press; 2020:488-503.
13. Jackowich RA, Poirier É, Pukall CF. A Comparison of Medical Comorbidities, Psychosocial, and Sexual Well-being in an Online Cross-Sectional Sample of Women Experiencing Persistent Genital Arousal Symptoms and a Control Group. J Sex Med 2020 Jan;17(1):69-82. [doi: 10.1016/j.jsxm.2019.09.016] [Medline: 31680008]
14. Jackowich R, Pink L, Gordon A, Pukall C. 059 Health Care Experiences of Women with Symptoms of Persistent Genital Arousal. J Sex Med 2017;14(6):e369 [FREE Full text] [doi: 10.1016/j.jsxm.2017.04.057]
15. Rosenbaum TY. Physical therapy treatment of persistent genital arousal disorder during pregnancy: a case report. J Sex Med 2010 Mar;7(3):1306-1310. [doi: 10.1111/j.1743-6109.2009.01654.x] [Medline: 20059652]
16. Goldstein AT, Pukall CF, Brown C, Bergeron S, Stein A, Kellogg-Spadt S. Vulvodynia: Assessment and Treatment. J Sex Med 2016 Apr;13(4):572-590. [doi: 10.1016/j.jsxm.2016.01.020] [Medline: 27045258]
17. Leiblum SR, Nathan SG. Persistent sexual arousal syndrome: a newly discovered pattern of female sexuality. J Sex Marital Ther 2001;27(4):365-380. [doi: 10.1080/009262301317081115] [Medline: 11441520]
18. Radloff LS. The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. Applied Psychological Measurement 1977 Jul 26;1(3):385-401. [doi: 10.1177/014662167700100306]
19. Spielberger CD. Manual for the State-Trait Anxiety Inventory. Palo Alto, CA: Consulting Psychologists Press; 1983.
20. Sullivan M, Bishop S, Pivik J. The pain catastrophizing scale: development and validation. Psychol Assess 1995:524 [FREE Full text] [doi: 10.1037/1040-3590.7.4.524]
21. Derogatis L, Clayton A, Lewis-D'Agostino D, Wunderlich G, Fu Y. Validation of the female sexual distress scale-revised for assessing distress in women with hypoactive sexual desire disorder. J Sex Med 2008 Feb;5(2):357-364. [doi: 10.1111/j.1743-6109.2007.00672.x] [Medline: 18042215]
22. Tabachnick B, Fidell L. Using multivariate statistics, 6th edition. Boston: Pearson; 2013.
23. Wantland DJ, Portillo CJ, Holzemer WL, Slaughter R, McGhee EM. The effectiveness of Web-based vs. non-Web-based interventions: a meta-analysis of behavioral change outcomes. J Med Internet Res 2004 Nov 10;6(4):e40 [FREE Full text] [doi: 10.2196/jmir.6.4.e40] [Medline: 15631964]
24. Bergeron S, Khalifé S, Dupuis M, McDuff P. A randomized clinical trial comparing group cognitive-behavioral therapy and a topical steroid for women with dyspareunia. J Consult Clin Psychol 2016 Mar;84(3):259-268. [doi: 10.1037/ccp0000072] [Medline: 26727408]
25. Goldfinger C, Pukall CF, Thibault-Gagnon S, McLean L, Chamberlain S. Effectiveness of Cognitive-Behavioral Therapy and Physical Therapy for Provoked Vestibulodynia: A Randomized Pilot Study. J Sex Med 2016 Jan;13(1):88-94. [doi: 10.1016/j.jsxm.2015.12.003] [Medline: 26755091]
26. Masheb RM, Kerns RD, Lozano C, Minkin MJ, Richman S. A randomized clinical trial for women with vulvodynia: Cognitive-behavioral therapy vs. supportive psychotherapy. Pain 2009 Jan;141(1-2):31-40 [FREE Full text] [doi: 10.1016/j.pain.2008.09.031] [Medline: 19022580]
27. Dargie EE, Chamberlain SM, Pukall CF. Provoked Vestibulodynia: Diagnosis, Self-Reported Pain, and Presentation During Gynaecological Examinations. J Obstet Gynaecol Can 2017 Mar;39(3):145-151. [doi: 10.1016/j.jogc.2017.01.001] [Medline: 28343555]
28. Reed BD, Haefner HK, Harlow SD, Gorenflo DW, Sen A. Reliability and validity of self-reported symptoms for predicting vulvodynia. Obstet Gynecol 2006 Oct;108(4):906-913. [doi: 10.1097/01.AOG.0000237102.70485.5d] [Medline: 17012453]

29. Reed BD, Harlow SD, Plegue MA, Sen A. Remission, Relapse, and Persistence of Vulvodynia: A Longitudinal Population-Based Study. J Womens Health 2016 Mar;25(3):276-283 [FREE Full text] [doi: 10.1089/jwh.2015.5397] [Medline: 26752153]
30. Pâquet M, Vaillancourt-Morel M, Jodouin J, Steben M, Bergeron S. Pain Trajectories and Predictors: A 7-Year Longitudinal Study of Women With Vulvodynia. J Sex Med 2019 Oct;16(10):1606-1614. [doi: 10.1016/j.jsxm.2019.07.018] [Medline: 31451398]
31. Nguyen RHN, Mathur C, Wynings EM, Williams DA, Harlow BL. Remission of vulvar pain among women with primary vulvodynia. J Low Genit Tract Dis 2015 Jan;19(1):62-67 [FREE Full text] [doi: 10.1097/LGT.0000000000000041] [Medline: 24859843]

Abbreviations

CES-D: Center for Epidemiologic Studies Depression Scale
FSDS-R: Female Sexual Distress Scale – Revised
ISSWSH: International Society for the Study of Women's Sexual Health
M-PCS: Modified Pain Catastrophizing Scale
PGAD/GPD: persistent genital arousal disorder/genito-pelvic dysesthesia
STAI: State-Trait Anxiety Inventory

Edited by G Eysenbach; submitted 12.07.20; peer-reviewed by A Patterson, M Plegue; comments to author 15.08.20; revised version received 29.09.20; accepted 25.10.20; published 11.01.21.

Please cite as:
Jackowich RA, Mooney KM, Hecht E, Pukall CF
Online Pelvic Floor Group Education Program for Women With Persistent Genital Arousal Disorder/Genito-Pelvic Dysesthesia: Descriptive Feasibility Study
JMIR Form Res 2021;5(1):e22450
URL: http://formative.jmir.org/2021/1/e22450/
doi:10.2196/22450
PMID:33427673

Original Paper

Mobile App–Based Remote Patient Monitoring in Acute Medical Conditions: Prospective Feasibility Study Exploring Digital Health Solutions on Clinical Workload During the COVID Crisis

Sachin Shailendra Shah[1], MBBS, BMedSci, MRCP; Andrew Gvozdanovic[1], MBBS, BSc; Matthew Knight[2], MBE, FRCP; Julien Gagnon[1], MSc

[1]Huma Therapeutics, London, United Kingdom

[2]West Hertfordshire NHS Trust, Watford, United Kingdom

Corresponding Author:
Sachin Shailendra Shah, MBBS, BMedSci, MRCP
Huma Therapeutics
13th Floor Millbank Tower
21-24 Millbank
London, SW1P 4QP
United Kingdom
Phone: 44 7875210783
Email: sachsshah@gmail.com

Abstract

Background: Digital remote patient monitoring can add value to virtual wards; this has become more apparent in the context of the COVID-19 pandemic. Health care providers are overwhelmed, resulting in clinical teams spread more thinly. We aimed to assess the impact of introducing an app-based remote patient monitoring system (Huma Therapeutics) on a clinician's workload in the context of a COVID-19–specific virtual ward.

Objective: This prospective feasibility study aimed to evaluate the health economic effects (in terms of clinical workload) of a mobile app on a telephone-based virtual ward used in the monitoring of patients with COVID-19 who are clinically ready for discharge from the hospital.

Methods: A prospective feasibility study was carried out over 1 month where clinician workload was monitored, and full-time equivalents savings were determined. An NHS hospital repurposed a telephone-based respiratory virtual ward for COVID-19. Patients with COVID-19 in the amber zone (according to the National Health Service definition) were monitored for 14 days postdischarge to help identify deteriorating patients earlier. A smartphone-based app was introduced to monitor data points submitted by the patients via communication over telephone calls. We then comparatively evaluated the clinical workload between patients monitored by telephone only (cohort 1) with those monitored via mobile app and telephone (cohort 2).

Results: In all, 56 patients were enrolled in the app-based virtual ward (cohort 2). Digital remote patient monitoring resulted in a reduction in the number of phone calls from a mean total of 9 calls to 4 calls over the monitoring period. There was no change in the mean duration of phone calls (8.5 minutes) and no reports of readmission or mortality. These results equate to a mean saving of 47.60 working hours. Moreover, it translates to 3.30 fewer full-time equivalents (raw phone call data), resulting in 1.1 fewer full-time equivalents required to monitor 100 patients when adjusted for time spent reviewing app data. Individual clinicians spent an average of 10.9 minutes per day reviewing data.

Conclusions: Smartphone-based remote patient monitoring technologies may offer tangible reductions in clinician workload at a time when service is severely strained. In this small-scale pilot study, we demonstrated the economic and operational impact that digital remote patient monitoring technology can have in improving working efficiency and reducing operational costs. Although this particular RPM solution was deployed for the COVID-19 pandemic, it may set a precedent for wider utilization of digital, remote patient monitoring solutions in other clinical scenarios where increased care delivery efficiency is sought.

(JMIR Form Res 2021;5(1):e23190) doi:10.2196/23190

KEYWORDS

mHealth; remote patient monitoring; digital health; COVID-19; service improvement; cost-effectiveness; monitoring

Introduction

Background

Over the last decade, the world has seen a surge in the creation of virtual wards, wherein patients are managed remotely [1]. With an aging population [2], the demand for health care has started to exceed supply, thereby causing significant strain on service provision [3]. The need for a solution that reduces the burden on both primary and secondary health care services is therefore paramount.

Economic Benefits of Remote Patient Monitoring

In the UK, as part of the National Health Service's (NHS) Long Term Plan, digital initiatives are being reviewed for their potential integration into the current national health care system, and primary care services are at the forefront of this movement [4]. It is thought that web-based, digital general practice consultations and redesigned hospital support will reduce outpatient appointments by up to a third [4]. This could reduce patient trips to the hospital by approximately 30 million each year and, in turn, save the NHS over £1 billion (~US $1.35 billion) annually in new expenditure averted [4]. Although the use of remote patient monitoring (RPM) has typically revolved around the management of chronic diseases (eg, diabetes) and perioperative care (eg, orthopedic surgery), the expansion of RPM to wider patient groups promises additional benefits to both health care professionals and patients. Evidence regarding clinical outcomes associated with RPM is limited; however, there is overwhelming evidence of its economic benefits; for example, a reduction in clinician workload allowing the redirection of services to more demanding environments.

In the South Eastern Trust of Northern Ireland, the use of virtual wards was found to create cost-avoided savings of £ 8,804,529 (~US $11,940,922) over a 3-year period. This case study evaluated a virtual ward operational in 3 locations encompassing patients with chronic health conditions such as respiratory disorders, heart failure, and diabetes [5]. Over the course of the 3-year review period, an estimated 812 hospital admissions were avoided, and 447 episodes of care were provided by the virtual ward service, thereby effecting a total saving of approximately 4547 bed days. The operational costs over the 3 years are approximated at £ 566,273 (~US $768,106), which equates to a total net saving of £ 8,238,256 (~US $11,175,433) [5].

Similar evidence was also reported by the Healthcare at Home team who, in 2014, effected a reduction in the number of bed days by 130,000 [1] via the implementation of virtual wards. As such, regardless of whether there are any clinical benefits, it is clear that RPM has significant economic benefits, with multiple studies showing cost-saving outcomes following the implementation of a virtual ward [1,5].

RPM in Disease States

Considerable evidence supports the use of RPM, although with varying levels of robustness, in the care of chronic diseases such as heart failure, chronic obstructive pulmonary disorder (COPD), and frailty [6,7]. However, to date, there has been little research on the management of acute illnesses or, specifically, respiratory illnesses in this setting. Differences between the management of chronic and acute illnesses can be striking. Therefore, when attempting to extrapolate the economic impact of virtual wards on chronic disease management to acute disease management, we may expect to encounter some disparity.

In a previous study evaluating the use of an RPM solution in the care of exacerbation of COPD and chronic heart failure, Isaranuwatchai et al [8] studied the economic impact of remote monitoring symptoms to detect early deterioration of patient health. It is, therefore, somewhat suited to an evaluation in the context of monitoring of acute conditions. This study showed a potential reduction of 68% and 35% in the number of emergency room (ER) visits and hospitalizations, respectively, between the 3-month pre– and post–RPM intervention periods. The average ER visit cost was reduced from CA $243 (~US $191) at the baseline to CA $67 (~US $53) during the 3-month follow-up and from CA $3842 (~US $3023) to CA $1399 (~US $1100) for hospitalization [8]. This result demonstrates that an RPM solution can not only free-up resources but also lead to a less resource-heavy and, therefore, less-expensive visit or admission in the context of acute care.

RPM During the COVID-19 Pandemic

Health systems worldwide are currently facing an unprecedented challenge that is rapidly transforming the ways in which clinical care is provided. As of July 2020, the total number of confirmed cases of COVID-19 worldwide was approximately 15,201,000, with an estimated total of 623,000 related deaths [9]. At the time of this research project, no current vaccine nor effective treatment was available, and minimizing the risk of exposure was the mainstay of intervention at a population level [10].

Thus, we find ourselves using health care systems that are fundamentally rooted in face-to-face interactions and managing a disease by implementing preventive measures such as social distancing and hygiene education—a challenging combination of contrasts. Many institutions are, therefore, turning to connected care or digital health solutions such as virtual wards, RPM, and telemedicine [11].

Traditionally, the speed at which these digital interventions have been introduced has been slow [12]. The COVID-19 pandemic, however, has led to a transformation in the mobile health (mHealth) landscape, with institutions choosing to quickly implement mHealth solutions and adapt rapidly [13]. A report by Mann et al [13] describes the feasibility and impact of a video-enabled telemedicine solution at the epicenter of the COVID-19 outbreak in New York. The study, carried out in conjunction with NYU Langone Health, demonstrated the impact of this mHealth solution across 25 locations. Among the various outcomes evaluated, they reported an increase in the number of daily telemedicine visits from 369.1 to 866.8 in urgent care settings and from 94.7 to 4209.3 in nonurgent care settings, of which 56.2% and 17.6% of the visits, respectively, were related to COVID-19. Overall, clinicians found the existing telemedicine solution to be a useful tool in diverting patients from the ER, in order to prioritize those who needed acute care and thus minimize the risk of COVID-19 exposure [13].

Owing to the unique impact of COVID-19 on health care systems, there is limited evidence to reflect on the economic effects of RPM on pandemics of this kind. Consequently, COVID-19 provides an opportunity to explore the economic impact of widespread implementation of RPM for acute care. These analyses may also be applicable to other pandemics and standard practices alike, as most research investigates the use of RPM for exacerbations of chronic disease states and their associated illnesses. These analyses are generally handled with well-established processes with known protocols in place. Such economic evaluations would, therefore, not take into account the potential impact of virtual wards on the following: reduction of viral transmission (and the subsequent knock-on effect of reduced sickness on the economy), increase in the efficiency of resource use (such as from those high-risk individuals who would not be available for front-line work), or reduction in the utilization of other resource such as personal protective equipment [14,15].

With the increasing popularity of digital technologies providing mobile app–based solutions for digital health, virtual wards have somewhat undergone an overhaul and are now able to provide clinical teams with real-time data to better manage their patients. Clinicians now have greater visibility over their patients and improved communication pathways with other health care professionals, all of which have been shown to improve the efficacy of virtual wards within mHealth solutions and, thus, enhance health care in general [16].

Study Aims

The aim of this study was to evaluate the operational and consequent economic benefits of app-based RPM as a supplement to the existing telephone-based virtual wards in the context of the COVID-19 pandemic.

Methods

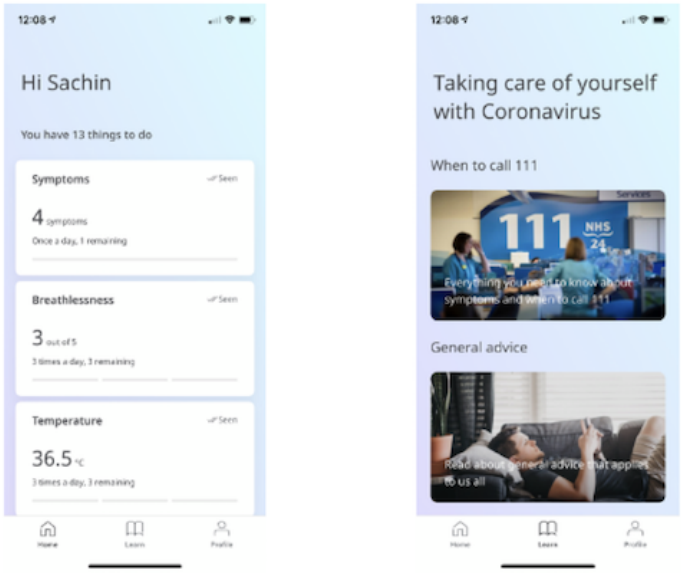
An NHS district hospital, in partnership with NHS-X and Huma Therapeutics, utilized its existing respiratory virtual ward as a temporary COVID-19 virtual ward. This virtual ward was initially designed to support inpatients who were experiencing exacerbated common respiratory conditions and met the criteria for early discharge. In the wake of COVID-19, this ward was used to monitor medium-risk patients (as per the NHS COVID-19 guidelines) who were ready to be discharged from the hospital. These patients needed to be monitored for 14 days to ensure no deterioration occurred but were deemed clinically safe to be discharged. Patients had the option of being monitored solely via telephone calls (cohort 1) or via a combination of mobile app and telephone calls (cohort 2).

Patients in cohort 1 followed a structured telephone call plan and would be followed-up at regular intervals via phone calls. Typically, a member of the respiratory team (ie, consultant, physiotherapist, or physiologist) would call patients in cohort 1 on days 1, 2, 3, 4, 5, 7, 9, 11, and 14. Patients were discharged after 14 days, if deemed clinically safe to do so. Phone calls were made to assess the patients' symptoms and functionality in relation to daily activities and, if possible, record any vital signs the patient may have taken.

Patients in cohort 2 were virtually onboarded to the Huma Therapeutics app. They were instructed via the app (Figure 1) to submit the following data on a daily basis: heart rate (obtained via photoplethysmography technology embedded in the app); oxygen saturation (obtained via a pulse oximeter wirelessly connected to the app or by manual entry); body temperature (obtained via a digital thermometer connected to the app or by manual entry); any symptoms (Textbox 1) experienced; and breathlessness measured using a single-question questionnaire ("How breathless are you when walking around or walking upstairs?") created by the clinicians involved in the project, which was scored on a scale of 1-5, with 1 being the least and 5 being the worst). This data was manually transcribed to a variety of electronic health records by populating a premade template. Based on the data submitted, a member of the clinical team would decide on whether a phone call was required or not. As per cohort 1, patients were monitored for 14 days, after which, if clinically safe, they were discharged.

The aim of this study was to investigate the operational and economic impact of an app-based RPM tool on clinician workload in a telephone-based virtual ward, in the context of managing patients with COVID-19. This prospective feasibility study was carried out over a 1-month period, during which clinician workload (ie, number and length of phone calls) was monitored and full-time equivalent (FTE) savings were equated. Clinical outcomes were defined as mortality and readmissions, simply as a basis to confirm noninferiority. Moreover, some informal qualitative information from clinician end-users was collected based on phone call quality and end-user feedback via unstructured telephone interviews.

Figure 1. Screenshots of the mobile app (Huma Therapeutics) used by the study patients.



Textbox 1. A list of symptoms available on the app for patients to choose from and submit to their care team.

List of symptoms:

- Fever
- Cough
- Shortness of breath
- Nausea
- Loss of taste
- Loss of smell
- Vomiting
- Chest pain/tightness
- Headache
- Heart palpitations
- Dizziness
- Loss of consciousness

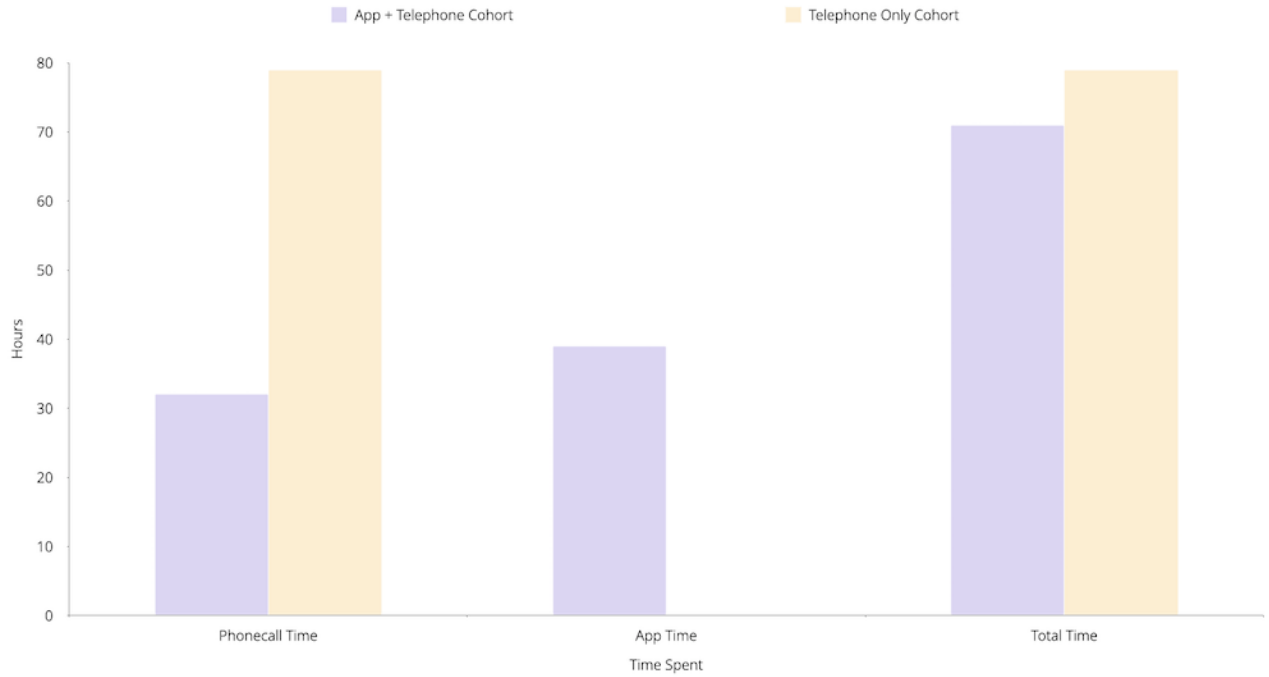
Results

Over a 1-month period, a total of 56 patients were enrolled into cohort 2 (app + telephone). These patients had a clinical or laboratory diagnosis of COVID-19. Of the 56 patients, 31 (55%) were male and 25 (45%) were female. The mean age of the study participants was 64 years (age range 21-79 years). Of these 56 patients, 40 (71%) patients used an iPhone, and the remainder 16 (29%) used an Android device. Clinical staff were hired for 12 hours a day and the service was available 7 days a week. We do not have the data on the number of patients enrolled into cohort 1 at this time.

Over the course of the monitoring period, we found that patients in cohort 2 (app + telephone) received significantly fewer phone calls (mean 4, SD 0.701) than those in cohort 1 (telephone only; mean 9, SD 1.13). In addition, the mean phone call time for patients in cohort 1 was 8.5 minutes, which was similar to the mean phone call time in cohort 2 (data not available).

The total time spent on phone calls for all 56 patients in cohort 2 was 31.73 hours. Based on the mean phone call time, the total time spent on phone calls for 56 patients in the cohort 1 model would be 79.33 hours. This equates to a 60% (47.60 hours) reduction from cohort 1 to cohort 2 (Figure 2). During this period, 7 clinicians monitored the 56 cohort patients. In terms of health economics, we observed a reduction of 3.30 FTE (ie, the number of clinicians reviewing these 56 patients). Each clinician spent an average of 10.9 minutes a day reviewing data of patients in cohort 2, resulting in a total time of 38.68 hours spent on the clinician dashboard. The FTE adjusted for time spent reviewing data was 1.1 per 100 patients; that is, for every 100 patients monitored in cohort 2, 1 less clinical personnel was needed compared to cohort 1.

Figure 2. Difference in time spent remotely monitoring 56 patients between the 2 cohorts.



Clinically, among the 56 patients reviewed, there were no mortalities and readmissions 14 days after discharge, indicating noninferiority of the model. It is worth noting that, due to the small patient sample size, the lack of observed mortalities and readmissions may not indicate a genuine impact of remote monitoring on these particular outcome metrics.

Informal qualitative feedback from clinicians held the solution in high regard. Qualitative data indicated that the app was emotionally well received by patients ("It's like an extension of the human touch.") and that the particular solution adopted was easy to use ("The Medopad patients have been much quicker and easier to deal with than non-Medopad virtual ward patients.")

Discussion

Summary

This prospective feasibility study highlights the impact a digital health solution can add to existing care services. In areas where large volumes of patients need to be monitored, an app-based tool can considerably reduce the time needed by the clinical team to manage the said cohort. Our small-scale, pilot study shows that the introduction of an app to supplement existing care pathways reduces the amount of time a clinician ends up spending to manage patients. Extrapolated economically, for every 100 patients enrolled into the virtual ward, you would need one less member of the clinical team to manage the group compared with the traditional phone call–only management style.

The COVID-19 pandemic has brought to light significant challenges concerning the infrastructure of health care provision, especially in acute, high-demand settings. Technology has advanced exponentially over the last few decades; however, digital solutions for monitoring health are still in their infancy. Digital, app-based solutions have historically shown large economic benefits with the additional potential for large-scale clinical benefit. They allow for greater reallocation of resources in terms of clinical personnel work distribution, allowing health professionals to better manage their already overburdened schedules.

Strengths and Limitations

Although this study was based on a small sample, the results clearly show the need for more robust and greater powered studies. This study depicts an improvement in service delivery; therefore, larger, more generalizable studies could not only compound these findings but also help demonstrate this positive effect in different care environments. If the study findings are generalizable, it may be concluded that an mHealth pathway could reduce the number of doctors needed to monitor these patients, if faced with a second wave or a similar pandemic. These studies need to accurately compare the time spent utilizing digital solutions with that spent monitoring patients by using in-person pathways, to ascertain a true economic representation of the change in the pathway. In addition, more formal, qualitative feedback needs to be obtained, not only from clinicians but also from patients, to review the psychological impacts of such digital health solutions. Further parallel studies should be run to accurately gauge clinical outcomes of mHealth solutions compared to traditional methods of patient monitoring.

Comparison With the Existing Literature

The existing literature clearly demonstrates the vast benefits across clinical and operational outcomes. However, majority of these studies are pilot studies and there are yet to be large landmark studies confirming the efficacy of digital solutions in the management of patients. In regard to the UK, the NHS clearly explains the need for an increase in digital services; however, evidence of this occurring, or even working, is nonexistent. COVID-19 has prompted fast deployment of mHealth solutions; over time, this evidence will start to come

to fruition. This study highlights the value of digital tools in novel disease states with clear potential advantages if applied to different care pathways.

Implications for Research or Practice

The findings of this study highlight the benefits of mHealth solutions and allow health care providers to introduce similar solutions to help manage future waves of COVID-19 and other pandemics. Clinical teams should heed the benefits displayed and implement these time- and cost-saving services into the everyday infrastructure. This would not only mean services are better prepared for future waves or pandemics but can also improve the operation of their day-to-day patient loads. As demand for health care continues to rise, any tool that can help reduce clinical team workload and allow for more patients to be seen with the same number of staff would be regarded as an invaluable tool.

Acknowledgments

The authors would like to thank the Clinical and Scientific team at Huma Therapeutics (Sophie Valentine and Dillon Obika) who reviewed the manuscript before submission.

Authors' Contributions

SS and MK were clinical leads in implementing and performing this study. AG assisted with data collection, and JG assisted with the economic calculations. All authors reviewed approved the final manuscript.

Conflicts of Interest

SS, AG, and JG were all employees of Huma Therapeutics during the study period. The study was a not-for-profit project with the sole aim of improving care delivery.

References

1. Virtual Wards: Bringing the hospital home. Healthcare at Home. URL: https://hah.co.uk/wp-content/uploads/2017/07/Virtual-Hospital-Report_AW_Final2.pdf [accessed 2021-01-07]
2. Age UK. 2019 May. URL: https://www.ageuk.org.uk/globalassets/age-uk/documents/reports-and-publications/later_life_uk_factsheet.pdf [accessed 2021-01-07]
3. The Complexities of Physician Supply and Demand: Projections from 2018 to 2033. Association of American Medical Colleges. 2020 Jun. URL: https://www.aamc.org/system/files/2020-06/stratcomm-aamc-physician-workforce-projections-june-2020.pdf [accessed 2021-01-07]
4. The NHS Long Term Plan. NHS. 2019 Jan. URL: https://www.longtermplan.nhs.uk/wp-content/uploads/2019/01/nhs-long-term-plan-june-2019.pdf [accessed 2021-01-07]
5. Colligan J. The Virtual Ward, Managing the Care of Patients with Chronic Conditions in the Community - An economic assessment of the South Eastern Trust Virtual Ward Internet. Royal College of Nursing. 2015. URL: https://www.rcn.org.uk/professional-development/research-and-innovation/innovation-in-nursing/case-studies-demonstrating-the-value-of-nursing/janice-colligan [accessed 2021-01-07]
6. Gagnon M, Ngangue P, Payne-Gagnon J, Desmartis M. m-Health adoption by healthcare professionals: a systematic review. J Am Med Inform Assoc 2016 Jan;23(1):212-220. [doi: 10.1093/jamia/ocv052] [Medline: 26078410]
7. Farias FACD, Dagostini CM, Bicca YDA, Falavigna VF, Falavigna A. Remote patient monitoring: a systematic review. Telemed J E Health 2020 May;26(5):576-583. [doi: 10.1089/tmj.2019.0066] [Medline: 31314689]
8. Isaranuwatchai W, Redwood O, Schauer A, Van Meer T, Vallée J, Clifford P. A remote patient monitoring intervention for patients with chronic obstructive pulmonary disease and chronic heart failure: pre-post economic analysis of the smart program. JMIR Cardio 2018 Dec 20;2(2):e10319 [FREE Full text] [doi: 10.2196/10319] [Medline: 31758770]
9. COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). Johns Hopkins University & Medicine - Coronavirus Resource Center. URL: https://coronavirus.jhu.edu/map.html [accessed 2021-01-07]
10. Coronavirus (COVID-19): Meeting with others safely (social distancing). GOV.UK. URL: https://www.gov.uk/government/publications/coronavirus-covid-19-meeting-with-others-safely-social-distancing/coronavirus-covid-19-meeting-with-others-safely-social-distancing [accessed 2021-01-13]
11. Schinköthe T, Gabri MR, Mitterer M, Gouveia P, Heinemann V, Harbeck N, et al. A web- and app-based connected care solution for COVID-19 in- and outpatient care: qualitative study and application development. JMIR Public Health Surveill 2020 Jun 01;6(2):e19033 [FREE Full text] [doi: 10.2196/19033] [Medline: 32406855]
12. Meskó B, Drobni Z, Bényei É, Gergely B, Győrffy Z. Digital health is a cultural transformation of traditional healthcare. Mhealth 2017;3:38 [FREE Full text] [doi: 10.21037/mhealth.2017.08.07] [Medline: 29184890]
13. Mann DM, Chen J, Chunara R, Testa PA, Nov O. COVID-19 transforms health care through telemedicine: evidence from the field. J Am Med Inform Assoc 2020 Jul 01;27(7):1132-1135 [FREE Full text] [doi: 10.1093/jamia/ocaa072] [Medline: 32324855]

14. Cavallo JJ, Forman HP. The economic impact of the COVID-19 pandemic on radiology practices. Radiology 2020 Sep;296(3):E141-E144 [FREE Full text] [doi: 10.1148/radiol.2020201495] [Medline: 32293225]
15. Nicola M, Alsafi Z, Sohrabi C, Kerwan A, Al-Jabir A, Iosifidis C, et al. The socio-economic implications of the coronavirus pandemic (COVID-19): A review. Int J Surg 2020 Jun;78:185-193 [FREE Full text] [doi: 10.1016/j.ijsu.2020.04.018] [Medline: 32305533]
16. 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]

Abbreviations

COPD: chronic obstructive pulmonary disorder
ER: emergency room
FTE: full-time equivalent
mHealth: mobile health
NHS: National Health Service
RPM: remote patient monitoring

Edited by G Eysenbach; submitted 05.08.20; peer-reviewed by T Schinkoethe, M Gabri; comments to author 25.09.20; revised version received 08.10.20; accepted 08.12.20; published 15.01.21.

Please cite as:
Shah SS, Gvozdanovic A, Knight M, Gagnon J
Mobile App–Based Remote Patient Monitoring in Acute Medical Conditions: Prospective Feasibility Study Exploring Digital Health Solutions on Clinical Workload During the COVID Crisis
JMIR Form Res 2021;5(1):e23190
URL: http://formative.jmir.org/2021/1/e23190/
doi:10.2196/23190
PMID:33400675

Original Paper

The Acceptability and Efficacy of Electronic Data Collection in a Hospital Neurodevelopmental Clinic: Pilot Questionnaire Study

Shrujna Patel[1,2*], PhD; Kelsie Ann Boulton[1,2*], PhD; Marie Antoinette Redoblado-Hodge[2,3], PhD; Angela Papanicolaou[3], BSocialWork; Diana Barnett[3], BAppSc; Beverley Bennett[3], MBBS; Suzi Drevensek[3], BAppSc; Jane Cramsie[3], MClinNeuropsych; Kalaichelvi Ganesalingam[3], PhD; Natalie Ong[3], MBBS; Magdalen Rozsa[3], BAppSc; Rebecca Sutherland[3], PhD; Marcia Williamsz[3], MA; Izabella Pokorski[1,2], MPubHealth; Yun Ju Christine Song[1], PhD; Natalie Silove[2,3*], MBBS; Adam John Guastella[1,2*], PhD

[1]Autism Clinic for Translational Research, Brain and Mind Centre, Children's Hospital Westmead Clinical School, Faculty of Medicine and Health, University of Sydney, Camperdown, Australia

[2]Child Neurodevelopment and Mental Health Team, Brain and Mind Centre, Children's Hospital Westmead Clinical School, Faculty of Medicine and Health, University of Sydney, Sydney, Australia

[3]Child Development Unit, The Children's Hospital at Westmead, Sydney Children's Hospital Network, Westmead, Australia

[*]these authors contributed equally

Corresponding Author:
Adam John Guastella, PhD
Autism Clinic for Translational Research
Brain and Mind Centre, Children's Hospital Westmead Clinical School
Faculty of Medicine and Health, University of Sydney
100 Mallett Street
Camperdown, 2050
Australia
Phone: 61 2 9351 0539
Email: adam.guastella@sydney.edu.au

Abstract

Background: There is a growing need for cost-efficient and patient-centered approaches to support families in hospital- and community-based neurodevelopmental services. For such purposes, electronic data collection (EDC) may hold advantages over paper-based data collection. Such EDC approaches enable automated data collection for scoring and interpretation, saving time for clinicians and services and promoting more efficient service delivery.

Objective: This pilot study evaluated the efficacy of EDC for the Child Development Unit, a hospital-based diagnostic assessment clinic in the Sydney Children's Hospital Network. Caregiver response rates and preference for EDC or paper-based methods were evaluated as well as the moderating role of demographic characteristics such as age, level of education, and ethnic background.

Methods: Families were sent either a paper-based questionnaire via post or an electronic mail link for completion before attending their first on-site clinic appointment for assessment. A total of 62 families were provided a paper version of the questionnaire, while 184 families were provided the online version of the same questionnaire.

Results: Completion rates of the questionnaire before the first appointment were significantly higher for EDC (164/184, 89.1%) in comparison to paper-based methods (24/62, 39%; $P<.001$). Within the EDC group, a vast majority of respondents indicated a preference for completing the questionnaire online (151/173, 87.3%), compared to paper completion (22/173, 12.7%; $P<.001$). Of the caregiver demographic characteristics, only the respondent's level of education was associated with modality preference, such that those with a higher level of education reported a greater preference for EDC ($P=.04$).

Conclusions: These results show that EDC is feasible in hospital-based clinics and has the potential to offer substantial benefits in terms of centralized data collation, time and cost savings, efficiency of service, and resource allocation. The results of this study therefore support the continued use of electronic methods to improve family-centered care in clinical practices.

(JMIR Form Res 2021;5(1):e18214) doi:10.2196/18214

KEYWORDS

electronic data collection; family-centered care; hospital-based data collection

Introduction

Electronic data collection (EDC) has been at the center of debate about the future of 21st century health care [1-3]. Such approaches have the potential to save billions in health care costs through improved data capture and clinical service responses that allow for more efficient patient-centered care [4-7]. There has, however, been a slow uptake of EDC approaches in clinical services globally, and limited evidence of successful technology integration in public hospital settings [5,8-10]. For instance, in most hospital-based clinics across Australia, data collection is largely paper based. Electronic medical records are being introduced in hospitals; however, this process has been slow and data entry into electronic medical records remains less than systematic [11]. As a result, a recent inquiry report by the Australian Government Productivity Commission suggested that Australia was falling behind in utilizing health care data for data linkage between health services for research purposes [12].

Despite these issues, EDC offers many benefits, warranting its evaluation in public health service settings. EDC provides the opportunity to engage families in more efficient services. For instance, patients can conveniently access forms and staff require less time to monitor, analyze, and interpret the gathered data, allowing for swift provision of feedback to patients and families [13,14]. In addition to this increased efficiency, EDC has been shown to result in fewer human errors in data processing and enables collection of data from a broader geography, increasing completeness of data collation and freeing clinical service resources for other needs, ultimately improving service outcomes [15,16]. The collection and integration of large amounts of data may then be better used to support clinical and research services that operate across rural and remote settings, where on-site attendance can be difficult [15-17].

One public health setting that could benefit considerably from EDC is child diagnostic and assessment services, specifically those clinics that assess children with neurodevelopmental concerns. These neurodevelopmental clinics aim to provide assessments at the earliest possible time in a child's development to increase the opportunity for earlier assessment, diagnosis, and intervention [18-20], with growing evidence that early intervention is associated with better long-term outcomes for the child and family [21,22]. Currently, however, public services are inundated with assessment requests, long wait lists, and limited resources to complete these tasks. These clinics typically do not use EDC, relying instead on pencil and paper for the vast majority of assessments. These public neurodevelopmental clinics are also more likely to provide services to a higher proportion of children and caregivers from disadvantaged backgrounds, those of lower socioeconomic status, and a higher proportion of linguistically diverse and indigenous communities in comparison to private clinical practices. It is, therefore, important to evaluate the utility of EDC in services that attend to these diverse patient populations.

Prior research has shown that demographic factors, such as age, socioeconomic status, level of education, language, and ethnicity may influence the completion of online data collection [23]. For example, socioeconomic deprivation (measured by the Scottish Index of Multiple Deprivation) and age (>70 years) have been associated with poorer completion of EDC in an orthopedic clinic in Scotland. Socioeconomic deprivation and age were both independently associated with lack of internet access [24], which may have contributed to the study findings. Similarly, a study of orthopedic surgery patients in California found that patients who were older (>75 years), of Hispanic or Black ethnicity, and had Medicare or Medicaid insurance were less likely to complete EDC patient-reported outcome surveys [25]. The authors argued that internet use is less prevalent among older patients, who formed much of the Medicare group. Additionally, Medicaid insurance includes low-income and vulnerable families who may not have had internet access to complete EDC surveys [25].

This study aimed to evaluate an initial pilot for EDC in one of Australia's busiest child diagnostic and assessment services, the Child Development Unit (CDU) at The Children's Hospital Westmead, part of the publicly funded Sydney Children's Hospital Network, Australia. The CDU assesses approximately 600 children per year, referred by pediatricians, who present with complex neurodevelopmental problems. The CDU provides multidisciplinary neurodevelopmental assessments to the state of New South Wales, with some families attending from regional and rural areas, and a high proportion of families from culturally and linguistically diverse backgrounds. Assessment in the CDU begins with gathering information on family and developmental history, via a questionnaire completed by caregivers before attending their first appointment. The CDU has traditionally mailed a paper version of this questionnaire to families prior to their appointment, with families asked to post the completed questionnaire back to the clinic ahead of their appointment so that clinicians can be prepared for the on-site assessments. Clinicians have noted, however, that response rates have been consistently low, with less than 50% of families returning the questionnaire. Such problems lead to delays in terms of clinicians needing to complete and interpret the questionnaire with the family during their appointment.

To the best of our knowledge, this is the first study to assess EDC in a child development clinic. In this pilot study, we digitized the CDU's caregiver questionnaire into a format that families could access via email and complete electronically in a secure, convenient, and efficient manner. We aimed to examine whether this would improve response rates for the questionnaire when compared to the paper version. We also investigated whether families preferred the electronic modality over the paper version, and the demographic characteristics that were associated with these preferences.

Methods

Setting

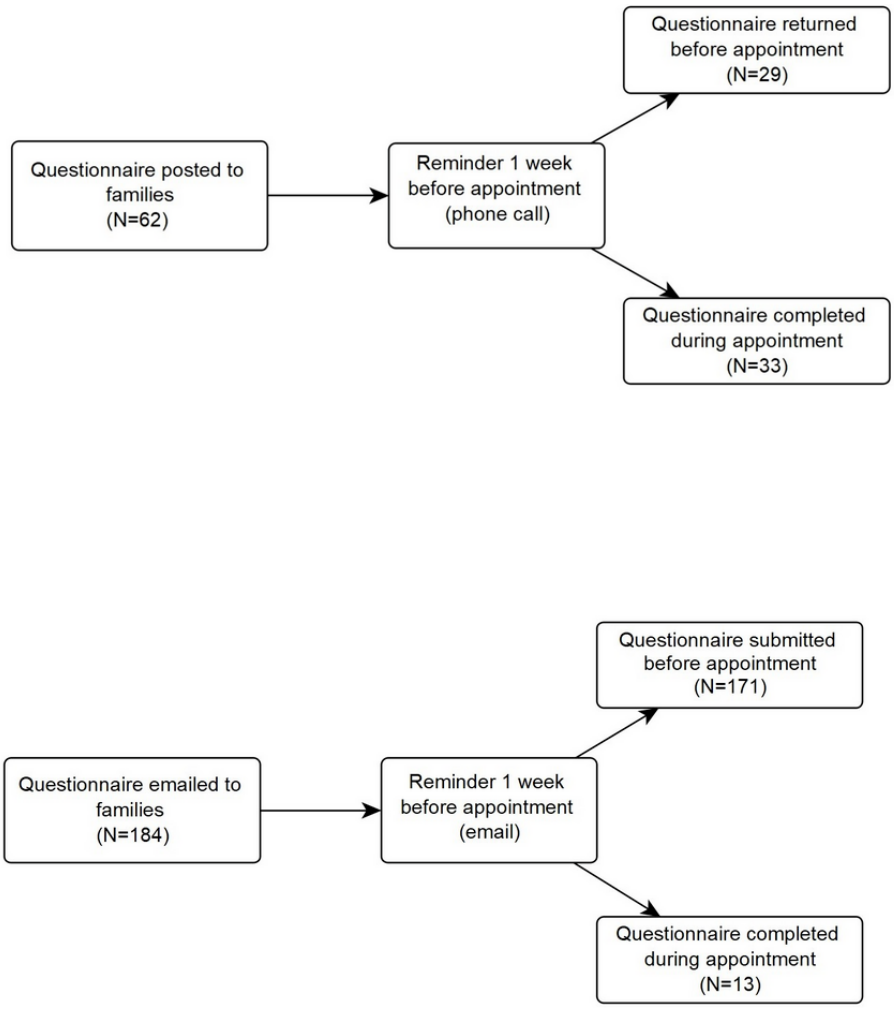
A total of 246 families who had an appointment in 2018-2019 with the CDU for the initial developmental assessment of their child were invited to participate in this study. Participants were consecutively recruited into this research study using opt-out informed consent methods. This study was approved by the Sydney Children's Hospital Network Ethics Committee (LNR/17/SCHN/293). The first 62 families entered into this study were sent the paper questionnaire by post. Subsequently, the service transitioned to EDC methods, and a further 184 families were sent the questionnaire via email. No family declined to participate in this study.

Children are referred to the CDU for assessment of complex neurodevelopmental difficulties, including possible autism spectrum disorder, intellectual disabilities, global developmental delay, or specific learning disorders. Prior to their appointment at the CDU, families are expected to complete a 6-page questionnaire, covering demographic information, family history, and child developmental history. The questionnaire has a Flesch Reading Ease score of 76.7, indicating fairly easy readability that should be understood by 12- to 15-year olds [26]. These data enable the CDU to assemble a suitable team and prepare the relevant assessment measures before the appointment. Originally completed by families on paper, we digitized the questionnaire, creating a digital form on the Research Enterprise Data CAPture (REDCap) platform. REDCap is an electronic data capture tool endorsed by the University of Sydney for the secure collection of all research data [27,28]. REDCap is designed to support data capture for research studies and allows questionnaires to be emailed to families ahead of their appointment. Families can click on the link provided in the email invitation to open the questionnaire. Data are automatically saved in REDCap as they are entered in the online form. Once the family exits the form, it is immediately available for clinicians to view.

Procedure

As per the existing CDU procedure, families were advised of their scheduled appointment via phone. During this phone call, they are informed that a questionnaire will be posted to them and are instructed to complete it and post it back to the clinic. A week before their appointment, families who have not returned the questionnaire via post receive a reminder phone call to do so. Those families who have not returned the questionnaire by the time of their appointment are required to complete it on the day of their appointment with a member of the clinical team (social worker or clinical nurse consultant). In this study, we implemented a pilot REDCap procedure for completion of the questionnaire. Families were advised by phone of their appointment confirmation and told to expect an email inviting them to complete the questionnaire on the REDCap platform. A week before their appointment, an email reminder was sent automatically via REDCap to families who had not completed the questionnaire. Families who had not completed the questionnaire by the time of their appointment were required to complete it on the day of their appointment with a member of the clinical team (social worker or clinical nurse consultant). The existing CDU procedure and pilot REDCap procedure are outlined in Figure 1.

Figure 1. Flow chart of procedures for questionnaire completion modes and response rates for each mode. Numbers in parentheses describe the response rate for each mode.



Statistical Analyses

Response rates were compared for questionnaires sent to families via post (paper completion rates) and those sent to families via email (EDC completion rates). Differences in questionnaire response rates for the postal (existing CDU procedure) and EDC (pilot REDCap procedure) groups were analyzed using chi-square tests.

Within the EDC group, we conducted additional chi-square tests to assess questionnaire modality preference (online or paper). To investigate the influence of demographic characteristics (age, primary language spoken, and highest level of education of caregiver completing questionnaire) on questionnaire modality preference in the EDC group, independent samples *t* tests and chi-square tests were used. Additionally, given that the CDU services families from diverse ethnic backgrounds, chi-square tests were used to assess questionnaire modality preference in respondents who requested an interpreter for the developmental assessment, and respondents who identified as being Aboriginal or Torres Strait Islander (ATSI).

Results

Study Population

Data were collected from 246 families across 2018 and 2019. Postal data were collected on 62 families seen in a 3-month period between June and August 2018, and online data were collected on 184 families seen between March and November 2019. Table 1 shows the distribution of responding across the postal and EDC groups.

Table 1. Questionnaire response rates for the postal (N=62) and EDC (N=184) groups.

Questionnaire completion status	Postal, n (%)	EDC[a], n (%)
Completed before appointment	24 (38.7)	164 (89.1)
Partially completed before appointment	5 (8.1)	7 (3.8)
Completed during appointment	33 (53.2)	10 (5.4)
Partially completed during appointment	0 (0.0)	3 (1.6)

[a]EDC: electronic data collection.

Differences in Response Rates Between the Postal and EDC Groups

As shown in Table 1, there was a significantly higher response rate in the EDC group (164/184, 89.1%) compared to the postal group (24/62, 39%; χ^2_3=78.8, *P*<.001). There was no variability in the number of partially completed responses between the EDC (10/184, 5.4%) and postal (5/62, 8%) groups (χ^2_1=0.2, *P*=.17).

Modality Preference

As shown in Table 2, families in the EDC group reported a significantly greater preference for completing and submitting the questionnaire online (151/173, 87.3%) compared to via post (22/173, 12.7%; χ^2_1=96.2, *P*<.001). This preference did not vary as a function of when the questionnaire was completed (ie, prior to or during appointment) or the amount of the questionnaire completed (ie, partial or full completion).

Table 2. Modality preference for questionnaire completion in the EDC[a] group.[b]

Questionnaire completion status	Paper preference, n (%)	Online preference, n (%)
Completed before appointment	21 (12.1)	143 (82.7)
Partially completed before appointment	0 (0.0)	1 (0.6)
Completed during appointment	1 (0.6)	7 (4.0)

[a]EDC: electronic data collection.

[b]Preference data missing for 11/184 families (6% of online sample). Percentages reported on the 173 respondents with completed preference data.

Influence of Demographic Characteristics on Questionnaire Modality Preference

Table 3 displays key demographic characteristics for individuals who completed the questionnaire in the EDC group. Caregiver ages ranged from 24 to 72 years (mean 37.41 [SD 7.05]) and most respondents (145/173, 83.8%) reported English as the main language spoken at home, either alone or in conjunction with a second language. Education level of caregivers who completed the questionnaires was stratified into nontertiary education (high school/vocational/trade) or tertiary education (undergraduate degree/postgraduate degree). In considering the influence of these characteristics on modality preference, the age of the person completing the questionnaire did not influence preference for online compared to paper completion, t_{167}=0.99, *P*=.32, nor did the primary language spoken by the person completing the questionnaire, χ^2_2=2.9, *P*=.24. However, our results revealed an association between education level and questionnaire modality preference, χ^2_2=4.2, *P*=.04. Overall, those individuals who had received tertiary education were less likely to report a preference for completing the questionnaire on paper, relative to those individuals who had received nontertiary education. Looking at the education levels of individuals who reported a preference for completing the online questionnaires, the opposite pattern was observed, such that a higher proportion of tertiary educated individuals reported a preference for online completion, compared to nontertiary educated individuals.

Table 3. Demographic characteristics by questionnaire modality preference in the EDC[a] group.

Characteristic	Paper preference	Online preference	*P* value[b]
Age (years), mean (SD)	36.02 (7.02)	37.62 (7.06)	.32
Primary language spoken, n (%)[c]			.24
English only	11 (6.4)	86 (49.7)	
English and other language	5 (2.9)	43 (24.9)	
Other language only	6 (3.5)	20 (11.6)	
Highest level of education, n (%)[d]			.04
Nontertiary	15 (8.7)	66 (38.2)	
Tertiary	7 (4.0)	81 (46.8)	

[a]EDC: electronic data collection.

[b]*P* value for independent samples *t* test (age) and chi-square test of independence (primary language spoken and highest level of education) for any group differences.

[c]Two respondents (2/173, 1.2%) did not provide information about primary language spoken at home.

[d]Four respondents (4/173, 2.3%) did not provide information about highest level of education.

Table 4 shows questionnaire modality preferences (online or paper) for caregivers who requested an interpreter for the assessment (10/173, 5.8%), and caregivers of ATSI origin (10/173, 5.8%). Within both subgroups of caregivers, there was no statistically significant difference in the number of families preferring online or paper completion (*P*=.53 and .21, respectively).

Table 4. Questionnaire modality preference for caregivers requesting an interpreter and ATSI caregivers in the EDC[a] group.[b]

Characteristic	Paper preference, n (%)	Online preference, n (%)	*P* value
Families requesting interpreter	4 (2.3)	6 (3.5)	.53
ATSI[c] origin	3 (1.7)	7 (4.0)	.21

[a]EDC: electronic data collection.

[b]Ten respondents (10/173, 5.8%) requested an interpreter for the assessment. A further 10 respondents (10/173, 5.8%) identified as being ATSI. There was no overlap between these subgroups of respondents.

[c]ATSI: Aboriginal or Torres Strait Islander.

Discussion

The results of this study show that EDC was associated with significantly increased questionnaire completion rates (*P*<.001) from caregivers prior to attending their first appointment. Response rates from EDC were more than double the rate from paper-based data collection methods. This overall superior completion rate was shown across families of different ethnic backgrounds and from caregivers with different education levels. Almost 90% (151/173, 87.3%) of caregivers who completed EDC reported a continued preference for using EDC over paper-based methods. This preference did not vary as a function of age, primary language spoken, or belonging to a minority subgroup. Consistent with previous findings, however, a higher level of education (tertiary compared to nontertiary) appeared to be associated with modality preference [23]. Of those caregivers who indicated a preference for the paper version, a greater proportion reported nontertiary education as their highest level. Overall, this pilot study supports the continued evaluation of EDC to improve efficiency, cost, and clinical and research services in public hospital–based child development clinics and supports its utility across diverse education levels and cultural groups [29].

Our finding of reduced questionnaire completion rates when paper-based data collection methods were used align with the clinical experiences of the CDU team, with staff reporting a long history of low response rates for the paper version of the questionnaire. This low response rate results in added clinical burden, as clinicians are required to complete the questionnaire with families at the time of their on-site appointment. This is far from ideal, given the logistics involved in preparing for each on-site assessment. For instance, the CDU carries out approximately 15 comprehensive assessments per week, spanning 1-3 full days. Assessments include tests of intelligence, developmental delay, language, neuropsychological assessments, comprehensive parent interviews, and medical examinations. Assessments are complex, requiring specific rooms, materials, and team members to be organized in advance. Without receiving the completed questionnaire prior to a family's appointment, the team of multidisciplinary clinicians are unable to adequately prepare for the type of assessment required in advance of the appointment. Our findings indicate that response rates are markedly improved when EDC is used, thereby giving the clinical team time to adequately prepare for assessments and optimizing time with families during assessments.

The increased response rates for the online questionnaire may be related to the increased preference seen for the electronic mode of completion [30-32]. Of note, in this study, we found that respondent age did not influence questionnaire preference for EDC. This finding is in contrast to previous studies that have reported a link between respondent age, response rates, and modality preferences [24,25,33]. Past studies have reported, however, that older age (eg, >60 years) is associated with greater preference for paper-based methods [33]. Given our sample principally comprised parents of young children, with a mean age of 37 and only 1 respondent above the age of 60, future studies may need to evaluate the utility of EDC in this service where primary caregivers are older (eg, grandparents). Our study also showed that the primary language spoken by the respondent did not influence questionnaire modality preference. Families who spoke a primary language other than English did not show differential preferences. Moreover, for 2 minority subgroups, namely, families who requested an interpreter for the assessment and families of ATSI origin, we did not observe an increased preference for paper-based methods compared to EDC. While these findings require replication in larger samples, they indicate that EDC may be suitable for the diverse populations typically serviced by developmental clinics such as the CDU. A more detailed investigation of language and ethnicity, and how these characteristics relate to socioeconomic status, may reveal differences and warrants further investigation [25].

In the small group of respondents who indicated a preference for the paper form, a majority (15/22, 68%) reported nontertiary education as their highest level. It has been shown that mothers with a high-school certificate level education or lower were less likely to use the internet for health-related purposes than those with a tertiary education. This may also be associated with lower socioeconomic status, lower household income, and lack of access to a computer or internet at home [34]. Education level is a known social determinant of health behavior and one that is difficult to address [35,36]. Publicly funded educational programs for vulnerable families may be a useful strategy to assist these families in better understanding their clinical care and options. Moreover, from a practical perspective, an understanding of the families likely to prefer paper forms will allow services such as the CDU to refocus their resources, by providing greater support at service entry to those families who cannot access EDC methods or require assistance from a team member. However, while we observed an association between education level and preference for a paper form, it should be noted that only a small minority of respondents (22/173, 12.7%) indicated preference for a paper version, highlighting the overall acceptability of EDC in this group.

There are some limitations in this study, namely, the relatively small sample size and uneven numbers in the postal and EDC groups. As the study aimed to explore tolerability of EDC, modality preference was only asked of online users. Additionally, the digitized questionnaire was a relatively short measure, taking approximately 15 minutes to complete. Results may differ for larger batteries of questionnaires and this would warrant further investigation in larger sample sizes. While we did not include an economic analysis in this study for EDC methods over paper-based approaches, this is clearly an avenue for future research. Such research would highlight the potential economic value of investing in high-quality internet-based health services for public settings. Moreover, future research would benefit from examining the feasibility and efficacy of EDC for clinician-collected data and evaluating staff satisfaction with these EDC methods. Such work would demonstrate the feasibility of extending EDC methods beyond patient-collected data in clinical health services such as the CDU.

Overall, this pilot study suggests that EDC is feasible and well accepted in a busy hospital-based clinic and has potential benefits for patient care, clinical practice, and clinical research. The increased response rates for online completion and the increased preference for EDC as opposed to paper forms suggest that EDC platforms may better suit the needs of families accessing these services.

Acknowledgments

This work was supported by NHMRC Project Grants (1125449 and 1043664) and a BUPA Foundation Grant to AG, a Research Training Program Scholarship to SP, and the Brain and Mind Centre Child Neurodevelopment and Mental Health Team.

Conflicts of Interest

None declared.

References

1. Klecun E. Transforming healthcare: policy discourses of IT and patient-centred care. European Journal of Information Systems 2016 Jan 1;25(1):64-76. [doi: 10.1057/ejis.2014.40]
2. Lavallee DC, Chenok KE, Love RM, Petersen C, Holve E, Segal CD, et al. Incorporating patient-reported outcomes into health care to engage patients and enhance care. Health Affairs 2016 Apr 01;35(4):575-582. [doi: 10.1377/hlthaff.2015.1362] [Medline: 27044954]
3. Ovretveit J, Keller C, Hvitfeldt FH, Essén A, Lindblad S, Brommels M. Continuous innovation: developing and using a clinical database with new technology for patient-centred care--the case of the Swedish quality register for arthritis. Int J Qual Health Care 2013 Apr;25(2):118-124. [doi: 10.1093/intqhc/mzt002] [Medline: 23360809]
4. Wang SJ, Middleton B, Prosser LA, Bardon CG, Spurr CD, Carchidi PJ, et al. A cost-benefit analysis of electronic medical records in primary care. Am J Med 2003 Apr 1;114(5):397-403. [Medline: 12714130]

5. Gardner W, Morton S, Tinoco A, Scholle SH, Canan BD, Kelleher KJ. Is It Feasible to Use Electronic Health Records for Quality Measurement of Adolescent Care? J Healthc Qual 2016;38(3):164-174. [doi: 10.1097/01.JHQ.0000462675.17265.db] [Medline: 26042752]
6. Staziaki PV, Kim P, Vadvala HV, Ghoshhajra BB. Medical Registry Data Collection Efficiency: A Crossover Study Comparing Web-Based Electronic Data Capture and a Standard Spreadsheet. J Med Internet Res 2016 Jun 08;18(6):e141 [FREE Full text] [doi: 10.2196/jmir.5576] [Medline: 27277523]
7. Highfill T. Do hospitals with electronic health records have lower costs? A systematic review and meta-analysis. International Journal of Healthcare Management 2019 May 21;13(1):65-71. [doi: 10.1080/20479700.2019.1616895]
8. Palacio C, Harrison JP, Garets D. Benchmarking electronic medical records initiatives in the US: a conceptual model. J Med Syst 2010 Jun 6;34(3):273-279. [doi: 10.1007/s10916-008-9238-5] [Medline: 20503611]
9. Park Y, Han D. Current Status of Electronic Medical Record Systems in Hospitals and Clinics in Korea. Healthc Inform Res 2017 Jul;23(3):189-198 [FREE Full text] [Medline: 28875054]
10. Ovretveit J, Scott T, Rundall TG, Shortell SM, Brommels M. Implementation of electronic medical records in hospitals: two case studies. Health Policy 2007 Dec;84(2-3):181-190. [doi: 10.1016/j.healthpol.2007.05.013] [Medline: 17624470]
11. Canaway R, Boyle DI, Manski-Nankervis JE, Bell J, Hocking JS, Clarke K, et al. Gathering data for decisions: best practice use of primary care electronic records for research. Med J Aust 2019 Apr;210 Suppl 6:S12-S16 [FREE Full text] [doi: 10.5694/mja2.50026] [Medline: 30927466]
12. Productivity Commission. Data Availability and Use, Inquiry Report. Canberra, ACT: Productivity Commission; 2017. URL: https://www.pc.gov.au/inquiries/completed/data-access/report/data-access.pdf [accessed 2020-01-16]
13. Dickinson FM, McCauley M, Madaj B, van den Broek N. Using electronic tablets for data collection for healthcare service and maternal health assessments in low resource settings: lessons learnt. BMC Health Serv Res 2019 May 27;19(1):336 [FREE Full text] [doi: 10.1186/s12913-019-4161-7] [Medline: 31133032]
14. Bodagh N, Archbold RA, Weerackody R, Hawking MKD, Barnes MR, Lee AM, et al. Feasibility of real-time capture of routine clinical data in the electronic health record: a hospital-based, observational service-evaluation study. BMJ Open 2018 Mar 08;8(3):e019790 [FREE Full text] [doi: 10.1136/bmjopen-2017-019790] [Medline: 29523565]
15. Ebert JF, Huibers L, Christensen B, Christensen MB. Paper- or Web-Based Questionnaire Invitations as a Method for Data Collection: Cross-Sectional Comparative Study of Differences in Response Rate, Completeness of Data, and Financial Cost. J Med Internet Res 2018 Jan 23;20(1):e24 [FREE Full text] [doi: 10.2196/jmir.8353] [Medline: 29362206]
16. Thriemer K, Ley B, Ame SM, Puri MK, Hashim R, Chang NY, et al. Replacing paper data collection forms with electronic data entry in the field: findings from a study of community-acquired bloodstream infections in Pemba, Zanzibar. BMC Res Notes 2012;5:113 [FREE Full text] [doi: 10.1186/1756-0500-5-113] [Medline: 22353420]
17. Bjertnaes O, Iversen HH, Skrivarhaug T. A randomized comparison of three data collection models for the measurement of parent experiences with diabetes outpatient care. BMC Med Res Methodol 2018 Sep 20;18(1):95 [FREE Full text] [doi: 10.1186/s12874-018-0557-z] [Medline: 30236067]
18. Barbaro J, Dissanayake C. Autism spectrum disorders in infancy and toddlerhood: a review of the evidence on early signs, early identification tools, and early diagnosis. J Dev Behav Pediatr 2009 Oct;30(5):447-459. [doi: 10.1097/DBP.0b013e3181ba0f9f] [Medline: 19823139]
19. Rutter M. Autism: its recognition, early diagnosis, and service implications. J Dev Behav Pediatr 2006 Apr;27(2 Suppl):S54-S58. [doi: 10.1097/00004703-200604002-00002] [Medline: 16685186]
20. Koegel LK, Koegel RL, Ashbaugh K, Bradshaw J. The importance of early identification and intervention for children with or at risk for autism spectrum disorders. Int J Speech Lang Pathol 2014 Feb;16(1):50-56. [doi: 10.3109/17549507.2013.861511] [Medline: 24328352]
21. McConachie H, Diggle T. Parent implemented early intervention for young children with autism spectrum disorder: a systematic review. J Eval Clin Pract 2007 Feb;13(1):120-129. [doi: 10.1111/j.1365-2753.2006.00674.x] [Medline: 17286734]
22. Ben Itzchak E, Zachor DA. Who benefits from early intervention in autism spectrum disorders? Research in Autism Spectrum Disorders 2011 Jan;5(1):345-350. [doi: 10.1016/j.rasd.2010.04.018]
23. Fan W, Yan Z. Factors affecting response rates of the web survey: A systematic review. Computers in Human Behavior 2010 Mar;26(2):132-139. [doi: 10.1016/j.chb.2009.10.015]
24. Jenkins P, Sng S, Brooksbank K, Brooksbank A. Socioeconomic deprivation and age are barriers to the online collection of patient reported outcome measures in orthopaedic patients. Ann R Coll Surg Engl 2016 Jan;98(1):40-44 [FREE Full text] [doi: 10.1308/rcsann.2016.0007] [Medline: 26688398]
25. Schamber EM, Takemoto SK, Chenok KE, Bozic KJ. Barriers to completion of Patient Reported Outcome Measures. J Arthroplasty 2013 Oct;28(9):1449-1453. [doi: 10.1016/j.arth.2013.06.025] [Medline: 23890831]
26. Flesch, R. A new readability yardstick. J Appl Psychol 1948 Jun;32(3):221-233. [doi: 10.1037/h0057532] [Medline: 18867058]
27. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]

28. 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]
29. Duffy JR, Kooken WC, Wolverton CL, Weaver MT. Evaluating patient-centered care: feasibility of electronic data collection in hospitalized older adults. J Nurs Care Qual 2012;27(4):307-315 [FREE Full text] [doi: 10.1097/NCQ.0b013e31825ba9d4] [Medline: 22677911]
30. Mlikotic R, Parker B, Rajapakshe R. Assessing the Effects of Participant Preference and Demographics in the Usage of Web-based Survey Questionnaires by Women Attending Screening Mammography in British Columbia. J Med Internet Res 2016 Mar 22;18(3):e70. [doi: 10.2196/jmir.5068] [Medline: 27005707]
31. Olson K, Smyth JD, Wood HM. Does Giving People Their Preferred Survey Mode Actually Increase Survey Participation Rates? An Experimental Examination. Public Opinion Quarterly 2012 Aug 02;76(4):611-635. [doi: 10.1093/poq/nfs024]
32. Smyth JD, Olson K, Millar MM. Identifying predictors of survey mode preference. Soc Sci Res 2014 Nov;48:135-144. [doi: 10.1016/j.ssresearch.2014.06.002] [Medline: 25131280]
33. Fitzgerald D, Hockey R, Jones M, Mishra G, Waller M, Dobson A. Use of Online or Paper Surveys by Australian Women: Longitudinal Study of Users, Devices, and Cohort Retention. J Med Internet Res 2019 Mar 14;21(3):e10672 [FREE Full text] [doi: 10.2196/10672] [Medline: 30869647]
34. Wen LM, Rissel C, Baur LA, Lee E, Simpson JM. Who is NOT likely to access the Internet for health information? Findings from first-time mothers in southwest Sydney, Australia. Int J Med Inform 2011 Jun;80(6):406-411. [doi: 10.1016/j.ijmedinf.2011.03.001] [Medline: 21493126]
35. Li J, Mattes E, Stanley F, McMurray A, Hertzman C. Social determinants of child health and well-being. Health Sociology Review 2014 Dec 17;18(1):3-11. [doi: 10.5172/hesr.18.1.3]
36. Li J, Powdthavee N. Does more education lead to better health habits? Evidence from the school reforms in Australia. Soc Sci Med 2015 Feb;127:83-91. [doi: 10.1016/j.socscimed.2014.07.021] [Medline: 25028347]

Abbreviations

ATSI: Aboriginal or Torres Strait Islander
CDU: Child Development Unit
EDC: electronic data collection
REDCap: Research Enterprise Data CAPture

Edited by G Eysenbach; submitted 12.02.20; peer-reviewed by C Middeldorp, A Vagelatos; comments to author 05.09.20; revised version received 13.10.20; accepted 04.11.20; published 19.01.21.

Please cite as:
Patel S, Boulton KA, Redoblado-Hodge MA, Papanicolaou A, Barnett D, Bennett B, Drevensek S, Cramsie J, Ganesalingam K, Ong N, Rozsa M, Sutherland R, Williamsz M, Pokorski I, Song YJC, Silove N, Guastella AJ
The Acceptability and Efficacy of Electronic Data Collection in a Hospital Neurodevelopmental Clinic: Pilot Questionnaire Study
JMIR Form Res 2021;5(1):e18214
URL: http://formative.jmir.org/2021/1/e18214/
doi:10.2196/18214
PMID:33464217

Original Paper

Cultural Adaptation of Digital Knowledge Translation Tools for Acute Otitis Media in Low- to Middle-Income Countries: Mixed Methods Usability Study

Salima Meherali[1*], PhD; Lisa Hartling[2], PhD; Shannon D Scott[1*], PhD

[1]Faculty of Nursing, University of Alberta, Edmonton, AB, Canada

[2]Department of Pediatrics, University of Alberta, Edmonton, AB, Canada

[*]these authors contributed equally

Corresponding Author:
Shannon D Scott, PhD
Faculty of Nursing
University of Alberta
Edmonton Clinic Health Academy
Edmonton, AB, T6G 1C9
Canada
Phone: 1 7804921037
Email: shannon.scott@ualberta.ca

Abstract

Background: Acute otitis media (AOM) is the most common pediatric bacterial ear infection. AOM presents challenges to parents who lack accurate information. Digital knowledge translation tools offer a promising approach to communicating complex health information. We developed AOM knowledge translation tools for Canadian parents and augmented them for Pakistani parent end users.

Objective: This pilot study aimed to (1) develop AOM knowledge translation tools for Canadian parents, (2) adapt the knowledge translation tools across cultural contexts, and (3) evaluate the usability of the adapted knowledge translation tools.

Methods: Parents' perceptions of the translated knowledge translation tools' usability were explored using a mixed-methods design. We recruited parent participants from a hospital in Pakistan to complete usability surveys (n=47) and focus group interviews (n=21). Descriptive statistics and content analysis were used to analyze data.

Results: Usability results showed the usefulness and effectiveness of both adapted knowledge translation tools. Parents reported preferring a digital media narrative format in their own language. Findings revealed that culturally adapted knowledge translation tools are effective in transferring health information to parents.

Conclusions: Digital knowledge translation tools offer a promising approach to improving health literacy and communicating complex health information to parents of children with AOM. Culturally adapting the tools generated important knowledge that will contribute to knowledge translation advancements. Evaluation of the tool effectiveness is a critical next step to exploring the impact of knowledge translation tools on child health outcomes.

(JMIR Form Res 2021;5(1):e13908) doi:10.2196/13908

KEYWORDS

acute otitis media; knowledge translation; pediatric; parent's experiences; information needs; global health

Introduction

Acute otitis media (AOM) is the most common pediatric bacterial infection, affecting up to 75% of children younger than 5 years [1,2]. AOM can cause pain in the ear, fever, and temporary hearing loss and is a leading cause of health care visits worldwide [2,3]. Despite the high incidence of AOM in children, it is often underrecognized and undertreated by clinicians [1]. Previous research identified that families want information about their child's illness, expected treatments, and post–emergency department [4] or clinic care. However, the gap between evidenced-based research and end user knowledge remains large. Knowledge translation (KT) is increasingly recognized as a solution to bridge this gap.

KT is an iterative process of synthesizing, disseminating, and ethically applying knowledge to improve health, health services, and health care systems globally [5]. In child health settings, emphasizing parents' role as partners in health care decision making reduces unnecessary health care use and ultimately improves health outcomes [5,6]. However, to be effective, parent health education should be multimodal and employ flexible, portable formats [7,8]. Incorporating illustrations and stories can also improve knowledge comprehension, retention, confidence, and compliance with discharge instructions [7-10]. Research has shown that innovative media (eg, digital and mobile technology, videos) are superior to traditional materials (eg, information sheets, pamphlets) in transferring information to consumers [8-14]. Previous research has illustrated the effectiveness of such tools in improving child health outcomes [8,11,13,15,16]. However, the scale-up of these digital KT tools for parents across different cultural contexts remains underdeveloped [16-18].

Closing the gap between research and practice has been consistently identified as a priority around the globe [19-21]. Despite this interest in KT, the best available research evidence is not consistently implemented in low- and middle-income countries (LMICs). Thus, the gap between research and practice are still increasingly wide in LMICs, where there are limited and scarce resources [21-24].

Digital or eHealth interventions have been identified as useful public health tools, particularly in underserved settings [25-29]. The availability and use of digital technologies, such as mobile phones, are increasing rapidly in LMICs [26,27]. eHealth interventions are useful in providing health information, reminders, emergency response, and monitoring [30]. In LMICs, digital interventions could reduce time, distance, and cost of information delivery, overcoming issues of inadequate financing, poor access to information, and limited human resources [31,32]. A wealth of research evidence is available from high-income countries on the effectiveness of digital health interventions. However, much less attention has been paid to how to augment these interventions to benefit LMICs.

The aim of this pilot project was to (1) develop digital KT tools on AOM for Canadian parents, (2) translate and augment the KT tools in a different cultural context, and (3) evaluate the usability of KT tools in a different cultural context (ie, Pakistani parents).

Methods

Overview

This study used a person-centered approach for the design and development of digital arts-based KT tools, with mixed methods for usability evaluation. We developed the digital arts-based KT tools in 5 stages (Figure 1).

In stage 1, we conducted a systematic review to determine the information needs of parents whose children have AOM. The findings revealed that parents' knowledge of AOM is generally limited. Further, parents were often poorly informed about AOM, resulting in uncertainties regarding how to help their children [33]. We conducted 16 individual qualitative interviews in stage 2 with parents who sought care for AOM in a hospital emergency department to understand their information needs and their experiences of having children with AOM [34]. Through thematic analysis, we found that AOM has considerable negative outcomes for both children and families (eg, pain, emotional strain for parents, etc) and that parents can benefit from evidence-based resources to meet their information needs [34].

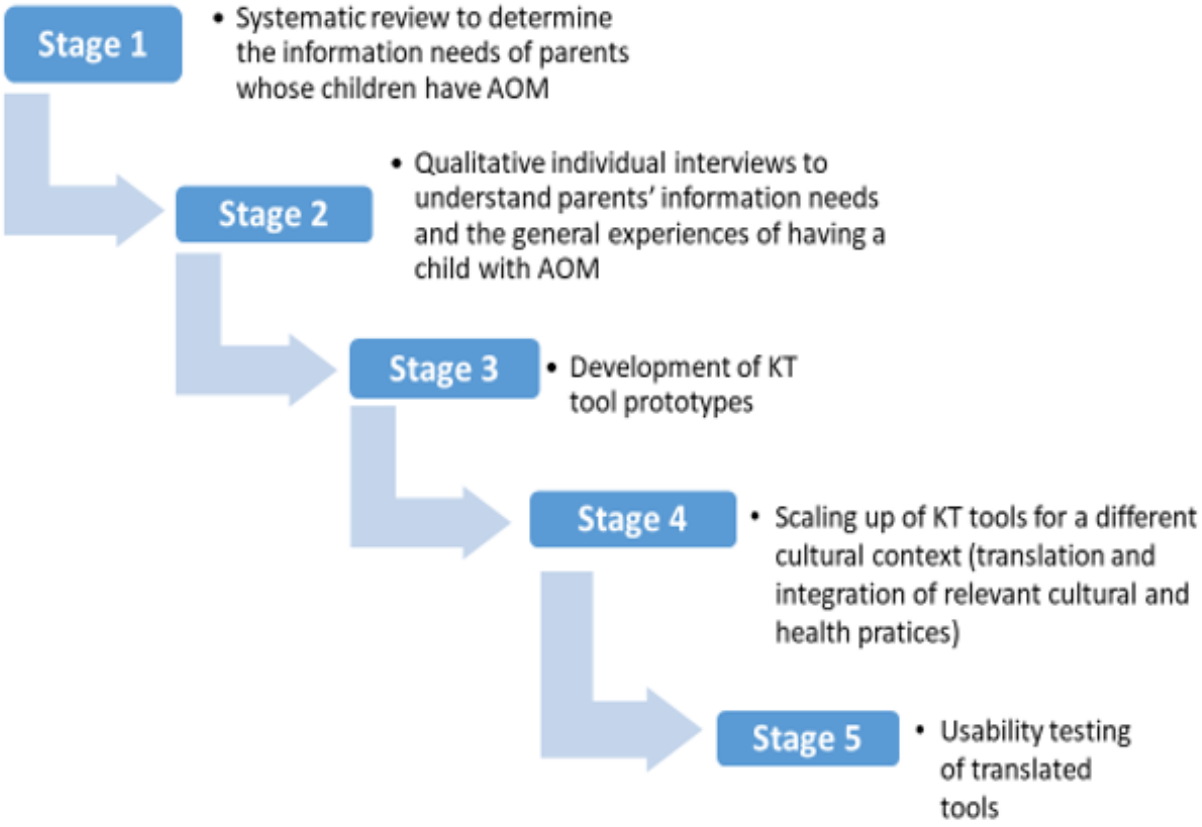
The data generated from stages 1 and 2 were used to develop digital KT tools for Canadian parents. In stage 3, we developed KT tools (whiteboard video and infographic) for Canadian parents. After sharing the stage 1 and 2 results with a creative team (storywriter, graphic designer, and editor), we collaboratively developed KT tool prototypes. This involved creating a composite narrative (a compilation of common themes from the parental interviews), ensuring that the narrative integrated the best available research, developing artwork to complement the composite narrative, and creating a graphic display of the narrative and artwork. We shared the prototypes with clinicians and content experts to ensure the accuracy and appropriate interpretation of the evidence.

To expand the use of evidence-based interventions across cultural contexts, we augmented and translated the digital AOM KT tools. In stage 4, we revised the tools to integrate relevant cultural and health practices for Pakistani parent end users. Specifically, we accommodated for language, literacy level, educational background, and the availability of technology. Research from many LMICs, including Pakistan, revealed that weak pharmaceutical regulations have allowed people to access antibiotics as over-the-counter drugs [35], which leads to antibiotic resistance among children. Thus, cultural norms for antibiotic prescriptions for AOM, health care professionals' roles in managing AOM, and parents' roles in making health care decisions for their children were incorporated. The evidence on treatment for AOM remained the same; the only changes were the names of drugs for symptomatic management and an increased emphasis in the video and infographics on the proper usage of antibiotics for AOM. After the revision, a professional translator translated the tools into Urdu, the national language of Pakistan. Considering the average literacy level in Pakistan, we used grade 5–level Urdu to enable the majority of parents to understand the information. The first author, who is fluent in Urdu, rechecked the translated version to ensure its accuracy and appropriateness. Multimedia Appendices 1 and 2 show the English and translated infographics, entitled "How to Help When Your Child Gets an Ear Infection." The whiteboard videos are available on YouTube in both English and Urdu (specially commended in the Institute of Human Development, Child and Youth Health Talks in 2018 and entitled "Mom! My ear hurts: What to do when your child has ear pain") [36,37].

Before we disseminated these tools in Pakistan, in stage 5 we conducted a mixed-methods study (surveys and focus groups) to determine their usability, usefulness, and cultural appropriateness for Pakistani parents. The study received ethics approval from the University of Alberta's Ethics Review Board. We also sought administrative approval from the medical director of the private hospital in Karachi, Pakistan, to recruit participants from that organization. We complied with all

national regulations and laws that apply to foreign researchers for data collection in Pakistan.

Figure 1. The 5-stage process of developing digital KT tools. AOM: acute otitis media; KT: knowledge translation.



Usability Survey

To evaluate the usability of digital KT tools (whiteboard video and infographic), we used a 9-item 5-point Likert survey. The key elements included were informed by a systematic search of more than 180 usability evaluations, such as usability, aesthetics, language, level of engagement, quality of information, length, preference of form over traditional dissemination venues, and value added [38]. The survey was translated into Urdu and then back translated into English to verify its consistency and accuracy.

Between July 17 and August 4, 2017, we shared the augmented and translated digital tools with parents at a large private hospital in Karachi. Two local research assistants approached parents waiting at pediatric clinics. The inclusion criteria for study participation were (1) being a parent or guardian of a child aged 1 to 16 years, with past experience taking care of a child with AOM, (2) fluency in Urdu (speaking, hearing, reading, and writing), and (3) agreement from parents to be contacted by a research team for iPad usability testing. First, the research assistants explained the research purpose and process, and upon participant agreement, they completed demographic information forms. Next, an iPad was provided to participants to assess the translated KT tools. Completion and submission of the survey indicated parents' consent. All of the participants viewed both tools in the same order, and the research assistants gave the paper-based usability survey forms to the parents. Parents also had an opportunity to provide free-text feedback on areas that required revisions or more information. The research assistants assisted parents who had difficulty understanding or filling out the survey forms. We cleaned and managed the data according to industry standards and entered them into IBM SPSS Statistics (version 23; IBM Corp) twice to ensure accuracy.

Qualitative Focus Groups

We conducted 3 focus group interviews with parents to augment the survey findings with rich detail. Using purposive and convenience sampling, we recruited different participants from the same private hospital in which we previously administered usability surveys due to time limitations and the unavailability of previous survey participants. The inclusion criteria for study participation were (1) being a parent or guardian of a child aged 1 to 16 years, with previous experience caring for a child with AOM; (2) fluency in Urdu (speaking, hearing, reading, and writing); and (3) being a parent of children aged 1 to 16 years, being interested in participating in a focus group discussion, and not necessarily having experience caring for a child with AOM. After consenting to participate, the parents and guardians completed demographic information forms. The first author conducted 3 semistructured focus group interviews with 21 parents (5 to 9 parents per group). The interviews lasted from 45 minutes to 1 hour. At the beginning of each focus group, all participants viewed both tools. A semistructured interview guide was used to explore the participants' perceptions of the translated KT tools (Multimedia Appendix 3). All the focus

groups were conducted in Urdu, audiorecorded, transcribed verbatim, and translated into English by a professional translator. Data analysis via NVivo 11 (QSR International) used a conventional content analysis technique [39]. The first author read each transcript carefully and highlighted text including participants' perceptions of the KT tools. All authors enhanced the analytic rigor by discussing the coding framework, analytic procedures, preliminary findings, and interpretations.

Results

Usability Survey Findings

We invited a total of 65 parents to participate in the usability survey, and 47 parents completed the survey forms. The majority (45/47, 96%) were mothers; 2 (4%) fathers participated. The ages ranged from 20 to 40 years (mean 30, SD 5.14 years). Table 1 shows the demographic characteristics of the study participants.

Table 1. Demographic characteristics of parents who participated in the usability testing survey and qualitative focus groups (N=68; 47 in usability survey and 21 in focus groups).

Variable	Participants, n (%)
Gender	
Male	4 (6)
Female	64 (94)
Parent's age (years)	
Younger than 20	1 (1)
20-30	30 (44)
31-40	29 (43)
41-50	8 (12)
Household income (RS)[a]	
Less than 15,000	6 (9)
15,000-30,000	16 (24)
31,000-60,000	20 (29)
61,000-1000,000	14 (21)
100,000 and above	5 (7)
I do not want to share this information	7 (10)
Highest level of education	
Some high school	12 (18)
Some postsecondary	7 (10)
Postsecondary certificate or diploma	3 (4)
Postsecondary degree	4 (6)
Graduate degree	35 (51)
Other	7 (10)
Total number of children in the house	
1	15 (22)
2	23 (34)
3	17 (25)
4	10 (15)
5	2 (3)
6	1 (1)

[a]A currency exchange rate of RS 160.65=US $1 is applicable.

Overall, the majority of the participants (40/47, 85%) reported that both of the KT tools were useful and effective in communicating health information. All participants (n=47) strongly agreed or agreed that the tool instructions were very simple and easy to use. The majority of the parents (43/47, 91%) strongly agreed or agreed that these tools would help them make health care decisions for their children with AOM, and most reported that they would use the tools in the future and

recommend them to their family members or friends (Table 2). The parents found these tools a great source of knowledge, stating that the tools raised their awareness of AOM. Additionally, participants reported that these tools would be useful for emergency cases.

Table 2. Frequency of participant answers on usability testing questionnaire (n=47).

Items	Strongly agree, n (%)	Agree, n (%)	Not sure, n (%)	Disagree, n (%)
The tools provide useful information	16 (34)	31 (66)	0 (0)	0 (0)
The tools provide information relevant to me	18 (38)	29 (62)	0 (0)	0 (0)
The tools are simple to use	17 (36)	28 (60)	2 (4)	0 (0)
I can use the tools without written instructions or additional help	13 (28)	19 (40)	7 (15)	8 (17)
Tools' lengths are appropriate	8 (17)	30 (64)	6 (13)	3 (6)
Tools are aesthetically pleasing	17 (36)	27 (58)	3 (6)	0 (0)
These tools help me to make decisions about my child's health	18 (38)	28 (60)	1 (2)	0 (0)
I would use these tools in the future	16 (34)	29 (62)	2 (4)	0 (0)
I would recommend these tools to a friend	20 (42)	24 (52)	3 (6)	0 (0)

Qualitative Focus Group Findings

A total of 21 parents participated in the focus group discussion. Table 1 shows the demographic characteristics of the participants. We identified 3 major codes: (1) parents' preference for KT tools (whiteboard video or infographics), (2) usability and feasibility of translated digital KT tools for parents, and (3) dissemination strategies.

Parents' Preference for KT Tools

Parents' reactions were generally positive. All parents in the focus groups preferred the whiteboard video to the infographic. They all considered the video more effective, as it provided details, including signs and symptoms and typical parental reactions. The parents also preferred the verbal information in the video, which made it possible to understand both audibly and visually. A few participants stated the infographic was useful, specifically in cases of limited technology access. A mother of a 4-year-old child stated:

> *Many people don't have smartphones, so not necessarily all of them can watch video. If such [a] community is targeted,…it would be handier that they read this pamphlet.*

Usability and Feasibility of Translated Digital KT Tools for Parents

All 21 focus group participants reported the information in both tools would be useful to families with young children. The majority agreed that the content was easy to understand. A mother of a 3-year-old child stated:

> *Information given in this video is in [a] very simple and easy [format] and in our own language [so] that everyone can comprehend it easily.*

The participants also acknowledged that using the character-plot-narrative format humanizes health information, evokes emotional responses, and creates connection to the subject matter. A mother of 2 children younger than 6 years remarked:

> *I really like the story format; it's really very interesting to see the parents, child, and doctor….Seems like I am seeing my personal experience.*

The participants also discussed the importance of using digital media to communicate health information to parents. As a mother of 3 children younger than 8 years stated:

> *Such videos will be helpful for the mother, as she will be aware that these are the immediate remedial steps which she can take to help her child with ear pain.*

Talking about Pakistani culture, one young mother said:

> *This kind of digital information is very beneficial for the girls who married at [an] early age. They don't like that their mother or mother-in-law teach[es] them about their children's healthcare. They believe more [in the] information available on social media or other digital platforms.*

The participants also liked that the video was translated into Urdu and will be available on YouTube, making it more relevant to their culture. A father of 3 children said:

> *The tools in our own language provide an accurate account of what we feel as parents when our child experience[s an] ear infection.*

They agreed that it is important to have health information in their local language.

Dissemination Strategies

The majority of parents preferred to have the KT tools disseminated through the private health organization's established social media platform, such as YouTube or Facebook. One mother of 2 children younger than 6 years stated:

> *This private organization has a big name, and if you just make a blog page or a video interactive page where this video can be uploaded for easy access, a pop-up message will attract the website user to watch these videos as soon as anybody visits the website.*

Some participants suggested that it would be helpful to show these kinds of videos in medical clinic waiting rooms to effectively use parents' time. In addition, one female participant suggested a more direct dissemination method:

> *[The] hospital administration can make a broadcast list and send videos officially to parents. Hence, in the presence of a broadcast list, you can easily direct messages to parents, and [their] phone number is not shared [with] other persons,. . . so, the privacy will remain intact too.*

Discussion

Principal Findings

Our findings demonstrate that culturally adapted translated KT tools permit parents' receptivity to information, reassure them, and foster confidence. The parents found the tools usable and appreciated receiving digital health information in a narrative form in their own language. They considered these tools a great source of knowledge that raised their awareness on AOM. We found that the culturally adapted digital KT tools we developed have the potential to increase participant engagement and help parents in health decision making.

There is little evidence on how to best scale up digital KT interventions developed in Western countries to reduce child morbidity and mortality in LMICs. This study was an effort to address this gap. The usability evaluation revealed that modifying KT tools to fit a different language, culture, and lifestyle would potentially increase the evidence-based information to reach a greater number of parents.

Although culturally adapted digital KT interventions show great promise in improving child health outcomes, these interventions are rarely implemented. The assumption is that effective KT interventions in the specific context of a Western industrialized setting will not necessarily work in LMICs [40]. The World Bank argues that the scaling up of KT interventions should be "driven by a universalist process of simplifying rules and procedures for use by many people on a larger scale" [41]. With this in mind, we ensured that people with low literacy can easily understand the KT tools that we developed, and because digital and web-based technology is growing quickly in Pakistan, the scale-up process was straightforward [41].

Digital media can also facilitate the dissemination of evidence-based health care information without requiring significant amounts of health and human resources. Many LMICs are grappling with a crisis in human resources for health care caused by factors such as underinvestment in health and the brain drain of health professionals [42]. Coupled with the increased burden of disease and lack of affordable health care, the human resource crisis means that it is not practical to deliver services only through physical interactions. Hence, a growing number of practitioners are leveraging advances in communications technology to strengthen health care systems in Pakistan. The findings reveal that digital KT interventions have the potential to improve patient knowledge and are the preferred method to receive health information.

Study Limitations

To the best of our knowledge, this is the first study conducted in an LMIC to evaluate the usability of culturally adapted digital KT tools. The findings of the study cannot be generalized to a broader population or other languages and cultures. Further, we recruited the sample from only one site, a large private hospital, so our sample potentially reflects parents who are educated and familiar with digital technology.

Study Implications

Global Health 2035 made a powerful case for increasing investments in the development of new KT health tools and scaling up new and existing tools. This study provides valuable insights into scaling up digital KT tools for a different culture than they were originally intended for. Scaling up digital KT tools for use in different cultures can change the trajectory of child health globally. However, very little funding is available to many LMICs to conduct this type of research. In the United States, the National Institutes of Health and the Bill & Melinda Gates Foundation reported that 97% of research funding is directed to the development of new health technologies and only 3% to research into implementation [43]. Commenting on what they called the "3/97 gap," the authors estimated that research into the development of new technologies could prevent about 22% of child deaths by scaling up the existing tools [44]. However, scaling up digital KT interventions and tools is not easy. Barriers include language and translation, literacy and education, culture, trust in Western interventions, cost of health care, public and private systems of care, use of unregulated private providers, unequal access to services, availability of digital technologies, and decision-making processes within families. Addressing these barriers and improving access to knowledge will improve the health outcomes. Our research team's next step is to augment other digital KT tools to address common acute conditions in LMICs; develop guidelines for the adaptation of KT tools for different cultures, countries, and contexts; and evaluate the effectiveness of these digital tools in improving the health outcomes of children in LMICs.

Conclusion

The process of scaling up digital KT tools discussed in this paper generated important new knowledge that contributes to the science of KT. On a global scale, several ongoing initiatives support scaling up successful digital health interventions. However, cultural adaptation in KT strategies and tools is critically important for the successful scale-up of digital health solutions. These novel findings highlight the potential for digital art-based KT tools, given their congruence with human communication and learning approaches. Our findings suggest that future research that involves digital art- and narrative-based tools for KT is needed and worthwhile; in particular, assessing these approaches with different types of clinical conditions (eg, acute vs chronic health conditions) and different types of parents (eg, demographics, educational levels, ethnic backgrounds) will be helpful. Future research to evaluate the acceptability of KT tools among local health care professionals and families as well as rigorous effectiveness evaluations of these tools are critical next steps to measuring the impact of KT tools on child health outcomes.

Acknowledgments

This project was supported by the Women's and Children's Health Research Institute Innovation grant (grant number RES0031220 awarded to SDS), Translating Emergency Knowledge for Kids (grant number SFR1164), and the Canadian Institutes of Health Research Foundation Grant (number 148411). SM is postdoctoral fellow and received a fellowship award from the Faculty of Nursing, University of Alberta. LH is supported by a Canada Research Chair. SDS is a Canada Research Chair (tier 2) for Knowledge Translation in Child Health. Both LH and SDS are Distinguished Researchers of Stollery Science Lab.

Authors' Contributions

SM conceived and co-designed this cultural adaptation study, led the analysis with support from SDS and LH, and drafted the manuscript. SDS co-designed the study and obtained research funding for this study. SDS and LH conceived, designed, led, and obtained funding for the intervention development in Canada. All authors commented on all drafts of the paper. All approved the final draft.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Infographic.
[PDF File (Adobe PDF File), 3519 KB - formative_v5i1e13908_app1.pdf]

Multimedia Appendix 2
Infographic in Urdu.
[PNG File , 1841 KB - formative_v5i1e13908_app2.png]

Multimedia Appendix 3
Focus group interview guide.
[DOCX File , 15 KB - formative_v5i1e13908_app3.docx]

References

1. Qureishi A, Lee Y, Belfield K, Birchall J, Daniel M. Update on otitis media - prevention and treatment. Infect Drug Resist 2014 Jan 10;7:15-24 [FREE Full text] [doi: 10.2147/IDR.S39637] [Medline: 24453496]
2. Venekamp R, Sanders S, Glasziou P, Del Mar CB, Rovers M. Antibiotics for acute otitis media in children. Cochrane Database Syst Rev 2015 Jun 23;2015(6):CD000219 [FREE Full text] [doi: 10.1002/14651858.CD000219.pub4] [Medline: 26099233]
3. Monasta L, Ronfani L, Marchetti F, Montico M, Vecchi Brumatti L, Bavcar A, et al. Burden of disease caused by otitis media: systematic review and global estimates. PLoS One 2012;7(4):e36226 [FREE Full text] [doi: 10.1371/journal.pone.0036226] [Medline: 22558393]
4. Scott S, Albrecht L, Given L, Hartling L, Johnson D, Jabbour M, et al. Pediatric information seeking behaviour, information needs, and information preferences of health care professionals in general emergency departments: Results from the Translating Emergency Knowledge for Kids (TREKK) Needs Assessment. CJEM 2018 Jan;20(1):89-99. [doi: 10.1017/cem.2016.406] [Medline: 28067181]
5. Sources of Potentially Avoidable Emergency Department Visits. Canadian IFHI. 2014. URL: https://www.longwoods.com/articles/images/HQ_Vol17_No4-Facts_at_a_Glance.pdf [accessed 2018-09-05]
6. Scott SD, Hartling L, O'Leary KA, Archibald M, Klassen TP. Stories – a novel approach to transfer complex health information to parents: A qualitative study. Arts Health 2012 Jun;4(2):162-173. [doi: 10.1080/17533015.2012.656203]
7. Gittelman MA, Mahabee-Gittens EM, Gonzalez-del-Rey J. Common medical terms defined by parents: are we speaking the same language? Pediatr Emerg Care 2004 Nov;20(11):754-758. [doi: 10.1097/01.pec.0000144918.00995.8a] [Medline: 15502657]
8. Hartling L, Scott S, Pandya R, Johnson D, Bishop T, Klassen TP. Storytelling as a communication tool for health consumers: development of an intervention for parents of children with croup. Stories to communicate health information. BMC Pediatr 2010 Sep 02;10:64 [FREE Full text] [doi: 10.1186/1471-2431-10-64] [Medline: 20813044]
9. Scott SD, Brett-MacLean P, Archibald M, Hartling L. Protocol for a systematic review of the use of narrative storytelling and visual-arts-based approaches as knowledge translation tools in healthcare. Syst Rev 2013 Mar 20;2:19 [FREE Full text] [doi: 10.1186/2046-4053-2-19] [Medline: 23514237]

10. Archibald M, Scott S, Hartling L. Mapping the waters: A scoping review of the use of visual arts in pediatric populations with health conditions. Arts Health 2013 Jan 29;6(1):5-23 [FREE Full text] [doi: 10.1080/17533015.2012.759980]
11. Graham ID, Tetroe J. How to translate health research knowledge into effective healthcare action. Healthc Q 2007;10(3):20-22 [FREE Full text] [doi: 10.12927/hcq..18919] [Medline: 17632905]
12. Boydell K, Gladstone B, Volpe T, Allemang B, Stasiulis E. The production and dissemination of knowledge: A scoping review of arts-based health research. Forum Qualitative Sozialforschung 2012;13(1):1711. [doi: 10.17169/fqs-13.1.1711]
13. Hartling L, Scott SD, Johnson DW, Bishop T, Klassen TP. A randomized controlled trial of storytelling as a communication tool. PLoS One 2013;8(10):e77800 [FREE Full text] [doi: 10.1371/journal.pone.0077800] [Medline: 24204974]
14. Houston TK, Allison JJ, Sussman M, Horn W, Holt CL, Trobaugh J, et al. Culturally appropriate storytelling to improve blood pressure: a randomized trial. Ann Intern Med 2011 Jan 18;154(2):77-84. [doi: 10.7326/0003-4819-154-2-201101180-00004] [Medline: 21242364]
15. Archibald MM. Developing a Patient-Driven Arts-Based Knowledge Translation Tool for Parents of Children with Asthma [thesis]. University of Alberta. 2016. URL: https://era.library.ualberta.ca/items/1be988bf-0020-40a4-9fa9-0c28f427ed75/view/6f8fc8e7-f9a6-4c44-8fa2-b1b24be9e564/Archibald_Mandy_M_201603_PhD.pdf [accessed 2021-01-08]
16. Aarons G, Sklar M, Mustanski B, Benbow N, Brown C. "Scaling-out" evidence-based interventions to new populations or new health care delivery systems. Implement Sci 2017 Sep 06;12(1):111 [FREE Full text] [doi: 10.1186/s13012-017-0640-6] [Medline: 28877746]
17. Milat A, Bauman A, Redman S. Narrative review of models and success factors for scaling up public health interventions. Implement Sci 2015 Aug 12;10:113 [FREE Full text] [doi: 10.1186/s13012-015-0301-6] [Medline: 26264351]
18. Scaling up projects and initiatives for better health: from concepts to practice. World Health Organization. 2016. URL: https://www.euro.who.int/__data/assets/pdf_file/0004/318982/Scaling-up-reports-projects-concepts-practice.pdf [accessed 2018-10-16]
19. Leach M, Tucker B. Current understandings of the research-practice gap in nursing: A mixed-methods study. Collegian 2018 Apr;25(2):171-179 [FREE Full text] [doi: 10.1016/j.colegn.2017.04.008]
20. Olswang LB, Prelock PA. Bridging the Gap Between Research and Practice: Implementation Science. J Speech Lang Hear Res 2015 Dec;58(6):S1818-S1826. [doi: 10.1044/2015_JSLHR-L-14-0305] [Medline: 26262822]
21. mHealth: New horizons for health through mobile technologies. World Health Organization. 2017. URL: http://www.who.int/goe/publications/goe_mhealth_web.pdf [accessed 2018-10-16]
22. Goyet S, Touch S, Ir P, SamAn S, Fassier T, Frutos R, et al. Gaps between research and public health priorities in low income countries: evidence from a systematic literature review focused on Cambodia. Implement Sci 2015 Mar 11;10:32 [FREE Full text] [doi: 10.1186/s13012-015-0217-1] [Medline: 25889672]
23. Lavis JN, Guindon GE, Cameron D, Boupha B, Dejman M, Osei EJA, Research to Policy and Practice Study Team. Bridging the gaps between research, policy and practice in low- and middle-income countries: a survey of researchers. CMAJ 2010 Jun 15;182(9):E350-E361 [FREE Full text] [doi: 10.1503/cmaj.081164] [Medline: 20439449]
24. Lavis J. How can we support the use of systematic reviews in policymaking? PLoS Med 2009 Nov;6(11):e1000141 [FREE Full text] [doi: 10.1371/journal.pmed.1000141] [Medline: 19936223]
25. McCannon CJ, Berwick DM, Massoud MR. The science of large-scale change in global health. JAMA 2007 Oct 24;298(16):1937-1939. [doi: 10.1001/jama.298.16.1937] [Medline: 17954547]
26. Agarwal S, Labrique A. Newborn health on the line: the potential mHealth applications. JAMA 2014 Jul 16;312(3):229-230. [doi: 10.1001/jama.2014.6371] [Medline: 24953141]
27. 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;20(1):4-34. [doi: 10.1080/10810730.2013.864735] [Medline: 24673171]
28. Clifford GD. E-health in low to middle income countries. J Med Eng Technol 2016;40(7-8):336-341 [FREE Full text] [doi: 10.1080/03091902.2016.1256081] [Medline: 27885920]
29. Qiang C, Yamamichi M, Hausman V, Miller R, Altman D. Mobile Applications for the Health Sector. World Bank. 2012. URL: http://documents.worldbank.org/curated/en/751411468157784302/pdf/726040WP0Box370th0report00Apr020120.pdf [accessed 2018-10-29]
30. Ormel H. MHealth and eHealth, innovations in public health and SRHR: state of evidence, opportunities and challenges. Royal Tropical Institute Amsterdam. 2013. URL: http://www.mhealthinfo.org/resources-database [accessed 2018-09-26]
31. Amplifying the Impact: Examining the Intersection of Mobile Health and Mobile Finance: A discussion guide for collaborative insight presented by the World Economic Forum, in partnership with the Health Alliance. World Economic Forum. 2011. URL: http://mhealthknowledge.org/resources/amplifying-impact-examining-intersection-mobile-health-and-mobile-finance [accessed 2018-09-20]
32. Labrique AB, Wadhwani C, Williams KA, Lamptey P, Hesp C, Luk R, et al. Best practices in scaling digital health in low and middle income countries. Global Health 2018 Nov 03;14(1):103 [FREE Full text] [doi: 10.1186/s12992-018-0424-z] [Medline: 30390686]
33. Meherali S, Hartling L, Campbell A, Robin F, Scott S. Parent information needs and experience regarding acute otitis media in children: A systematic review. Patient Educ Couns 2020 Sep 21. [doi: 10.1016/j.pec.2020.09.013] [Medline: 32998839]

34. Meherali S, Campbell A, Hartling L, Scott S. Understanding Parents' Experiences and Information Needs on Pediatric Acute Otitis Media: A Qualitative Study. J Patient Exp 2019 Mar;6(1):53-61 [FREE Full text] [doi: 10.1177/2374373518771362] [Medline: 31236452]
35. Khalid L, Mahsood N, Ali I. The public health problem of OTC antibiotics in developing nations. Res Social Adm Pharm 2016;12(5):801-802. [doi: 10.1016/j.sapharm.2016.03.003] [Medline: 27067233]
36. Mom! My ear hurts: What to do when your child has ear pain. YouTube. URL: https://www.youtube.com/watch?v=OFEJhUoqkVA&index=15&list=PLxWz0fEGuv6riED2 [accessed 2021-01-07]
37. Mom! My ear hurts: What to do when your child has ear pain. Translated Version. YouTube. URL: https://www.youtube.com/watch?v=P5Be5yEX0UA [accessed 2021-01-05]
38. Hornbæk K. Current practice in measuring usability: Challenges to usability studies and research. Int J Hum Comput Stud 2006 Feb;64(2):79-102 [FREE Full text] [doi: 10.1016/j.ijhcs.2005.06.002]
39. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
40. Miranda JJ, Zaman MJ. Exporting 'failure': why research from rich countries may not benefit the developing world. Rev Saude Publica 2010 Feb;44(1):185-189 [FREE Full text] [doi: 10.1590/s0034-89102010000100020] [Medline: 20140343]
41. Scaling up the impact of good practices in rural development: a working paper to support implementation of the World Bank's rural development strategy. World Bank. 2003. URL: http://documents.worldbank.org/curated/en/203681468780267815/pdf/260310White0co1e1up1final1formatted.pdf [accessed 2018-10-20]
42. Misau YA, Al-Sadat N, Gerei AB. Brain-drain and health care delivery in developing countries. J Public Health Afr 2010 Sep 01;1(1):e6 [FREE Full text] [doi: 10.4081/jphia.2010.e6] [Medline: 28299040]
43. Kuruvilla S, Schweitzer J, Bishai D, Chowdhury S, Caramani D, Frost L, Success Factors for Women's and Children's Health study groups. Success factors for reducing maternal and child mortality. Bull World Health Organ 2014 Jul 01;92(7):533-44B [FREE Full text] [doi: 10.2471/BLT.14.138131] [Medline: 25110379]
44. Kruk M, Yamey G, Angell S, Beith A, Cotlear D, Guanais F, et al. Transforming Global Health by Improving the Science of Scale-Up. PLoS Biol 2016 Mar;14(3):e1002360 [FREE Full text] [doi: 10.1371/journal.pbio.1002360] [Medline: 26934704]

Abbreviations

AOM: acute otitis media
KT: knowledge translation
LMIC: low- and middle-income country

Edited by G Eysenbach; submitted 04.03.19; peer-reviewed by K Baker, T Giardina, R Grace, R Poluru; comments to author 03.10.19; revised version received 14.11.19; accepted 21.12.20; published 20.01.21.

Please cite as:
Meherali S, Hartling L, Scott SD
Cultural Adaptation of Digital Knowledge Translation Tools for Acute Otitis Media in Low- to Middle-Income Countries: Mixed Methods Usability Study
JMIR Form Res 2021;5(1):e13908
URL: http://formative.jmir.org/2021/1/e13908/
doi:10.2196/13908
PMID:33470942

Original Paper

A Decision Aid Intervention for Family Building After Cancer: Developmental Study on the Initial Steps to Consider When Designing a Web-Based Prototype

Catherine Benedict[1,2], PhD; Katherine L Dauber-Decker[3], PhD; D'Arcy King[3], PhD; Alexandria Hahn[4], MSc; Jennifer S Ford[5], PhD; Michael Diefenbach[3], PhD

[1]Stanford University School of Medicine, Palo Alto, CA, United States

[2]Stanford Cancer Institute, Palo Alto, CA, United States

[3]Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, Manhasset, NY, United States

[4]Albert Einstein College of Medicine, The Bronx, NY, United States

[5]Hunter College and The Graduate Center, City University of New York, New York, NY, United States

Corresponding Author:
Catherine Benedict, PhD
Stanford University School of Medicine
401 Quarry Road
Palo Alto, CA, 94305-5718
United States
Phone: 1 2012077117
Email: cbenedict@stanford.edu

Abstract

Background: An important aspect of patient-centered care involves ensuring that patient-directed resources are usable, understandable, and responsive to patients' needs. A user-centered design refers to an empathy-based framework and an iterative design approach for developing a product or solution that is based on an in-depth understanding of users' needs, values, abilities, and limitations.

Objective: This study presents the steps taken to develop a prototype for a patient resource for young women who have completed treatment for gonadotoxic cancer to support their decision making about follow-up fertility care and family building.

Methods: User-centered design practices were used to develop *Roadmap to Parenthood*, a decision aid (DA) website for family building after cancer. A multidisciplinary steering group was assembled and input was provided. Guidelines from the International Patient DA Society and the Ottawa Decision Support Framework were used throughout the development process. In addition, guidelines for developing health DAs with respect to patient diversity and health literacy were also followed.

Results: The *Roadmap to Parenthood* DA website prototype was systematically and iteratively developed. An extensive process of designing and developing solutions from the perspective of the end user was followed. The steps taken included formative work to identify user needs; determining goals, format, and delivery; design processes (eg, personas, storyboards, information architecture, user journey mapping, and wireframing); and content development. Additional design considerations addressed the unique needs of this patient population, including the emotional experiences related to this topic and decision-making context wherein decisions could be considered iteratively while involving a multistep process.

Conclusions: The design strategies presented in this study describe important steps in the early phases of developing a user-centered resource, which will enhance the starting point for usability testing and further design modifications. Future research will pilot test the DA and a planning tool, and evaluate improvement in the decisional conflict regarding family building after cancer. Consistent with a patient-centered approach to health care, the strategies described here may be generalized and applied to the development of other patient resources and clinical contexts to optimize usability, empathy, and user engagement.

(JMIR Form Res 2021;5(1):e20841) doi:10.2196/20841

KEYWORDS

patient-centered care; user-centered design; decision support techniques; decision aid; cancer; fertility; internet-based intervention; web-based intervention; mobile phone; psychosocial intervention

Introduction

Background

Patient-centered care is well established as an important aspect of health care quality. As put forth by the Institute of Medicine, all care should respect and be responsive to patients' preferences, needs, and values [1]. Patients should have access to education and support to act as informed decision makers and participate in shared decision making with providers to ensure that their individual values are reflected in the treatment plans [2]. Operationally, an important aspect of patient-centered care involves ensuring that patient-directed information, education, and communication are usable, understandable, and responsive to patients' needs. To support value-based decision making, it is important to develop patient resources with the target user group in mind.

This study focuses on oncofertility as an example of a clinical context in which there is an unmet need for patient-centered support. For young adult survivors of cancer (ie, aged 18-39 years), fertility is ranked among the most important survivorship issues [3,4]. Patient-centered resources are needed to inform patients about infertility risks and family-building options, support their decision making, guide their involvement in seeking care, and prepare them for potential future challenges. This paper describes the first phase of the development process of a patient decision aid (DA) and planning tool for family building after cancer.

Family Building After Cancer

Owing to gonadotoxic treatments, many women experience reduced ovarian function or are unable to safely carry a pregnancy to term after cancer. The prevalence of primary ovarian insufficiency in female survivors of pediatric, adolescent, and young adult cancers ranges from 2% to 82%, based on patient factors, cancer diagnosis, and treatment exposures [5]. Alternative family-building options include the use of assisted reproductive technology, such as in vitro fertilization (IVF) and surrogacy, or adoption or fostering. With assisted reproduction, options comprise the use of fresh, frozen, or donated gametes to achieve pregnancy in the survivor or a gestational carrier. Adoption may be domestic or international. Each of these family-building options comes with a number of physical, emotional, financial, legal, and logistical challenges that need proper consideration; hence, decision making can be complex. For many patients, there may be benefits of an *early action* even if desired family building may be years away, including undergoing a fertility evaluation posttreatment to better understand their reproductive options and expected reproductive timeline, undergoing egg/embryo freezing posttreatment if they are at a risk for early menopause but not yet ready to start their family, or financial planning. Family-building decisions are based on values, and survivors must weigh the pros and cons of their options regarding risk-benefit tradeoffs. Given the emotional salience of motherhood desires, many women report high levels of uncertainty and distress when prompted to consider fertility and family-building decisions after cancer [6].

Decision Support

Young female survivors of cancer report unmet support needs related to posttreatment fertility care in survivorship and want to be provided informational resources to help them understand their options for pursuing future parenthood [7,8]. Patient DAs are effective for improving tailored decision-making quality such that the users are more likely to be informed, gain clarity about how their values align with their decision options, and take a more active role in decision making [9]. Advantages of delivering patient DAs over the internet include an increasing reach and potential effectiveness [10]. Multiple patient DAs exist for young women diagnosed with cancer who are considering fertility preservation before treatment [11,12]. Although these studies support the use of DAs for fertility-related decisions in the context of cancer care [11,12], to our knowledge, there are no decision support resources that address the posttreatment reproductive survivorship care and family-building decisions that must be made after the completion of treatment.

User-Centered Design

This study used user-centered design principles to develop a patient resource that supports decision making about family building after cancer treatment. A user-centered design is an empathy-based framework and an iterative design approach for developing a product or solution based on an in-depth understanding of users' needs, values, abilities, and limitations. This iterative process is effective and essential because it places end users at the center of every stage of development—in this case young adult female survivors of cancer—to ensure that the end product reflects and addresses their needs [13]. Conversely, the failure to consider end users' insights, feedback, and needs results in products and solutions that are less likely to achieve optimal adoption, retention, and advocacy [14]. Technology acceptance models and theories on telemedicine adoption highlight the importance of co-design with end users to develop products that are perceived as useful, easy to use, and responsive to needs [15,16].

Study Objectives

To address a critical gap in young adult cancer survivorship care, we set out to develop a web-based patient DA and planning tool to support young women interested in family building after cancer. This study presents the steps taken to develop the prototype of the website, *Roadmap to Parenthood*, based on user-centered design practices and guidelines for developing DAs and health care resources for diverse patient groups and health literacy levels. This work was guided by a theoretical approach grounded in the self-regulation theory [17,18] and further developed in our preliminary work, which is described elsewhere [19,20]. In this paper, we review the *initial* design steps and process to develop a DA prototype before conducting formal usability testing. These steps aim to optimize usability, empathy, and user engagement to ensure universal applicability across patient subgroups. Our intention in this paper is to thoroughly describe the prototype design process, which allows us to enter a formal usability testing phase that considers key design issues and user feedback.

Methods

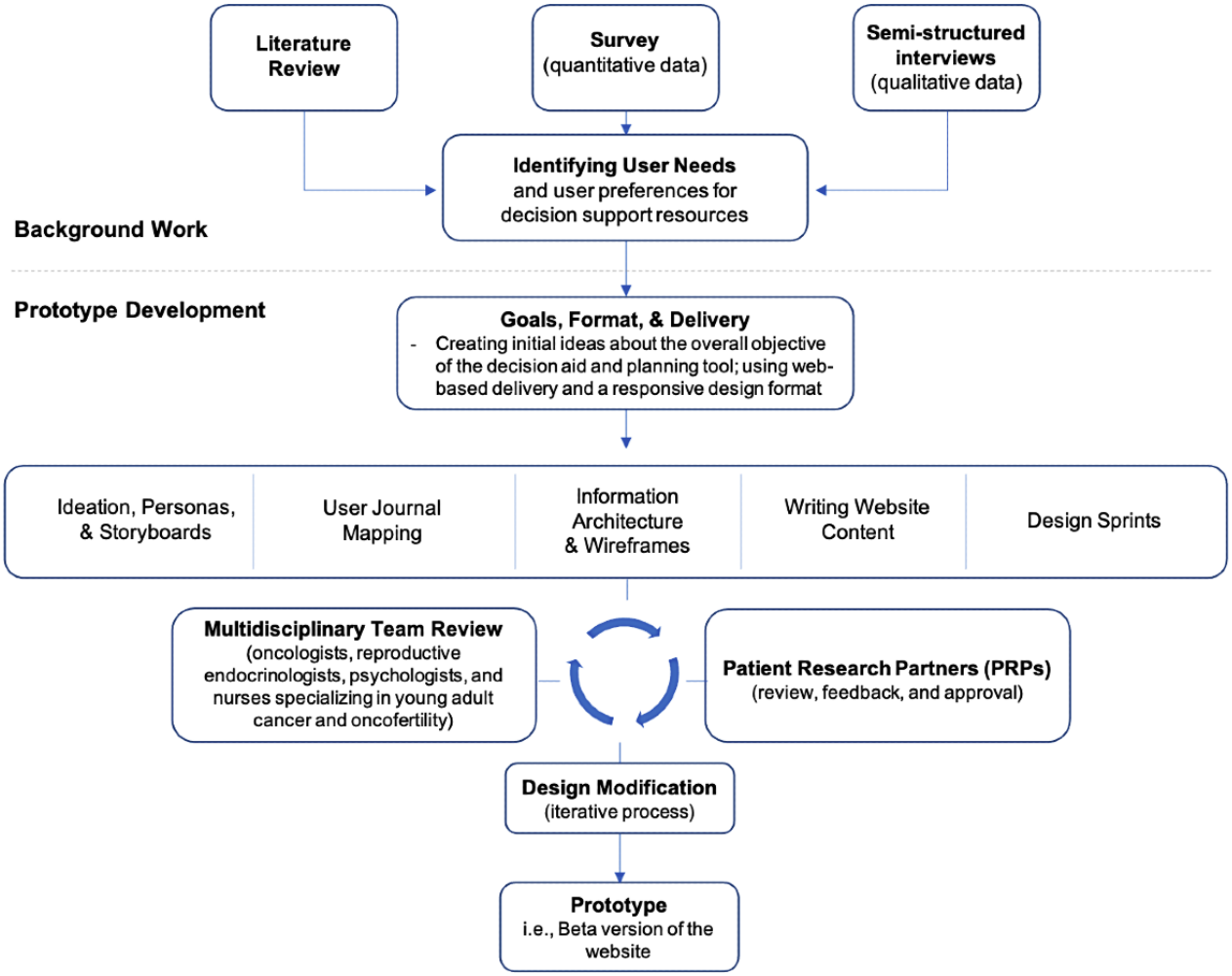
Preliminary Studies

The study team led several oncofertility studies focusing on young adult female cancer survivors' experiences related to fertility and family building posttreatment and identified unmet decision-making needs and patient preferences for support. Our national survey of posttreatment reproductive concerns and decision making uncertainty identified the areas of decisional conflict about family building after cancer (eg, lack of information, clarification of values, and lack of emotional support) [6]. Two additional qualitative studies explored posttreatment fertility concerns [21] and family-building experiences [22] and informed our understanding of user needs. On the basis of this work, semistructured interviews (N=25) were conducted with young adult female survivors of cancer (aged 15-45 years) who received gonadotoxic treatment and were either interested in future family building or uncertain about their family-building plans [19]. Briefly, women reported high rates of unmet information needs, including uncertainty about reproductive survivorship care and where to obtain trusted information. They felt overwhelmed and distressed by the prospect of pursuing family building and its expected, associated challenges [19]. When asked about support preferences, they indicated a desire for *step-by-step* instructions to learn about their options and guide decision making and follow-up care [20]. They also reported a preference for web-based resources for self-education, which they envisioned would prepare them for and provide complementary support to in-person counseling with a clinician [20]. Notably, although the definition of *young adult* per the National Cancer Institute (NCI) is defined as an individual aged between 18 and 39 years, our work included women aged 15 to 45 years, as fertility and family-building concerns are highly relevant at somewhat younger and older ages [23,24].

Study Design

The *Roadmap to Parenthood* DA and planning tool (website) was developed by following the steps depicted in Figure 1. The website was designed and tailored considering the shared experiences, emotions, and support needs of young adult female survivors of cancer identified in our previous work. All procedures followed user-centered design methods such that users' needs, contexts, and points of view were key drivers in the iterative design decisions throughout product development [15]. For this stage of the development process, decisions were made with input from patient research partners representing the target user population with the goal of optimizing the prototype design to best prepare for usability testing (which is currently underway). The development process followed the guidelines set forth by Coulter et al [25] and was consistent with the International Patient Decision Aid Society (IPDAS) and the Ottawa Decision Support Framework guidelines for patient DAs [25-28]. For the purposes of building a website, guidelines from the Department of Health and Human Services were followed for best practices of a user-centered web design and digital communication [29].

Figure 1. Steps taken to develop a patient decision aid and planning tool prototype using user-centered design strategies. Adapted from the user-centered design process map from the National Institutes of Health.



Exploratory Work

The research team reviewed and discussed oncofertility patient DAs, web-based oncofertility open access resources, and websites targeting young women such as those focused on women's health and fertility to explore ideas about structure, tonality, and appealing visual identity and design aesthetic for this demographic (Multimedia Appendix 1). The research team then completed a *discovery worksheet* to ensure alignment with the web developer regarding the goals of the project (Multimedia Appendix 2).

Responsive Design Format

We selected a digital format to optimize user access, flexibility, and convenience, aligned with the stated preference of the target user group [20]. Internet use is nearly ubiquitous in the United States among young adults (eg, 97%-100%) with 77% of adults aged 18 to 29 years having home broadband service and 96% owning a smartphone [30-32]. A responsive design website was developed for the decision tool. This choice was made given the flexibility of adapting the layout and content across digital devices and the relative ease and low cost of website updates. A responsive design also provides a consistent user experience regardless of the operating system or device—desktop computer/laptop, tablet, or mobile phone.

Steering Group

A multidisciplinary steering group was assembled, which included clinicians and researchers with expertise in oncofertility and developing patient DAs (ie, oncologists, reproductive endocrinologists, psychologists, and nurses), a digital health communication researcher, an expert in user-centered design and usability testing, and a web developer. The team also included 4 patient research partners, who were asked for advice, provided feedback, and reviewed design decisions and content throughout the ideation phases and the entire development process.

Results

The following steps were taken to develop the prototype website of *Roadmap to Parenthood*. The tool was designed to be used by young adult female survivors of cancer who completed gonadotoxic treatment and were interested in future family building or were uncertain of their family-building plans. The primary purpose of the tool was to educate users about options to achieve parenthood after cancer (ie, natural conception, IVF or surrogacy with fresh/frozen/donated gametes, and adoption or fostering) and to guide value-based decision making and preparatory action toward family-building goals.

Identifying User Needs

The first phase of user-centered design processes involves exploratory work to fully understand and define the problem, comprising literature reviews, end-user interviews and surveys, and team brainstorming [13]. We did much of this work previously, and user needs are described under *preliminary studies.* We also conducted a scoping review of the literature [11,12] and discussed our understanding of user needs with our patient research partners. Common themes (eg, lack of information, uncertainty about reproductive survivorship care options, and a lack of awareness about high costs and legal complications) were reviewed by the research team, which led to brainstorming about how a web-based decision support and planning tool could address user needs (described in the following sections).

Determination of Goals, Format, and Delivery

From the start, the overall objective of the site was to help users become informed, clarify values and priorities with respect to family-building goals, and consider options for actionable preparatory behaviors (eg, pursuing a fertility evaluation or accessing social support). Family-building decision options included the possibility of natural conception and alternative options, that is, IVF, surrogacy, and adoption with subsidiary options (ie, use of fresh, frozen, or donor gametes and domestic or international adoption or fostering). Notably, personalized information about infertility risk and likelihood of success with family-building options could not be provided. Instead, the tool was built to increase awareness of the potential for challenges and benefits of early action and to prompt decisions about pursuing *next steps* aligned with parenthood goals. At the same time, we aimed to create a website that would *feel* empowering and would be usable, engaging, and effective. The tool was designed to be used independently by young adult female survivors of cancer and delivered via internet access using a responsive design format.

Design Process

We designed the DA website using an agile development process, which provided a nimble system for ongoing revision and iterative design decision making based on team review and input from the web developer, usability experts, and patient research partners. Modified beginning stage *design sprints* (ie, a rapid cycle user-centered prototype development and testing process [33]) were undertaken to generate and test ideas, obtain feedback, and iterate features of the prototype. Patient research partners were asked via email and phone/video communication for feedback and recommendations.

Ideation Phase

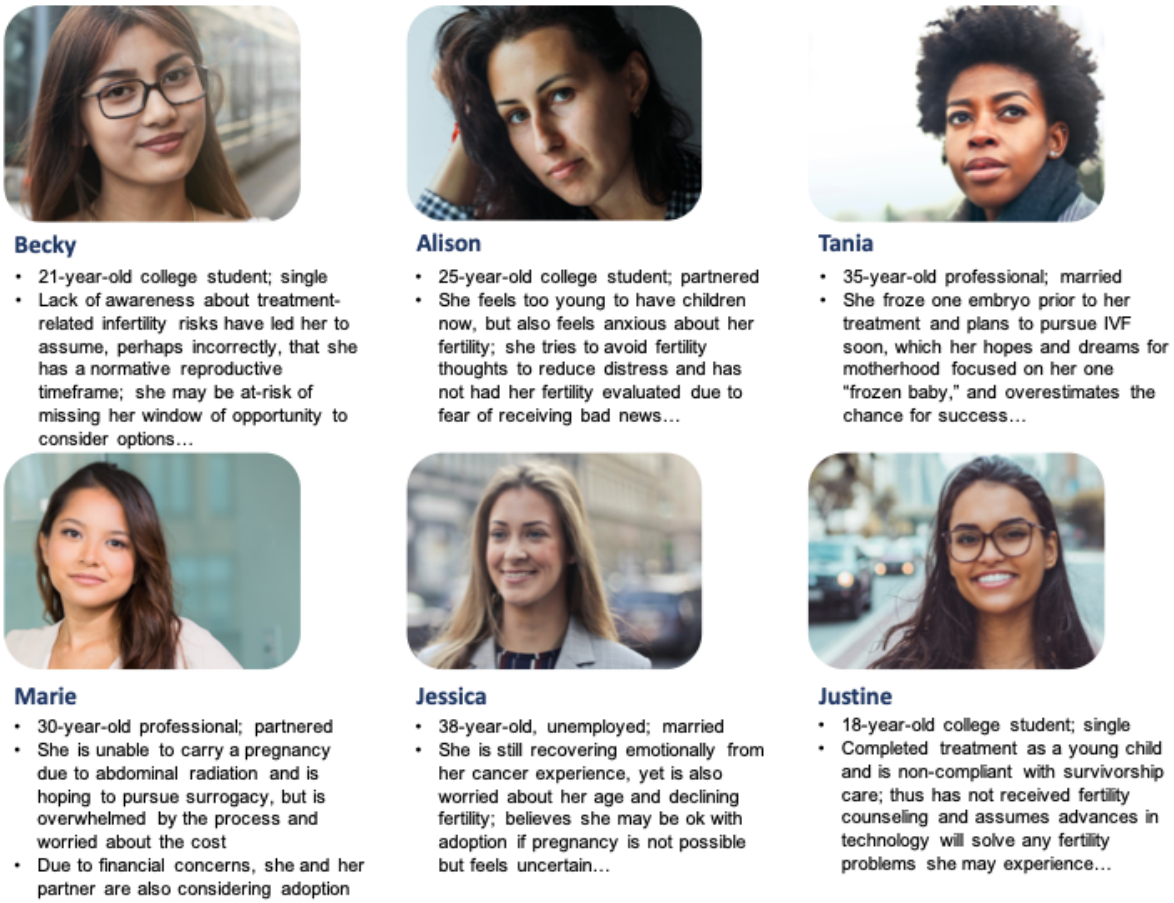
Ultimately, we wanted the website to be empowering for young women by supporting their family-building decision-making processes. A period of research and brainstorming was undertaken by the team with input from the web developer and patient research partners. We reviewed 9 publications reporting on 7 oncofertility patient DAs and were able to access 5 of the DAs available on the web (Multimedia Appendix 1). We looked for IPDAS DA components and descriptions of development processes. We discussed aspects we believed were useful, appealing, and should be considered for our own design and, conversely, aspects that we felt could be improved upon. For example, very dense text and long paragraphs prompted discussions about layout and content organization. It was also our aim to create a website that would feel approachable while conveying trustworthiness and reliability (ie, the *personality* of the website). A review of web-based oncofertility resources and women-targeted websites (Multimedia Appendix 1) focused on the esthetic appeal of designs and tonality. The websites rated most positively were those that felt *clean* and easy to use, with clear text and appealing use of graphics and white space. Fewer favorably reviewed websites included elements that felt stereotypically feminine (eg, *too pink*), content that felt crowded, or pages that had distracting visual designs such as overlaying text on busy backgrounds. Patient research partners were asked for feedback about the likes and dislikes and ideas for an appealing *look and feel* of DA and website esthetics. Tonality across DAs and websites ranged from professional and more business-like to friendly and more conversational. We aimed to strike a balance between friendly and approachable, yet informative and trustworthy.

We completed the *discovery worksheet* to facilitate communication and a shared understanding among team members about the overall objective for the site and initial stylistic ideas (Multimedia Appendix 2). For example, the 4 stylistic descriptors of the ideal website were empowering, informative, friendly, and clean. The web developer used the worksheet and descriptions of our likes/dislikes of the DA/website examples to understand the overall objective for the website and design, combined with user personas that provided further guidance.

Personas and Storyboards

On the basis of our background work [19-22], literature review [34], and clinical experience (Figure 2), 6 personas were created representing *user types*. The personas varied to represent different patient situations based on sociodemographic characteristics, cancer, and reproductive health factors. These factors were used to construct an overall picture of user archetypes and the kinds of questions, concerns, intentions, and goals they would have to interact with the website. Each persona was given a name, sociodemographic descriptors, a cancer story, social context, and description of values, priorities, and goals related to family building. Personas depicted the types of users for which the website was being designed and the scope of user intentions and needs. Decisions about design, features, and navigation were made to meet the needs of all user personas, in combination with the discovery worksheet.

Figure 2. Personas depicting "user types" for the decision aid and planning tool website. Shortened versions of personas are depicted. IVF: in vitro fertilization.



Storyboarding was used in conjunction with persona mapping to envision the end-to-end user experience of someone engaging with the site over time. This process mapped how the website would fit into a user's life as well as in what context they would be seeking out and engaging with the site for the first time and in subsequent viewings. We based our storyboards on the lives of actual patients depicted through the personas, imagining their life experiences leading up to and after viewing the website. For example, a young woman in her early 20s, not yet ready to have children but aware that she may be at risk for fertility problems in the future, may approach the website with curiosity to learn about her options and recommendations for reproductive health as a survivor of cancer. Alternatively, an older woman in her late 30s, who is ready to have a child and fearful that known fertility problems will prevent her from becoming a mother, may approach the site with greater anxiety and fear, concern about reproductive time pressure, and looking for guidance for immediate action and resources.

User Journey Mapping

User journey mapping involved envisioning the different ways in which users might navigate through our website. During the user journey mapping, it is important to consider what the user will be thinking, feeling, and doing as they engage with the website. The personas we created guided our vision for user journey options. For example, given our conceptualization of users having different levels of knowledge and decision-making readiness at the outset, we debated various options for progressing through the website. The goal of user journey mapping was to plan and optimize how users would move through the website, identify gaps in the user experience, and iteratively pivot to correct errors [35].

User Content Control

It was important to design the website giving users control over their user journey with freedom to access web pages that best matched their needs, as opposed to a more rigid user journey with a single preconceived path through content (ie, similar to paper-based resources in which there is only one path to access content by turning pages). Content control is intended to provide users with control over the order, detail, and type of evidence presented. Providing users greater content control is related to improved quality of decision making [36]. Conversely, tailoring content via preset frameworks has been shown to reduce decision-making quality, despite the intention that more personalized information will be delivered to the user [36,37]. On the basis of this research and as depicted in the personas and storyboarding, we sought to give users a greater control over their user journey to explore content and review material relevant to their situations and interests.

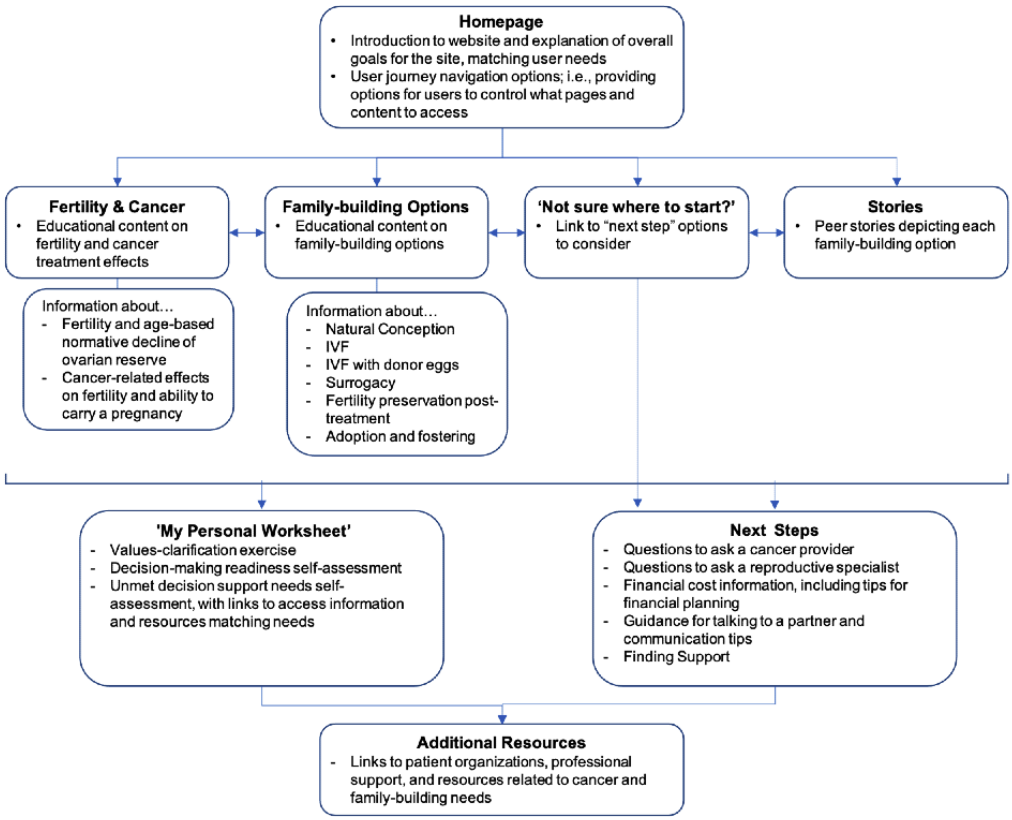
We imagined that some users would need to move linearly through the DA components, starting with education about fertility and cancer treatment effects and moving on to review information about family-building options, whereas others may be quite informed and ready for *next steps* (eg, questions to ask your doctor) and still others may wish to avoid information that

is irrelevant or even upsetting if particular family-building options are no longer possible. We wanted to allow flexibility to navigate through the website to best meet different users' needs and motivations. One idea was to prompt users to answer a set of questions on the homepage with a branching logic to guide them to the best landing page, but this was discarded after initial mockups because of its complexity. Ultimately, we created an omnichannel user journey, or *choose your own adventure* design, in which users had control over the user journey and could easily see options for *next step* recommendations and click on webpages to access content that matched their needs, preferences, and decision-making readiness. The DA components were marked across the top navigation bar. Many pages included links at the bottom that suggested the next pages to visit but these could be ignored, and the user had control over which pages they visited. This was an iterative process designed to match users at different stages of decision making. We identified gaps in our initial design ideas and developed solutions to optimize the user experience.

Information Architecture

Following the development of user journeys, we moved into the information architecture phase of the project (Figure 3). This process involved leveraging the user flows to decide how content should be organized, structured, and labeled across the website pages. The main components of the information architecture process included finalizing decisions about categorizing and structuring information, labeling systems (ie, how information is represented), navigation systems (ie, how users would browse or move through information), and search systems (ie, how users would look for and find information). Various ways of organizing content and implications for the user journey have been discussed and debated. For example, multiple options were considered for how to best introduce and categorize family-building options and how to organize decision support content. As depicted in Figure 3, users had multiple options to move on from the homepage. The *next step* options were grouped together and introduced to users on a single page with links through which they could click for more content on each topic. Decisions about the information architecture guided content strategy and informed the design of the user interface, which was later used in wireframing and prototyping.

Figure 3. Depiction of the information architecture of the decision aid website. IPDAS: International Patient Decision Aid Society.



Sketches and Wireframes

The next step in our website development was to create sketches of our ideas and then wireframes. All appropriate web standards were used to develop content for the site. Our overarching goal with the content was to make it user-friendly and helpful to the reader. We placed special emphasis on using concrete examples that would be highly relevant to the reader. Initial sketches were made by the study team via pen and paper and dry-erase white boards, and ideas were discussed with the web developer. Wireframes (ie, two-dimensional models of the website interface) focused on content presentation and space allocation, functionalities of the site, and intended behaviors of the user [38]. They were used to give the team a sense of how the site would be organized and function once it was fully developed, without focusing on styling, color, and graphics. These digital wireframes allowed us to collect early feedback from collaborators and patient research partners and led to multiple

iterations of website wireframes. Patient research partners were critical at this juncture to clearly understand the purpose of the website, comment on available content, design esthetic upfront, and appropriate navigation that would support their decision-making needs. Once the basic wireframes were set, more detailed illustrations were created using static images to depict the *look and feel* of the site, such as color palettes, font choices, and pictures. They discussed the emotional experiences that users might have when approaching the website, based on the stress and uncertainty of fertility and family-building futures, and cited a desire for the website to feel *calming* and *hopeful*. Wireframes were informed by this feedback and the work done during the *discovery phase*.

Aligning on a Design Aesthetic

It is recommended to use images that end users will find to be realistic and relatable when designing health- and medicine-related content [29]. Research suggests users of digital health tools may prefer photographs of *real* people, as opposed to illustrations or no photographs at all [39]. It is also important to show people of diverse backgrounds, allowing more people to *find themselves* and relate to the content [40]. Accordingly, we included photographic images of young women representing different races, ethnicities, and ages throughout the site. There were no medical or fertility-related photographs. To facilitate understanding, whenever possible, we accompanied written text with graphs (eg, a line graph depicting the decline in ovarian reserve over time) and comparison charts (eg, a table with the costs of family-building options listed side by side). Each family-building option had a different icon to guide the user's journey and comprehension. The research team developed an initial conceptualization of photos, icons, and graphs based on the dual goals of optimizing usability and achieving the desired stylistic feel, and patient research partners were asked for their impressions. Generally, feedback was positive; however, designs were modified based on specific suggestions (eg, to use a different photo or improve labeling). One key issue was how to best depict potentially distressing information. In particular, patient research partners told us that seeing the downward slope of the line graph showing a declining ovarian reserve over time was a powerful and potentially upsetting image. The pros and cons of conveying this information in text or graphs were discussed. Ultimately, with agreement from our patient research partners, we decided to keep the graph for its effectiveness in displaying the critical information. We will test these design decisions and their emotional impact during usability testing.

Content Development

In developing the narrative for the website, guidelines for developing web-based informational content were followed [41]. Writing user-friendly content for a website involves consideration of word choice (eg, use of *plainlanguage* and keywords known to users and an active voice), use of short sentences and paragraphs, chunking content and presenting information in bullets and numbered lists, use of clearly distinguished headlines and subheadings, placement of key informational pieces on the page, and use of white space [41,42]. We followed the plain language checklist for writing website content (Textbox 1; adapted from the checklist developed by the National Institutes of Health [NIH; 43]). Definitions of medical terminology were provided, and simpler medical terms were used whenever possible. For example, the title *ask a fertility specialist* replaced *ask a reproductive endocrinologist*, and the definition for a reproductive endocrinologist was provided for reference.

Textbox 1. Plain language checklist for writing for the web.

1. Be concise; eliminate unnecessary words
2. Break information up into separate topics
3. Use short paragraphs (ie, shorter paragraphs than when writing for printed materials)
4. Use short lists and bullets to organize information
5. Use headings and subheadings that are descriptive, with limited text under each heading
6. Consider using questions as headings
7. Present each topic or point separately
8. Keep the information on each page to no more than two levels
9. Use white space to allow users to easily scan the page for key information
10. Write using the same words users would use when doing a web search for the information, particularly for page titles and headings
11. Clearly explain things such that each page can stand on its own; ie, don't assume users will have knowledge of the subject or have read other content/pages on the site
12. Use language to guide the user journey that describes what the user will get if they click on the link; ie, never use "click here" as a link
13. Eliminate unnecessary words as much as possible

Content was also written to be all-inclusive with respect to user diversity, particularly regarding partnership status (ie, single vs coupled users), sexual orientation (ie, users identifying as lesbian, gay, bisexual, transgender, and queer/questioning [LGBTQ]), and definitions of family makeup (eg, same-sex parents and single women pursuing parenthood). Users were not assumed to have a partner (now or when pursuing family building), and partners were not assumed to be of a specific gender. Listed resources provided access to more detailed information (eg, state-by-state laws regulating LGBTQ and

same-sex couple adoption, legal advocacy, and financial grant opportunities for LGBTQ prospective parents). Religious and cultural factors that may impact users' decision making, particularly with respect to the use of reproductive medicine (eg, transvaginal procedures and creation of embryos in the lab), were addressed in a limited way at several points throughout the website. For example, users are prompted to consider religious, cultural, and ethical beliefs in relation to their decision-making options in the values clarification exercise and can also reflect on these factors when answering open-ended questions.

Additional Design Considerations

Patient research partners discussed the emotional experiences that users might have when approaching the website, based on the stress and uncertainty of fertility and family-building futures, and cited a desire for the website to feel *calming* and *hopeful*. Several design considerations were made to reflect the emotional experiences and health literacy levels of users interacting with the site.

Designing for Iterative Decision Making

Most patient DAs for health care decisions involve discrete periods in which a single decision about treatment options must be made [43]. DAs developed in cancer and fertility have almost exclusively focused on pretreatment fertility preservation in which there are 2 decision options (yes or no) and a limited time window, as cancer treatment must be initiated [12,44]. Conversely, decision making about family building may involve a more complex set of decision points. For example, for some users, the focus of the decision may be about seeing a fertility specialist, and decision options may change based on feedback about reproductive viability and the likelihood of success with natural pregnancy, IVF, or surrogacy, thus changing their informational and support needs. Many survivors may first prioritize having a biologically related child, but if given a low chance for success, they may re-evaluate their priorities and choose to spend financial resources on pursue adoption. Others may restart the decision process if they are unsuccessful, such as after failed IVF attempts, or if the challenges become too great. Single women may change their preferences when they involve a decision partner. On the basis of this conceptualization of longitudinal decision-making processes, the design of the website included support for iterative engagement such that the information architecture was set up to allow users to have maximum control over the user journey and easily circle back to content about alternative family-building options.

Emotional Design

Our previous work suggested that women who experienced more intense emotions of distress and fear often described lower self-efficacy to manage risks and, at the most extreme, avoidance of fertility information and disengagement from decision-making processes [19]. These findings are consistent with the research delineating the impact of affective states on medical decision making and behavior, such that anxiety and fear tend to lead individuals to prioritize short-term gratification over long-term goals [45,46]. In this case, young women who are distressed about infertility risks or fear of receiving bad news may avoid information to avoid further distress (thus prioritizing short-term relief), diminishing their chances of achieving long-term goals for parenthood. One of the objectives for the website was to be empowering for young women, for example, to guide users in becoming informed and setting realistic expectations about potential challenges, while inspiring hope and optimism that parenthood may be achieved. With consultation and input from experts on the team, we aimed to achieve this emotional experience for users through design decisions about tonality, color, language, and picture selection. Acknowledging that information on the website may be upsetting for users, we made decisions about design and photo images to create a positive user experience (eg, facial expressions of women in photos that suggest confidence, hope, and optimism), without negating the difficulties and negative emotions users may experience as a part of this journey. We also used color and design to facilitate comprehension and guide engagement with the site. We attempted to avoid design elements that might overwhelm users, perhaps leading them to abandon the website. For example, large blocks of text can be difficult to read and comprehend, which may be even more challenging for cancer survivors with lasting treatment side effects such as fatigue or cognitive impairment [47], and our patient research partners corroborated concerns about information overload and text-heavy pages. Efforts to reduce the cognitive load and emotional distress included using short text blocks, white space, clearly identified and defined terminology, and graphs, charts, and icons.

Guidelines and Standards

The website was designed to meet varying health literacy and reading levels of users and in accordance with the IPDAS guidelines and the Ottawa Decision Support Framework for developing patient DAs [25,26,48], and health literacy guidelines, including those set by the NIH [49,50]. The Centers for Medicaid and Medicare Services *toolkit for making written material clear and effective* was also employed [51]. The IPDAS checklist is presented in Table 1, whereas the Health Literacy Online Strategies Checklist is presented in Table 2. Standards required for the design and development of websites affiliated with the US Department of Health and Human Services were also reviewed and used to guide design decisions [52]. Guidelines from the Office of Disease Prevention and Health Promotion [49], NIH [50], and the Centers for Medicare and Medicaid Services [51] for designing digital health websites and information tools for low health literacy and culturally diverse populations were also followed.

Table 1. Review of the Roadmap to Parenthood patient decision aid using the International Patient Decision Aid Standards quality checklist.

Criteria	Answer
Criteria to be defined as a patient DA[a]	
1. The DA describes the condition related to the decision	Yes
2. The DA describes the decision that needs to be considered	Yes
3. The DA identifies the target audience	Yes
4. The DA lists the options (health care or other)	Yes
5. The DA has information about the positive features of the options (eg, benefits or advantages)	Yes
6. The DA has information about the negative features of the options (eg, harms, side effects, or disadvantages)	Yes
7. The DA helps patients clarify their values for outcomes of options by (a) asking people to think about which positive and negative features of the options matter most to them AND/OR (b) describing each option to help patients imagine the physical, social, and/or psychological effects[b]	Yes
8. The DA makes it possible to compare the positive and negative features of the available options	Yes
9. The DA shows the negative and positive features of the options with equal detail	Yes
10. The DA compares probabilities (eg, chance of a disease, benefit, harm, or side effect) of options using the same denominator	N/A[c,d]
11. The DA (or available technical documents) reports funding sources for development	Yes
12. The DA reports whether authors of the DA or their affiliations stand to gain or lose by choices people make after using the DA	Yes
13. The DA includes authors/developers' credentials or qualifications	Yes
14. The DA reports the date when it was last updated	Yes
15. The DA (or available technical document) reports readability levels	Yes
16. The DA provides references to scientific evidence used	Yes
Other criteria for Das about screening or testing	
17. The DA has information about what the test is designed to measure	Yes[e]
18. The DA describes possible next steps based on the test results	Yes[e]
19. The DA has information about the chances of disease being found with and without screening	Yes[e]
20. The DA has information about detection and treatment of disease that would never have caused problems if screening had not been done	Yes[e]
Other criteria indicating quality	
21. The DA describes what happens in the natural course of the condition (health or other) if no action is taken	Yes
22. The DA has information about the procedures involved (eg, what is done before, during, and after the health care option)	Yes
23 The information about outcomes of options (positive and negative) includes the changes that may happen	N/A[d]
24. The DA presents probabilities using event rates in a defined group of people for a specified time	N/A[d]
25. The DA compares probabilities of options over the same period of time	N/A[d]
26. The DA uses the same scales in diagrams comparing options	Yes
27. Users (people who previously faced the decision) were asked what they need to prepare them to discuss a specific decision	Yes
28. The DA was reviewed by people who previously faced the decision and were not involved in the DA's development and field testing	Yes
29. People who were facing the decision field tested the DA	No[f]
30. Field testing showed that the DA was acceptable to users (the general public and practitioners)	No[f]
31. Field testing showed that people who were undecided felt that the information was presented in a balanced way	No[f]
32. There is evidence that the DA (or one based on the same template) helps people know about the available options and their features	N/A[g]

Criteria	Answer
33. There is evidence that the DA (or one based on the same template) improves the match between the features that matter most to the informed person and the option that is chosen	N/A[g]

[a]DA: decision aid.

[b]We expanded this definition to also include financial effects of decision option outcomes.

[c]N/A: not applicable.

[d]The primary purpose of the decision aid is to educate and support patients facing limited family-building options in which it is not possible to predict the likelihood of success or failure with in vitro fertilization, surrogacy, and adoption.

[e]The decision aid and planning tool provides information on infertility risk due to gonadotoxic cancer treatment, options to test fertility, and possible next steps for family building based on the results of fertility testing and indications of reproductive potential; however, this is only one aspect of the entire decision-making process encompassing family building after cancer.

[f]This criterion was not yet relevant, as the patient decision aid was still in development. Usability and field testing are currently underway.

[g]Study of efficacy will begin after the completion of usability and field testing and once the design of the decision aid is finalized.

Table 2. Review of the Roadmap to Parenthood patient decision aid using the Health Literacy Online Strategies Checklist from the National Institutes of Health.

Criteria	Answer
Write actionable content	
1. Identify user motivations and goals	Yes
2. Put the most important information first	Yes
3. Describe the health behavior [information] – just the basics	Yes
4. Stay positive. Include the benefits of taking action	Yes
5. Provide specific action steps	Yes
6. Write in plain language	Yes
7. Check content for accuracy	Yes
Display content clearly on the page	
8. Limit paragraph size. Use bullets and short lists	Yes
9. Use meaningful headings	Yes
10. Use readable font	Yes
11. Use white space and avoid clutter	Yes
12. Keep the most important content above the fold – even on mobile	Yes
13. Use links effectively	Yes
14. Use color or underline to identify links	Yes
15. Use Images that help people learn	Yes
16. Use appropriate contrast	Yes
17. Make web content printer-friendly	Yes
18. Make your site accessible to people with disabilities	Yes[a]
19. Make websites responsive	Yes
20. Design mobile content to meet mobile users' needs	Yes
Organize content and simplify navigation	
21. Create a simple and engaging homepage	Yes
22. Label and organize content with your users in mind	Yes
23. Create linear information paths	Yes[b]
24. Give buttons meaningful labels	Yes
25. Make clickable elements recognizable	Yes
26. Make sure the browser "back" button works	Yes
27. Provide easy access to home and menu pages	Yes
28. Give users options to browse	Yes
29. Include a simple search function	No[c]
30. Display search results clearly	No[c]
Engage users	
31. Share information through multimedia	No
32. Design intuitive interactive graphics and tools	Yes
33. Provide tailored information	Yes[d]
34. Create user-friendly forms and quizzes	Yes
35. Consider social media sharing options	N/A[e,f]
Test your site with users with limited literacy skills	

Criteria	Answer
36. Recruit users with limited literacy skills and limited health literacy skills	No[f]
37. Identify and eliminate logistical barriers to participation	Yes
38. Create plain language testing materials	Yes
39. Test whether your content is understandable and actionable	No[f]
40. Use moderators who have experience with users with limited literacy skills	N/A[g]

[a]Design decisions were made within the scope of the project to make content accessible to people with disabilities, including using large font and white space.

[b]Beginning sections of the decision aid tool were designed so that users would follow a linear path to obtain information about fertility, infertility risks associated with cancer, and family-building options. Following this, users were prompted to choose their own path with respect to which content was most applicable to their situation and needs (eg, finding support vs financial planning strategies).

[c]We were unable to include a search function due to the limitations of web development resources.

[d]Information was tailored to the extent that users had control over content they viewed because of the choose your own path click-through user journey design of the website, use of a drawer design to hide/reveal content based on user interest, and available drop-down features.

[e]N/A: not applicable.

[f]This criterion was not yet relevant, as the patient decision aid was still in development.

[g]The tool is designed to be used by young adult female cancer survivors independently, without help from moderators, clinicians, or professionals to provide guidance or decision support. Future work will explore options for use of the tool during patient-provider interactions and for shared decision making.

Discussion

User-centered design practices involve an extensive and iterative process of designing and developing solutions from the perspective of the end user. The development of the *Roadmap to Parenthood* prototype was based on pilot work to understand the experiences and needs of young adult female survivors of cancer related to family building after cancer, combined with a collaborative, multidisciplinary team approach to making initial design decisions that would best meet their needs. Ultimately, usability testing with members of the target patient population is necessary to determine whether design decisions have been made to optimize the ease of use, comprehension, and usefulness of the product or whether improvements are needed. Once completed, we hope that this DA for family building after cancer will be a complementary resource to the DAs and resources that exist for pretreatment fertility preservation [11,12]. We followed similar development procedures to those reported for other DAs in the literature, including adherence to IPDAS guidelines, review of previously published DAs, use of a multidisciplinary team approach, iterative design with feedback from target users, and digital platforms [53-57]. Fertility preservation DAs have been shown to be acceptable and beneficial to young adult female cancer survivors [53,55,57], suggesting the DA presented here may similarly support patients through the next steps of family building.

The early design considerations presented here are important steps for developing a user-centered prototype that is a good starting point for usability testing. We are currently underway in conducting usability testing with target end users to obtain feedback about the website prototype. This process involves quantitative and qualitative data analytic approaches using standardized usability testing procedures including think-aloud sessions and validated usability assessment measures. Design modifications and additional testing will be conducted until user feedback indicates that we have optimized the product design with the appropriate degree of compassion and empathy. Following usability testing, we will conduct a single-arm pilot study to test the tool as a DA intervention for family building after cancer [58]. For this study, we will follow the SUNDAE Checklist (Standards for UNiversal reporting of patient Decision Aid Evaluations) for reporting patient DA evaluative studies [48]. Future directions of this research will also explore how the tool may be used for dyadic decision making, including users' partners, and as a part of cancer survivorship care to support patient-provider communication and shared decision making.

Limitations and Future Directions

There were limitations to this study. One of our most difficult tasks in creating the website was to balance the amount of informational content with concerns about *information overload*, an issue that was brought up by our patient research partners. Although the website provides a comprehensive overview of fertility and family-building topics and multiple aspects of decision-making support, some subsidiary topics were not as thoroughly reviewed. For example, while users are prompted to self-reflect and explore personal factors that are most relevant to their decisions, in-depth content specific to cultural and religious factors was limited. Similarly, this version of the website does not mention step-motherhood as a family-building option. Future usability testing will determine whether more comprehensive information is needed on these topics. In addition, while we decided to build a web-based resource to increase access and convenience among the target user group of young adult women who reported a preference for digital resources, we recognize that some members of the target population may not have regular or dependable access to the internet. However, we do not believe this is a widespread issue, based on the data gathered on internet accessibility across various demographic cohorts [30-32]. In order to accommodate users who may prefer a paper-based version, we included a print

button on the top of each page that autoformats the website content for the ease of printing.

Conclusions

Following this first phase of the development process of a patient DA and planning tool for family building after cancer, our subsequent usability testing phase will guide modifications and finalization of the design. This clinical research will contribute to a priority area set forth by the NCI and the American Society of Clinical Oncology to develop age-specific resources for young adult cancer survivors while leveraging the advantages of digital communication technology [59,60].

Acknowledgments

Research reported in this publication was supported by the NCI of the NIH under award number K07CA229186. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. The authors would also like to thank their patient research partners and young adult cancer patient organizations for their collaboration (Stupid Cancer, Lacuna Loft, The Samfund, and Gryt Health).

Conflicts of Interest

None declared.

Multimedia Appendix 1
List of oncofertility patient decision aids, online resources, and websites reviewed during the ideation phase of website development.
[DOCX File , 33 KB - formative_v5i1e20841_app1.docx]

Multimedia Appendix 2
Discovery worksheet for initial website design.
[DOCX File , 14 KB - formative_v5i1e20841_app2.docx]

References

1. Institute of Medicine Committee on Quality of Health Care in America. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: National Academies Press; 2001.
2. Rathert C, Wyrwich MD, Boren SA. Patient-centered care and outcomes: a systematic review of the literature. Med Care Res Rev 2013 Aug;70(4):351-379. [doi: 10.1177/1077558712465774] [Medline: 23169897]
3. Gorman JR, Bailey S, Pierce JP, Su HI. How do you feel about fertility and parenthood? The voices of young female cancer survivors. J Cancer Surviv 2012 Jun;6(2):200-209 [FREE Full text] [doi: 10.1007/s11764-011-0211-9] [Medline: 22179785]
4. Armuand GM, Wettergren L, Rodriguez-Wallberg KA, Lampic C. Women more vulnerable than men when facing risk for treatment-induced infertility: a qualitative study of young adults newly diagnosed with cancer. Acta Oncol 2015 Feb;54(2):243-252. [doi: 10.3109/0284186X.2014.948573] [Medline: 25140859]
5. Gargus E, Deans R, Anazodo A, Woodruff TK. Management of primary ovarian insufficiency symptoms in survivors of childhood and adolescent cancer. J Natl Compr Canc Netw 2018 Sep;16(9):1137-1149 [FREE Full text] [doi: 10.6004/jnccn.2018.7023] [Medline: 30181423]
6. Benedict C, Thom B, N Friedman D, Diotallevi D, M Pottenger E, J Raghunathan N, et al. Young adult female cancer survivors' unmet information needs and reproductive concerns contribute to decisional conflict regarding posttreatment fertility preservation. Cancer 2016 Jul 1;122(13):2101-2109 [FREE Full text] [doi: 10.1002/cncr.29917] [Medline: 27213483]
7. Murphy D, Klosky JL, Reed DR, Termuhlen AM, Shannon SV, Quinn GP. The importance of assessing priorities of reproductive health concerns among adolescent and young adult patients with cancer. Cancer 2015 Aug 1;121(15):2529-2536 [FREE Full text] [doi: 10.1002/cncr.29466] [Medline: 26054052]
8. Gorman JR, Drizin JH, Smith E, Flores-Sanchez Y, Harvey SM. Patient-centered communication to address young adult breast cancer survivors' reproductive and sexual health concerns. Health Commun 2020 Jul 23:1-16. [doi: 10.1080/10410236.2020.1794550] [Medline: 32703034]
9. Stacey D, Kryworuchko J, Belkora J, Davison BJ, Durand M, Eden KB, et al. Coaching and guidance with patient decision aids: a review of theoretical and empirical evidence. BMC Med Inform Decis Mak 2013;13(Suppl 2):S11 [FREE Full text] [doi: 10.1186/1472-6947-13-S2-S11] [Medline: 24624995]
10. Hoffman AS, Volk RJ, Saarimaki A, Stirling C, Li LC, Härter M, et al. Delivering patient decision aids on the internet: definitions, theories, current evidence, and emerging research areas. BMC Med Inform Decis Mak 2013;13(Suppl 2):S13 [FREE Full text] [doi: 10.1186/1472-6947-13-S2-S13] [Medline: 24625064]
11. Speller B, Sissons A, Daly C, Facey M, Kennedy E, Metcalfe K, et al. An evaluation of oncofertility decision support resources among breast cancer patients and health care providers. BMC Health Serv Res 2019 Feb 6;19(1):101 [FREE Full text] [doi: 10.1186/s12913-019-3901-z] [Medline: 30728004]

12. Wang Y, Anazodo A, Logan S. Systematic review of fertility preservation patient decision aids for cancer patients. Psychooncology 2019 Mar;28(3):459-467. [doi: 10.1002/pon.4961] [Medline: 30523651]
13. Improving the User Experience. US Department of Health & Human Services. 2013. URL: https://www.usability.gov/ [accessed 2019-05-24]
14. Slater H, Campbell JM, Stinson JN, Burley MM, Briggs AM. End user and implementer experiences of mhealth technologies for noncommunicable chronic disease management in young adults: systematic review. J Med Internet Res 2017 Dec 12;19(12):e406 [FREE Full text] [doi: 10.2196/jmir.8888] [Medline: 29233804]
15. Harst L, Lantzsch H, Scheibe M. Theories predicting end-user acceptance of telemedicine use: systematic review. J Med Internet Res 2019 May 21;21(5):e13117 [FREE Full text] [doi: 10.2196/13117] [Medline: 31115340]
16. Darlow S, Wen K. Development testing of mobile health interventions for cancer patient self-management: a review. Health Informatics J 2016 Sep;22(3):633-650 [FREE Full text] [doi: 10.1177/1460458215577994] [Medline: 25916831]
17. Leventhal H, Phillips LA, Burns E. The Common-Sense Model of Self-Regulation (CSM): a dynamic framework for understanding illness self-management. J Behav Med 2016 Dec;39(6):935-946. [doi: 10.1007/s10865-016-9782-2] [Medline: 27515801]
18. Diefenbach MA, Leventhal H. The common-sense model of illness representation: theoretical and practical considerations. J Social Distress Homeless 2016 Jul 7;5(1):11-38. [doi: 10.1007/bf02090456]
19. Benedict C, Hahn AL, McCready A, Kelvin JF, Diefenbach M, Ford JS. Toward a theoretical understanding of young female cancer survivors' decision-making about family-building post-treatment. Support Care Cancer 2020 Oct;28(10):4857-4867. [doi: 10.1007/s00520-020-05307-1] [Medline: 31993754]
20. Benedict C, Nieh J, Hahn AL, McCready A, Diefenbach M, Ford JS. Looking at future cancer survivors, give them a roadmap: addressing fertility and family-building topics in post-treatment cancer survivorship care. Support Care Cancer 2020 Sep 5:- epub ahead of print. [doi: 10.1007/s00520-020-05731-3] [Medline: 32889581]
21. Benedict C, Shuk E, Ford JS. Fertility issues in adolescent and young adult cancer survivors. J Adolesc Young Adult Oncol 2016 Mar;5(1):48-57 [FREE Full text] [doi: 10.1089/jayao.2015.0024] [Medline: 26812452]
22. Benedict C, McLeggon J, Thom B, Kelvin JF, Landwehr M, Watson S, et al. 'Creating a family after battling cancer is exhausting and maddening': Exploring real-world experiences of young adult cancer survivors seeking financial assistance for family building after treatment. Psychooncology 2018 Dec;27(12):2829-2839 [FREE Full text] [doi: 10.1002/pon.4898] [Medline: 30238545]
23. Halliday LE, Boughton MA. Exploring the concept of uncertain fertility, reproduction and motherhood after cancer in young adult women. Nurs Inq 2011 Jun;18(2):135-142. [doi: 10.1111/j.1440-1800.2011.00532.x] [Medline: 21564394]
24. Mac Dougall K, Beyene Y, Nachtigall RD. Age shock: misperceptions of the impact of age on fertility before and after IVF in women who conceived after age 40. Hum Reprod 2013 Feb;28(2):350-356 [FREE Full text] [doi: 10.1093/humrep/des409] [Medline: 23203214]
25. Coulter A, Stilwell D, Kryworuchko J, Mullen PD, Ng CJ, van der Weijden T. A systematic development process for patient decision aids. BMC Med Inform Decis Mak 2013;13(Suppl 2):S2 [FREE Full text] [doi: 10.1186/1472-6947-13-S2-S2] [Medline: 24625093]
26. Decision Aid Toolkit - Patient Decision Aids. Ottawa Hospital Research Institute. URL: https://decisionaid.ohri.ca/resources.html [accessed 2019-06-04]
27. Elwyn G, O'Connor A, Stacey D, Volk R, Edwards A, Coulter A, International Patient Decision Aids Standards (IPDAS) Collaboration. Developing a quality criteria framework for patient decision aids: online international Delphi consensus process. Br Med J 2006 Aug 26;333(7565):417 [FREE Full text] [doi: 10.1136/bmj.38926.629329.AE] [Medline: 16908462]
28. Durand M, Witt J, Joseph-Williams N, Newcombe RG, Politi MC, Sivell S, et al. Minimum standards for the certification of patient decision support interventions: feasibility and application. Patient Educ Couns 2015 Apr;98(4):462-468. [doi: 10.1016/j.pec.2014.12.009] [Medline: 25577469]
29. HHS Web Standards Usability Guidelines. US Department of Health and Human Services. URL: https://www.hhs.gov/ [accessed 2019-05-24]
30. Demographics of Internet and Home Broadband Usage in the United States. Pew Research Center. 2019. URL: https://www.pewresearch.org/internet/fact-sheet/internet-broadband/ [accessed 2020-08-13]
31. Demographics of Mobile Device Ownership and Adoption in the United States. Pew Research Center. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2019-12-04]
32. Pew Research Center. URL: https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/ [accessed 2019-11-24]
33. Banfield R, Lombardo C, Wax T. Design Sprint: A Practical Guidebook for Building Great Digital Products. Canada: O'Reily Media, Inc; 2015.
34. Logan S, Perz J, Ussher JM, Peate M, Anazodo A. A systematic review of patient oncofertility support needs in reproductive cancer patients aged 14 to 45 years of age. Psychooncology 2018 Feb;27(2):401-409. [doi: 10.1002/pon.4502] [Medline: 28734119]
35. Journey Mapping 101. Nielsen Norman Group. URL: https://www.nngroup.com/articles/journey-mapping-101/ [accessed 2020-10-15]

36. Syrowatka A, Krömker D, Meguerditchian AN, Tamblyn R. Features of computer-based decision aids: systematic review, thematic synthesis, and meta-analyses. J Med Internet Res 2016 Jan 26;18(1):e20 [FREE Full text] [doi: 10.2196/jmir.4982] [Medline: 26813512]
37. Scherer LD, Ubel PA, McClure J, Greene SM, Alford SH, Holtzman L, et al. Belief in numbers: When and why women disbelieve tailored breast cancer risk statistics. Patient Educ Couns 2013 Aug;92(2):253-259 [FREE Full text] [doi: 10.1016/j.pec.2013.03.016] [Medline: 23623330]
38. Improving the User Experience: Wireframing. Department of HealthHuman Services Digital Communications Division. URL: https://www.usability.gov/how-to-and-tools/methods/wireframing.html [accessed 2019-08-07]
39. Cyr, Head, Larios, Pan. Exploring human images in website design: a multi-method approach. MIS Q 2009;33(3):539. [doi: 10.2307/20650308]
40. Health Literacy Online: A Guide to Simplifying the User Experience. Office of Disease Prevention and Health Promotion (ODPHP). 2016. URL: https://health.gov/healthliteracyonline/ [accessed 2019-08-07]
41. Improving the User Experience: Writing for the Web. Office of Disease Prevention and Health Promotion (ODPHP). URL: https://www.usability.gov/how-to-and-tools/methods/writing-for-the-web.html [accessed 2019-08-07]
42. How to Present Scientific Findings Online. Nielsen Norman Group. URL: https://www.nngroup.com/articles/scientific-findings-online/ [accessed 2019-08-07]
43. Stacey D, Légaré F, Lewis K, Barry MJ, Bennett CL, Eden KB, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2017 Apr 12;4:CD001431 [FREE Full text] [doi: 10.1002/14651858.CD001431.pub5] [Medline: 28402085]
44. Speller B, Micic S, Daly C, Pi L, Little T, Baxter NN. Oncofertility decision support resources for women of reproductive age: systematic review. JMIR Cancer 2019 Jun 6;5(1):e12593 [FREE Full text] [doi: 10.2196/12593] [Medline: 31199289]
45. Loewenstein G. Hot-cold empathy gaps and medical decision making. Health Psychol 2005 Jul;24(4S):S49-S56. [doi: 10.1037/0278-6133.24.4.S49] [Medline: 16045419]
46. Loewenstein GF, Weber EU, Hsee CK, Welch N. Risk as feelings. Psychol Bull 2001 Mar;127(2):267-286. [doi: 10.1037/0033-2909.127.2.267] [Medline: 11316014]
47. Micaux Obol C, Lampic C, Wettergren L, Ljungman L, Eriksson LE. Experiences of a web-based psycho-educational intervention targeting sexual dysfunction and fertility distress in young adults with cancer-A self-determination theory perspective. PLoS One 2020;15(7):e0236180 [FREE Full text] [doi: 10.1371/journal.pone.0236180] [Medline: 32697801]
48. Sepucha KR, Abhyankar P, Hoffman AS, Bekker HL, LeBlanc A, Levin CA, et al. Standards for UNiversal reporting of patient Decision Aid Evaluation studies: the development of SUNDAE Checklist. BMJ Qual Saf 2018 May;27(5):380-388 [FREE Full text] [doi: 10.1136/bmjqs-2017-006986] [Medline: 29269567]
49. Hou S. Health literacy online: a guide to writing and designing easy-to-use health web sites. Health Promot Pract 2012 Sep;13(5):577-580. [doi: 10.1177/1524839912446480] [Medline: 22763891]
50. Health Literacy. National Institutes of Health (NIH). 2015. URL: https://www.nih.gov/institutes-nih/nih-office-director/office-communications-public-liaison/clear-communication/health-literacy [accessed 2019-05-24]
51. Toolkit for Making Written Material Clear and Effective. Centers for Medicare Medicaid Services. 2012. URL: https://www.cms.gov/Outreach-and-Education/Outreach/WrittenMaterialsToolkit/index.html [accessed 2019-05-24]
52. US Department of Health and Human Services. URL: https://www.hhs.gov/ [accessed 2019-08-07]
53. Woodard TL, Hoffman AS, Covarrubias LA, Holman D, Schover L, Bradford A, et al. The Pathways fertility preservation decision aid website for women with cancer: development and field testing. J Cancer Surviv 2018 Feb;12(1):101-114. [doi: 10.1007/s11764-017-0649-5] [Medline: 29034438]
54. Gonçalves V, Travado L, Ferreira PL, Quinn G. Protocol for the development and acceptability of a fertility-related decision aid for young women with breast cancer in Portugal. BMJ Open 2019 Jul 24;9(7):e030690 [FREE Full text] [doi: 10.1136/bmjopen-2019-030690] [Medline: 31345986]
55. Ehrbar V, Urech C, Rochlitz C, Dällenbach RZ, Moffat R, Stiller R, et al. Fertility preservation in young female cancer patients: development and pilot testing of an online decision aid. J Adolesc Young Adult Oncol 2018 Feb;7(1):30-36. [doi: 10.1089/jayao.2017.0047] [Medline: 28759303]
56. Benoit A, Grynberg M, Morello R, Sermondade N, Grandazzi G, Moutel G. Does a web-based decision aid improve informed choice for fertility preservation in women with breast cancer (DECISIF)? Study protocol for a randomised controlled trial. BMJ Open 2020 Feb 10;10(2):e031739 [FREE Full text] [doi: 10.1136/bmjopen-2019-031739] [Medline: 32047010]
57. Peate M, Meiser B, Friedlander M, Saunders C, Martinello R, Wakefield CE, et al. Development and pilot testing of a fertility decision aid for young women diagnosed with early breast cancer. Breast J 2011;17(1):112-114. [doi: 10.1111/j.1524-4741.2010.01033.x] [Medline: 21129095]
58. Benedict C, Ford JS, Schapira L, Simon P, Spiegel D, Diefenbach M. Family-building decision aid and planning tool for young adult women after cancer treatment: protocol for preliminary testing of a web-based decision support intervention in a single-arm pilot study. BMJ Open 2019 Dec 29;9(12):e033630 [FREE Full text] [doi: 10.1136/bmjopen-2019-033630] [Medline: 31888941]

59. McCabe MS, Bhatia S, Oeffinger KC, Reaman GH, Tyne C, Wollins DS, et al. American Society of Clinical Oncology statement: achieving high-quality cancer survivorship care. J Clin Oncol 2013 Feb 10;31(5):631-640 [FREE Full text] [doi: 10.1200/JCO.2012.46.6854] [Medline: 23295805]
60. Smith AW, Seibel NL, Lewis DR, Albritton KH, Blair DF, Blanke CD, et al. Next steps for adolescent and young adult oncology workshop: an update on progress and recommendations for the future. Cancer 2016 Apr 1;122(7):988-999 [FREE Full text] [doi: 10.1002/cncr.29870] [Medline: 26849003]

Abbreviations

ASCO: American Society of Clinical Oncology
DA: decision aid
IPDAS: International Patient Decision Aid Society
IVF: in vitro fertilization
LGBTQ: lesbian, gay, bisexual, transgender, and queer/questioning
NCI: National Cancer Institute
NIH: National Institutes of Health

Edited by G Eysenbach; submitted 05.06.20; peer-reviewed by A Bradford, A Hoffman; comments to author 15.07.20; revised version received 29.10.20; accepted 15.11.20; published 22.01.21.

Please cite as:
Benedict C, Dauber-Decker KL, King D, Hahn A, Ford JS, Diefenbach M
A Decision Aid Intervention for Family Building After Cancer: Developmental Study on the Initial Steps to Consider When Designing a Web-Based Prototype
JMIR Form Res 2021;5(1):e20841
URL: http://formative.jmir.org/2021/1/e20841/
doi:10.2196/20841
PMID:33480848

Original Paper

A Mobile Health Platform for Self-Management of Pediatric Cystic Fibrosis: Qualitative Study of Adaptation to Stakeholder Needs and Integration in Clinical Settings

Sarah B Rutland[1], PhD; Rikard Palmer Bergquist[2], MSc; Andreas Hager[3], LLM; Robin Geurs[1], BSc; Cathy Mims[4], RN, MSHQS; Hector H Gutierrez[1], MD; Gabriela R Oates[1], PhD

[1]Pediatric Pulmonary and Sleep Medicine, The University of Alabama at Birmingham, Birmingham, AL, United States

[2]Motivo Management LLC, Reno, NV, United States

[3]Upstream Dream AB, Stockholm, Sweden

[4]Children's of Alabama, Birmingham, AL, United States

Corresponding Author:
Gabriela R Oates, PhD
Pediatric Pulmonary and Sleep Medicine
The University of Alabama at Birmingham
Lowder 620
1600 7th Ave S
Birmingham, AL, 35233
United States
Phone: 1 2056389583
Email: goates@uab.edu

Abstract

Background: Cystic fibrosis (CF) is an inherited chronic condition that requires extensive daily care and quarterly clinic visits with a multidisciplinary care team. The limited exchange of information outside of the quarterly clinic visits impedes optimal disease self-management, patient engagement, and shared decision making.

Objective: The aim of this study is to adapt a mobile health (mHealth) app originally developed in Sweden to the needs of patients, families, and health care providers in a CF center in the United States and to test it as a platform for sharing patient-generated health data with the CF health care team.

Methods: Focus groups with health care providers of patients with CF, adolescents with CF, and caregivers of children with CF were conducted to determine what modifications were necessary. Focus group data were analyzed using a thematic analysis, and emergent themes were ranked according to desirability and technical feasibility. The mHealth platform was then modified to meet the identified needs and preferences, and the flow of patient-generated health data to a secure Research Electronic Data Capture database was tested. Protocols for data management and clinical follow-up were also developed.

Results: A total of 5 focus groups with 21 participants were conducted. Recommended modifications pertained to all functionalities of the mHealth platform, including tracking of symptoms, treatments, and activities of daily care; creating and organizing medication lists and setting up reminders; generating reports for the health care team; language and presentation; sharing and privacy; and settings and accounts. Overall, health care providers recommended changes to align the mHealth platform with US standards of care, people with CF and their caregivers requested more tracking functionalities, and both groups suggested the inclusion of a mental health tracker as well as more detailed response options and precise language. Beta testers of the modified platform reported issues related to translatability to US environment and various bugs.

Conclusions: This study demonstrated the importance of identifying the needs and preferences of target users and stakeholders before adopting existing mHealth solutions. All relevant perspectives, including those of clinicians, patients, and caregivers, should be thoroughly considered to meet both end users' needs and evidence-based practice recommendations.

(JMIR Form Res 2021;5(1):e19413) doi:10.2196/19413

KEYWORDS

cystic fibrosis; mHealth

Introduction

Background

Cystic fibrosis (CF) is the second most common inherited disorder in the United States. The disease affects multiple systems, including the respiratory, digestive, endocrine, and reproductive systems, and involves a complex, time-consuming daily care routine with multiple oral and inhaled medications, airway clearance therapy, diet, and exercise [1,2]. In the United States, clinical care guidelines recommend quarterly multidisciplinary CF clinic visits. During such visits, patients and families are asked to remember and communicate to the clinical team the most relevant aspects of their disease experience from the previous 3 months. From that snapshot, clinicians are expected to gather enough detail to make optimal treatment recommendations [3]. Although the majority of care takes place at home, exchange of information between the patient and the clinical team outside of the clinical setting is currently limited and occurs mostly via phone calls. This traditional, episodic model of care delivery is not well suited for a chronic condition such as CF and does not support optimal disease self-management, patient engagement, and shared decision making [3-6]. For optimal CF care, a bidirectional patient-clinician communication that takes place between visits is necessary. Such communication is particularly important for children and adolescents who are learning to transition from clinician-regulated care to autonomous self-management of their disease [7,8]. Ideally, the bidirectional communication would occur with electronic tracking of patient-reported symptoms, real-time sharing of this patient-generated health data controlled by the patient but matching the requirements of clinical workflows, and timely clinical feedback to such data [3,9].

Research in other complex chronic conditions has shown that patients benefit from routine collection of patient-generated health data. Studies have reported associations between the use of patient-generated health data and improved symptom control and quality of life [10-14], patient-clinician communication and satisfaction [15,16], reduced health care utilization, and increased survival [17-20]. Such an approach is not yet used in CF clinical practice, despite a strongly stated need by the CF patient community [21-23]. A previous assessment of preferences for remote collection and sharing of patient-generated health data among patients with CF or their caregivers reported that the majority are willing to share such data [24].

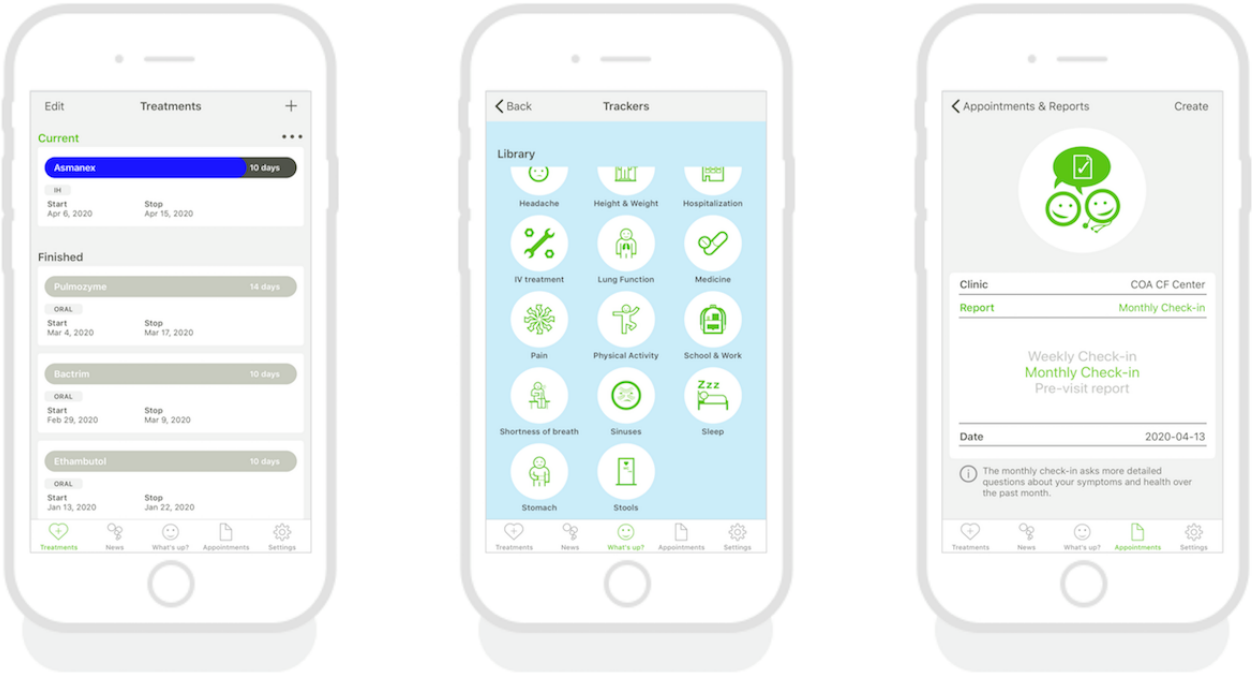
The opportunities for collaborative, continuous, and data-driven chronic care have increased exponentially with the application of mobile health (mHealth) technologies. Mobile apps allow patients to collect, monitor, and report symptoms and events outside of an office visit and to make these data available to providers for shared decision making. mHealth interventions have been widely used in the management of diabetes, hypertension, and chronic pain, among others [13,14,20]. Although technology has been used to facilitate earlier detection of pulmonary exacerbations and adherence monitoring in CF care [25-27], mHealth solutions for CF self-management have not been adopted [28,29]. CF care is uniquely suitable for mHealth solutions because of well-established clinical guidelines, multidisciplinary care teams, national patient registry, and an expansive network of accredited CF centers; however, barriers related to the integration of patient-generated health data with information technology (IT) systems and electronic medical records have been documented [21]. In addition, the few existing CF apps only partially satisfy the user needs reported in previous research [29,30]. mHealth solutions are even more important in the context of the COVID-19 pandemic, which has imposed limitations on in-person visits for chronic care. As such, mHealth approaches may be used to supplement telehealth as a risk-reduction strategy [31].

Genia

Genia is a mHealth platform designed to facilitate the coproduction of CF care (Figure 1). Developed by Upstream Dream in collaboration with the Swedish CF community at Karolinska Institute in 2014, it was first introduced to pediatric CF centers in Sweden and has been well received by patients, with 65% to 87% of families using the app at the most active clinics to track symptoms, activities, and aspects of daily care and share them with the clinical team [32].

Genia supports the upload of images (eg, a patient may take a photo of their sputum and upload it via the app) and can connect to other apps or devices, such as electronic spirometers and smart medication dispensers. A proof-of-concept test of the Genia platform in juvenile idiopathic arthritis reported that it improves patient engagement, patient-centered care, and practice-based learning and recommended the uptake of the platform in other chronic conditions [33,34].

Figure 1. Genia dashboards: (A) treatments, (B) trackers, and (C) appointments and reports.



Objectives

This study adapted Genia to the needs and preferences of the CF community in a pediatric CF center in the United States. A proof-of-concept testing was conducted to optimize the platform, remove glitches, and ensure the security and privacy of patient-generated health data.

Methods

Theoretical Model

The project is informed by the model of health care service coproduction [5] whose central tenet is that health care services are coproduced by patients and providers in systems that support and constrain effective partnerships.

Study Design

Determining the Needs and Preferences of the CF Community

We used focus groups to assess the needs and preferences of the pediatric patients with CF, their families, and their health care providers regarding the Genia mHealth platform and to determine what modifications were necessary. Patient focus groups were conducted via videoconferencing in keeping with infection control guidelines [35]. For accessibility and convenience, caregiver focus groups were conducted in 3 different locations in the catchment area of the University of Alabama at Birmingham/Children's of Alabama (UAB/COA) CF center. UAB/COA is the only pediatric CF program in Alabama, serving families from across the state. Patients and familial caregivers were English-speaking and were recruited during routine clinic visits and by phone. Participants provided informed consent and received a US $25 gift card to compensate for their time. The clinical focus group was conducted on site at the UAB/COA CF center during a routine meeting of the CF clinical team. As such, clinical participants were not compensated for their time.

During caregiver focus groups, participants were first shown a 2-minute video that demonstrated the functionalities of the existing Genia app and then asked to explore the app on their own using study iPads preloaded with a beta version of the app in English. Moderators were available to answer questions about the app. Patients who participated in a virtual focus group were given access to the app on their personal devices up to 3 days before the focus group. As they had the opportunity to explore the app beforehand, less time was devoted to app demonstration. Focus group sessions were led by 3 research team members (GO, SR, and RG) who used a semistructured guide (Table 1) to facilitate discussions. Each session lasted approximately 90 minutes. Focus groups were recorded, transcribed verbatim, and deidentified for analysis. The study was approved by the Institutional Review Board of the University of Alabama at Birmingham (Protocol IRB-300001749).

Table 1. Semistructured guide for focus group discussions.

Question type	Questions
General	1. What are your first impressions of the app? 2. How do you see yourself using the app? 3. What part is most interesting to you? 4. What part is most useful to you? 5. What do you think about the visual presentation? 6. What could be improved in the app?
Overarching	1. How should the Genia platform be adapted to meet your needs and preferences?
Feature specific	1. Trackers: Are these the topics you would track and talk to your doctor about? Are there other important topics that are not included? 2. Previsit reports: What do you think about the questions? Is there something that is not clear or could be improved? 3. Treatments: What do you think about this feature? Is there anything that could be improved? 4. Appointments: What do you think about this feature? Is there anything that could be improved? 5. News: What do you think about this feature? Is there anything that could be improved? 6. Settings: Is there anything that could be improved in the app settings?

Data from all focus groups were combined and analyzed together using a thematic analysis [36], an interpretative research approach that uses a purely qualitative account of data rather than frequency of codes for theme development [37]. Thematic analysis identifies themes from both latent and manifest content, without considering code frequency. Thus, the number of focus groups, participants, or categories had no consequence on the results. A constant comparative method [38] was employed to generate categories and themes. Transcriptions were coded independently by 3 research team members (GO, RR, and RG), and then, codes were discussed by the entire research team and the final coding scheme was decided jointly. The Genia app developers did not participate in the data analysis and were blinded to the identity of the participants.

Adapting the Genia App to Identified Needs and Preferences

Focus group data were first categorized according to app features. The research team then ranked participant recommendations in 3 categories (highly desired, moderately desired, and less desired) based on participant feedback and sent the recommendations to the Genia developers. The developers categorized these recommendations according to technical feasibility (immediately feasible, feasible in the future, and not feasible) based on app infrastructure, time required to make the changes, and cost. The research team and Genia developers met to discuss what recommendations were both highly or moderately desired and immediately feasible. The mHealth platform was then modified according to these recommendations and prepared for the proof-of-concept testing phase.

Testing the Modified Platform

In this phase, we tested the Health Insurance Portability and Accountability Act (HIPAA)–compliant data flow from the app to a secure study server housed at UAB/COA and beta tested the modified platform. To ensure the security and privacy of patient-generated health data, we tested the transmission of patient-generated reports sent as HIPAA-compliant PDF and .xls files to a secure Research Electronic Data Capture (REDCap) server at our academic institution. REDCap, which uses cellular networks and encrypted point-to-point communication, is a secure web app specifically designed to support and house web-based data capture for research studies. We enrolled 6 participants from the caregiver focus groups to test data flow to the server and to report bugs and other problems in the mobile app for 30 days. Participants provided informed consent and were compensated US $100 for their time. Additional beta testing of the modified app was performed by research team members (GO, RR, RG, and CM).

Development of Data Management and Clinical Follow-Up Protocols

In this phase, we developed protocols for the management of patient-generated health data submitted to REDCap and for clinical follow-up in response to submitted data. Clinical protocols were developed with input from the COA CF Nurse Coordinator and the UAB/COA CF Center Director. We also trained the CF care team on the use of the platform in a one-day face-to-face session.

Results

Determining the Needs and Preferences of the CF Community

A total of 5 focus groups with 21 participants were conducted: 1 with pediatric clinical team members (n=5), 1 with adolescent patients with CF treated at COA (n=5), and 3 with familial caregivers of pediatric patients with CF (n=11). Clinical team members represented the specialties included in a multidisciplinary CF clinic (pulmonology, nursing, nutrition, respiratory therapy, and social work). Patients with CF, mean age 17.8 years (SD 1.7; range 15-20 years) and 80% (4/5) female, resided in 4 different geographic areas of the state. Familial caregivers, mean age 42 years (SD 6.9; range 33-59 years), 90% (10/11) female, and 90% (10/11) parents, represented 3 different geographic areas of residence.

Trackers

The Genia app allows users to keep track of their symptoms using a library of developed trackers. Clinical, patient, and caregiver respondents prompted developers to create or improve trackers for a variety of health measures and physical symptoms. Overall, all groups requested that existing trackers be more specific. For example, they recommended that the pain tracker be modified to ask about the location of the pain, its severity, and how long it has been occurring. In addition, all groups recommended a tracker for stool composition and frequency. Caregivers shared that trackers for seemingly embarrassing topics such as stools could help older children better communicate their symptoms:

> *…because when they get to the age of 17 and 13, they don't wanna discuss it at all. That is gross, and they don't wanna talk about it...When my children were younger, I changed their diapers. I went in the bathroom with them, and you could check. That don't happen any longer, so to have an app that [...] they can put in there when they see, and then it's not necessarily "I've got to go tell mom this is what it looked like."* [Caregiver]

Overall, patients and caregivers requested more detailed clinical trackers, whereas the clinical team considered such trackers unnecessary. For example, contrary to the opinion of the clinical team, patients and caregivers wanted to have a lung function tracker to input the latest pulmonary function test results from each clinic visit and track changes over time, a hospitalization tracker to keep record of when and for how long they are hospitalized, and a vaccinations tracker to help keep track of flu shots and other vaccinations. They also requested a tracker for sinus symptoms, shortness of breath and oxygen saturation, and CF-related diabetes. In addition, caregivers requested that the separate height and weight trackers be combined into one for convenience and ease of use and that the nighttime cough and daytime cough trackers be combined as well. Patients recommended that the exercise tracker be renamed to *physical activity*, with a drop-down menu of options, including various sports, yoga, dancing, hiking, or playing outside, among others.

When asked what trackers were missing from the current list, all respondents highlighted the importance of tracking mental health:

> *Yeah, a mental health question would be very, very helpful ... for the parents and child, as well, because it gives them a better insight into what's going on...* [Caregiver]

> *I would like to see if we could add in mental health [...] anxiety or stuff like that because we— that's something that some of us get asked at clinic and stuff, and sometimes we don't remember it, so that would be a really good one to add to it.* [Patient]

Caregivers hoped that children who are not comfortable talking to their parents about their mental health would tell the app how they are feeling and share these data with the CF care team:

> *I'm okay with leaving it and let him feel comfortable. I'm not gonna try to look and see what he's writing.* [Caregiver]

Clinical participants recommended including a mental health tracker with questions about depression and anxiety from the 4-item Patient Health Questionnaire [39]. In addition, patients and caregivers recommended creating an emotions tracker, where users can choose from a drop-down list of emotions and elaborate on why they feel that way:

> *Call the button daily emotions, because you want them to put it in every day so we can keep track of it, so I think that's under emotions.* [Caregiver]

> *The one I'm thinking is just your mood [...] Were you happy today? Was there anything that you made feel sad? Something like that.* [Patient]

Creating a Treatment

The Genia app allows patients to track medications and treatments, including frequency and duration, and to set treatment reminders. Clinical participants requested updating the treatment and medication list to the US context, to match current CF clinical care guidelines. Caregivers requested including a medication dosage in addition to the medication name and an option to print their medication list from the app for sharing with others, such as pediatricians, school nurses, or daycare providers. They wanted to be able to set multiple daily reminders for the same treatment and reminders for treatments that cycle on and off (eg, antibiotics taken on 28 days on/28 days off schedule). Caregivers also asked for the ability to enter the prescription date and to set refill reminders for medications, medical supplies (eg, nebulizer cups), or treatment-related tasks (eg, changing air filters and flushing a port). Finally, caregivers wanted the flexibility to reorder their medication lists as necessary for different purposes, for example, alphabetically, by time of treatment (morning vs evening), by type of treatment (inhaled vs oral), or by organ system (digestive vs respiratory).

Exporting Reports to the Clinical Setting

One of the most important features of the Genia patient support system is sending patient-generated health data to the care team in the form of reports in a PDF format before clinic appointments. Caregivers and patients wanted to make sure that there was a place to keep a list of questions for an upcoming visit. Caregivers also strongly suggested appending the previsit reports to also include the pen-and-paper questionnaire traditionally administered in the clinic waiting room:

> *[...] 'cause I've been to the point where I'm like I'm just not gonna fill it out.* [Caregiver]

> *I find the papers extremely stressful. [...] I'm trying to have [my kids] not kill each other, and I'm trying to fill out— then the people keep coming in, and I'm trying to answer questions. I just get stressed out about trying to fill out the papers.* [Caregiver]

Caregivers particularly appreciated the opportunity to take notes during the clinic visit and keep them in the app as a personal record of what was discussed or what changes were made during

that visit, along with vital signs, laboratory test results, and results from pulmonary function testing.

Language and Communication

Both caregivers and clinicians suggested multiple changes to the current app language. Some of these changes reflected differences in terminology between the Swedish and US context (eg, breathing treatments vs airway clearance), whereas others pertained to the use of a more idiomatic English language. Several caregivers expressed that there should always be a free-text option or a comment box to encourage patients to share details about their observations and encourage open communication. Caregivers of adolescents were particularly worried that their children would not communicate enough information in the app without detailed prompts.

Sharing and Privacy

Respondents had different opinions about sharing and privacy within the app by role (patient vs caregiver) and age. Adolescent patients thought that their caregivers did not need to be heavily involved in their app usage:

> *I'm 15, and I think it's a great way for growing up—a great way to learn out on your own. [...] I would learn when to take my medications and to refill them and different things like that.*

However, some caregivers had mixed feelings: although they wanted to give their children privacy, they also wanted to be aware of what their children were entering in the app:

> *[...] still, I hate the fact that there's things I just don't know.* [Caregiver]

Other caregivers believed that children should not be able to keep anything private in the app if they are not legal-age adults:

> *[...] as a parent, my child can't keep anything from me until they turn 18.*

Caregivers recommended making all notes private by default to encourage people to record their true thoughts in the app. Users can then choose what notes to share with the CF care team when they generate their previsit report.

Settings and Account

These features were consistently important points of discussion for caregivers and patients. Caregivers of multiple children with CF wanted to be able to track their different children seamlessly, without having to log off from one profile and log on to another:

> *Eventually if they could pick out their color scheme or they could add their own picture. That way at a glance I know if I'm looking at [Son's] or I'm lookin' at [Daughter's], and I don't have to go back and be like, whose is this?*

Caregivers of younger patients wanted both the parent and the child to be able to track and view symptoms and treatments. Both caregivers and patients requested the ability to personalize their settings with profile pictures, avatars, and color schemes.

News

Caregivers suggested that the news feature of the app should be used to maintain current information about the CF care team, including their names, photos, role in the team, and contact information. Caregivers also wanted to receive frequent updates about clinical trials, research news, medication discounts, and CF educational events and resources:

> *I don't think you can ever get too much information. As parents, we can never get too much information.*

Adapting the Genia App to Identified Needs and Preferences

The results of the focus groups highlighted both overlapping and differing preferences among clinical staff, caregivers, and patients. Overall, when making decisions about modifications, the research and development teams prioritized the perspectives of patients and caregivers—the app's end users—over those of clinical team members, which resulted in a greater number of trackers. Although not every recommendation made by participants was feasible, most suggestions were successfully incorporated into the modified app, highlighting the necessity of identifying the needs and preferences of the end users before entering a test environment. A summary of the modifications made as a result of focus group recommendations is presented in Table 2. A full list of all changes is provided in Multimedia Appendix 1.

Table 2. Summary of modifications.

Type	Extent		
	New	Revision	Planned for 2.0 release
Trackers	• Emotions • Hospitalizations • Intravenous treatments • Mental health • Shortness of breath • Sinuses • Stomach • Stools	• Airway clearance • Cough • Energy • Headache • Height and weight • Lung function • Medicine • Pain • Physical activity • School and work • Sleep	• CF[a]-related diabetes • Diagnosis and mutations • Goals, with reminders and rewards • Vaccinations
Treatments	• "Other" option • Printing list of medications • Reorganizing medications	• US-specific medications and treatments • Multiple reminder options	• Equipment reminders • Medication dosage • Refill reminders
Reports	• Adding clinical history form • Notes for upcoming visit • Notes during visit • Weekly and monthly check-in	• N/A[b]	• N/A
Language and presentation	• N/A	• Personalization with a user name • Idiomatic English	• Avatars and color options
Privacy	• N/A	• Default "do not share"	• N/A
Settings	• Sync other health apps	• Parent and child track under the same account • Technical support contact	• Calendar • Clinical team contact information • Within-app messages to the clinical team

[a]CF: cystic fibrosis.

[b]N/A: not applicable.

Testing the Modified Platform for Technical Issues

Testing of the HIPAA-compliant data flow from the patient-facing app to a secure US-based cloud server to the REDCap study database was necessary to ensure the protection and privacy of patient-generated health data. The research and development teams also tested that patient-generated reports in the mobile app were correctly uploaded to the REDCap database. The issue of data readability for the clinical team and the research team had to be addressed as well. The development team successfully converted the incoming patient-generated reports into PDF files with charts, graphs, and text boxes of tracked symptoms and raw data in .xls files. This ensured that the clinical team had an easy way to interpret the output of patient symptoms and that the research team had easily accessible data for data analysis.

Beta testing of the modified Genia platform for technical issues was conducted for 30 days with 6 participants (mean age 41 years, SD 2.4; range 43-51) from the pool of caregiver focus group participants, as they were already familiar with the platform. Specifically, we asked participants to test every tracker, submit notes with and without pictures, test if the treatment reminders feature is functional, and make note of anything that needs to be corrected. The research team informed testers that data entered into the app would not be acted upon by the clinical team during the beta testing period and that they should contact the clinical team through usual channels if health concerns arose.

Beta testers reported bugs and issues to a contact person on the research team who then funneled that information to the development team. The most common issues reported by beta testers included metric units for height and weight, unnecessary caps on information input, medication reminders not going off, and leftover Swedish language.

The developers removed the reported glitches and worked collaboratively with the research and clinical team members to make adaptations that ensure smooth user experience. For example, beta testers reported a medication missing from the menu of options. The CF nurse coordinator reviewed and updated the medication list, and the developers updated the app accordingly. In addition, beta testers reported an issue with attaching images to previsit reports, and the research team detected a glitch in the upload of previsit reports to the REDCap database. All these issues were fixed by app developers. A

complete list of optimizations and removed glitches is provided in Multimedia Appendix 1.

Development of Data Management and Clinical Follow-Up Protocols

The clinical and research teams met with the development team during a site visit in November 2019. This meeting was instrumental in laying the groundwork of clinical protocols. For example, clinical providers had concerns about how the patient-generated health data from the app would fit into clinical planning and flow. To address these concerns, a clinical management protocol was developed that summarizes patient reports from the app and incorporates them into the weekly previsit overview of patients with an upcoming clinical appointment that week. To ease strain on the providers, an appointed research team member organizes the patient reports from REDCap and provides the PDFs to the CF nurse coordinator on a weekly basis. The CF nurse coordinator then follows a patient support protocol that directs what needs to be done with specific types of information (eg, >3 increased symptoms for >2 days→report to treating pulmonologist and add antibiotic and increased depression reports→initiate contact with a social worker or mental health counselor).

The effect of the platform for disease self-management, patient-reported outcomes, and patient-centered care is being assessed in an ongoing clinical trial (NCT03910881).

Discussion

Principal Findings

In this study, we adapted an mHealth platform developed for the CF community in Sweden to the needs and preferences of the US CF community. We also conducted proof-of-concept testing to optimize the platform, remove glitches, and ensure the security and privacy of patient-generated health data. The project demonstrated the importance of identifying the needs and preferences of end users before using existing mHealth solutions and the need to consider all relevant perspectives, including those of clinicians, patients, caregivers, and app developers.

Although end user preferences guided the modification of app functionalities and presentation, clinician perspectives governed the development of clinical protocols that detailed how patient-generated health data are used in clinical settings. Successful coproduction of health care services for managing a chronic health condition [5] requires the perspectives of both parties. Engaging end users in the development or modification of mHealth apps can help prevent implementation challenges and increase the likelihood of app uptake by the target audience [40]. Engaging clinicians can ensure that mHealth solutions are practical in clinical settings and can be integrated into the clinical data flow within existing IT environments [21]. Other barriers and challenges also exist. For example, not all end users may be able or willing to self-monitor and report symptoms or activities, particularly those with fewer resources or lower literacy, whereas clinicians may find the use of patient-generated health data burdensome or having little measurable effect on health outcomes. Addressing the priorities and concerns of both parties will facilitate realistic and attainable goals in the context of coproduction facilitated by technology-enabled data sharing [5]. For successful mHealth interventions, patient-generated health data must be both meaningful to patients and useful to clinicians [22].

The preferences of end users of Genia mirrored the preferences of patients with CF and their caregivers reported in a previous study [30]. Specifically, end users preferred an app with multiple functions that facilitates access to information, automates disease management, integrates with other apps, facilitates communication with the health care team, and is highly customizable to meet individual goals and preferences. Involving end users in all stages of mHealth development and collaborating with clinicians and health care system experts may result in apps that maintain engagement, improve coproduction of services, and ultimately impact self-management and health outcomes [30].

The described modification process represents a scalable approach to adapting mHealth solutions to local needs and changing contexts. Feedback from patients, caregivers, health care providers, and experts gave input to developers on how to prepare the mobile platform for further adaptations to local needs and changing contexts. For example, a file-based approach was developed to support a multilanguage platform, and medication and treatment lists were prepared to allow for dynamic updates and changes in response to clinician requests and patient preferences. The collection of patient-generated health data was implemented according to a structure that could be used to create new trackers and new report questions. The result was a more scalable and dynamic platform, facilitating further adaptations. Such an approach can be replicated in further modification of this and other mHealth platforms to track home spirometry, home intravenous antibiotics, oral and nebulized medications, or airway clearance therapy via smart devices.

App repositories include hundreds of apps that claim to improve disease management, health outcomes, and health-related behaviors. However, few are evidence-based solutions developed with the involvement of health care professionals, patients, caregivers, and behavioral scientists [33,34]. Even fewer are tailored to the components of daily CF care and the symptoms typically experienced by people with CF. The few apps that collect CF-specific patient-generated health data are not integrated in the CF care plan or the clinical workflow. Without such integration, mHealth technologies can have only a minimal impact on chronic disease care and management [41]. At the same time, integration of patient-generated health data in clinical care via mHealth solutions needs to be done in a way that safeguards the privacy and integrity of patients in their domestic domain [42]. mHealth platforms and patient support systems, such as Genia, provide new opportunities for improved self-care and clinical management of CF. However, for the integration of mHealth solutions in clinical care, their feasibility, their acceptability to patients and providers, and their clinical effectiveness need to be tested in rigorous clinical trials.

Limitations

Participants included English-speaking patients, caregivers, and health care providers recruited from a single pediatric CF center in the United States. The perspectives of this convenience sample may not be representative of all CF centers or geographic areas in the country. However, as CF disease management and care delivery are highly protocolized across the network of accredited CF centers, vast differences between CF stakeholders from other centers and regions are unlikely. In addition, focus group participants had an opportunity to use the app only during focus group discussions, which may have limited the scope and depth of their experience and feedback. Finally, the project did not address data mapping of patient-generated health data to interoperable standards to integrate Genia-collected patient-generated health data in the electronic health record, a step that will be addressed in a future study phase.

Future Directions

A currently ongoing pilot clinical trial (NCT03910881) assesses whether the use of Genia over 6 months enhances patient-centered care (shared decision making, patient activation, patient satisfaction, and patient self-efficacy) and improves patient-reported outcomes (symptom scales and quality-of-life domains of the Cystic Fibrosis Questionnaire-Revised [43]). A randomized clinical trial will evaluate the impact of Genia on respiratory and nutritional outcomes. Informed by the model of health care service coproduction [5], these study phases are predicated on the hypotheses that (1) use of patient-generated health data will affect providers' actions, which will have downstream benefits, and (2) use of patient-generated health data will affect patients' disease management behaviors and interaction with the health system.

Conclusions

mHealth offers new opportunities to support self-management of CF, facilitate patient-physician communication, and promote coproduction of care. Broad and successful uptake of mHealth solutions in CF care requires careful consideration of patient and family perspectives and active involvement of CF clinical teams.

Acknowledgments

This study was sponsored by a grant from the Kaul Pediatric Research Institute of the Alabama Children's Hospital Foundation (KPRI 000522124).

Conflicts of Interest

AH is the founder and CEO of Upstream Dream, the company providing the Genia patient support system. The remaining authors declare no conflict of interest.

Multimedia Appendix 1
Detailed description of changes based on focus groups and beta testing: implemented, not implemented, and planned.
[DOCX File , 19 KB - formative_v5i1e19413_app1.docx]

References

1. Sawicki GS, Sellers DE, Robinson WM. High treatment burden in adults with cystic fibrosis: challenges to disease self-management. J Cyst Fibros 2009 Mar;8(2):91-96 [FREE Full text] [doi: 10.1016/j.jcf.2008.09.007] [Medline: 18952504]
2. Konstan MW, VanDevanter DR, Rasouliyan L, Pasta DJ, Yegin A, Morgan WJ, Scientific Advisory Group, InvestigatorsCoordinators of the Epidemiologic Study of Cystic Fibrosis. Trends in the use of routine therapies in cystic fibrosis: 1995-2005. Pediatr Pulmonol 2010 Dec;45(12):1167-1172 [FREE Full text] [doi: 10.1002/ppul.21315] [Medline: 20717935]
3. Claudia G, W Alexander G, LeighAnne O, J Michael M. Engineering a Learning Healthcare System: A Look at the Future. Workshop Summary. Washington, D.C: National Academies Press; 2011.
4. Wagner EH, Austin BT, Von KM. Organizing care for patients with chronic illness. Milbank Q 1996;74(4):511-544. [Medline: 8941260]
5. Batalden M, Batalden P, Margolis P, Seid M, Armstrong G, Opipari-Arrigan L, et al. Coproduction of healthcare service. BMJ Qual Saf 2016 Jul;25(7):509-517 [FREE Full text] [doi: 10.1136/bmjqs-2015-004315] [Medline: 26376674]
6. Kaplan HC, Thakkar SN, Burns L, Chini B, Dykes DM, McPhail GL, et al. Protocol of a pilot study of technology-enabled coproduction in pediatric chronic illness care. JMIR Res Protoc 2017 Apr 28;6(4):e71 [FREE Full text] [doi: 10.2196/resprot.7074] [Medline: 28455274]
7. West NE, Mogayzel PJ. Transitions in health care: what can we learn from our experience with cystic fibrosis. Pediatr Clin North Am 2016 Dec;63(5):887-897. [doi: 10.1016/j.pcl.2016.06.010] [Medline: 27565366]
8. Patel A, Dowell M, Giles BL. Current concepts of transition of care in cystic fibrosis. Pediatr Ann 2017 May 01;46(5):e188-e192. [doi: 10.3928/19382359-20170425-02] [Medline: 28489224]
9. Boland P. The emerging role of cell phone technology in ambulatory care. J Ambul Care Manage 2007;30(2):126-133. [doi: 10.1097/01.JAC.0000264602.19629.84] [Medline: 17495681]

10. Valderas JM, Kotzeva A, Espallargues M, Guyatt G, Ferrans CE, Halyard MY, et al. The impact of measuring patient-reported outcomes in clinical practice: a systematic review of the literature. Qual Life Res 2008 Mar;17(2):179-193. [doi: 10.1007/s11136-007-9295-0] [Medline: 18175207]
11. Chen J, Ou L, Hollis SJ. A systematic review of the impact of routine collection of patient reported outcome measures on patients, providers and health organisations in an oncologic setting. BMC Health Serv Res 2013;13:211 [FREE Full text] [doi: 10.1186/1472-6963-13-211] [Medline: 23758898]
12. Berry DL, Blumenstein BA, Halpenny B, Wolpin S, Fann JR, Austin-Seymour M, et al. Enhancing patient-provider communication with the electronic self-report assessment for cancer: a randomized trial. J Clin Oncol 2011 Mar 10;29(8):1029-1035 [FREE Full text] [doi: 10.1200/JCO.2010.30.3909] [Medline: 21282548]
13. Farzandipour M, Nabovati E, Sharif R, Arani MH, Anvari S. Patient self-management of asthma using mobile health applications: a systematic review of the functionalities and effects. Appl Clin Inform 2017 Oct;8(4):1068-1081 [FREE Full text] [doi: 10.4338/ACI-2017-07-R-0116] [Medline: 29241254]
14. 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;18(5):e97 [FREE Full text] [doi: 10.2196/jmir.4883] [Medline: 27185295]
15. Cleeland CS, Wang XS, Shi Q, Mendoza TR, Wright SL, Berry MD, et al. Automated symptom alerts reduce postoperative symptom severity after cancer surgery: a randomized controlled clinical trial. J Clin Oncol 2011 Mar 10;29(8):994-1000 [FREE Full text] [doi: 10.1200/JCO.2010.29.8315] [Medline: 21282546]
16. Gilbert JE, Howell D, King S, Sawka C, Hughes E, Angus H, et al. Quality improvement in cancer symptom assessment and control: the Provincial Palliative Care Integration Project (PPCIP). J Pain Symptom Manage 2012 Apr;43(4):663-678. [doi: 10.1016/j.jpainsymman.2011.04.028] [Medline: 22464352]
17. Kotronoulas G, Kearney N, Maguire R, Harrow A, Di DD, Croy S, et al. What is the value of the routine use of patient-reported outcome measures toward improvement of patient outcomes, processes of care, and health service outcomes in cancer care? A systematic review of controlled trials. J Clin Oncol 2014 May 10;32(14):1480-1501. [doi: 10.1200/JCO.2013.53.5948] [Medline: 24711559]
18. Basch E, Deal AM, Kris MG, Scher HI, Hudis CA, Sabbatini P, et al. Symptom monitoring with patient-reported outcomes during routine cancer treatment: a randomized controlled trial. J Clin Oncol 2016 Feb 20;34(6):557-565. [doi: 10.1200/JCO.2015.63.0830] [Medline: 26644527]
19. Denis F, Basch E, Septans A, Bennouna J, Urban T, Dueck AC, et al. Two-year survival comparing web-based symptom monitoring vs routine surveillance following treatment for lung cancer. J Am Med Assoc 2019 Jan 22;321(3):306-307 [FREE Full text] [doi: 10.1001/jama.2018.18085] [Medline: 30667494]
20. Martinez R, Rogers AD, Numanoglu A, Rode H. The value of WhatsApp communication in paediatric burn care. Burns 2018 Jun;44(4):947-955. [doi: 10.1016/j.burns.2017.11.005] [Medline: 29395403]
21. Van Citters AD, Gifford AH, Brady C, Dunitz JM, Elmhirst M, Flath J, et al. Formative evaluation of a dashboard to support coproduction of healthcare services in cystic fibrosis. J Cyst Fibros 2020 Apr 27. [doi: 10.1016/j.jcf.2020.03.009] [Medline: 32354650]
22. Dixon-Woods M, Campbell A, Chang T, Martin G, Georgiadis A, Heney V, et al. A qualitative study of design stakeholders' views of developing and implementing a registry-based learning health system. Implement Sci 2020 Mar 06;15(1):16 [FREE Full text] [doi: 10.1186/s13012-020-0976-1] [Medline: 32143678]
23. Oates G, Dunn K, Frederick C, Hall S, Maggs J, Niranjan S. Managing daily life with cystic fibrosisxploration of patient and parent perspectives. In: Pediatric Pulmonology. 2019 Presented at: 33rd Annual North American Cystic Fibrosis Conference; Oct 31-Nov 2, 2019; Nashville, TN p. S400 URL: https://www.researchgate.net/publication/337923724_Managing_Daily_Life_with_Cystic_Fibrosis_Exploration_of_Patient_and_Parent_Perspectives
24. Rowbotham NJ, Smith S, Leighton PA, Rayner OC, Gathercole K, Elliott ZC, et al. The top 10 research priorities in cystic fibrosis developed by a partnership between people with CF and healthcare providers. Thorax 2018 Apr;73(4):388-390 [FREE Full text] [doi: 10.1136/thoraxjnl-2017-210473] [Medline: 28778919]
25. Lechtzin N, Mayer-Hamblett N, West NE, Allgood S, Wilhelm E, Khan U, eICE Study Team. Home monitoring of patients with cystic fibrosis to identify and treat acute pulmonary exacerbations. eICE study results. Am J Respir Crit Care Med 2017 Nov 01;196(9):1144-1151 [FREE Full text] [doi: 10.1164/rccm.201610-2172OC] [Medline: 28608719]
26. Sarfaraz S, Sund Z, Jarad N. Real-time, once-daily monitoring of symptoms and FEV in cystic fibrosis patients--a feasibility study using a novel device. Clin Respir J 2010 Apr;4(2):74-82. [doi: 10.1111/j.1752-699X.2009.00147.x] [Medline: 20565480]
27. Wood J, Jenkins S, Putrino D, Mulrennan S, Morey S, Cecins N, et al. A smartphone application for reporting symptoms in adults with cystic fibrosis improves the detection of exacerbations: Results of a randomised controlled trial. J Cyst Fibros 2020 Mar;19(2):271-276. [doi: 10.1016/j.jcf.2019.09.002] [Medline: 31522924]
28. Calthorpe RJ, Smith S, Gathercole K, Smyth AR. Using digital technology for home monitoring, adherence and self-management in cystic fibrosis: a state-of-the-art review. Thorax 2020 Jan;75(1):72-77. [doi: 10.1136/thoraxjnl-2019-213233] [Medline: 31594802]

29. Martinez-Millana A, Zettl A, Floch J, Calvo-Lerma J, Sevillano JL, Ribes-Koninckx C, et al. The potential of self-management mHealth for pediatric cystic fibrosis: mixed-methods study for health care and app assessment. JMIR Mhealth Uhealth 2019 Apr 18;7(4):e13362 [FREE Full text] [doi: 10.2196/13362] [Medline: 30998222]
30. Hilliard ME, Hahn A, Ridge AK, Eakin MN, Riekert KA. User preferences and design recommendations for an mHealth app to promote cystic fibrosis self-management. JMIR Mhealth Uhealth 2014;2(4):e44 [FREE Full text] [doi: 10.2196/mhealth.3599] [Medline: 25344616]
31. Kemp MT, Williams AM, Alam HB. eClinic: increasing use of telehealth as a risk reduction strategy during the covid-19 pandemic. Trauma Surg Acute Care Open 2020;5(1):e000481 [FREE Full text] [doi: 10.1136/tsaco-2020-000481] [Medline: 32377571]
32. Longacre M, Grande S, Hager A, Montan M, Bergquist RP, Martensson M, et al. Clinical adoption of mHealth technology to support pediatric cystic fibrosis care in Sweden: qualitative case study. JMIR Pediatr Parent 2018 Dec 10;1(2):e11080 [FREE Full text] [doi: 10.2196/11080] [Medline: 31518297]
33. Nelson E, Batalden P, Corrigan J. Enabling uptake of a registry-supported care and learning system in the United States: a report to the Robert Wood Johnson Foundation. The Dartmouth Institute for Health Policy and Clinical Practice. Hannover, NH; 2014. URL: http://srq.nu/srqny/wp-content/uploads/2014/10/Summary-Report-4-30-14-FINAL.pdf [accessed 2021-01-06]
34. Grande SW, Longacre MR, Palmblad K, Montan MV, Berquist RP, Hager A, et al. Empowering young people living with juvenile idiopathic arthritis to better communicate with families and care teams: content analysis of semistructured interviews. JMIR Mhealth Uhealth 2019 Feb 22;7(2):e10401 [FREE Full text] [doi: 10.2196/10401] [Medline: 30794202]
35. Saiman L, Siegel JD, LiPuma JJ, Brown RF, Bryson EA, Chambers MJ, Cystic Fibrous Foundation, Society for Healthcare Epidemiology of America. Infection prevention and control guideline for cystic fibrosis: 2013 update. Infect Control Hosp Epidemiol 2014 Aug;35(Suppl 1):S1-S67. [doi: 10.1086/676882] [Medline: 25025126]
36. Vaismoradi M, Turunen H, Bondas T. Content analysis and thematic analysis: implications for conducting a qualitative descriptive study. Nurs Health Sci 2013 Sep;15(3):398-405. [doi: 10.1111/nhs.12048] [Medline: 23480423]
37. Vaismoradi M, Snelgrove S. Theme in qualitative content analysis and thematic analysis. Forum Qual Soc Res 2019;20(3). [doi: 10.17169/fqs-20.3.3376]
38. Strauss A, Corbin J. Basics of Qualitative Research: Procedures and Techniques for Developing Grounded Theory. Thousand Oaks, CA: Sage; 1998.
39. Kroenke K, Spitzer RL, Williams JB, Löwe B. An ultra-brief screening scale for anxiety and depression: the PHQ-4. Psychosomatics 2009;50(6):613-621. [doi: 10.1176/appi.psy.50.6.613] [Medline: 19996233]
40. Bendixen RM, Fairman AD, Karavolis M, Sullivan C, Parmanto B. A user-centered approach: understanding client and caregiver needs and preferences in the development of mHealth apps for self-management. JMIR Mhealth Uhealth 2017 Sep 26;5(9):e141 [FREE Full text] [doi: 10.2196/mhealth.7136] [Medline: 28951378]
41. Hamine S, Gerth-Guyette E, Faulx D, Green BB, Ginsburg AS. Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. J Med Internet Res 2015;17(2):e52 [FREE Full text] [doi: 10.2196/jmir.3951] [Medline: 25803266]
42. Hager A, Brommels M, Salomonsson S, Lindblad S. Design of a conceptual model for knowledge commons in the patient world arena: a case study approach. JMIR Preprints (forthcoming). [doi: 10.2196/preprints.16842]
43. Quittner AL, Buu A, Messer MA, Modi AC, Watrous M. Development and validation of The Cystic Fibrosis Questionnaire in the United States: a health-related quality-of-life measure for cystic fibrosis. Chest 2005 Oct;128(4):2347-2354. [doi: 10.1378/chest.128.4.2347] [Medline: 16236893]

Abbreviations

CF: cystic fibrosis
COA: Children's of Alabama
HIPAA: Health Insurance Portability and Accountability Act
mHealth: mobile health
REDCap: Research Electronic Data Capture
UAB: University of Alabama at Birmingham

Edited by G Eysenbach; submitted 16.04.20; peer-reviewed by K Sabadosa, J Amann; comments to author 08.06.20; revised version received 03.07.20; accepted 17.12.20; published 26.01.21.

Please cite as:
Rutland SB, Bergquist RP, Hager A, Geurs R, Mims C, Gutierrez HH, Oates GR
A Mobile Health Platform for Self-Management of Pediatric Cystic Fibrosis: Qualitative Study of Adaptation to Stakeholder Needs and Integration in Clinical Settings
JMIR Form Res 2021;5(1):e19413
URL: http://formative.jmir.org/2021/1/e19413/
doi:10.2196/19413
PMID:33496667

Original Paper

Association Between Care Utilization and Anxiety Outcomes in an On-Demand Mental Health System: Retrospective Observational Study

Sarah Kunkle[1], MPH; Manny Yip[1], MSc; Justin Hunt[1], MSc, MD; Watson Ξ[1], PhD; Dana Udall[1], PhD; Patricia Arean[2], PhD; Andrew Nierenberg[3], MD; John A Naslund[4], PhD

[1]Ginger, San Francisco, CA, United States

[2]Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, United States

[3]Massachusetts General Hospital, Boston, MA, United States

[4]Department of Global Health and Social Medicine, Harvard Medical School, Boston, MA, United States

Corresponding Author:
Sarah Kunkle, MPH
Ginger
116 New Montgomery Street
San Francisco, CA
United States
Phone: 1 7175197355
Email: sarah@ginger.io

Abstract

Background: Anxiety is an extremely prevalent condition, and yet, it has received notably less attention than depression and other mental health conditions from a research, clinical, and public health perspective. The COVID-19 pandemic has only exacerbated growing concerns about the burden of anxiety due to the confluence of physical health risks, economic stressors, social isolation, and general disruption of daily activities.

Objective: This study examines differences in anxiety outcomes by care modality (coaching, teletherapy and telepsychiatry, and combined care) within an on-demand mental health system. We also explore the association between levels of engagement within each care modality and odds of improvement in symptoms of anxiety.

Methods: We conducted a retrospective observational study of individuals who accessed Ginger, an on-demand mental health system. Data were collected from 1611 Ginger members between January 1, 2018, and December 31, 2019. We used logistic regression to assess the association between care modality and improvement in anxiety symptoms. Within each modality, we assessed the association between level of engagement and improvement.

Results: Of 1611 Ginger members, 761 (47.0%) experienced a decrease in anxiety symptoms, as measured by a change from a positive to a negative 2-item Generalized Anxiety Disorder (GAD-2) screen. Among members who still screened positive at follow-up (865/1611, 53%), a total of 192 members (11.9%) experienced a clinically significant score reduction in the full GAD-7 (ie, a score reduction of >5 points), even though their GAD-2 scores were still positive. All modalities showed increased odds of improvement compared to those who were not engaged with coaching or clinical services (“app-only”). Higher GAD-7 intake scores were also associated with decreased odds of improvement.

Conclusions: This study found increased odds of anxiety improvement for all care modalities compared to those who did not engage in care, with larger effect sizes for higher utilization within all care modalities. Additionally, there is a promising observation that those engaged in combined care (teletherapy and text-based coaching) had the greatest odds of anxiety improvement. Future directions include more detailed classifications of utilization patterns and an exploration of explanations and solutions for lower-utilization members.

(JMIR Form Res 2021;5(1):e24662) doi:10.2196/24662

KEYWORDS

mental health; digital health; anxiety; telehealth; virtual care; utilization; outcome; retrospective; observational

Introduction

Background

Anxiety disorders, including generalized anxiety disorder, panic disorder, and various phobia-related disorders, are prevalent mental health conditions in the United States and globally; large population surveys have shown that about one-third of individuals are affected by an anxiety disorder during their lifetimes [1]. In 2010, anxiety disorders were the sixth leading cause of disability in high-income and low- and middle-income countries [2]. Despite these far-reaching effects, anxiety has received notably less attention than depression from a research, clinical, and public health perspective and often goes unreported or untreated [3,4]. A study by the World Health Organization (WHO) found that only a fifth (20.6%) of participants with an anxiety disorder sought help from health care services, and of those individuals who sought help, 23.2% received no treatment at all [1]. Similarly, in the United States, anxiety disorders are the most common mental health condition, but a large portion of those affected (36.9%) are estimated to go untreated [5].

The COVID-19 pandemic has only exacerbated growing concerns about the burden of anxiety due to the confluence of physical health risks, economic stressors, social isolation, and general disruption of daily activities [6]. A preprint showed that as of April 2020, 1 out of 4 US adults meet the criteria for serious mental distress, 8 times more than a demographically similar sample from 2018 [7]. Data from the Centers for Disease Control and Prevention (CDC) show an increasing percentage of adults experiencing symptoms of anxiety disorder since early April; approximately 35% of adults reported symptoms of anxiety in July 2020 compared to 8.2% in January 2019 to June 2019 [8].

Treatment

Various treatment options have been shown to be effective for anxiety, including psychotherapy and medication [9]. Psychotherapy techniques include Cognitive Behavioral Therapy (CBT), which teaches people different ways of thinking, behaving, and reacting to anxiety-producing objects and situations. These evidence-based treatments have been effectively delivered as telemedicine offerings, with several systematic reviews indicating that treatment delivered in this manner increases access and reach to care [10-12]. Other benefits of teletherapy include more convenient access, reduced stigma, and greater scalability compared with traditional in-person therapy [13]. Beyond teletherapy, there is also evidence that smartphone interventions can reduce anxiety symptoms [12]. Given the current environment with COVID-19, the ability to reach people in need with virtual care is critical.

Health coaching has also emerged as a potential solution to overcoming traditional shortages of specialist mental health providers. Specific to mental health and anxiety, coaching can work similarly to psychotherapy in addressing symptoms through positive psychology, mindfulness, motivational interviewing, strength, and solution-oriented focuses, among other techniques [14]. A systematic review concluded that health coaching could motivate change in the lifestyle behavior of patients with chronic illness, leading to improvements in both physical and mental health status [15]. Recent studies focused on text-based coaching interventions have shown significant improvements in mental health outcomes equivalent to treatment as usual, namely, in-person and telephone counseling [16,17]. Because there can be significant heterogeneity in these types of interventions, researchers and clinicians have published guidance on developing protocols for text-based coaching in digital mental health interventions [17].

Study Objectives

There is an established evidence base for anxiety treatments. However, there is a need to understand what is happening “in the wild” versus in controlled settings to understand if members are achieving expected outcomes, to further our understanding of how evidence-based interventions are implemented within new care delivery models, and to potentially uncover new areas for future study [18-20]. Thus, this study examines differences in anxiety outcomes by care modality (coaching, teletherapy and telepsychiatry, and combined care) within an on-demand mental health system. We also explore the association between levels of engagement within each care modality and odds of improvement in symptoms of anxiety. We hypothesize that those engaged in multiple modalities will show greater odds of improvement compared to those engaged in a single care modality, and that within each care modality, more consistent and regular engagement will be associated with greater odds of improvement.

Methods

Overview

This is a retrospective observational study of individuals who accessed Ginger, an on-demand mental health system. Data were collected from members of Ginger between January 1, 2018, and December 31, 2019. This time period was chosen because it reflects the approximate timing of when Ginger began to provide care to members via its employer business.

Participants

Study participants have access to the Ginger system as part of their employer or health plan benefits. Internal clinical protocols include the following exclusionary criteria where self-directed telehealth is likely not appropriate and where more specialized and urgent psychiatric services are required: (1) active suicidal ideation; (2) active high-risk self-harm behavior; (3) 2 or more hospitalizations within the past 6 months, or 1 hospitalization in the past month for psychiatric reasons; (4) certain symptoms of psychosis that are poorly managed (eg, member is not medication compliant or symptoms are unresponsive to treatment) and are likely incompatible with telehealth; (5) a primary diagnosis of a substance use disorder, or moderate-to-severe substance abuse issues, due to the high complexity, severity, and risk frequently associated with such members, as well as the need for specialized care; (6) active eating disorders with symptoms considered to be high risk; (7) ongoing grave disability, including certain patients who are bipolar with active mania/hypomania or mixed episodes who are unmedicated or who have poor compliance with a medication regimen over time; (8) 2 or more medical hospitalizations in

the last month, due to the high likelihood that the individual has a poorly controlled medical condition that requires close monitoring.

For this study, we included Ginger users aged 18 years or older who downloaded the app during the study data collection period and screened positive for anxiety on the 2-item Generalized Anxiety Disorder (GAD-2) intake survey.

Procedures

The Ginger System

Ginger provides members with access to behavioral health coaching, teletherapy, telepsychiatry, and self-guided content and assessments, primarily via a mobile app platform. After downloading the mobile app, users are able to start texting with a behavioral health coach within minutes of requesting to connect. Ginger coaches are full-time employees who have an advanced degree in a field related to mental health or have accredited coach certification [21,22]. While many users are solely engaged with text-based coaching services, some will request or require escalation to clinical services (teletherapy or telepsychiatry), depending on preference or clinical severity. Examples of situations that require escalation include individuals with chronic mental illness and severe trauma, the potential to harm oneself or others, and significant mental instability (hallucinations, delusions, extreme mood swings, etc). When members are escalated to therapy or psychiatry, they may continue working with a coach, provided that they also seek additional specialized care concurrently. Coaching can continue supporting them in addressing day-to-day goals and challenges and act as an adjunct to the care plan put in place by their therapist or psychiatrist [23].

Data Collection

Ginger uses the 2-item and 7-item Generalized Anxiety Disorder questionnaires (GAD-2 and GAD-7) to assess and track anxiety, both of which have been validated and shown to have good operating characteristics for all anxiety disorders [24]. Additionally, the GAD-2 has been shown to be sensitive to treatment change and, thus, an efficient measure of treatment progress and outcomes [25].

There is no strict guidance on administration protocol; however, an accepted approach is to use the GAD-2 (the first 2 questions of the GAD-7) to screen for anxiety disorders in clinical practice, followed by the remaining 5 items of the full GAD-7 for patients who receive positive results on screening with the first 2 items [3]. Ginger users are prompted to answer the GAD-2 questionnaire through the platform interface when they begin using the Ginger app. Individuals that score a 2 or above for either question are then prompted to complete the full GAD-7. Ginger administers the survey every 2 weeks to users who screen positive and every 3 months for users who screen negative to monitor symptom response and assess if additional care is warranted. Survey completion is not required so as not to withhold support from members who decline to answer the questions but are still interested in accessing the Ginger services.

Measures

Anxiety Symptoms

Symptom improvement is assessed using the GAD-2. For the purpose of this study, we defined *improvement* as individuals who experienced a change from a positive screen to a negative screen at follow-up. A negative screen was defined as a response score for each question of less than 2 (ie, a response of “not at all” or “several days”). A positive screen was defined as a response score for either question of 2 or greater (ie, a response of “more than half the days” or “nearly every day”). Thus, this improvement can be interpreted as an individual’s reduced frequency of reporting key anxiety symptoms (“feeling nervous, anxious, or on edge” or “not being able to stop or control worrying”); more specifically, it reflects a change in experiencing these symptoms for more than half the days or nearly every day to not experiencing them at all or for only several days over the past 2 weeks. As a secondary analysis for users who screened positive at follow-up and completed a full GAD-7 survey, we also looked at a clinically significant reduction in score, defined as a score reduction of >5 points [26-28].

Utilization

We calculated utilization based on product user behavior data, including the number of coaching sessions and messages, number of clinical (therapy and psychiatry) appointments completed, and length of time engaged. Prior studies have similarly used these metrics on virtual interactions as utilization measures [29-31].

We categorized utilization levels based on Ginger clinical protocols, external guidelines, and supporting literature. In general, a typical treatment period consists of 8-12 weeks, with weekly interactions to check in on progress. However, there are certain situations in which the duration of treatment is modified (either extended or shortened) based on member-specific circumstances (clinical presentation, covered sessions, the goodness of fit between clinician and member, and life circumstances), which is consistent with recommendations from the literature [17,32]. For example, Ginger clinical protocols state that maintenance and termination can be considered if there is a response by 4-8 weeks. As Ginger members achieve improvements in their GAD scores, they could be moved to a lower level of intervention to ensure therapists are working at the “top of their license” and the system is efficiently using scarce clinical resources. Additionally, some serious and persistent mental illnesses require ongoing and chronic medication management and, thus, a longer treatment duration.

Care Modality

Our first set of models assesses the association between care modality and improvement. We considered 5 categories of members: (1) only accessed the Ginger app, (2) interacted with a coach or clinician but did not meet a minimum threshold for engagement, (3) only interacted with a behavioral health coach via text, (4) only interacted with a clinician (therapist or psychiatrist) via video, (5) interacted with both a behavioral health coach and clinician. Table 1 summarizes care modality categorization and rationale.

Table 1. Care modality definitions and rationale.

Category	Definition	Rationale
App only	<14 messages sent to coach and 0 clinical sessions	These are members who have not attended a clinical session or completed a full coaching interaction. The 14-message threshold is based on internal analyses of what constitutes a "typical" intake coaching session.
Minimal care utilization	<4 weeks of coaching, or the average days between interactions is 14 days, and <2 clinical sessions	These are members who have completed a coaching or clinical session but received minimal therapeutic intervention based on their length of engagement. For coaching, it is not uncommon for an initial consultation to take place over multiple sessions and weeks.
Coaching	≥4 weeks of coaching, and average days between interactions is ≤14 days	Coaches work to get members on a weekly schedule. Prior to 4 weeks of engagement, members are unlikely to achieve a meaningful reduction in symptoms.
Clinical	≥2 clinical visits	The first clinical session is generally considered an "information gathering" intake session. Members generally begin receiving active intervention during their second session.
Combined	≥4 weeks of coaching, and average days between interactions is ≤14 days, and ≥2 clinical sessions	These are members who are engaged with both a coach and clinician, meeting the criteria for both coaching and clinical engagement.

Coaching Utilization

In our coaching-only analyses, we considered members who had exchanged at least 14 messages with a coach and created categories based on both the length and the frequency of their interaction (Textbox 1). In general, this categorization allows us to understand associations by months of engagement, from less than a month ("minimal") to more than 3 months ("high").

Textbox 1. Categorizations for different care modality utilization.

Coaching utilization:

- Minimal:

  less than 4 weeks, or the average days between interactions is greater than 14 days
- Low:

  4-8 weeks, with average days between interactions <15 days
- Medium:

  9-12 weeks, with average days between interactions <15 days
- High:

  13+ weeks, with average days between interactions <15 days

Clinical utilization:

- Minimal:

  1 session
- Low:

  2-6 sessions
- Medium:

  7-12 sessions
- High:

  13+ sessions

Combined (coaching + clinical) utilization:

- Minimal:

  total score=0
- Low:

  total score=1-3
- Medium:

  total score=4-5
- High:

  total score=6

Clinical Utilization

In our clinical-only analyses, we created 4 categories based on the member's number of sessions (Textbox 1). We decided on these cutpoints because the first session is generally considered an intake session, with minimal therapeutic intervention, and since many commercial contracts and Employee Assistance Programs (EAPs) cover up to 6 sessions, this categorization allows us to understand differences in outcomes for those who are on either side of this cutpoint. Finally, 12 sessions is generally considered the upper limit for the recommended course of treatment, although sessions may be extended for serious and persistent mental illnesses that require ongoing and chronic medication management.

Combined (Coaching + Clinical) Utilization

In our coaching with clinical care (combined) analyses, we considered both a member's coaching and clinical utilization level to calculate an overall utilization level. Based on the coaching-only and clinical-only model specifications, we assigned the following values to calculate both a coach and clinical score: minimal=0, low=1, medium=2, and high=3. Finally, we calculated a total score (coach score + clinical score) to categorize combined coaching and clinical utilization (Textbox 1).

Data Management and Analysis

Data for this study were processed using Looker (Looker Data Sciences Inc), a business intelligence and data analytics platform. Data were analyzed in Python and exported to spreadsheets for final analysis and review. We first looked at descriptive statistics of our measures for users who completed GAD questionnaires, segmented by individuals who experienced an improvement versus those who did not. We performed chi-square tests for categorical variables, *t* tests for continuous variables, and Mood median tests for medians to assess differences in characteristics between those who improved versus those who did not.

Given the binary nature of our dependent outcome variable ("improved" vs "did not improve"), we used logistic regression modeling, a common statistical method for quantifying the relationship between various factors and a binary clinical outcome (ie, dependent variable). Our data further meet the assumptions of logistic regression, including independent observations and little or no multicollinearity among the independent variables [33]. Our first set of models assessed the association between care modality and improvement in anxiety symptoms. We assessed the association between level of engagement and improvement within each modality, and we adjusted for baseline anxiety score with the reference group denoted as the "app-only" group (ie, those not engaged with a coach or clinician) as this is the lowest intensity intervention of all options. For each model, we also calculated the Hosmer-Lemeshow goodness-of-fit statistic [34].

Ethics Statement

This is a secondary analysis of pre-existing de-identified data. The study team does not have access to participant identifying information and does not intend to recontact participants. Ginger's research protocols and supporting policies have been reviewed and approved by Advarra's institutional review board in accordance with the US Department of Health and Human Services regulations at 45 CFR 46 [35].

Results

Participant Demographics and Characteristics

Based on our inclusion criteria, 4369 individuals were eligible for this study. Of users who screened positive for anxiety, 1611 users (36.9%) completed a follow-up survey at least 14 days after intake and were included in our analysis.

Table 2 shows descriptive characteristics of Ginger platform users, categorized by individuals who experienced a change in GAD screen (ie, from a positive screen to a negative follow-up screen) and those who did not (positive follow-up screen). A total of 1611 individuals initially screened GAD-2 positive for anxiety symptoms, completed a full GAD-7, completed a follow-up screen, and were included in this analysis. Of these 1611 individuals, 756 (46.9%) experienced a decrease in anxiety symptoms as measured by a change from a positive to a negative GAD-2 screen. Among members who still screened positive at follow-up (855/1611, 53.1%), a total of 192 members (11.9%) experienced a clinically significant score reduction in the full GAD-7 (ie, a score reduction of 5 points), even though their GAD-2 scores were still positive.

Gender and age data were missing for a large portion of the sample, as this is optional information provided in employer eligibility files. For those users who had reported gender (560/1611, 34.5%), 371 (66.3%) were female and 187 (33.3%) were male. For those users who had available date of birth information (996/1611, 61.8%), 131 (13.2%) were 18-24 years of age, 522 (52.4%) were 25-34 years of age, 220 (22.1%) were 35-44 years of age, 121 (12.1%) were 45-64 years of age, and 2 (0.2%) were 65 years of age or older.

In addition to demographics, Table 2 also reports care modality, levels of utilization, satisfaction scores, and clinical [Patient Health Questionnaire (PHQ) and GAD] intake scores for those who screened negative at follow-up and those who screened positive at follow-up. Compared to those who screened positive at follow-up, those who screened negative at follow-up (ie, "improved") were significantly less likely to have only engaged with the app or have minimal care utilization. They also tended to have higher levels of utilization within each modality and lower levels of anxiety at intake.

Table 2. Characteristics of the study cohort (n=1611).

Characteristic	All participants (N=1611)	Negative follow-up screen (n=756)	Positive follow-up screen (n=855)	*P* value
Intake year, n (%)				.49
2018	581 (36.06)	266 (35.19)	315 (36.84)	
2019	1030 (63.94)	490 (64.81)	540 (63.16)	
Gender, n (%)				.26
Female	371 (23.03)	174 (23.02)	197 (23.04)	
Male	187 (11.61)	98 (12.96)	89 (10.41)	
Other	2 (0.12)	2 (0.26)	0 (0.00)	
No response	1051 (65.24)	482 (63.76)	569 (66.55)	
Age in years, n (%)				.36
18-24	131 (8.13)	57 (7.54)	74 (8.65)	
25-34	522 (32.40)	263 (34.79)	259 (30.29)	
35-44	220 (13.66)	102 (13.49)	118 (13.80)	
45-64	121 (7.51)	57 (7.54)	64 (7.49)	
≥65	2 (0.12)	2 (0.26)	0 (0.00)	
No response	615 (38.18)	275 (36.38)	340 (39.77)	
Care modality, n (%)				<.001
App only	144 (8.94)	50 (6.61)	94 (10.99)	
Minimal care utilization	625 (38.80)	271 (35.85)	354 (41.40)	
Coaching only	366 (22.72)	183 (24.21)	183 (21.40)	
Clinical only	385 (23.90)	197 (26.06)	188 (21.99)	
Combined (coaching + clinical)	91 (5.65)	55 (7.28)	36 (4.21)	
Engagement, median (IQR)				
Coaching sessions	3 (2-7)	4 (2-8)	3 (2-6)	.01
Coaching messages	54 (21-127)	62 (26-161)	48 (18-105)	.003
Clinical appointments	6 (3-12)	7 (3-13)	5 (2-11)	.003
Therapy appointments	6 (2.5-11)	6 (3-12.75)	5 (2-10)	.06
Psychiatry appointments	3 (2-5)	2 (1.25-4.75)	3 (2-5)	.30
Days from intake to last interaction	56 (28-105)	70 (29-126)	43 (16-98)	<.001
Member satisfaction, mean (SD)				
Coach star rating	4.60 (0.68)	4.64 (0.62)	4.55 (0.73)	.04
Clinical star rating	4.75 (0.66)	4.80 (0.61)	4.70 (0.70)	.14
GAD[a] intake, n (%)				<.001
0-4: minimal anxiety	4 (0.25)	3 (0.40)	1 (0.12)	
5-9: mild anxiety	226 (14.03)	134 (17.72)	92 (10.76)	
10-14: moderate anxiety	705 (43.76)	367 (48.54)	338 (39.53)	
15-21: severe anxiety	676 (41.96)	252 (33.33)	424 (49.59)	
PHQ[b] intake, n (%)				<.001
Negative screen	680 (42.21)	348 (46.03)	332 (38.83)	
0-4 minimal or none	3 (0.19)	2 (0.26)	1 (0.12)	
5-9 mild	58 (3.60)	35 (4.63)	23 (2.69)	

Characteristic	All participants (N=1611)	Negative follow-up screen (n=756)	Positive follow-up screen (n=855)	*P* value
10-14 moderate	305 (18.93)	155 (20.50)	150 (17.54)	
15-19 moderately severe	333 (20.67)	133 (17.59)	200 (23.39)	
20-27 severe	232 (14.40)	83 (10.98)	149 (17.43)	

[a]GAD: Generalized Anxiety Disorder.

[b]PHQ: Patient Health Questionnaire.

Table 3 reports the results of our primary model examining the association between care modality and anxiety symptom improvements. All modalities (coaching, clinical, combined) showed increased odds of improvement compared to those who were not engaged with coaching or clinical services ("app only"). A higher GAD-7 intake score was also associated with decreased odds of improvement.

Table 3. Associations between care modality and improvement (n=1611).

Modality	Model 1	
	Odds ratio[a]	95% CI
App only	N/A[b]	N/A
Minimal care utilization	1.45	0.98-2.12
Coaching only	1.90	1.27-2.86
Clinical	1.97	1.32-2.96
Combined	3.26	1.87-5.68
GAD-7 intake score	0.90	0.88-0.92

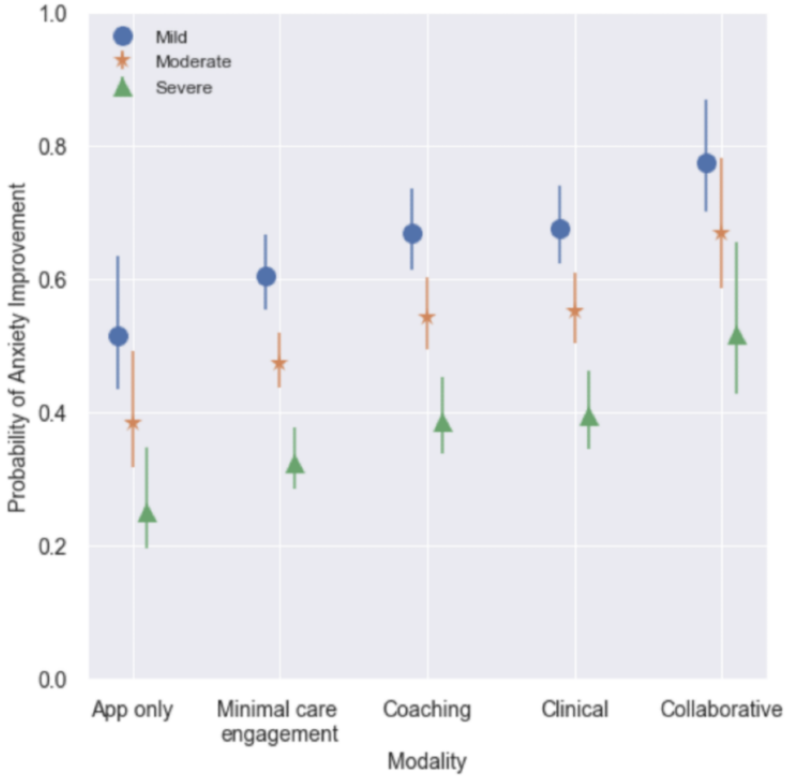
[a]Odds ratio obtained by exponentiation of the regression coefficients.

[b]N/A: not applicable.

A Hosmer-Lemeshow test failed to reject the null hypothesis, indicating goodness of fit, X^2 (8, N=1611)=8.9, P=.35.

These results are shown graphically as probability of anxiety improvement by care modality and levels of intake severity in Figure 1. For all figures, shapes represent the expected mean probability of improvement by modality and intake severity; lines represent the corresponding 95% confidence interval.

Figure 1. Probability of anxiety improvement by care modality and level of intake severity. Shapes represent the expected mean probability of improvement by modality and intake severity; lines represent the corresponding 95% confidence interval.



Coaching-Only Cohort

Table 4 reports outputs for the model examining the association between utilization and anxiety symptom improvement for the text-based coaching-only cohort. Compared to the minimal-utilization reference group (those who sent less than 14 messages to a coach, engaged less than 4 weeks, or the average days between interactions were greater than 14 days), only the high-utilization group had significantly increased odds of improvement. This association remained after adjustment for baseline severity (GAD-7 intake score).

A Hosmer-Lemeshow test failed to reject the null hypothesis, indicating goodness of fit, X^2 (8, N=1080)=9.14, P=.33.

These results are shown graphically as probability of anxiety improvement by coaching utilization and intake severity in Figure 2.

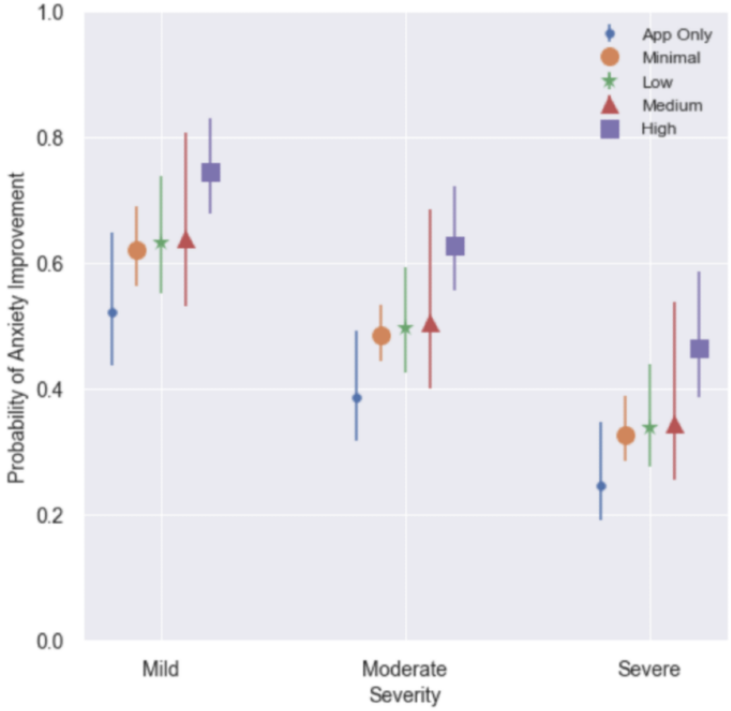
Table 4. Associations between utilization and improvement among the text-based coaching-only cohort (n=936).

Modality	Model 2	
	Odds ratio[a]	95% CI
App Only	N/A[b]	N/A
Minimal	1.50	1.02-2.21
Low	1.57	0.98-2.52
Medium	1.63	0.87-3.08
High	2.70	1.65-4.41
GAD-7 intake score	0.89	0.87-0.92

[a]Odds ratio obtained by exponentiation of the regression coefficients.

[b]N/A: not applicable.

Figure 2. Probability of anxiety improvement by utilization and intake severity among the text-based coaching-only cohort (n=936). Shapes represent the expected mean probability of improvement by modality and intake severity; lines represent the corresponding 95% confidence interval.



Clinical-Only Cohort

Table 5 reports outputs for the model examining the association between utilization and anxiety symptom improvement for the clinical-only cohort. Compared to the minimal-utilization reference group (those who only attended 1 therapy session), those in the low, medium, and high utilization categories had increased odds of improvement, with odds ratios increasing ordinally with each category. After adjusting for GAD-7 intake score, only the high utilization coefficient remained significant.

Table 5. Associations between clinical utilization and improvement among the clinical-only cohort (n=205).

Modality	Model 3	
	Odds ratio[a]	95% CI
App Only	N/A[b]	N/A
Minimal (1 session)	0.75	0.29-1.92
Low (2-6 sessions)	1.91	1.13-3.21
Medium (7-12 sessions)	2.24	1.04-4.83
High (≥13 sessions)	2.30	1.07-4.93
GAD-7 intake score	0.90	0.85-0.96

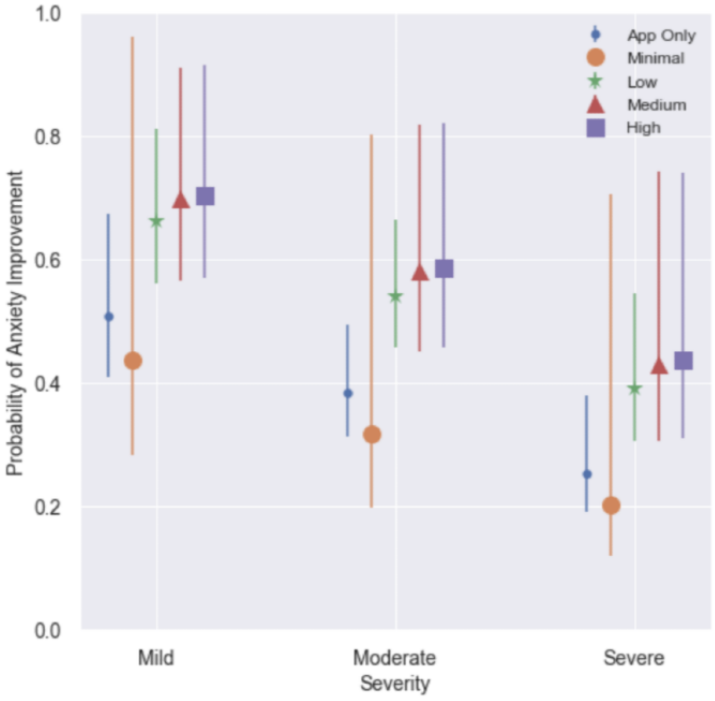
[a]Odds ratio obtained by exponentiation of the regression coefficients.

[b]N/A: not applicable.

A Hosmer-Lemeshow test failed to reject the null hypothesis, indicating goodness of fit, X^2 (8, N=349)=6.44, P=.50.

These results are shown graphically as probability of anxiety improvement by clinical utilization and intake severity in Figure 3.

Figure 3. Probability of anxiety improvement by clinical utilization and intake severity among the clinical-only cohort (n=205). Shapes represent the expected mean probability of improvement by modality and intake severity; lines represent the corresponding 95% confidence interval.



Combined (Coaching and Clinical Services) Cohort

Finally, Table 6 reports outputs for the model examining the association between utilization and anxiety symptom improvement for individuals engaged in both coaching and clinical services. Compared to the low-utilization group, there were significantly increased odds of improvement for the high utilization group but not for the low- and medium-utilization group.

A Hosmer-Lemeshow test failed to reject the null hypothesis, indicating goodness of fit, X^2 (8, N=470)=7.11, P=.53.

These results are shown graphically as probability of anxiety improvement by coaching and clinical utilization and intake severity in Figure 4.

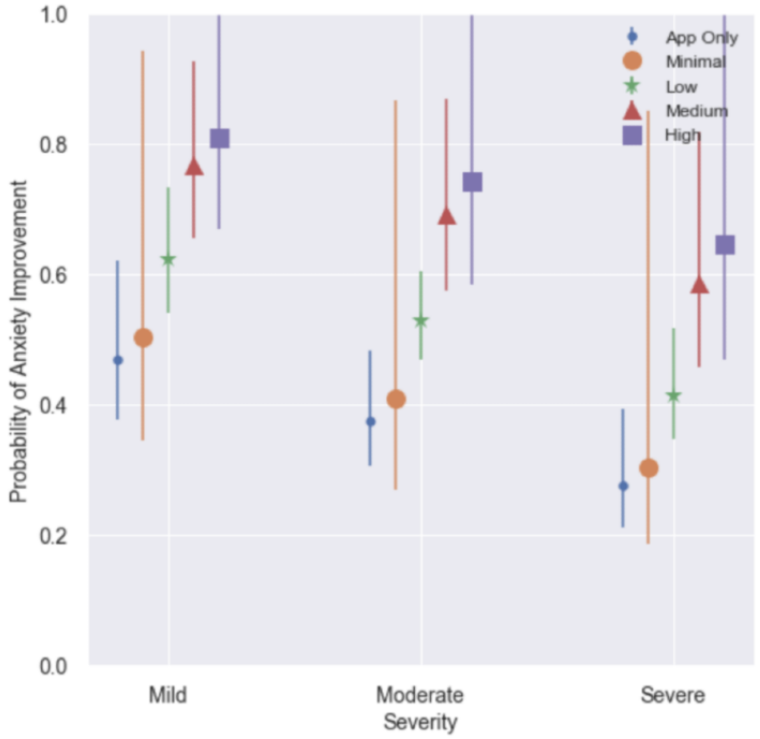
Table 6. Associations between utilization and improvement among the combined-care cohort (n=326).

Modality	Model 4	
	Odds ratio[a]	95% CI
App Only	N/A[b]	N/A
Minimal	1.15	0.44-2.99
Low	1.87	1.22-2.87
Medium	3.76	1.79-7.89
High	4.85	1.6-14.7
GAD-7 intake score	0.93	0.88-0.97

[a]Odds ratio obtained by exponentiation of the regression coefficients.

[b]N/A: not applicable.

Figure 4. Probability of anxiety improvement by utilization and intake severity among the combined-care cohort (n=326). Shapes represent the expected mean probability of improvement by modality and intake severity; lines represent the corresponding 95% confidence interval.



Discussion

Principal Findings

In this study, we examined differences in anxiety outcomes by care modality—coaching, clinical (teletherapy and telepsychiatry), and combined (coaching and clinical) care—within an on-demand mental health system, as well as the association between levels of utilization within each care modality and improvement in anxiety symptoms. Our primary model examining the association between improvement and care modality found increased odds of improvement for all care modalities compared to the reference group (those who did not engage in any coaching or clinical services). This aligns with existing literature that finds that most forms of treatment are better than nothing, further highlighting the need to get even low-intensity treatments to individuals who need help. Our outcomes are also in line with prior observational research estimating GAD recovery rates of 30%-60%, depending on treatment and individual characteristics [36,37].

We also found the largest effect size for the combined-care (coaching and clinical) group, suggesting that engaging in multiple levels of care might be more effective for treating anxiety. More specifically, coaching might provide an added benefit of longitudinal support toward goals between episodic clinical visits. This is a novel finding given the limited research focused on text-based coaching and this form of combined care. It also suggests that more intensive forms of digital mental health services appear to contribute to greater improvement in outcomes, which is important for considering the scalability of these programs. It is important to note that a relatively small percentage of our study cohort was engaged in combined care, suggesting a need to promote this modality more broadly.

Our data suggest that while all treatment modalities appeared to offer comparable benefits in managing anxiety, the largest effects were observed among those who engaged in services for at least 13 weeks of care. This mirrors data from clinical trials of in-person care for anxiety, where the largest effects are for those who receive more frequent sessions. This is also consistent with most evidence-based protocols of 8-12 sessions, as these sessions are not always completed at a weekly cadence [17].

Strengths of this study include the real-world setting and relatively large sample size, which allow us to observe varying levels of digital mental health support for anxiety among individuals seeking care for their mental concerns. While limited in several ways, this design has an important benefit in not being constrained by the strict requirements for controlled clinical trials. Additionally, due to the virtual nature of the system, we have detailed data on coaching utilization (eg, messaging volume and frequency) that likely would not exist for in-person care settings, and the ability to compare multiple modalities of care. This study is also novel in its ability to analyze data for people engaged in multiple modalities of care (ie, teletherapy and text-based coaching).

Limitations

There are several limitations to this study. As our dataset is limited to people who completed surveys, our results are not necessarily generalizable to all members (ie, those who drop off or do not complete surveys). Furthermore, our results cannot generalize to individuals who do not have access to this system. We also had a relatively large amount of missing data for gender and age, which limited our ability to stratify analyses by key demographics. Due to the survey design of this system and efforts to maintain anonymity and protect the privacy of

members, we were unable to study race/ethnicity, socioeconomic status, living situation, history of trauma or other mental illnesses, and other factors that could affect treatment response. However, because the Ginger platform is offered through employers, we know that the survey respondents are working-age adults, suggesting that these findings may generalize to the professional workforce and those enrolled in health benefits through their employer.

Another potential limitation is that we had to rely on GAD-2 rather than GAD-7 to assess anxiety symptom improvement. This is due to the survey system design, which aims to avoid response burnout among users. This limits our ability to assess certain anxiety symptoms; however, GAD-2 represents 2 core anxiety symptoms and has been shown to have good sensitivity and specificity in the diagnosis of the most common anxiety disorders [4].

It is also important to note that we are likely underestimating the number of individuals who experienced a clinically meaningful improvement in their anxiety symptoms, as some members screened positive at follow-up (and those were not classified as "improved" in our models) but experienced a reduction in GAD-7 score. If we include the commonly accepted definition of a 5-point reduction in score, an additional 11.9% (192/1617) of the cohort would be classified as "improved." Finally, we lack a control group to understand what would have happened in the absence of Ginger and to attribute causality, although we are able to understand these associations relative to defined reference groups.

This study segues into many directions of future research. We can build upon our understanding of these associations by increasing the collection of demographic data to enhance our understanding of member utilization patterns and user personas, and by adding content analysis of coaching messages. For example, we might consider looking more specifically and in greater depth at the frequency and intensity of coaching utilization and different patterns of coaching and clinical utilization (sequential vs concurrent). Additionally, we plan to explore how the different treatment modalities contribute to improved outcomes, digging into the "mechanism of action" to better understand how to replicate aspects of the platform that work well to support the larger rollout of these services to users.

Conclusion

This study found increased odds of anxiety improvement for all care modalities compared to those who did not engage in care, with larger effect sizes for higher utilization within all care modalities. Additionally, there is a promising observation that those engaged in combined care (teletherapy and text-based coaching) have the greatest odds of anxiety improvement. Future directions include more detailed classifications of utilization patterns and exploring explanations and solutions for lower utilization members.

Acknowledgments

The authors would like to acknowledge Alexis Hahn and Erica Hayes for their literature review input and support.

Authors' Contributions

SK authored this manuscript. MY conducted the data analysis. JH, WX, DU, AN, PA, and JN provided clinical and technical expertise and supported manuscript review and edits.

Conflicts of Interest

SK, MY, JH, WX, and DU are current employees of Ginger. AN and PA are paid scientific advisors. JN has no conflicts of interest to declare.

References

1. Bandelow B, Michaelis S. Epidemiology of anxiety disorders in the 21st century. Dialogues Clin Neurosci 2015 Oct;17(3):327-335. [doi: 10.31887/dcns.2015.17.3/bbandelow]
2. Baxter AJ, Vos T, Scott KM, Ferrari AJ, Whiteford HA. The global burden of anxiety disorders in 2010. Psychol. Med 2014 Jan 22;44(11):2363-2374. [doi: 10.1017/s0033291713003243]
3. Kroenke K, Spitzer RL, Williams JB, Monahan PO, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. Ann Intern Med 2007 Mar 06;146(5):317-325. [doi: 10.7326/0003-4819-146-5-200703060-00004] [Medline: 17339617]
4. Sapra A, Bhandari P, Sharma S, Chanpura T, Lopp L. Using Generalized Anxiety Disorder-2 (GAD-2) and GAD-7 in a Primary Care Setting. Cureus 2020 May 21;12(5):e8224 [FREE Full text] [doi: 10.7759/cureus.8224] [Medline: 32582485]
5. Anxiety and Depression Association of America. 2016 Aug. URL: https://adaa.org/about-adaa/press-room/facts-statistics [accessed 2020-09-26]
6. Mental Health and Coping During COVID-19. Centers for Disease Control and Prevention. 2020. URL: https://www.cdc.gov/coronavirus/2019-ncov/daily-life-coping/managing-stress-anxiety.html [accessed 2020-09-27]
7. Twenge J, Joiner T. Mental distress among U.S. adults during the COVID-19 pandemic. PsyArXiv . [doi: 10.31234/osf.io/wc8ud]

8. Anxiety and Depression Household Pulse Survey. Centers for Disease Control and Prevention. 2020. URL: https://www.cdc.gov/nchs/covid19/pulse/mental-health.htm [accessed 2020-09-28]
9. Anxiety Disorders. National Institute of Mental Health. 2020. URL: https://www.nimh.nih.gov/health/topics/anxiety-disorders/index.shtml#part_145338 [accessed 2020-09-28]
10. Osenbach JE, O'Brien KM, Mishkind M, Smolenski DJ. Synchronous telehealth technologies in psychotherapy for depression: a meta-analysis. Depress Anxiety 2013 Nov 06;30(11):1058-1067. [doi: 10.1002/da.22165] [Medline: 23922191]
11. Andrews G, Basu A, Cuijpers P, Craske M, McEvoy P, English C, et al. Computer therapy for the anxiety and depression disorders is effective, acceptable and practical health care: An updated meta-analysis. J Anxiety Disord 2018 Apr;55:70-78 [FREE Full text] [doi: 10.1016/j.janxdis.2018.01.001] [Medline: 29422409]
12. 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 Dec 15;218:15-22 [FREE Full text] [doi: 10.1016/j.jad.2017.04.046] [Medline: 28456072]
13. Rollman BL, Herbeck Belnap B, Abebe KZ, Spring MB, Rotondi AJ, Rothenberger SD, et al. Effectiveness of Online Collaborative Care for Treating Mood and Anxiety Disorders in Primary Care: A Randomized Clinical Trial. JAMA Psychiatry 2018 Jan 01;75(1):56-64 [FREE Full text] [doi: 10.1001/jamapsychiatry.2017.3379] [Medline: 29117275]
14. Hill B, Richardson B, Skouteris H. Do We Know How to Design Effective Health Coaching Interventions: A Systematic Review of the State of the Literature. Am J Health Promot 2015 May;29(5):e158-e168 [FREE Full text] [doi: 10.4278/ajhp.130510-lit-238]
15. Kivelä K, Elo S, Kyngäs H, Kääriäinen M. The effects of health coaching on adult patients with chronic diseases: a systematic review. Patient Educ Couns 2014 Nov;97(2):147-157. [doi: 10.1016/j.pec.2014.07.026] [Medline: 25127667]
16. Hoermann S, McCabe KL, Milne DN, Calvo RA. Application of Synchronous Text-Based Dialogue Systems in Mental Health Interventions: Systematic Review. J Med Internet Res 2017 Jul 21;19(8):e267 [FREE Full text] [doi: 10.2196/jmir.7023] [Medline: 28784594]
17. Lattie EG, Graham AK, Hadjistavropoulos HD, Dear BF, Titov N, Mohr DC. Guidance on defining the scope and development of text-based coaching protocols for digital mental health interventions. Digit Health 2019 Dec 16;5:2055207619896145 [FREE Full text] [doi: 10.1177/2055207619896145] [Medline: 31897306]
18. Frieden TR. Evidence for Health Decision Making — Beyond Randomized, Controlled Trials. N Engl J Med 2017 Aug 03;377(5):465-475. [doi: 10.1056/nejmra1614394]
19. Safavi K, Cohen A. The Oversell And Undersell Of Digital Health. Health Aff;(January) 2019:1-9. [doi: 10.1377/hblog20190226.63748]
20. Cohen AT, Goto S, Schreiber K, Torp-Pedersen C. Why do we need observational studies of everyday patients in the real-life setting?: Table 1. Eur Heart J Suppl 2015 Jul 10;17(suppl D):D2-D8. [doi: 10.1093/eurheartj/suv035]
21. National Board for Health and Wellness Coaching. URL: https://nbhwc.org/ [accessed 2020-09-28]
22. Introduction to the Coaching Profession. Institute of Coaching. URL: https://www.instituteofcoaching.org/resources/introduction-coaching-profession [accessed 2020-09-28]
23. Kunkle S, Yip M, Ξ W, Hunt J. Evaluation of an On-Demand Mental Health System for Depression Symptoms: Retrospective Observational Study. J Med Internet Res 2020 Jun 18;22(6) [FREE Full text] [doi: 10.2196/17902] [Medline: 32554387]
24. Löwe B, Decker O, Müller S, Brähler E, Schellberg D, Herzog W, et al. Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. Med Care 2008 Mar;46(3):266-274. [doi: 10.1097/MLR.0b013e318160d093] [Medline: 18388841]
25. Staples LG, Dear BF, Gandy M, Fogliati V, Fogliati R, Karin E, et al. Psychometric properties and clinical utility of brief measures of depression, anxiety, and general distress: The PHQ-2, GAD-2, and K-6. Gen Hosp Psychiatry 2019 Jan;56:13-18 [FREE Full text] [doi: 10.1016/j.genhosppsych.2018.11.003] [Medline: 30508772]
26. GAD Info Sheet. VA Center for Integrated Healthcare (CIH). URL: https://www.mirecc.va.gov/cih-visn2/Documents/Clinical/GAD_with_Info_Sheet.pdf [accessed 2020-09-28]
27. Toussaint A, Hüsing P, Gumz A, Wingenfeld K, Härter M, Schramm E, et al. Sensitivity to change and minimal clinically important difference of the 7-item Generalized Anxiety Disorder Questionnaire (GAD-7). J Affect Disord 2020 Mar 15;265:395-401. [doi: 10.1016/j.jad.2020.01.032] [Medline: 32090765]
28. Bischoff T, Anderson SR, Heafner J, Tambling R. Establishment of a reliable change index for the GAD-7. Psychol. Community Health 2020 Apr 08;8(1):176-187. [doi: 10.5964/pch.v8i1.309]
29. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res 2011 Aug 05;13(3):e52 [FREE Full text] [doi: 10.2196/jmir.1772] [Medline: 21821503]
30. Fuhr K, Schröder J, Berger T, Moritz S, Meyer B, Lutz W, et al. The association between adherence and outcome in an Internet intervention for depression. J Affect Disord 2018 Mar 15;229:443-449. [doi: 10.1016/j.jad.2017.12.028] [Medline: 29331706]
31. Marcelle ET, Nolting L, Hinshaw SP, Aguilera A. Effectiveness of a Multimodal Digital Psychotherapy Platform for Adult Depression: A Naturalistic Feasibility Study. JMIR Mhealth Uhealth 2019 Jan 23;7(1):e10948 [FREE Full text] [doi: 10.2196/10948] [Medline: 30674448]

32. Clinical Practice Guideline for the Treatment of Posttraumatic Stress Disorder: How Long Will It Take for Treatment to Work? American Psychological Association. URL: https://www.apa.org/ptsd-guideline/patients-and-families/length-treatment [accessed 2020-09-28]
33. Tolles J, Meurer WJ. Logistic Regression: Relating Patient Characteristics to Outcomes. JAMA 2016 Aug 02;316(5):533-534. [doi: 10.1001/jama.2016.7653] [Medline: 27483067]
34. Hosmer D, Lemeshow S. Applied Logistic Regression. 2nd ed. John Wiley & Sons 2000:143-164. [doi: 10.1002/0471722146]
35. U.S. Department of Health and Human Services. 45 CFR 46. Office for Human Research Protections. 2016. URL: https://www.hhs.gov/ohrp/regulations-and-policy/regulations/45-cfr-46/index.html [accessed 2020-09-28]
36. Rodriguez BF, Weisberg RB, Pagano ME, Bruce SE, Spencer MA, Culpepper L, et al. Characteristics and predictors of full and partial recovery from generalized anxiety disorder in primary care patients. J Nerv Ment Dis 2006 Feb;194(2):91-97 [FREE Full text] [doi: 10.1097/01.nmd.0000198140.02154.32] [Medline: 16477186]
37. Fisher PL, Durham RC. Recovery rates in generalized anxiety disorder following psychological therapy: an analysis of clinically significant change in the STAI-T across outcome studies since 1990. Psychol Med 1999 Nov 01;29(6):1425-1434. [doi: 10.1017/s0033291799001336] [Medline: 10616949]

Abbreviations

GAD: Generalized Anxiety Disorder
PHQ: Patient Health Questionnaire

Edited by G Eysenbach; submitted 30.09.20; peer-reviewed by MW Haun; comments to author 22.10.20; revised version received 16.11.20; accepted 20.12.20; published 26.01.21.

Please cite as:
Kunkle S, Yip M, Hunt J, Ξ W, Udall D, Arean P, Nierenberg A, Naslund JA
Association Between Care Utilization and Anxiety Outcomes in an On-Demand Mental Health System: Retrospective Observational Study
JMIR Form Res 2021;5(1):e24662
URL: http://formative.jmir.org/2021/1/e24662/
doi:10.2196/24662
PMID:33496679

Original Paper

A Smartphone App to Support Sedentary Behavior Change by Visualizing Personal Mobility Patterns and Action Planning (SedVis): Development and Pilot Study

Yunlong Wang[1*], PhD; Laura M. König[2*], PhD; Harald Reiterer[1], PhD

[1]Department of Computer and Information Science, University of Konstanz, Konstanz, Germany

[2]Department of Psychology, University of Konstanz, Konstanz, Germany

[*]these authors contributed equally

Corresponding Author:
Yunlong Wang, PhD
Department of Computer and Information Science
University of Konstanz
Universitätsstraße 10
Konstanz, 78457
Germany
Phone: 49 7531 88 3704
Email: yunlong.wang@uni-konstanz.de

Abstract

Background: Prolonged sedentary behavior is related to a number of risk factors for chronic diseases. Given the high prevalence of sedentary behavior in daily life, simple yet practical solutions for behavior change are needed to avoid detrimental health effects.

Objective: The mobile app SedVis was developed based on the health action process approach. The app provides personal mobility pattern visualization (for both physical activity and sedentary behavior) and action planning for sedentary behavior change. The primary aim of the study is to investigate the effect of mobility pattern visualization on users' action planning for changing their sedentary behavior. The secondary aim is to evaluate user engagement with the visualization and user experience of the app.

Methods: A 3-week user study was conducted with 16 participants who had the motivation to reduce their sedentary behavior. Participants were allocated to either an active control group (n=8) or an intervention group (n=8). In the 1-week baseline period, none of the participants had access to the functions in the app. In the following 2-week intervention period, only the intervention group was given access to the visualizations, whereas both groups were asked to make action plans every day and reduce their sedentary behavior. Participants' sedentary behavior was estimated based on the sensor data of their smartphones, and their action plans and interaction with the app were also recorded by the app. Participants' intention to change their sedentary behavior and user experience of the app were assessed using questionnaires.

Results: The data were analyzed using both traditional null hypothesis significance testing (NHST) and Bayesian statistics. The results suggested that the visualizations in SedVis had no effect on the participants' action planning according to both the NHST and Bayesian statistics. The intervention involving visualizations and action planning in SedVis had a positive effect on reducing participants' sedentary hours, with weak evidence according to Bayesian statistics (Bayes factor, BF_{+0}=1.92; median 0.52; 95% CI 0.04-1.25), whereas no change in sedentary time was more likely in the active control condition (BF_{+0}=0.28; median 0.18; 95% CI 0.01-0.64). Furthermore, Bayesian analysis weakly suggested that the more frequently the users checked the app, the more likely they were to reduce their sedentary behavior (BF_{-0}=1.49; r=−0.50).

Conclusions: Using a smartphone app to collect data on users' mobility patterns and provide real-time feedback using visualizations may be a promising method to induce changes in sedentary behavior and may be more effective than action planning alone. Replications with larger samples are needed to confirm these findings.

(JMIR Form Res 2021;5(1):e15369) doi:10.2196/15369

KEYWORDS

sedentary behavior; data visualization; mobile app; action planning; human mobility patterns; mobile phone

Introduction

Background

Sedentary behavior refers to any waking behavior characterized by an energy expenditure ≤1.5 metabolic equivalents while in a sitting, reclining, or lying posture [1,2]. Studies have shown evidence of the detrimental effects of prolonged sedentary behavior, which is ubiquitous in daily life, especially when at work. For instance, a study [3] tracking 425 adults for 10 years (2002-2004 to 2012-2014) showed that a greater increase in sedentary behavior was associated with detrimental changes in cardiometabolic risk factors, such as waist circumference, high-density lipoprotein cholesterol, and triglycerides, independent of the change in moderate-to-vigorous physical activity. In other words, exercising after sitting for a prolonged time while at work might not reduce the health risk caused by prolonged sitting. Moreover, a study [4] involving 168 participants in Australia showed that the total number of breaks in sedentary time was associated with improved health parameters, such as significantly lower waist circumference, BMI, triglycerides, and 2-hour plasma glucose. Consequently, several governments (eg, Australia [5] and Canada [6]) have released guidelines to specifically reduce people's sedentary behavior for improved health. For example, people with a sedentary lifestyle could introduce light physical activity (eg, short walking) throughout the day to reduce the risk of many chronic diseases.

The high prevalence of sedentary activities in daily life leads to a stronger habit of sedentary behavior [7], that is, a high degree of automaticity owing to frequent repetition in a stable context [8], which makes it difficult to change in the long term [9]. Interventions are, therefore, needed to support individuals to reduce their sedentary time. In their review, Chu et al [10] divided intervention strategies for reducing sedentary behavior into 3 categories: (1) educational or behavioral (eg, goal setting, action planning, and self-monitoring), (2) environmental changes (eg, sit-stand workstation and treadmill desk), and (3) multi-component (eg, sit-stand workstation plus goal setting). Environmental and multi-component interventions might require policy support and additional facilities, which might hinder their immediate application on a larger scale. Therefore, simple yet practical solutions are needed.

Mobile devices, including smartphones and wearables (eg, smartwatches and fitness wristbands), might be useful platforms for sedentary behavior interventions. First, the prevalence of both smartphone and wearable device ownership is increasing globally [11]. As smartphones include sensors that allow for the collection of physical activity data [12], smartphone owners do not need additional devices to collect data and receive interventions, thus making the solution simple to deliver and practical to use. Second, interest in mobile apps targeting lifestyle behaviors such as physical inactivity is high [13]. Accordingly, research on digital solutions for the promotion of physical activity and the reduction of sedentary behavior is increasing [14]. However, compared with the number of apps targeting physical activity, there are only a few apps specifically targeting sedentary behavior [15]. Moreover, as previous reviews noted, both commercially available apps and apps developed for research are often not grounded in theory, which might limit their effectiveness [16]. This study, therefore, sought to develop a mobile app for sedentary behavior change that is grounded in behavioral theory. Specifically, it sought to integrate the parameters of action planning and the visualization of users' sedentary behavior patterns to better support sedentary behavior change.

Action Planning for Sedentary Behavior Change

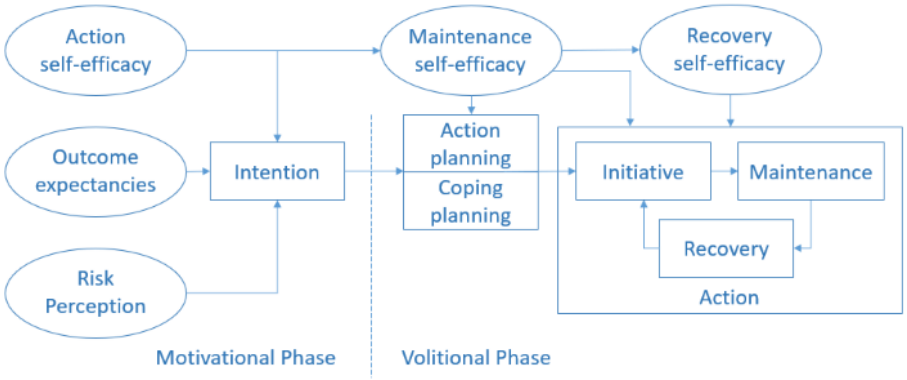
Wang et al [17] recently proposed a holistic framework for developing digital health behavior interventions, drawing from several classic theories of health behavior change in psychology, such as social-cognitive theory [18,19] and the health action process approach (HAPA; Figure 1) [18]. The latter theory is especially important for the design of health behavior interventions as it bridges the intention-behavior gap through action planning [20]. Several meta-analyses have shown that action planning is positively related to goal attainment and health behavior change [21-23] and thus might be an effective behavior change technique. Accordingly, action planning was included in the taxonomy of behavior change techniques [24]. An action plan combines specific situation parameters (*when* and *where*) and a sequence of actions (*how*) for a target behavior [25]. In this vein, it is suggested that behavior will be triggered automatically when encountering specific situations [26].

Although action planning is an effective behavior change technique, there are only a few studies that included action planning in digital interventions targeting sedentary behavior [27]. In a recent systematic review of digital technologies supporting health behavior change [28], only 2 out of the 45 studies reviewed involved action planning related to sedentary behavior change. On the basis of the step counts at baseline, Aittasalo et al [29] offered participants visual feedback to facilitate action planning, whereas De Cocker et al [30] used several motivational questions to stimulate the participants to make action plans. In both studies, sedentary behavior was successfully reduced. However, both used action planning as one of the several behavior change techniques, and it is, therefore, unclear whether the change can be solely attributed to action planning. Maher and Conroy [7], on the other hand, specifically tested the main effect of action planning on reducing sedentary behavior and found that daily action planning did not induce sedentary behavior change. This study, however, has limitations. First, sedentary behavior was only assessed subjectively, which might not correspond to objectively measured behavior [31]. Second, the quality of the action plans was not evaluated, which might preclude important insights into why the intervention was not successful.

The quality of an action plan can be evaluated based on plan characteristics such as the specificity of the situational parameters; plan instrumentality, that is, the degree to which a

plan is helpful to achieve the desired outcome; and viability, that is, how realistic an action plan is. Fleig et al [32] showed that specificity of when to perform a behavior and instrumentality of the action plan were related to an increased likelihood of plan enactment. Quality of action plans, therefore, might be an important variable to consider when evaluating interventions. Although none of the aforementioned studies on sedentary behavior change investigated the quality of action plans, this study aims to test the effect of action planning on sedentary behavior change quantitatively and additionally include a qualitative analysis of the action plans to determine their specificity, instrumentality, and viability.

Figure 1. The model of the health action process approach.



Visualizations of Mobility Patterns

Human mobility patterns reflect the spatial and temporal periodicity or routines of human activities in their daily lives [33-35]. Mobile devices allow for the passive monitoring of human mobility, including physical activity and sedentary behavior. When fed back to the user, the data might help them to generate meaningful insights about their activity patterns and subsequently induce behavior change. Self-monitoring and feedback based on the collected data are frequently used to change physical activity and sedentary behavior [36-39]. However, the feedback is often numerical or uses simple static visualizations such as bar charts or line graphs to display step counts or energy expenditure (eg, Google Fit and Fitbit). On the basis of this information, it might be difficult to extract all relevant information needed to formulate effective action plans defining the *when*, *where*, and *how*. It could be hypothesized that map-based visualizations, such as visualizations provided by apps to track running, might be more effective, as they provide information about where activities took place [40].

Building upon the idea that visualizations of sedentary behavior data might facilitate action planning [29], a novel tool to support action planning for reducing sedentary behavior using interactive visualization was developed. This study thus extends previous mobile sedentary behavior interventions by using an interactive visualization of sedentary behavior data to specifically support daily action planning, which in turn was hypothesized to reduce sedentary time in daily life. A mobile app, SedVis, was implemented by the study team. Mobility patterns were determined based on objective data collected by the app: using internal sensors of the smartphone and existing services provided by the operating system, SedVis automatically tracks and classifies users' activity (eg, walking, biking, and being in a vehicle), step count, and location. In this vein, it determines locations and time windows in which users are sedentary. The visualization elements thus correspond to the aforementioned action planning factors—when, where, and how (ie, the planned activity). By specifically highlighting situations in which users are sedentary, visualizations can serve as a visual aid for formulating action plans. To the best of our knowledge, SedVis is the first app targeting visualizations and action planning on mobile devices for sedentary behavior change.

Study Objectives

This paper reports on the results of a 3-week user study of SedVis (N=16). Specifically, the study aims to answer 4 research questions (RQs). The first aim is to examine the effect of SedVis on users' action planning for their sedentary behavior change (RQ1). Specifically, we tested whether using the visualization improved 3 characteristics of action plans that have been identified as potentially impacting the effectiveness of the plans for behavior change [32]: (1) specificity, that is, the level of detail the plan provided on when and where the behavior was to be shown; (2) instrumentality, that is, the degree of helpfulness of the plan for behavior change; and (3) viability, that is, the degree of control an individual has over plan enactment of formulated action plans. Second, we tested whether the intervention involving visualizations and action planning is effective in reducing sedentary behavior compared with action planning without visualizations (RQ2). Third, because the designed visualizations could also serve as a self-monitoring tool, users' engagement with the visualizations in SedVis and its impact on users' sedentary behavior change was investigated (RQ3). Four, user acceptance and experience of SedVis as a simple intervention tool for the daily use of the sedentary population were studied (RQ4).

Methods

SedVis App

SedVis was developed and implemented by the study team. Specifically, YW developed the app concept and programmed the app. LK and HR tested the app and provided critical feedback. App development was guided by the following requirements that were derived from the literature: (1) develop an app to reduce sedentary behavior, (2) grounded in HAPA with a focus on action planning, (3) using mobility pattern

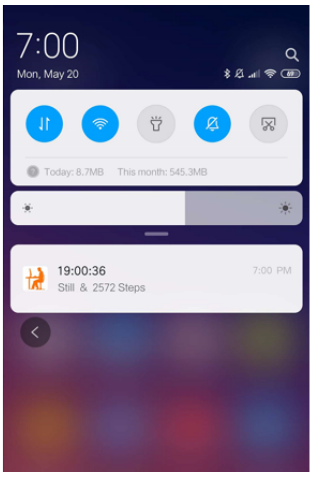
visualizations, and (4) that is simple and practical to use to ensure user acceptance and use in daily life.

Data Collection

SedVis was developed for Android smartphones and pretested internally by the study authors. It collected the data of physical activities (via Google Activity Recognition application programming interface [API]) [41], geolocation (via Google Maps API), steps (via Google Fit API), screen states (turned on or off), users' interaction within the app, users' action plans, and time stamps. On the basis of the built-in sensor data, the Google Activity Recognition service on the Android platform could recognize physical activities, including running, walking, cycling, being in a vehicle, and being still. As high battery consumption (eg, through constant geolocation tracking) or large disk-space requirements might lead to users abandoning the app, geolocation was only updated when movements were detected based on activity recognition and steps counting every 5 seconds. In addition, a new data point was only recorded when a change in the activity state was detected (eg, the steps increase or the physical activity changes). This strategy minimizes energy consumption and data storage without losing information on users' mobility [42].

To improve power consumption, Google imposes limitations on background services since Android 8.0. Some original equipment manufacturer versions of Android (eg, Xiaomi MIUI [43] and Huawei EMUI [44]) additionally introduced limitations on background services to optimize the battery life. The operating system might automatically kill the background service. Therefore, the logged data might not indicate the difference between true sedentary periods and the periods during which the background service was not running. Therefore, a timer was added to the background service to log a timestamp to the local database every 20 min. To improve the data collection quality, a data collection service was bound to a notification showing the latest update time, steps, and activity in the notification bar (as shown in Figure 2). A system clock was used to monitor if the background service was running and, if necessary, to initiate a restart. Users could also manually restart the data collection service if the notification disappeared.

Figure 2. The always-on notification of SedVis on a user's smartphone.



Mobility Pattern Detection

Mobility patterns refer to when, where, and how the user moves or is sedentary, which directly corresponds to the 3 elements in action planning (ie, when, where, and how). In SedVis, this involved tracking of users' moving trajectory and sedentary place detection. The trajectories showed the routes the user had taken and related information on step counts and time windows. The app detected the users' physical activity every 5 seconds, which enabled a high temporal resolution for trajectory tracking. Modern smartphones use high-precision and low-power movement sensors, which make physical activity recognition and step tracking both accurate and efficient [45]. Google Play services provide fused location tracking by using GPS, Wi-Fi, and cellular signals to allow for precise positioning even in some indoor environments (accuracy depends on the strength of the indoor Wi-Fi and cellular signals).

Custom programmed sedentary place detection was used to detect the participants' sedentary places based on the users' geolocation data. Many office workers spend the day in a limited number of locations (eg, home, office, and lab) where they spend much time sitting. Existing services, such as the Places software development kit for Android [46], only recognize public places (eg, the university), which could not enable personalized place detection in other places such as at home. Therefore, a spatio-temporal data clustering algorithm [47] was used to detect the places based on each user's data. These detected sedentary places, which were displayed in mobility pattern visualizations, provide users with intuitive cues on where to reduce their sedentary behavior.

Mobility Pattern Visualization

Within SedVis, users could access 2 visualizations of data on their sedentary and active hours that were generated based on the collected mobility pattern data. An hour was labeled sedentary if the user took fewer than 250 steps per hour as in the Fitbit mobile app and according to recent evidence suggesting that 2-min walking (about 250 steps) per hour might lower the risk of premature death [48].

Participants could access the single-day visualization via the dashboard or by clicking the always-on notification (Figure 2). In the daily visualization, the tracked trajectories and the

detected sedentary places are shown on a map, and the corresponding temporal information is shown using a bar chart (as shown in Figure 3) for a single day. Specifically, sedentary hours were marked by orange bars, and sedentary locations were marked with orange triangles on a map to highlight situations in which users were sedentary. Participants could interact with the visualization by tapping on the bar chart or on the locations and trajectories displayed on the map. Specifically, they could see (1) the active hours and the corresponding routes on the map once clicking on a blue bar and (2) the sedentary hours and the corresponding locations once clicking on an orange bar. Likewise, clicking the sedentary location on the map highlighted the corresponding sedentary hours in the bar chart. Although the bar chart illustrated temporal patterns, the map demonstrated spatial patterns. Participants could switch between days by tapping on the arrows at the bottom of the screen.

In the multi-day visualization, data were aggregated across multiple study days. Sedentary places were determined based on aggregated data from the user-selected days. Differing from the daily visualization, the bar chart in the multi-day visualization showed the frequency of the user being sedentary in each hour during the selected days for all the places or one selected place (Figure 4).

Figure 3. The mobility patterns in the daily visualization mode.

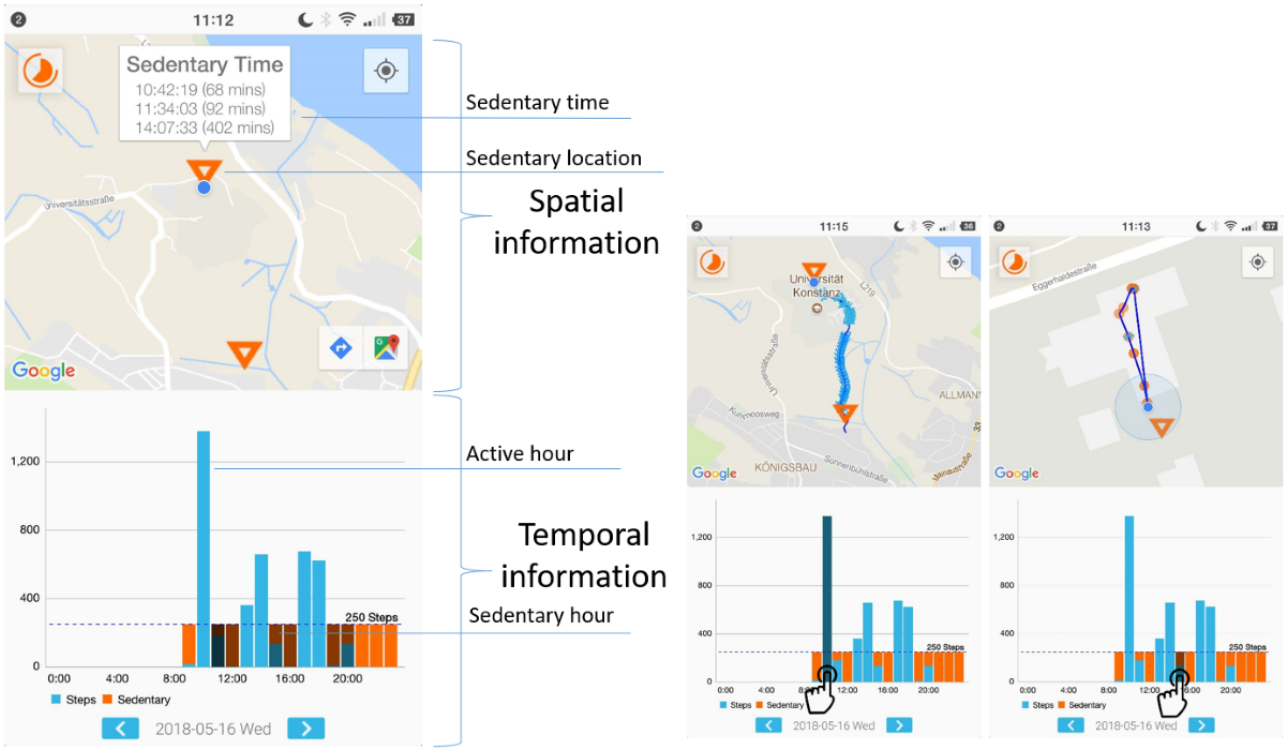
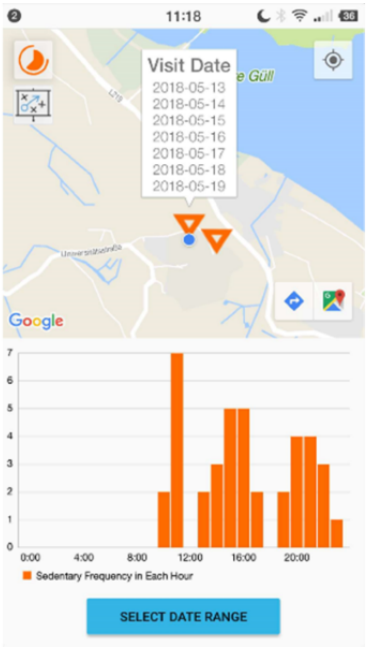


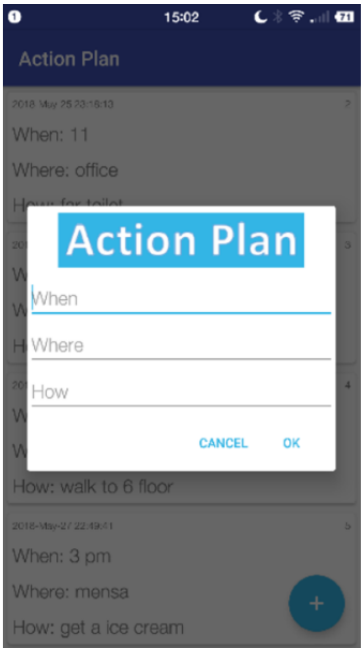
Figure 4. The multi-day visualization.



Action Planning

The user could enter the action planning view from the dashboard, the daily notification, or the shortcuts in the visualization views (see the second button on the left-side corner of Figure 3 and Figure 4). All action plans that the user had made were shown in a list view. The action plans were shown chronologically and could not be deleted. When adding an action plan, the user was asked to specify the *when*, *where*, and *how* elements (Figure 5).

Figure 5. The action planning function in the app.



Dashboard

From the dashboard of SedVis, participants could access all the functionalities of the app, as shown in Figure 6. In the settings tab, the study staff could enable intervention functions. Passwords were used to restrict users' access to these functions during the study.

Figure 6. The dashboard of SedVis.

Study Design and Procedure

The study deployed a mixed design with one between-subject factor *group* (with vs without visualization) and one within-subject factor *time* (baseline vs intervention; Figure 7). Participants were assigned to 1 of the 2 groups, which determined the intervention they received: group A (intervention group), for which the visualization functions were enabled, and group B (active control group), for which the visualization functions were disabled. Participants were assigned to the groups according to the enrollment time (ie, every odd number was assigned to group A, whereas every even number was assigned

to group B). This strategy enables fast study deployment for each participant while maintaining the balance of sample size in both groups [49]. As the sequence of the participants enrolled in the study was random, this strategy preserved the randomization of group allocation.

The study included 3 interviews (ie, the entry interview before starting data collection, the after-baseline interview after week 1, and the exit interview after week 3) on day 1, day 9, and day 25 for each participant. The data collected on these 3 days in the app were excluded from the data analysis because they were incomplete and could not be compared between participants because of appointments being scheduled throughout the day.

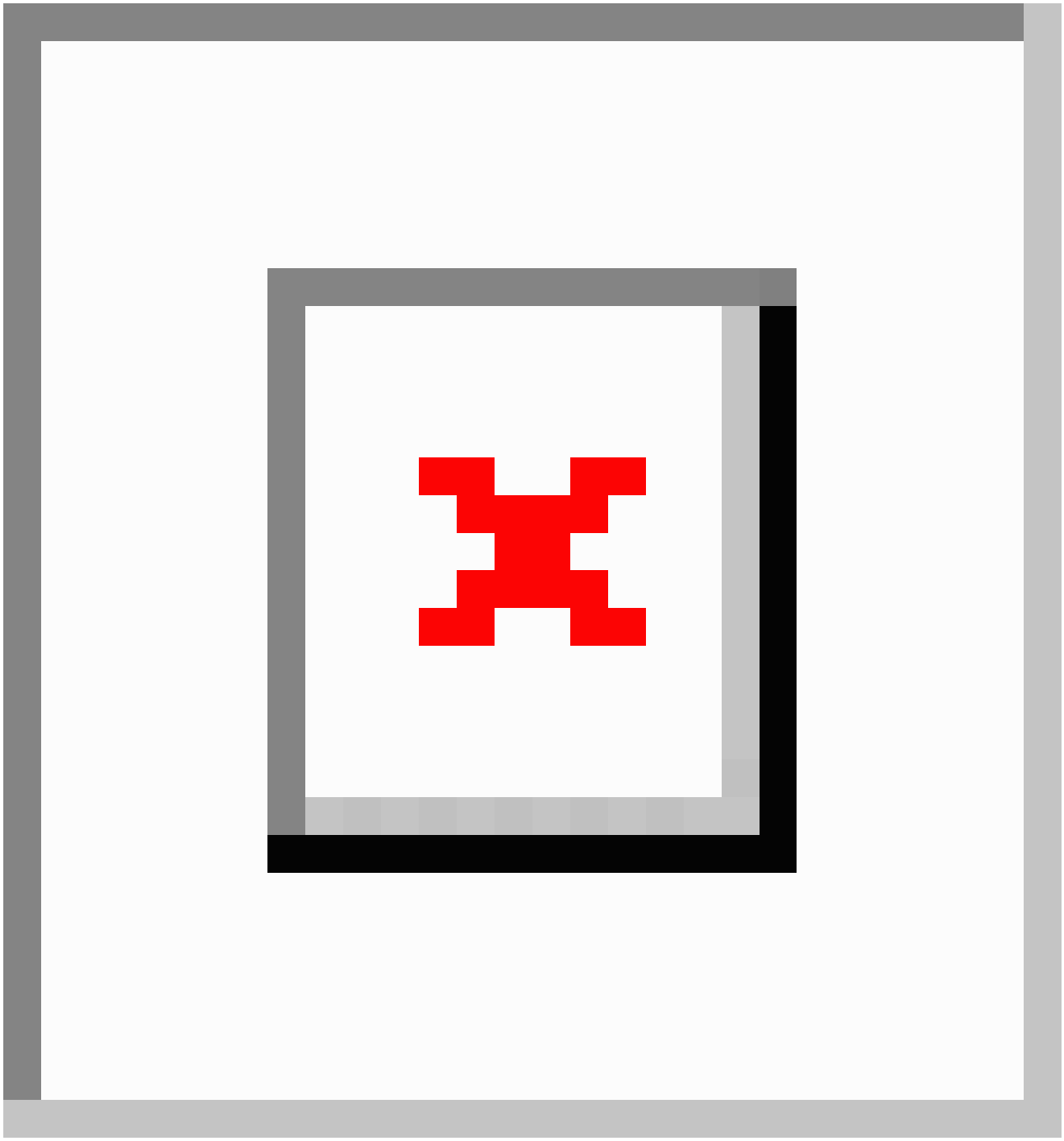
During the entry interview, the participants were informed about the purpose of the study, signed the consent form, and filled out questionnaires on demographics and psychosocial variables related to sedentary behavior (intention, risk perception, and self-efficacy based on HAPA; only results for intention are reported as it is the only construct directly associated with action planning) [25]. A member of the study staff then installed SedVis on their smartphones.

The after-baseline interview took place on the first day after the baseline week. Participants again filled out the questionnaire on psychosocial variables before watching an educational video about the risks of prolonged sedentary behavior [50]. Subsequently, the study staff showed them a flyer to explain behavior change theory [51] and emphasized the importance of action planning. The participants were asked to make at least one plan per day to reduce their sedentary behavior for the following 2 weeks. Finally, the study staff introduced the functions of the app, depending on which group participants of the session were assigned to. For group A, all the functions were activated, including daily visualization; multi-day visualization, which allowed for displaying mobility patterns for multiple days; and action planning. For group B, only the action planning function was enabled. Participants were demonstrated how to make an action plan in the app with dummy examples (eg, *10 am, office, take a walk*). For both groups, participants were asked to set a daily reminder within the settings of the app when they used it for the first time, which served as a prompt to make action plans.

After 2 more weeks, the participants returned to the lab for the exit interview, when they again completed a questionnaire on psychosocial variables as well as an additional questionnaire on user experience. They further transferred the data stored on their smartphones to the study team by email and took part in a short, semistructured interview. Participants were asked questions about their current health status, especially regarding acute infections that might have limited their physical activity; changes in daily routines that might have affected their physical activity or sedentary behavior; and divergences from their sleeping habits, for example, having slept longer or shorter than usual. In addition, they were provided with a list of their action plans and asked to rate them. The participants' answers were written on printed forms and archived into digital forms after the study. Each participant received €20 (US $25) after completing the study.

The ethics committee of the University of Konstanz approved the study protocol. For privacy reasons, only data related to the study were collected. To ensure transparency of data collection, data were recorded and stored on the participants' smartphones until the study was completed. The participants were shown the data details when they transferred the data via email to the study staff. The data remained anonymous and stored on the encrypted server hosted in the university.

Figure 7. The mixed design of the 3-week study.



Participants

Participants were recruited through university mailing lists, the authors' social media profiles, and posters in the university. A total of 16 participants expressed interest in taking part in the study. Participants were eligible for participation if they (1) had the intention to change their sedentary behavior, (2) had no injuries that precluded them from being physically active, (3) were able to speak English fluently, (4) had a smartphone with Android 6.0 and above, (5) did not use a standing desk, and (6) had no travel plans during the study period. The fifth criterion was used to filter out people who had already started to change their sedentary behavior. The other criteria were used to control the motivation and objective ability for using the app, communicating with the study staff, and changing sedentary behavior. The criteria were listed in the study advertisement, and potential participants self-evaluated whether they fit the inclusion criteria. In addition, the intention to change sedentary behavior was assessed in the entry interview as a control measure.

All 16 participants were students (9 out of 16 PhD students and 9 out of 16 females) at the university. Group A comprised 5 females and 3 males. Their mean age was 26.6 years (SD 3.8). Group B comprised 4 females and 4 males. Their mean age was 27.0 years (SD 4.0). Among the 16 participants, one was overweight (ie, BMI>25 kg/m^2), one was underweight (ie, BMI<18.5 kg/m^2), and the remaining had a normal weight (mean BMI of 22.0 kg/m^2, SD 2.8).

Measures

Sedentary Behavior

Sedentary hours were assessed throughout the 3 weeks of the study and calculated based on the step counts assessed by the SedVis app, which were again determined based on the Google Activity Recognition API native to Android smartphones. Studies have shown that off-the-shelf smartphones and smartwatches could provide a reliable estimation of users' physical activity [45,52]. Sedentary behavior was quantified per hour: an hour was labeled as sedentary if less than 250 steps were recorded.

It should be noted that the sedentary hours the app estimated included the participants' sleeping time. It was assumed that the participants' sleeping time did not change over the 3-week study period, which was confirmed by the participants in the exit interview. Thus, the difference in the daily sedentary hours between the baseline and intervention weeks should not be influenced by the sleeping hours. The estimated sedentary hours will be used to reflect sedentary behavior in the rest of the paper.

Number of Action Plans

The total number of action plans formulated during the 2-week intervention phase was counted automatically by the SedVis app. As participants were allowed to repeat the plans of previous days, the number of unique action plans was calculated additionally.

Quality of Action Plans

To evaluate the quality of the action plans, the specificity of the when, where, and how of the plans were coded. The rating criteria for 3 levels of specificity (ie, vague, medium specific, and highly specific) were adapted from Fleig et al [32] (see Table 1 for coding criteria). In addition, participants were asked to evaluate the viability (how realistic) and instrumentality (how useful) of their action plans based on the plan characteristics used by Fleig et al [32]. For viability, participants were asked to rate each action plan on a scale from 1 (not realistic at all) to 4 (very realistic). For instrumentality, participants were asked to rate each action plan on a scale from 1 (not helpful at all) to 4 (very helpful).

Table 1. Coding criteria for specificity.

Specificity type	Vague (1[a])	Medium specific (2[a])	Highly specific (3[a])
"When"	Empty; "Now"; "Anytime"; "Today"	"Every Hour"; "After Lunch"	Timepoint (eg, "13:00")
"Where"	Empty; "Out"	Large area (eg, "City," "University")	Places (eg, "Post," "Lab," "Office," "Home," "Library")
"How"	Empty	"Going to the park"	Activity (eg, "Walk," "Yoga," "Cycle," "Push-ups," "Stretch," "Stand up")

[a]The numbers are the rating levels corresponding to vague, medium specific, and highly specific.

Engagement With the App

Participants' interaction with the app was quantified by recording all operations in the app during the study, including how often the participants checked the visualizations. In addition, the timestamps of when participants made action plans were logged, which were then used as the basis for discussing the users' experience with the app during the exit interview.

Intention to Change Sedentary Behavior

The participants' intention was measured using a scale from 1 ("I do not plan to reduce my sedentary behavior at all") to 4 ("I do exactly plan to reduce my sedentary behavior") following the example in HAPA [25]. The intention was used as a control measure, as participants were required to be motivated to reduce their sedentary behavior instead of other factors (eg, receiving monetary compensation).

User Experience: Quantitative Measure

Using the user experience questionnaire (UEQ) [53], the user experience of the app was quantified at the exit interview.

User Experience: Qualitative Interviews

In addition, closed- and open-ended questions were used to explore the participants' attitudes toward the app and the study as well as their desired features missing in the app: (1) Would you like to receive a reminder for performing the action plans?, (2) Do you want to continue using this app? Why?, (3) Do you think that the logged data on sedentary time and location were accurate? (only group A), and (4) Did you always take your phone with you during work? Replies were recorded as written notes by the interviewer; most replies were either yes/no or statements of 1 to 2 sentences, for example, "The app underestimated the number of steps because I cannot take my phone with me during experiments." For questions that were usually answered with yes or no, the number of participants replying with either option is reported. Owing to the small sample size and limited number of statements exceeding yes/no, responses were only aggregated if they addressed the exact same issue (eg, the smartphone's sensor not being sensitive enough to properly capture nonsedentary periods); otherwise, individual statements are reported.

Qualitative Control Measures

In addition, participants were asked to report unexpected issues that might have affected the data quality during the study: (1) Were your daily routines during the study, including your sleep, typical or not? and (2) Did you have to complete urgent tasks (eg, related to PhD thesis) during the study? This information was recorded to potentially inform interpretation of the data, for example, to explain divergences in step counts between weeks that may mask intervention effects.

Statistical Analysis

Data were analyzed using both traditional null hypothesis significance testing (ie, nonparametric Mann-Whitney U tests, Wilcoxon signed-rank tests, and Spearman rank-order correlations to account for the small sample size) and equivalent Bayesian statistics to provide Bayes factors (BFs). In addition, descriptive statistics (ie, median, mean, and SD) were reported, as suggested by Lee et al [54]. RQ1 was evaluated using the Mann-Whitney U test with the independent variable group and dependent variables total and unique number of action plans and measures of action plan quality. RQ2 was evaluated using a Wilcoxon signed-rank test with study conditions as the independent variable and sedentary hours as the dependent variable. RQ3 was evaluated using Spearman rank-order correlation test to examine the relationship between the frequency of checking the visualization and the sedentary hours. RQ4 was evaluated using the Wilcoxon test with the independent variable group and the UEQ scores as dependent variables. JMP Pro (version 14.1.0 [55]) was used for statistical analysis. The normalized statistics (ie, the *Z* scores) of the nonparametric null-hypothesis significance tests will be reported along with the *P* values.

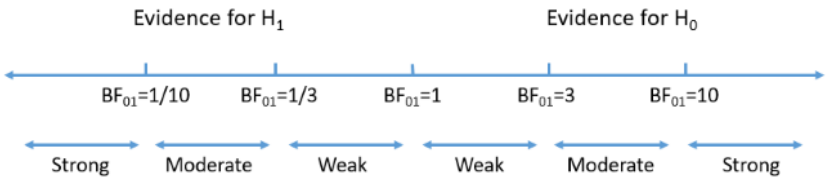
We adopted BFs as complementary statistics. The conventional null hypothesis significance tests provide little information when the result is not statistically significant; only the alternative hypothesis is tested [56]. Nonsignificant results might support a null hypothesis over the alternative, or the data are just insensitive. In contrast, BFs [57] compare the extent to which the samples support 2 hypotheses (eg, equal or different). Moreover, Bayesian methods allow more principled conclusions from studies with a small sample size of novel techniques in the field of human-computer interaction [58]. Therefore, BF was used in addition to the *P* value [59] and Cohen *d* [60] to report and interpret the results. JASP (Jeffreys' Amazing Statistics Program; version 0.9.2) [61] was used for Bayesian analysis.

BF is the ratio of likelihood probabilities.  is the probability of the null hypothesis (H_0) given the data, whereas  is the probability of the alternative hypothesis (H_1) given the data. The definition of BF is shown in formula 1 below:

BF indicates which hypothesis is supported more by the data. Figure 8 shows the BF classification and the adapted interpretation [62]. The default prior distributions of the alternative hypothesis and the calculation methods for different study designs can be found in the study by Rouder et al [63,64].

The default Cauchy distribution,  was used as the prior distribution when estimating the effect size. Following the JASP guidelines [62], the posterior median and the 95% CI of the effect size are also reported. For correlation analysis, the Bayesian Pearson correlation test was used with the default prior distribution suggested by Rouder and Morey [65]. Depending on the context, the one-side BF (BF_{-0} or BF_{+0}) or the two-side BF (BF_{01}) is reported.

Figure 8. A graphical representation of a Bayes factor classification and interpretation. BF: Bayes factor; H0: null hypothesis; H1: alternative hypothesis.



Results

Data Collection

All participants completed the study. First, data quality was checked based on the actual running duration of the app to ensure that all participants had access to the app as expected. The missing duration may be caused by the smartphone being switched off or the background service being shut down for battery optimization. Only the data from one participant in group A (A8) showed a relatively low coverage (65.61% of the study duration); for the other participants, the mean coverage was 93.88% (SD 5.44%). After checking the data of the participant A8, it was found that they had the habit of shutting down the phone during the night. Therefore, the missing data did not limit conclusions about the mobility of the participant, and the participant's data were analyzed as planned. No participant reported urgent tasks or long-term travel that might have impacted their daily routines during the study. Therefore, it is unlikely that the potential sedentary behavior change during the study was owing to reasons unrelated to the intervention.

Participants' Intention (Control Measure)

The participants' intention to reduce sedentary behavior was generally high (median 3.00, mean 3.20, SD 0.59) in both groups. No significant difference was found between groups at each appointment according to Mann-Whitney U tests (appointment 1: Z=0.06, P=.95; appointment 2: Z=0.35, P=.72; appointment 3: Z=0.06, P=.95). The BFs showed evidence preferring H_0 (appointment 1: BF_{01}=2.34; appointment 2: BF_{01}=2.23; appointment 3: BF_{01}=2.34). The results indicated that the participants in both groups had similar strong intentions.

RQ1: Effect of Visualization on Participants' Action Planning

The first aim was to investigate the effect of visualizations on participants' action planning. Both the quantity and quality of the action plans were evaluated.

The Mann-Whitney U test showed no significant group difference regarding the total number of action plans made in the 2 groups (Z=−0.37; P=.71; $median_{group\ A}$ 8.5; $mean_{group\ A}$

8.9, $SD_{group\ A}$ 5.69; $median_{group\ B}$ 5.5; $mean_{group\ B}$ 7.8, $SD_{group\ B}$ 6.76). BF (BF_{01}=2.24; median 0.11; 95% CI −0.66 to 0.92) showed weak evidence toward no difference (H_0). This was the same case for the number of unique action plans (Z=−1.06; P=.29; BF_{01}=1.81; median 0.28; 95% CI −0.50 to 1.14; $median_{group\ A}$ 5.0; $mean_{group\ A}$ 3.8, $SD_{group\ A}$ 2.19; $median_{group\ B}$ 2.5; $mean_{group\ B}$ 2.8, $SD_{group\ B}$ 2.38).

The quality of the action plans showed mixed results, as shown in Table 2. The means of the perceived viability and instrumentality were slightly higher in group B than in group A. The results of the Mann-Whitney U test suggested a statistically significant difference in perceived viability between group A and B. BF also showed weak evidence toward difference (H_1) for perceived viability, whereas it suggested no difference (H_0) for perceived instrumentality. In addition, no meaningful group differences were found regarding specificity (*When* and *Where*). The means of the specificity of the response activity (*How*) were both very high in the 2 groups because most of the users simply specified the activity as walking.

Table 2. The measurements of the quality of the action plans.

Measure	Group A, mean (SD)	Group B, mean (SD)	Z^a	P value[b]	BF^d_{01}	Median	95% CI
Perceived viability	3.28 (0.68)	3.81 (0.37)	1.98	*.048*[c]	0.71	−0.67	−1.72 to 0.17
Perceived instrumentality	3.10 (0.55)	3.22 (0.73)	0.53	.60	2.13	−0.15	−1.00 to 0.59
Specificity (when)	2.55 (0.70)	1.88 (1.00)	−1.11	.27	1.05	0.51	−0.30 to 1.50
Specificity (where)	2.21 (0.82)	2.10 (0.91)	−0.11	.91	2.24	0.10	−0.68 to 0.94
Specificity (how)	2.99 (0.04)	3.00 (0.00)	0.88	.38	—[e]	—	—

[a]Z refers to the normalized statistic of Mann-Whitney U test.

[b]P value of Mann-Whitney U test.

[c]An italicized P value indicates significant difference (P<.05).

[d]BF: Bayes factor.

[e]—: For specificity (*how*), no results are reported for BF because the SD in group B was 0.

Therefore, regarding RQ1, no statistically significant effects of the visualizations in SedVis on the participants' action planning were found, except for perceived viability. BFs indicated weak evidence toward no difference between the 2 groups, except for perceived viability.

Regarding the specificity (*When*) of action plans, some unexpected patterns were observed, especially in group B. Two participants (B3 and B6) in group B always entered the current time when they made the plan. They explained in the exit interview that each of their plans was actually what they were about to do at the moment when they logged the plan. Participant B6 further commented that she found it difficult to make action plans for the future because she was not sure about her behavior patterns. In addition, another 3 participants always used vague cues to specify the *When*: participant B1 used *today*, participant B4 used *anytime*, and participant A5 used *today*.Table 3 shows a summary of the *When*, *Where*, and *How* in the participants' action plans.

Table 3. Summary of "When," "Where," and "How" components identified in the participants' action plans.

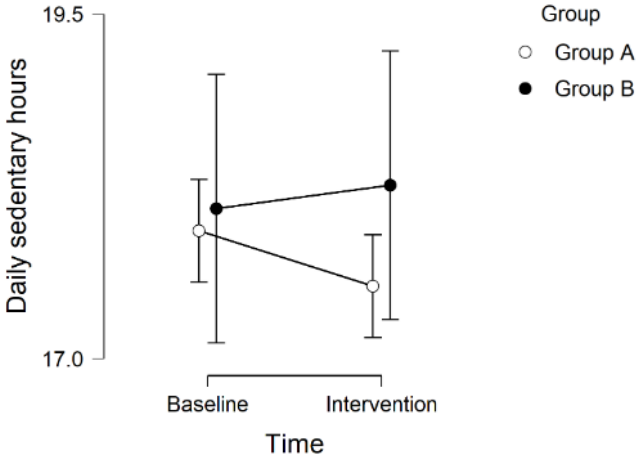
When	Where	How
• Timepoint (eg, 4 AM) • Now • Vague time (eg, today, tomorrow, and anytime)	• Workplace (eg, university, lab, library, campus, garden, office, building Z, and outdoor) • Home/dormitory/kitchen • City • Park	• Walk (eg, tea walk, walk to post, walk between lectures, walk after lecture/meeting/lunch/dinner, 5-min walk, 6000 steps, and 250 steps per hour) • Yoga • Cycle instead of the bus • Push-ups • Get up and stretch • Stand up for 5 min every 30 min • Jump

RQ2: Changes in Participants' Sedentary Behavior

Wilcoxon signed-rank tests showed a marginally significant decrease in daily sedentary time in group A (Z=−11.50; P=.06) and no significant difference in daily sedentary time in group B (Z=2.50; P=.59). The median change in sedentary time was −0.19 hours per day in group A and 0.07 hours per day in group B. The results of the Bayesian paired samples t test suggested (with weak evidence) that the daily sedentary hours decreased from the baseline week to the intervention weeks in group A (BF_{+0}=1.92; median 0.52; 95% CI 0.04-1.25). This is also mirrored in the descriptive statistics plotted in Figure 9 (mean −0.40, SD 0.63). In contrast, in group B, it was more likely that the intervention had no effect than a positive effect with moderate evidence (BF_{+0}=0.28; median 0.18; 95% CI 0.01-0.64; mean 0.17, SD 1.65).

Therefore, regarding RQ2, the intervention involving visualizations and action planning in SedVis had a positive effect on reducing participants' sedentary hours, with weak evidence. Meanwhile, action planning alone had no effect on reducing participants' sedentary hours, with moderate evidence.

Figure 9. The participants' daily sedentary hours based on the app-logged data during the baseline week and the intervention weeks. The bars refer to the confidence intervals with 95% confidence level.



RQ3: Participants' Interaction With SedVis

The frequency of checking the visualizations per day reflects the participants' strength of self-monitoring, which might also act as a cue for self-reminding of sedentary behavior change. Figure 10 shows the daily frequency of participants checking the visualizations in SedVis. Participants were more likely to check the visualizations from the notification bar (302/442, 68.3%) than from the dashboard (140/442, 31.7%).

To test the assumption that the participants' engagement with the app is positively associated with the effect of the app on their behavior, the daily frequency of participants checking the visualization in SedVis was correlated with their change of sedentary hours, calculated as daily sedentary hours during the intervention weeks minus the counterparts during the baseline week (Figure 11). A Spearman rank-order correlation test did not show a statistically significant correlation between participants' checking of visualizations in SedVis with the change in daily sedentary hours ($\rho=-0.37$; $P=.15$) [66]. Then, a Bayesian Pearson correlation with the alternative hypothesis of negative correlation was calculated. BF (BF =1.49; $r=-0.50$) weakly suggested that the 2 factors were more likely to be negatively related than unrelated. To some extent, this confirmed that the participants' engagement was positively related to the effect of reducing sedentary hours.

Figure 10. The daily frequency of participants checking the visualizations in SedVis through the notification and the home screen.

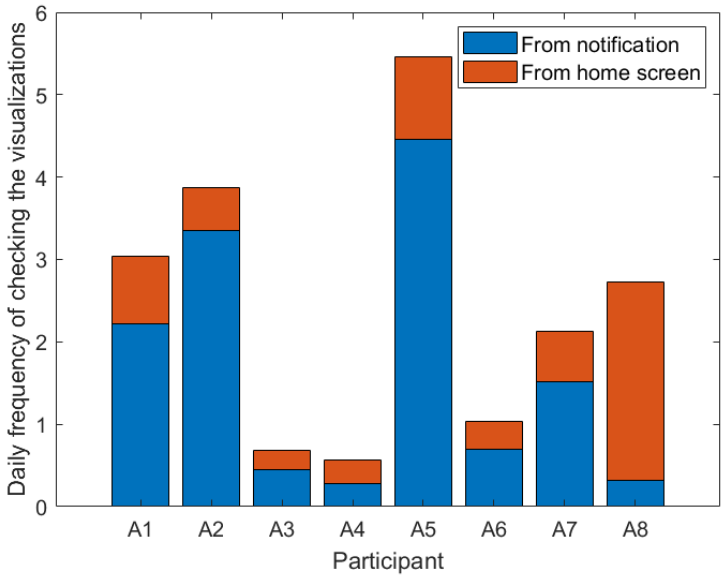
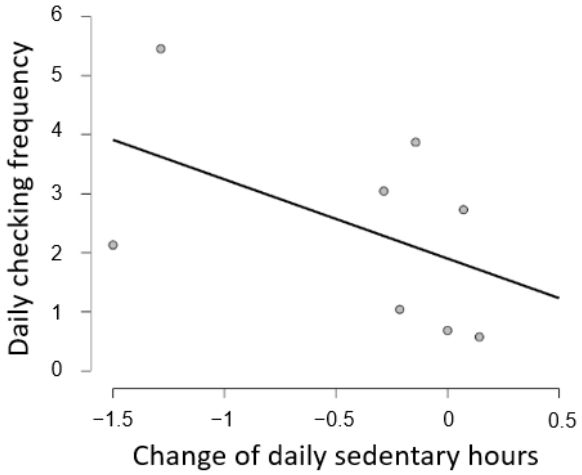


Figure 11. The scatter plot of the participants' daily frequency of checking the visualizations and their change in daily sedentary hours in group A. The change of daily sedentary hours (x-axis) equals to the daily sedentary hours during the intervention weeks minus the counterparts during the baseline week. Thus, negative values indicate a reduction in sedentary behavior.



RQ4: User Experience

Quantitative Results

User experience was investigated both quantitatively and qualitatively. By comparing the ratings with the benchmark provided by the UEQ toolkit [10], the participants' scores of user experience were mapped to quality levels, as shown in Table 4.

According to the results of the Bayesian *t* test with the alternative hypothesis that group A scored higher than group B, visualizations yielded more perceived stimulation (BF =1.99; median 0.62; 95% CI 0.05-1.62) and novelty (BF =7.44; median 1.03; 95% CI 0.16-2.16). For other aspects, the scores tended to be equivalent. It was observed that the previewed dependability is only *above average*, which indicates that the participants did not think the data shown in the app were very accurate.

Table 4. The user experience scores based on user experience questionnaire.

User experience questionnaire scales	Group A, mean (SD)	Comparison with benchmark (group A)	Group B, mean (SD)	Comparison with benchmark (group B)	Z^a	*P* value[b]
Attractiveness	1.65 (0.69)	Good	1.44 (0.71)	Above average	−0.74	.46
Perspicuity	2.25 (0.81)	Excellent	2.10 (0.46)	Excellent	−0.95	.34
Efficiency	1.91 (0.48)	Excellent	1.84 (0.79)	Excellent	0.00	>.99
Dependability	1.34 (0.55)	Above average	1.22 (0.95)	Above average	−0.53	.60
Stimulation	1.66 (0.53)	Good	1.16 (0.63)	Above average	−1.56	.12
Novelty	1.22 (0.65)	Good	0.44 (0.48)	Below average	−2.28	*.02* [c]

[a]*Z* refers to the normalized statistic of Mann-Whitney U test.

[b]*P* value of Mann-Whitney U test.

[c]An italicized *P* value indicates significant difference (*P*<.05).

Qualitative Results

Although participants were asked to make at least one action plan every day during the 2-week intervention phase, the average number of daily action plans was only 0.59, which hints that participants might not have used the app regularly. According to the feedback in the exit interview, no participant complained about interruptions of daily activities through using the app, although one participant commented that making action plans every day was boring.

The topics of continued use of the app and reminders were often related to participants' responses. Of the 16 participants, 8 wanted to continue using the app to reduce sedentary behavior. The reported reasons for continuing to use the app included "self-monitoring is helpful/could increase self-awareness" (n=4), "I wanted to frequently check step counts" (n=2), "I wanted to see the change" (n=1), and "writing down the action plans are important" (n=1). On the other hand, 4 participants said they did not want to continue to use the app for the following reasons: (1) "it underestimates my steps," (2) "I do not want to always keep the GPS on," (3) "the app provided too little new information compared to other devices like a smartwatch," (4) "I need a reminder for enacting my plans." The remaining 4 participants were undecided about future use. Two participants stated that they would continue to use the app if reminders were implemented; one participant desired a greater range of functionalities, and one participant would have considered continued use if the app would consume less battery and would

not require daily action planning. In contrast to the participants who suggested implementing reminders, 6 out of the 8 participants who wanted to keep using the app reported that they did not need a reminder. One of these participants explicitly gave the reason that they did not want to be interrupted during work.

In addition, the topics of accuracy and constantly carrying the phone were linked. Among the 8 participants in group A who had access to the visualization in the app, 4 reported that the data shown in the app seemed accurate. Two participants felt that the app underestimated their steps. This may have been because they did not carry their phone at all times (eg, working in their lab) or because they thought that the sensor in their phone was not sensitive enough. Interestingly, one participant reported that the underestimated steps did not influence their perception of accuracy as they knew the reason, whereas another participant thought that the underestimated steps were disappointing. Therefore, future versions of the app should consider integrating more data sources (eg, wristband or manual adjustment) to improve the users' perceived truthfulness.

Discussion

Principal Findings

This paper presents a pilot test of SedVis, an app-based sedentary behavior intervention that aims to reduce sedentary behavior through a combination of mobility pattern visualization and daily action planning. Specifically, it was hypothesized that mobility pattern visualization would lead to improved action plans, which would, in turn, lead to a reduction in sedentary hours.

Contrary to this expectation, the visualizations did not impact the participants' action planning (see the *Results* section). However, these results are in line with those of Maher and Conroy [7], who also found no effect of daily action planning on reducing sedentary behavior in the short term among college students. Furthermore, the data suggested that sedentary behavior change did not correlate with the quantity and quality action plans. As explained by Maher and Conroy [7], one reason for the ineffectiveness could be that the cue-to-action response expected by action planning relies much on the conscious self-regulatory process, which is difficult for highly habitual behavior, such as sedentary behavior. Another explanation could be based on prospective memory [67], inspired by the work of Grundgeiger et al [68]: prospective memory tasks, which require us to remember to do something at a future time, are very difficult, especially when focusing on other tasks. As sedentary behavior is often coupled with other tasks demanding attention, the action plans for reducing sedentary behavior might be easily forgotten.

Still, SedVis may be effective in reducing sedentary behavior: when having access to mobility pattern visualizations, the intervention group slightly reduced sedentary hours compared with baseline. At the same time, the control group did not show a reduction in sedentary hours. It could thus be concluded that visualizations might have impacted sedentary behavior by promoting awareness when self-monitoring sedentary behavior [69]. This idea is supported by the association between the change in sedentary time and the participants' engagement with SedVis. Engagement with the app, in turn, might have been strengthened by the stimulation and novelty of the visualizations. As Perski et al [70] pointed out in their review on engagement with behavior change interventions, novelty is positively related to engagement as it prevents boredom. The inclusion of novel and stimulating visualizations may thus indirectly influence behavior change.

The participants' evaluations of SedVis with visualizations were good or excellent regarding attractiveness, perspicuity, efficiency, stimulation, and novelty. Only perceived dependability was above average. This may reflect some participants' concerns that SedVis underestimated their steps. At the exit interview, several participants mentioned that they believed the app missed part of their daily steps because they did not take the smartphone with them for certain activities (eg, working in the lab). This limitation of this study could be avoided in future studies by using wearable sensors (eg, wristbands and posture monitors) [71].

The results of this study support the notion that smartphone apps might be an effective tool to reduce sedentary behavior in daily life [15,72]. However, they also indicate that behavior change techniques might differ in their effectiveness to induce changes in sedentary behavior. Three commonly used behavior change techniques were used in this study, that is: self-monitoring, feedback, and action planning [36,72]. Interestingly, action planning was not sufficient to induce changes in sedentary hours in the active control group, whereas additional feedback visualizations induced a small reduction in sedentary hours in the intervention group. Thus, it could be concluded that engaging visualizations to provide feedback on behavior might be more effective in inducing a change in sedentary behavior than action planning. However, as the sample of this study was small, further studies are needed to identify which behavior change techniques are most effective for sedentary behavior change.

Implications for Future Work

Rethinking Action Plans

Although most participants made action plans in accordance with the format of specifying *When*, *Where*, and *How* to reduce their sedentary time, one participant additionally enclosed other contextual cues in their plans, for example, "15:00, lab, take a walk in between experiments" and "13:00, university, walk between lectures." Due to the additional cues—experiment and lectures—the plans might be easier to remember. These plans are in line with the *if-then* format of implementation intentions, which emphasize the contextual cues linking to the goal-directed behavioral response [26,73]. As sedentary behavior is prevalent, the cues of *When* and *Where* might provide limited strength of conditional links to the response behavior. Owing to the requirement of less self-regulatory resources, the more contextual plans in the *if-then* format might be more effective than the plans in *when, where, and how* format [7,73]. However, no prior studies have assessed potential differential effects in sedentary behavior change.

Relating to SedVis, future work might explore how the app could support personalized implementation intentions and their effectiveness on sedentary behavior change, such as generating recommendations of plans based on users' mobility patterns and context, which they might not even notice. Several heuristic rules can be used, for example, going to the restroom downstairs instead of the nearest one or more frequently going to the kitchen to drink water. Armitage [74] found that experimenter-provided and self-generated implementation intentions could be equally effective in reducing alcohol consumption. It is worth investigating this effect on sedentary behavior change following the study design. Some participants commented that making plans every day was boring, so generating plan recommendations might also increase user acceptance in the long term.

Rethinking Self-Monitoring, Feedback, and Reminders

As this study suggests that a higher interaction frequency could lead to a greater reduction of sedentary behavior, future work might need to study more convenient and intuitive user interfaces (eg, glanceable feedback [75]) to simplify self-monitoring and interaction with the app even further. In the current version of SedVis, the easiest way to access the daily visualization was to swipe down the notification bar and click on the notification. In a future version, the app could display real-time sedentary information using an always-on progress bar [76] embedded in the notification or the app widget on the smartphone's home screen.

Future work should also consider the users' need for reminders. Participants expressed differential attitudes toward reminders: some of them expressed that reminders for the action plans they made would be helpful because they sometimes forgot the plans; others thought that reminders would be unnecessary because of the potential interruption. Although fixed-time reminders (eg, prompts on PC screens) were frequently used in prior interventions to reduce sedentary behavior at work [77], no studies have explored the effectiveness and user experience of reminders for personalized action plans.

Limitations

This study has several limitations. First, this study determined sedentary hours based on activity tracked with the smartphone, which may be less rigorous than using dedicated activity trackers (eg, activPAL and ActiGraph) [45]. However, having to wear additional devices might be inconvenient for users (eg, charging the device and attaching the device to the thigh) and may create bias. Moreover, the sedentary hours based on the app-logged data might underestimate the participants' movements. One reason for this might be that participants might not take the smartphone with them during certain activities, such as going to the washroom. Another reason could be that some activities could not be recognized and counted as steps. For example, one participant made an action plan to perform push-ups at home, which cannot be recognized and recorded using a smartphone's sensors. We consider integrating more data sources (eg, smartwatches) in the following version of the app.

Second, the app did not differentiate sleeping time from the sedentary time, and it was assumed that the participants' sleeping time was consistent during the study. Although participants were asked if their sleeping time was normal in the exit interview, their recall might not be accurate. Moreover, the app was not able to distinguish between prolonged periods of sitting and standing. Although using a standing desk at work was an exclusion criterion for participation, it cannot be excluded that participants stood for longer periods, for example, while cooking at home. Future studies, therefore, need to employ more accurate measures for body position to distinguish sitting from lying down and standing (eg, using several sensors [78]).

Third, the small sample size and the relatively large between-subjects variances of the measurements may have reduced the statistical power of the null hypothesis significance tests and may hinder the generalization of our study results. Finally, the study period is relatively short, which limits the validation of the results in short-term scenarios. Therefore, future studies should replicate the results of this study in larger samples and with longer study durations.

Conclusions

This paper presents the results of a user study in which the effect of a novel visualization within a mobile app on users' action planning and sedentary behavior change was evaluated. The results suggest that using a smartphone app to collect mobility data and provide real-time feedback using visualizations is a promising method to induce changes in sedentary behavior and may be more effective than action planning alone. Future research should thus further explore the potential of the visualizations of users' sedentary behavior to induce behavior change.

Conflicts of Interest

None declared.

References

1. Tremblay MS, Aubert S, Barnes JD, Saunders TJ, Carson V, Latimer-Cheung AE, SBRN Terminology Consensus Project Participants. Sedentary Behavior Research Network (SBRN) - Terminology Consensus Project process and outcome. Int J Behav Nutr Phys Act 2017 Jun 10;14(1):75 [FREE Full text] [doi: 10.1186/s12966-017-0525-8] [Medline: 28599680]
2. Sedentary Behaviour Research Network. Letter to the editor: standardized use of the terms. Appl Physiol Nutr Metab 2012 Jun;37(3):540-542 [FREE Full text] [doi: 10.1139/h2012-024] [Medline: 22540258]
3. Knaeps S, Bourgois JG, Charlier R, Mertens E, Lefevre J, Wijndaele K. Ten-year change in sedentary behaviour, moderate-to-vigorous physical activity, cardiorespiratory fitness and cardiometabolic risk: independent associations and

mediation analysis. Br J Sports Med 2018 Aug;52(16):1063-1068 [FREE Full text] [doi: 10.1136/bjsports-2016-096083] [Medline: 27491779]
4. Healy G, Dunstan D, Salmon J, Cerin E, Shaw JE, Zimmet PZ, et al. Breaks in sedentary time: beneficial associations with metabolic risk. Diabetes Care 2008 Apr;31(4):661-666. [doi: 10.2337/dc07-2046] [Medline: 18252901]
5. Australia's physical activity and sedentary behaviour guidelines and the Australian 24-hour movement guidelines. Australian Government: Department of Health. 2019 Apr 12. URL: http://www.health.gov.au/internet/main/publishing.nsf/Content/health-pubhlth-strateg-phys-act-guidelines [accessed 2021-01-07]
6. Canadian 24-hour movement guidelines: an integration of physical activity, sedentary behaviour, and sleep. Canadian Society for Exercise Physiology. URL: https://csepguidelines.ca/ [accessed 2021-01-07]
7. Maher J, Conroy D. Habit strength moderates the effects of daily action planning prompts on physical activity but not sedentary behavior. J Sport Exerc Psychol 2015 Feb;37(1):97-107. [doi: 10.1123/jsep.2014-0258] [Medline: 25730895]
8. Orbell S. Habit Strength. In: Encyclopedia of Behavioral Medicine. New York, NY: Springer; 2013.
9. Dunstan DW, Healy GN, Sugiyama T, Owen N. Too much sitting and metabolic risk—has modern technology caught up with us? European Endocrinology 2010;6:19-23 [FREE Full text] [doi: 10.17925/EE.2010.06.00.19]
10. Chu AHY, Ng SHX, Tan CS, Win AM, Koh D, Müller-Riemenschneider F. A systematic review and meta-analysis of workplace intervention strategies to reduce sedentary time in white-collar workers. Obes Rev 2016 May;17(5):467-481. [doi: 10.1111/obr.12388] [Medline: 26990220]
11. Number of connected wearable devices worldwide from 2016 to 2022 Internet. statista. 2019. URL: https://www.statista.com/statistics/487291/global-connected-wearable-devices/ [accessed 2019-06-26]
12. Consolvo S, Klasnja P, McDonald DW, Landay JA. Designing for healthy lifestyles: design considerations for mobile technologies to encourage consumer health and wellness. In: Foundations and Trends® in Human–Computer Interaction. Boston - Delft: Now Foundations and Trends; 2014.
13. König LM, Sproesser G, Schupp H, Renner B. Describing the process of adopting nutrition and fitness apps: behavior stage model approach. JMIR Mhealth Uhealth 2018 Mar 13;6(3):e55 [FREE Full text] [doi: 10.2196/mhealth.8261] [Medline: 29535078]
14. Müller AM, Maher C, Vandelanotte C, Hingle M, Middelweerd A, Lopez M, et al. Physical activity, sedentary behavior, and diet-related ehealth and mhealth research: bibliometric analysis. J Med Internet Res 2018 Apr 18;20(4):e122 [FREE Full text] [doi: 10.2196/jmir.8954] [Medline: 29669703]
15. Schoeppe S, Alley S, Van Lippevelde W, Bray N, Williams S, Duncan M, et al. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. Int J Behav Nutr Phys Act 2016 Dec 07;13(1):127 [FREE Full text] [doi: 10.1186/s12966-016-0454-y] [Medline: 27927218]
16. Borrelli B, Ritterband LM. Special issue on eHealth and mHealth: challenges and future directions for assessment, treatment, and dissemination. Health Psychol 2015 Dec;34S:1205-1208. [doi: 10.1037/hea0000323] [Medline: 26651461]
17. 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]
18. Bandura A. Social cognitive theory: an agentic perspective. Annu Rev Psychol 2001;52(1):1-26. [doi: 10.1146/annurev.psych.52.1.1] [Medline: 11148297]
19. 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]
20. Sutton S. How does the health action process approach (HAPA) bridge the intention–behavior gap? An examination of the model's causal structure. App Psychol 2008 Jan;57(1):66-74. [doi: 10.1111/j.1464-0597.2007.00326.x]
21. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. Health Psychol 2009 Nov;28(6):690-701. [doi: 10.1037/a0016136] [Medline: 19916637]
22. Olander EK, Fletcher H, Williams S, Atkinson L, Turner A, French DP. What are the most effective techniques in changing obese individuals' physical activity self-efficacy and behaviour: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 2013 Mar 03;10:29 [FREE Full text] [doi: 10.1186/1479-5868-10-29] [Medline: 23452345]
23. Gollwitzer PM, Sheeran P. Implementation intentions and goal achievement: a meta ‐ analysis of effects and processes. Advances in Experimental Social Psychology 2006;38:69-119. [doi: 10.1016/s0065-2601(06)38002-1]
24. Michie S, Wood C, Johnston M, Abraham C, Francis J, Hardeman W. Behaviour change techniques: the development and evaluation of a taxonomic method for reporting and describing behaviour change interventions (a suite of five studies involving consensus methods, randomised controlled trials and analysis of qualitative data). Health Technol Assess 2015 Nov;19(99):1-188 [FREE Full text] [doi: 10.3310/hta19990] [Medline: 26616119]
25. Schwarzer R. Modeling health behavior change: how to predict and modify the adoption and maintenance of health behaviors. App Psychol 2008 Jan;57(1):1-29. [doi: 10.1111/j.1464-0597.2007.00325.x]
26. Gollwitzer PM. Implementation intentions: strong effects of simple plans. American Psychologist 1999;54(7):493-503. [doi: 10.1037/0003-066x.54.7.493]

27. Stephenson A, McDonough S, Murphy M, Nugent C, Mair J. Using computer, mobile and wearable technology enhanced interventions to reduce sedentary behaviour: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 2017 Aug 11;14(1):105 [FREE Full text] [doi: 10.1186/s12966-017-0561-4] [Medline: 28800736]
28. Wang Y, Fadhil A, Reiterer H. Health behavior change in HCI: trends, patterns, and opportunities. arXiv. 2019. URL: http://arxiv.org/abs/1901.10449 [accessed 2019-02-18]
29. Aittasalo M, Livson M, Lusa S, Romo A, Vähä-Ypyä H, Tokola K, et al. Moving to business - changes in physical activity and sedentary behavior after multilevel intervention in small and medium-size workplaces. BMC Public Health 2017 Apr 17;17(1):319 [FREE Full text] [doi: 10.1186/s12889-017-4229-4] [Medline: 28415993]
30. De Cocker K, De Bourdeaudhuij I, Cardon G, Vandelanotte C. Theory-driven, web-based, computer-tailored advice to reduce and interrupt sitting at work: development, feasibility and acceptability testing among employees. BMC Public Health 2015 Sep 24;15:959 [FREE Full text] [doi: 10.1186/s12889-015-2288-y] [Medline: 26404915]
31. Hagstromer M, Ainsworth B, Oja P, Sjostrom M. Comparison of a subjective and an objective measure of physical activity in a population sample. J Phys Act Health 2010 Jul;7(4):541-550. [doi: 10.1123/jpah.7.4.541] [Medline: 20683097]
32. Fleig L, Gardner B, Keller J, Lippke S, Pomp S, Wiedemann AU. What contributes to action plan enactment? Examining characteristics of physical activity plans. Br J Health Psychol 2017 Nov;22(4):940-957. [doi: 10.1111/bjhp.12263] [Medline: 28833976]
33. González MC, Hidalgo CA, Barabási A. Understanding individual human mobility patterns. Nature 2008 Jun 05;453(7196):779-782. [doi: 10.1038/nature06958]
34. Lima A, Stanojevic R, Papagiannaki D, Rodriguez P, González MC. Understanding individual routing behaviour. J R Soc Interface 2016 Mar;13(116) [FREE Full text] [doi: 10.1098/rsif.2016.0021] [Medline: 26962031]
35. Song C, Qu Z, Blumm N, Barabási AL. Limits of predictability in human mobility. Science 2010 Feb 19;327(5968):1018-1021 [FREE Full text] [doi: 10.1126/science.1177170] [Medline: 20167789]
36. Schoeppe S, Alley S, Rebar A, Hayman M, Bray N, Van Lippevelde W, et al. Apps to improve diet, physical activity and sedentary behaviour in children and adolescents: a review of quality, features and behaviour change techniques. Int J Behav Nutr Phys Act 2017 Jun 24;14(1):83 [FREE Full text] [doi: 10.1186/s12966-017-0538-3] [Medline: 28646889]
37. Middelweerd A, Mollee J, van der Wal CN, Brug J, Te Velde SJ. Apps to promote physical activity among adults: a review and content analysis. Int J Behav Nutr Phys Act 2014 Jul 25;11:97 [FREE Full text] [doi: 10.1186/s12966-014-0097-9] [Medline: 25059981]
38. 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]
39. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
40. Epstein D, Cordeiro F, Bales E, Fogarty J, Munson S. Taming Data Complexity in Lifelogs: exploring Visual Cuts of Personal Informatics Data. USA: ACM; 2014 Presented at: Proceedings of the 2014 conference on Designing interactive systems; June, 2014; Vancouver BC Canada. [doi: 10.1145/2598510.2598558]
41. ActivityRecognitionClient. Google APIs for Android. URL: https://developers.google.com/android/reference/com/google/android/gms/location/ActivityRecognitionClient [accessed 2021-01-07]
42. Wang Y, Duan L, Mueller J, Butscher S, Reiterer H. Fingerprints: detecting meaningful moments for mobile health intervention. New York, NY, United States: ACM; 2016 Presented at: MobileHCI '16: Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct; 2016; Florence, Italy. [doi: 10.1145/2957265.2965006]
43. MIUI 12. MIUI 12. URL: https://en.miui.com/ [accessed 2021-01-07]
44. EMUI 11. HUAWEI Consumer. URL: https://consumer.huawei.com/en/emui/ [accessed 2021-01-07]
45. Hekler EB, Buman MP, Grieco L, Rosenberger M, Winter SJ, Haskell W, et al. Validation of Physical Activity Tracking via Android Smartphones Compared to ActiGraph Accelerometer: Laboratory-Based and Free-Living Validation Studies. JMIR Mhealth Uhealth 2015 Apr 15;3(2):e36 [FREE Full text] [doi: 10.2196/mhealth.3505] [Medline: 25881662]
46. Places SDK for Android: Overview. Google Maps Platform. URL: https://developers.google.com/places/android-sdk/intro [accessed 2021-01-07]
47. Wang Y, Sommer B, Schreiber F, Reiterer H. Clustering with Temporal Constraints on Spatio-Temporal Data of Human Mobility. arXiv. 2018. URL: https://arxiv.org/abs/1807.00546 [accessed 2019-02-18]
48. Beddhu S, Wei G, Marcus RL, Chonchol M, Greene T. Light-intensity physical activities and mortality in the United States general population and CKD subpopulation. Clin J Am Soc Nephrol American Society of Nephrology 2015 Apr 30;10(7):1145-1153. [doi: 10.2215/cjn.08410814]
49. Schäfer H, Willemsen M. Rasch-based tailored goals for nutrition assistance systems. In: Proceedings of the 24th International Conference on Intelligent User Interfaces. New York, NY, USA: ACM Press; 2019 Presented at: IUI '19: Proceedings of the 24th International Conference on Intelligent User Interfaces; March, 2019; Marina del Ray California. [doi: 10.1145/3301275.3302298]
50. Why sitting is bad for you - Murat Dalkilinç. YouTube. URL: https://youtu.be/wUEl8KrMz14 [accessed 2021-01-07]

51. Schwarzer R. Health action process approach (HAPA) as a theoretical framework to understand behavior change. AP 2016 Dec 05;30(121):119-130. [doi: 10.15517/ap.v30i121.23458]
52. Mortazavi B, Nemati E, VanderWall K, Flores-Rodriguez HG, Cai JYJ, Lucier J, et al. Can smartwatches replace smartphones for posture tracking? Sensors (Basel) 2015 Oct 22;15(10):26783-26800 [FREE Full text] [doi: 10.3390/s151026783] [Medline: 26506354]
53. Schrepp M, Hinderks A, Thomaschewski J. Applying the user experience questionnaire (UEQ) in different evaluation scenarios. In: Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience. 2014 Presented at: DUXU 2014: Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience; June 22-27, 2014; Heraklion, Crete, Greece p. 392. [doi: 10.1007/978-3-319-07668-3_37]
54. Lee EC, Whitehead AL, Jacques RM, Julious SA. The statistical interpretation of pilot trials: should significance thresholds be reconsidered? BMC Med Res Methodol 2014 Mar 20;14:41 [FREE Full text] [doi: 10.1186/1471-2288-14-41] [Medline: 24650044]
55. SAS Institute Inc. URL: https://www.jmp.com/ [accessed 2021-01-11]
56. Dienes Z. Using Bayes to get the most out of non-significant results. Front Psychol 2014;5 [FREE Full text] [doi: 10.3389/fpsyg.2014.00781] [Medline: 25120503]
57. Kass RE, Raftery AE. Bayes Factors. J of the Amer Stati Asso 1995 Jun;90(430):773-795. [doi: 10.1080/01621459.1995.10476572]
58. Kay M, Nelson GL, Hekler EB. Researcher-centered design of statistics: why bayesian statistics better fit the cultureincentives of hci. 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; 2016; San Jose California USA. [doi: 10.1145/2858036.2858465]
59. Greenland S, Senn S, Rothman K, Carlin J, Poole C, Goodman S, et al. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. Eur J Epidemiol 2016 Apr;31(4):337-350 [FREE Full text] [doi: 10.1007/s10654-016-0149-3] [Medline: 27209009]
60. Fritz CO, Morris PE, Richler JJ. Effect size estimates: current use, calculations, and interpretation. J Exp Psychol Gen 2012 Feb;141(1):2-18. [doi: 10.1037/a0024338] [Medline: 21823805]
61. A fresh way to do statistics. JASP. URL: https://jasp-stats.org/ [accessed 2021-01-07]
62. van Doorn J, van den Bergh D, Böhm U, Dablander F, Derks K, Draws T, et al. The JASP guidelines for conducting and reporting a Bayesian analysis. Psychon Bull Rev 2020 Oct 09. [doi: 10.3758/s13423-020-01798-5] [Medline: 33037582]
63. Rouder JN, Speckman PL, Sun D, Morey RD, Iverson G. Bayesian t tests for accepting and rejecting the null hypothesis. Psychonomic Bulletin & Review 2009 Apr;16(2):225-237. [doi: 10.3758/pbr.16.2.225]
64. Rouder JN, Morey RD, Speckman PL, Province JM. Default Bayes factors for ANOVA designs. J of Math Psychol 2012 Oct;56(5):356-374. [doi: 10.1016/j.jmp.2012.08.001]
65. Rouder JN, Morey RD. Default bayes factors for model selection in regression. Multivariate Behav Res 2012 Nov;47(6):877-903. [doi: 10.1080/00273171.2012.734737] [Medline: 26735007]
66. Cohen J. A power primer. Psychol Bull 1992 Jul;112(1):155-159. [doi: 10.1037//0033-2909.112.1.155] [Medline: 19565683]
67. McDaniel MA, Einstein GO. Prospective memory: an overview and synthesis of an emerging field. Thousand Oaks, California, United States: SAGE Publications, Inc; 2007.
68. Grundgeiger T, Pichen J, Häfner J, Wallmann-Sperlich B, Löffler D, Huber S. Combating sedentary behavior: an app based on a distributed prospective memory approach. In: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems. New York, NY: ACM; 2017 Presented at: CHI EA '17: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems; May, 2017; Denver Colorado USA. [doi: 10.1145/3027063.3053094]
69. Carver CS, Scheier MF. Control theory: a useful conceptual framework for personality-social, clinical, and health psychology. Psychol Bull 1982 Jul;92(1):111-135. [Medline: 7134324]
70. 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]
71. Atkin A, Gorely T, Clemes S, Yates T, Edwardson C, Brage S, et al. Methods of measurement in epidemiology: sedentary behaviour. Int J Epidemiol 2012 Oct;41(5):1460-1471 [FREE Full text] [doi: 10.1093/ije/dys118] [Medline: 23045206]
72. Direito A, Carraça E, Rawstorn J, Whittaker R, Maddison R. mHealth technologies to influence physical activity and sedentary behaviors: behavior change techniques, systematic review and meta-analysis of randomized controlled trials. Ann Behav Med 2017 Apr;51(2):226-239. [doi: 10.1007/s12160-016-9846-0] [Medline: 27757789]
73. Hagger MS, Luszczynska A. Implementation intention and action planning interventions in health contexts: state of the research and proposals for the way forward. Appl Psychol Health Well Being 2014 Mar;6(1):1-47. [doi: 10.1111/aphw.12017] [Medline: 24591064]
74. Armitage CJ. Effectiveness of experimenter-provided and self-generated implementation intentions to reduce alcohol consumption in a sample of the general population: a randomized exploratory trial. Health Psychol 2009 Sep;28(5):545-553. [doi: 10.1037/a0015984] [Medline: 19751080]

75. Gouveia R, Pereira F, Karapanos E, Munson S, Hassenzahl M. Exploring the design space of glanceable feedback for physical activity trackers. In: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing. New York, NY: ACM; 2016 Presented at: UbiComp '16: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing; September, 2016; Heidelberg Germany. [doi: 10.1145/2971648.2971754]
76. Wang Y, Reiterer H. The point-of-choice prompt or the always-on progress bar?: a pilot study of reminders for prolonged sedentary behavior change. In: Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems. 2019 Presented at: CHI EA '19: Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems; May, 2019; Glasgow Scotland UK. [doi: 10.1145/3290607.3313050]
77. Wang Y, Wu L, Lange J, Fadhil A, Reiterer H. Persuasive technology in reducing prolonged sedentary behavior at work: a systematic review. Smart Health 2018 Jun;7-8:19-30. [doi: 10.1016/j.smhl.2018.05.002]
78. Bassett DR, John D, Conger SA, Rider BC, Passmore RM, Clark JM. Detection of lying down, sitting, standing, and stepping using two activPAL monitors. Med Sci Sports Exerc 2014 Oct;46(10):2025-2029. [doi: 10.1249/MSS.0000000000000326] [Medline: 24598698]

Abbreviations

API: application programming interface
BF: Bayes factor
HAPA: health action process approach
NHST: null hypothesis significance testing
RQ: research question
UEQ: user experience questionnaire

Edited by G Eysenbach; submitted 04.07.19; peer-reviewed by A Stephenson, J Mair, X Ren, X Ding; comments to author 02.12.19; revised version received 22.04.20; accepted 18.11.20; published 27.01.21.

Please cite as:
Wang Y, König LM, Reiterer H
A Smartphone App to Support Sedentary Behavior Change by Visualizing Personal Mobility Patterns and Action Planning (SedVis): Development and Pilot Study
JMIR Form Res 2021;5(1):e15369
URL: http://formative.jmir.org/2021/1/e15369/
doi:10.2196/15369
PMID:33502322

Original Paper

Development and Feasibility of a Family-Based Health Behavior Intervention Using Intelligent Personal Assistants: Randomized Controlled Trial

Angela Carlin[1], BSc, PhD; Caomhan Logue[2], PhD; Jonathan Flynn[1], MSc; Marie H Murphy[1], PhD; Alison M Gallagher[2], PhD

[1]Centre for Exercise Medicine, Physical Activity and Health, Sport and Exercise Sciences Research Institute, Ulster University, Newtownabbey, United Kingdom

[2]Nutrition Innovation Centre for Food and Health, Biomedical Sciences Research Institute, Ulster University, Coleraine, United Kingdom

Corresponding Author:
Angela Carlin, BSc, PhD
Centre for Exercise Medicine, Physical Activity and Health
Sport and Exercise Sciences Research Institute
Ulster University
Shore Road
Newtownabbey
United Kingdom
Phone: 44 2871675037
Email: a.carlin1@ulster.ac.uk

Abstract

Background: Intelligent personal assistants such as Amazon Echo and Google Home have become increasingly integrated into the home setting and, therefore, may facilitate behavior change via novel interactions or as an adjunct to conventional interventions. However, little is currently known about their potential role in this context.

Objective: This feasibility study aims to develop the Intelligent Personal Assistant Project (IPAP) and assess the acceptability and feasibility of this technology for promoting and maintaining physical activity and other health-related behaviors in both parents and children.

Methods: This pilot feasibility study was conducted in 2 phases. For phase 1, families who were attending a community-based weight management project were invited to participate, whereas phase 2 recruited families not currently receiving any additional intervention. Families were randomly allocated to either the intervention group (received a smart speaker for use in the family home) or the control group. The IPAP intervention aimed to promote positive health behaviors in the family setting through utilization of the functions of a smart speaker and its linked intelligent personal assistant. Data were collected on recruitment, retention, outcome measures, intervention acceptability, device interactions, and usage.

Results: In total, 26 families with at least one child aged 5 to 12 years were recruited, with 23 families retained at follow-up. Across phase 1 of the intervention, families interacted with the intelligent personal assistant a total of 65 times. Although device interactions across phase 2 of the intervention were much higher (312 times), only 10.9% (34/312) of interactions were coded as relevant (related to diet, physical activity or well-being). Focus groups highlighted that the families found the devices acceptable and easy to use and felt that the prompts or reminders were useful in prompting healthier behaviors. Some further intervention refinements in relation to the timing of prompts and integrating feedback alongside the devices were suggested by families.

Conclusions: Using intelligent personal assistants to deliver health-related messages and information within the home is feasible, with high levels of engagement reported by participating families. This novel feasibility study highlights important methodological considerations that should inform future trials testing the effectiveness of intelligent personal assistants in promoting positive health-related behaviors.

Trial Registration: ISRCTN Registry ISRCTN16792534; http://www.isrctn.com/ISRCTN16792534

(JMIR Form Res 2021;5(1):e17501) doi:10.2196/17501

KEYWORDS

children; parent; physical activity; healthy eating; technology; mobile phone

Introduction

The high incidence of childhood obesity has been well documented, with 29% of children aged 2 to 15 years in England [1] and one-fourth of children living in Northern Ireland [2] classified as overweight or obese. Furthermore, approximately one-fifth of children in the United Kingdom meet the recommended guideline of at least 60 minutes per day of moderate-to-vigorous physical activity [3,4]. The associated risk of developing obesity-related comorbidities earlier in life means that schoolchildren are a key target population for the promotion of sustainable healthy behaviors [5].

Interventions to promote healthy behaviors in children have largely focused on the school setting [6,7], with only 7% of randomized controlled trials (RCTs) targeted at the home setting [6]. The influence of parents and other family members on health behaviors in children is well established [8,9]. Research has highlighted the need for interventions that target children within the home environment [10], which encourages positive behaviors before children progress into adolescence and develop more autonomy over their health choices and the influence of the family context wanes [10]. Family-based interventions typically involve target children and at least one other family member, typically a parent [11]. Without the involvement of family members in interventions, long-term behavior change is unlikely to be sustained in children [12]. A recent meta-analysis identified 19 family-based interventions targeting physical activity, with 66% of included studies reporting a positive effect on physical activity [12]. This is in contrast to the lower levels of effect noted in reviews of school-based interventions [13,14]. Family-based interventions that target diet alongside physical activity appear to be more effective in reducing BMI z-score when compared with diet-only or physical activity–only interventions; however, the evidence is considered to be of low certainty [6]. Furthermore, interventions that target the family psychosocial environment and emphasize the child as the agent of change warrant further investigation [12].

Alongside family involvement, incorporating technology within the family setting has been identified as a potential means of enhancing the effectiveness of interventions targeting childhood obesity [15,16] and may also present further opportunities to increase the reach of interventions [16]. There has been a rapid increase in interventions adopting technology, as it can provide a cost-effective means of providing information and feedback alongside existing interventions or can function as a stand-alone intervention [17,18]. Children and adolescents have been described as *digital natives*, having been exposed to technology for most of their lives [19]. This coupled with high levels of smartphone ownership (78% of adults) and broadband connections (80% of homes) [20] highlights the potential of internet-based technologies for changing health behaviors.

Researchers and practitioners have used technology to change how we deliver interventions (eg, moving from print-based information to web-based resources) and how we incorporate behavior change techniques within interventions. To date, interventions using interactive electronic media [18] or web-based management programs [21] have demonstrated some potential for weight management; however, studies were generally of a lower quality and largely conducted in the United States [18]. A recent systematic review identified 8 eHealth interventions (comprising internet-based interventions, voice prompts, or telemedicine) whereby parents or guardians were the agents of change [16]. Included studies did not report a significant effect on BMI or BMI z-score; however, half of the interventions reviewed found significant improvements in physical activity–related or dietary-related outcome measures [16].

There is a strong need for research studies to target the family setting [22]. Innovative interventions are required [23], with the aim of improving both parents' and children's behaviors. In addition, there is a need for interventions to include more detailed process evaluation with their methodology to further understand the reasons why certain interventions are, or are not, effective in this setting [6]. Intelligent personal assistants (eg, Amazon Alexa) represent an efficient, low-cost method of delivering individualized behavioral interventions, with the potential for scaling at the population level [24]. Unlike other technologies such as wearable devices (pedometers, Fitbit, etc), which have been a primary focus for research studies in recent years, little is known about the potential role that intelligent personal assistants can play in positively influencing health-related behaviors [25].

This study (1) outlines the development of the GetAMoveOn+ Intelligent Personal Assistant Project (IPAP), (2) compares the acceptability of intelligent personal assistants alongside an existing intervention or as a stand-alone intervention, and (3) evaluates the potential of intelligent personal assistants for promoting and maintaining physical activity and other health-related behaviors in both parents and children.

Methods

Study Design

IPAP was a 12-week RCT conducted in 2 phases. Phase 1 was an RCT that evaluated the effect of a home-based intelligent personal assistant intervention on obesity-related behaviors (diet and physical activity) in families attending a community-based weight management project.

Phase 2 was an RCT that evaluated the effect of the home-based intelligent personal assistant in families not attending a weight management project. Randomization for both phases of recruitment took place at the family level, with families (a parent and 1 or 2 children) randomly allocated to an intervention or control group. Randomization was performed by a university staff member who was independent of the research team. Sealed, opaque envelopes were used to randomly assign families to a study arm.

Participants

Families were eligible to participate when at least one child (aged 5-12 years) and one parent or adult responsible for their care consented to participate in the study. Given the nature of the intervention, access to internet connection with their home (Wi-Fi) and ownership of one smart device within the home (eg, a tablet or smartphone) or access to a computer or laptop to enable the family members to interact with the home-based intelligent personal assistant was required. The adult and child or children taking part in the study also had to live within the same household. No restrictions were placed on the family type. No inclusion criteria were placed on parents or children in relation to any medical condition. Participants were asked to notify the research team of any related issues that might affect participation in the intervention. No issues that limited or affected participation or resulted in adverse events were reported.

Recruitment

Phase 1

All families (n=16) attending a community-based obesity prevention project, Safe Wellbeing Eating & Exercise Together (SWEET) as a family, were invited to participate in the study. The SWEET project is a community-based obesity prevention and management program aimed at children and families across a number of sites (community organizations, healthy living centers, etc) in the Western Trust area of Northern Ireland. It aims to work with families in areas of high economic deprivation and targets lifestyle characteristics, such as dietary habits, physical activity, and mental well-being. Families are recruited to the SWEET project via social media sites, flyer distributions in schools, and local paper advertisements. Before approaching families, permission was obtained from the Healthy Lifestyle Coordinator of the Healthy Living Centre where the project was being delivered. Members of the research team attended the first session of the project and provided a verbal overview of the research study. Written informed consent was obtained from all parents or guardians, and written parental consent and child assent were obtained for each child. Phase 1 of the study was conducted from January to April 2019.

Phase 2

Phase 2 was subsequently undertaken to further assess the acceptability of intelligent personal assistants as a stand-alone intervention. Potentially eligible families (as mentioned earlier) were invited to take part in the study (not restricted to those attending the SWEET project) through a number of recruitment strategies. Local community group leaders were contacted and asked to provide permission for a member of the research team to approach families (parents) at relevant events, for example, parent or child groups, youth club, sports training sessions etc. Similar to phase 1, prospective families were provided with a verbal overview of the study and detailed written information on the study. Written informed consent was obtained from all parents or guardians, and written parental consent and child assent were obtained for each child. Efforts were made by the research team to ensure families in phase 1 and phase 2 were recruited from similar community groups to avoid any potential sampling bias. Phase 2 of the study was conducted from May to August 2019. Families were only able to participate in one phase, that is, families who took part in phase 1 were not eligible to take part in phase 2.

Intervention Selection

A smart speaker (Amazon Echo) and its linked intelligent personal assistant (Amazon Alexa) were chosen as the tools for intervention delivery in this study. A market survey (n=2274) highlighted that 33% of respondents based in the United States and the United Kingdom owned a smart speaker [26]. Among these, Amazon's devices were the most popular.

Intelligent personal assistants can perform a range of basic home assistant functions, including playing music, setting alarms, checking the weather, and searching for information. Users can also personalize the devices by adding apps or *Skills*, which further the device's capabilities [25]. Research has shown that *health and fitness* apps are readily available for devices, with health education and fitness training apps the most common types of *health and fitness* apps [25]. The IPAP intervention involved using the existing features and skills developed for Amazon Echo devices.

Intervention Description and Protocol

Following the completion of baseline measurements, families recruited to both phase 1 and phase 2 of the study were randomly allocated to either the intervention group (receive an intelligent personal assistant) or the control group (continue as usual without the provision of additional technology within the home). The IPAP intervention aimed to promote positive health behaviors in the family setting through the utilization of the functions of a smart speaker and its linked intelligent personal assistant. Each family in the intervention arm of the study received a smart speaker (Echo Dot, third generation, Amazon 2018 release) for use in the family home for the duration of the intervention (12 weeks).

The research team set up an individual user account for each family, creating a new email and password, not linked to the family's other email accounts (for security purposes). Each family was provided with their log-in detail, meaning that the research team and family members could both access the accounts during the intervention period. Each family was provided with a detailed information sheet on how to set up and use the device and were instructed to contact a member of the research team for support or troubleshooting throughout the intervention period.

The research team was able to remotely access the devices and set weekly tasks, prompts, and reminders for family members. The prompts and reminders provided by the research team were developed in line with recommendations for the management of childhood obesity [27] and based on current public health recommendations in relation to physical activity [28] and dietary habits [29]. Examples of weekly prompts or reminders and potential ways in which the family could interact with the device are shown in Table 1. For phase 1, the intervention content from the device was aligned to the topics covered at each week of the SWEET program, ensuring that the message was appropriate for the target population. Families received one specific

reminder or prompt per day, which was repeated at a number of times throughout the day to maximize reach. Reminders or prompts were delivered in the morning (before work or school) and in the evening. Families were asked to advise the research team of the most convenient times to receive the prompts or reminders. Families were also encouraged to inform the research team if they were missing the prompts or reminders. In these instances, the timings were revised.

Table 1. Examples of intervention components delivered by the intelligent personal system.

Intervention component and type of interaction	Interaction content
Diet	
Skill	Ask *Vitality* [device-based skill] to give you a recipe—pick a simple meal and have a go cooking with Alexa
Task	Plan your shopping list for the week and add foods to your list using Alexa
Tip	Fruit and vegetables that are fresh, frozen, or tinned all count toward your 5-a-day
Reminder	How much water have you had today?
Physical activity	
Skill	Use Alexa to find some fun games that can help you be active
Task	Kids, do 10 star jumps every morning
Tip	You should aim to be active daily—try going for a 30-min walk on most days this week
Reminder	Have you been for a walk as a family this week?

In addition, families were informed that the devices were to be used as a health promotion tool within the home setting and were free to add their own reminders at times convenient to them and had complete autonomy over what *Skills* (apps) they wanted to enable on their devices. A specific *Skill* was not developed for this intervention; rather, families were signposted to search for *Skills* under the topics of Health and Fitness, Lifestyle, Sport, Cooking, and Recipes. Within this, families could choose the skills most suitable for their children's age and interests. In addition to the preprogrammed messages controlled by the research team, families were instructed that they were free to use the devices for other general functions, not specific to the research project.

Families were informed during the recruitment and throughout the intervention that the research team would also be able to view and manage their user accounts. Families were also made aware that all interactions with the device would be noted by the research team, including interactions that may not be linked to the goals of the intervention, for example, asking the intelligent system nonrelated questions.

Outcome Evaluation Measures

Within this pilot feasibility study, we aimed to evaluate the potential of intelligent personal assistants for promoting and maintaining physical activity and other health-related behaviors in both parents and children. Data collection was carried out at local community centers or at the university by trained researchers, and all participant outcome measures were assessed at baseline and follow-up (12 weeks).

Physical Activity

Physical activity was measured using an ActiGraph GT3 accelerometer (ActiGraph LLC). Participants (parent and child or children) were instructed to wear the device on the waist for 7 consecutive days, removing it only for bathing, water-based activities such as swimming, and when asleep. During the measurement periods, participants were asked to keep a family log of when they wore the accelerometer and took it off. A sampling epoch of 15 seconds was used for data collection. Periods of ≥60 minutes of zero counts were classified as nonwear and were removed. Cutpoints were used to estimate time spent in sedentary behavior and light-, moderate-, and vigorous-intensity physical activity for adults [30] and children [31]. The primary outcome was total physical activity (light, moderate, and vigorous physical activity combined). Secondary accelerometer outcomes included data provision and the proportion of participants meeting the recommended guidelines for physical activity [28]. Participants who provided at least three weekdays of at least 480 minutes of data between 5 AM and 11.59 PM were included in the analysis. Families were given an incentive at each time point for returning the devices (GBP £20 [US $27] One4All voucher).

Health Outcomes

Height (nearest 0.1 cm) and weight (nearest 0.1 kg) were measured according to standardized protocols. BMI was calculated and converted to BMI z-scores using the World Health Organization AnthroPlus software (version 1.0.4).

Family Eating and Activity Habits

Behaviors related to eating and activity habits were assessed using the Revised Family Eating and Activity Habits Questionnaire (FEAHQ-R). The FEAHQ-R is a 32 item self-report instrument designed to assess changes in eating and activity habits of family members as well as obesogenic factors in the overall home environment (stimulus and behavior patterns) related to energy balance [32]. The questionnaire was completed by one parent on behalf of themselves, their spouse, and their child. Summary scores were calculated for physical activity, eating style, stimulus exposure (eg, unhealthy snacks at home), and eating related to hunger. A reduction in scores signifies improvements across all domains.

Process Evaluation

Device Interactions and Usage

The research team was able to access each family's account via their log-in details and view each interaction with the device across the intervention period. An interaction was defined as any engagement with the device made by a parent or child, in addition to the reminders and information provided by the device from the research team. A copy of all interactions was downloaded from the device website and anonymously stored. The research team recorded the number of interactions and the type of interaction. Interactions were primarily coded as *relevant* (related to physical activity or diet or well-being) or *nonrelevant* (ie, not related to the intended purpose of the intervention), with relevant interactions further coded based on their theme. For example, "How many portions of fruit and vegetables per day should I eat?" was recorded as a relevant interaction and subcoded under *Healthy-eating question*. *Waking up* the device, controlling volume, and prompts such as *Next song* were not recorded as interactions for the purposes of this study. In instances where the device was not able to provide a transcript of the voice command received, the device registered this interaction as, "Text not found. Click here to listen to the recording." The research team did not listen to the voice recordings or include these in the interaction analysis. It was not possible for the research team to distinguish whether a parent or child interacted with the device.

Intervention Acceptability

A record of any technical issues in relation to the smart speaker was held by the research team. All parents in the intervention arm of phase 1 and phase 2 were invited to participate in focus group discussions. These discussions focused on the acceptability of the intelligent personal assistants, intervention fidelity, any challenges that arose during the intervention, and suggestions for future improvements. Owing to practical issues (timing and location), it was not possible to facilitate focus groups with all parents, so these were replaced with semistructured interviews. One focus group (n=4 parents) and 3 semistructured interviews (n=3 parents) were conducted with participating parents in the intervention arm of the study. All discussions were audiorecorded. The mean duration of the recordings was 26 (SD 20) minutes.

Ethical Considerations

Participants were provided with detailed instructions on the use of the device and the functionality of the device, that is, what the device is capable of doing and picking up. The mute or temporality disable functions of the device were also highlighted to families. These instructions were developed using the manufacturer's instructions. As these devices were present within the home and accessible to both parents and children, a protocol was developed to consider the potential issue of disclosure and unintended collection of data. No such issues were observed during the intervention period. The search history of the device was kept confidential, and the device was not used for any other purpose during the intervention, for example, recording information or conversations within the home. This pilot feasibility study was approved by the Ulster University Research Ethics Committee and was registered retrospectively (ISRCTN16792534).

Data Analysis

Quantitative Analysis

Frequencies, percentages, means, and SDs were used to describe data related to recruitment, retention, outcome measures, intervention acceptability, device interactions, and usage. Data analysis was conducted using SPSS for Windows (version 25; SPSS Inc).

Qualitative Analysis

Focus groups and semistructured interviews were transcribed verbatim and analyzed thematically, following a deductive approach [33]. Following familiarization with the data, each transcript was reviewed for meaningful quotes and systematically coded by a member of the research team. Potentially relevant codes were grouped together to develop themes, which were reviewed to ensure representativeness. These themes were then reviewed by a member of the research team to ensure that the themes were representative of the coded excerpts. Coding and reviewing of themes were repeated independently by a second member of the research team.

Results

Recruitment and Retention

Phase 1

A total of 16 families attending the SWEET project were invited to participate in the IPAP study (Figure 1). Of the 16 families approached, one family was excluded for not meeting the inclusion criteria and 4 families failed to respond to the initial invitation. Of the 6 families allocated to the intervention, 2 families did not set up the device. Of those allocated to the control arm, one family was absent for follow-up measurements, whereas a further 2 families discontinued the SWEET project and subsequently this study as well. Participant characteristics are described in Table 2. All adult participants were categorized as overweight or obese at baseline.

Figure 1. Consolidated Standards of Reporting Trials 2010 flow diagram for phase 1 participants. SWEET: Safe Wellbeing Eating & Exercise Together.

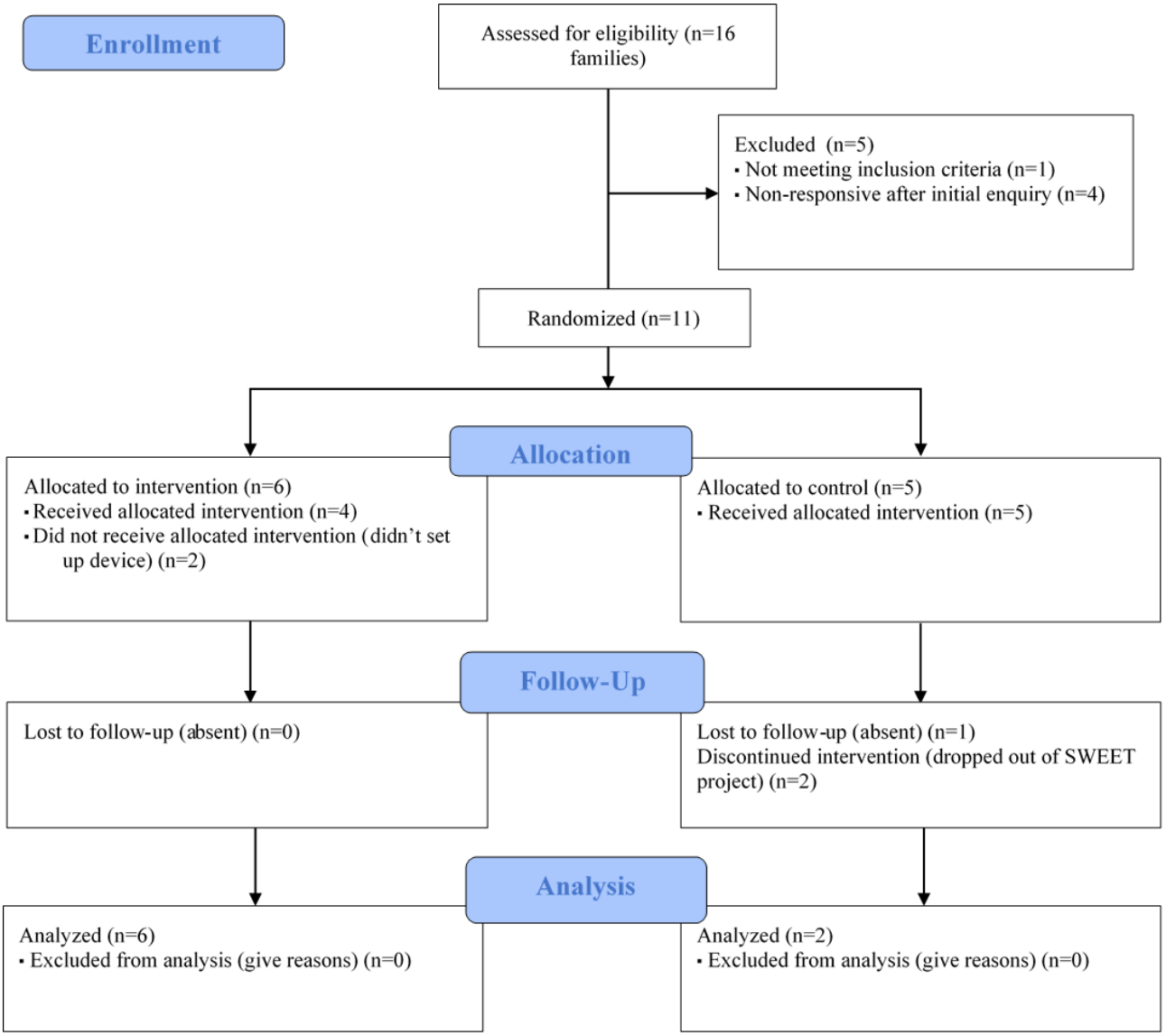


Table 2. Individual participant characteristics at baseline.

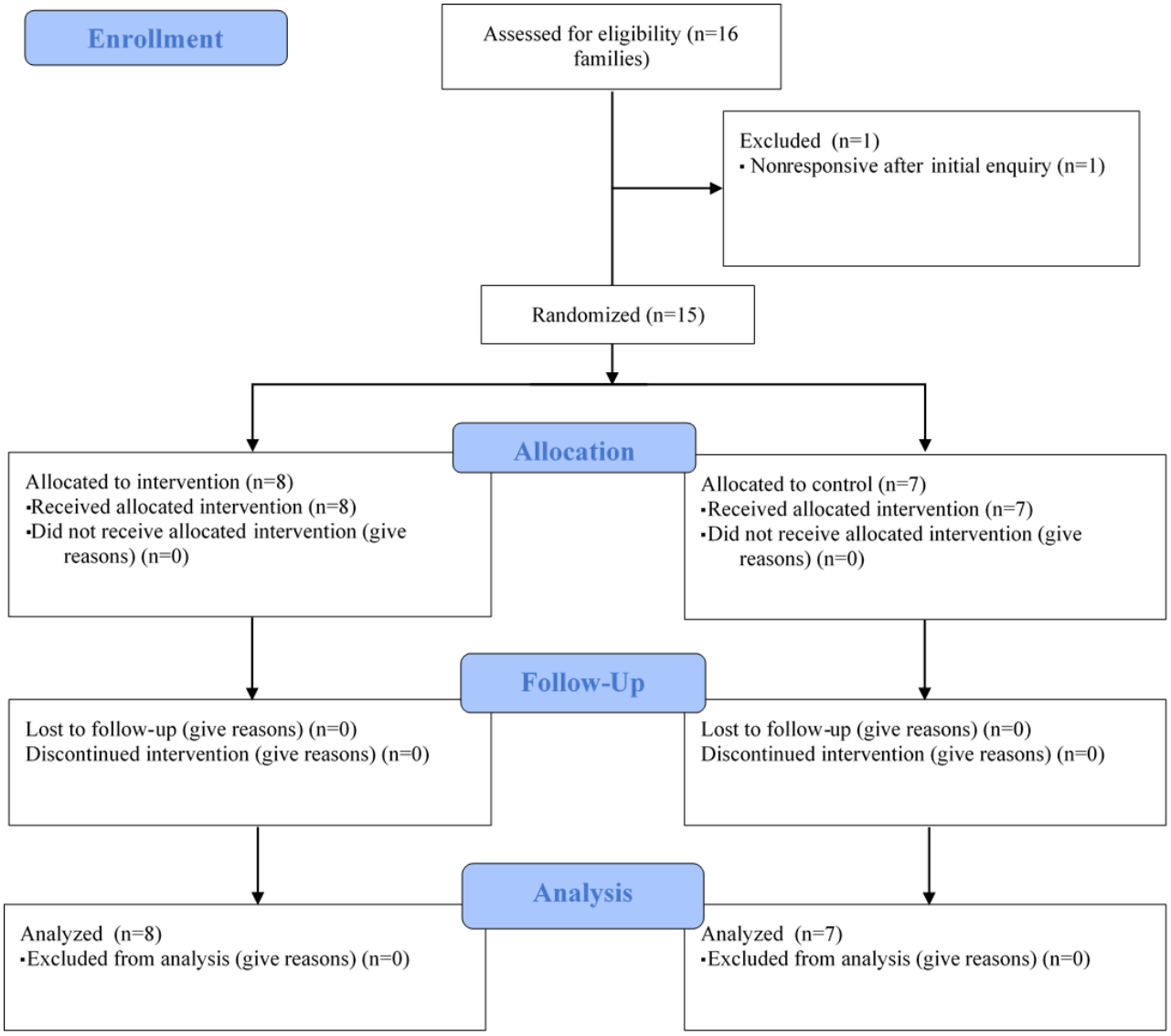
Characteristic	Phase 1		Phase 2	
	Adults (n=11)	Children (n=16)	Adults (n=15)	Children (n=18)
Sex, female, n (%)	10 (91)	9 (56)	11 (73)	8 (44)
Age (years), mean (SD)	40.5 (5.4)	9.1 (2.0)	38.9 (5.2)	7.9 (2.0)
Height (cm), mean (SD)	166.0 (6.2)	141.1 (14.5)	166.9 (8.5)	130.0 (12.8)
Weight (kg), mean (SD)	97.0 (22.8)	49.5 (15.4)	81.4 (15.8)	28.3 (7.7)
BMI (kg/m^2)	35.0 (6.4)	N/A[a]	29.1 (4.9)	N/A
BMI, z-score	N/A	2.61 (1.23)	N/A	0.02 (1.17)

[a]N/A: not applicable.

Phase 2

A total of 20 families from local community groups were approached to take part, of which 16 were assessed for eligibility (Figure 2). Of 20 families, 15 were enrolled in the IPAP study, with all families retained at follow-up. Participant characteristics are described in Table 2. Overall, 80% (12/15) of adult participants were categorized as overweight or obese at baseline.

Figure 2. Consolidated Standards of Reporting Trials 2010 flow diagram for phase 2 participants.



Outcome Evaluation Measures

Physical Activity

In phase 1, 91% (10/11) of adults and 69% (11/16) of children met the minimum inclusion criteria for accelerometer wear time. At baseline, the mean valid wear time was 720 (SD 90.3) and 657.2 (SD 47.8) minutes per day for adults and children, respectively. At follow-up, the proportion of participants meeting the minimum inclusion wear time dropped to 55% (6/11) of adults and 19% (3/16) of children. In phase 2, 86% (13/15) of adults and 89% (16/18) of children met the minimum inclusion criteria for accelerometer wear time. At baseline, mean valid wear time was 782.1 (SD 63.2) and 695.4 (SD 36.3) minutes per day for adults and children, respectively. At follow-up, the proportion of participants meeting the minimum inclusion wear time remained at 87% (13/15) of adults and dropped to 72% (13/18) of children, indicating greater compliance to the accelerometer outcome measure in phase 2 of the IPAP study.

Of those who fulfilled the minimum wear time criteria, 70% (7/10) of adults and 36% (4/11) of children achieved the recommended physical activity guidelines at baseline for phase 1, compared with 77% (10/13) of adults and 38% (6/16) of children in phase 2 of the study. Owing to the small sample size, statistical testing was not undertaken to assess changes in physical activity before and after intervention (Table 3). Adherence to the accelerometer protocol may have been affected by the timing of the intervention and follow-up measurements coinciding with school holidays.

Table 3. Change in accelerometer measured physical activity and sedentary behavior across the Intelligent Personal Assistant Project study (adults).

Physical activity and sedentary behavior	Intervention, mean (SD)	Control, mean (SD)
Phase 1		
Baseline (n=10)		
Daily physical activity (minute per day)	268.5 (35.3)	234.2 (67.4)
Sedentary behavior (minute per day)	440.5 (115.5)	492.8 (52.5)
Follow-up (n=6)		
Daily physical activity (minute per day)	293.7 (57.8)	201.1 (9.5)
Sedentary behavior (minute per day)	587.6 (132.8)	531.4 (26.9)
Phase 2		
Baseline (n=14)		
Daily physical activity (minute per day)	260.7 (35.6)	241.8 (47.7)
Sedentary behavior (minute per day)	562.3 (10.1)	492.7 (56.6)
Follow-up (n=12)		
Daily physical activity (minute per day)	218.9 (40.7)	244.8 (33.1)
Sedentary behavior (minute per day)	513.9 (65.1)	498.3 (21.4)

Family Eating and Activity Habits

Questionnaire data were provided by 85% (22/26) of adult participants at all time points. In phase 1, positive improvements in scores for eating style were observed for adults (−1.75, SD 2.06) and children (−0.50, SD 2.81) in the intervention group, with increases observed in the control group. In phase 2, there was a slight improvement in both the activity level score and stimulus exposure and control for children in the intervention group, with all other summary scores increasing across the intervention period (Table 4).

Table 4. Change in scores for Family Eating and Activity Habits Questionnaire for adults and children in phase 2.

Characteristics	Adults, mean (SD)		Children, mean (SD)	
	Intervention (n=6)	Control (n=5)	Intervention (n=7)	Control (n=4)
Activity level	0.75 (2.72)	1.70 (2.11)	−1.07 (8.23)	−0.25 (6.65)
Eating style	1.80 (8.56)	5.33 (1.15)	3.33 (6.65)	−1.00 (3.00)
Eating related to internal cues	0.83 (1.33)	0.00 (2.00)	1.14 (1.46)	−0.13 (1.55)
Stimulus exposure and control	1.80 (4.09)	1.25 (4.99)	−0.25 (6.65)	0.00 (4.63)

Process Evaluation

Device Interactions and Usage

Across phase 1 of the intervention, families who received a smart speaker on average interacted with the intelligent personal assistant (Alexa) 65 times. *Waking up* the device, controlling volume, and prompts such as *Next song* were not recorded as interactions for the purposes of this study. *Other* (including general knowledge questions and jokes) and *Music* were the most frequently observed interactions across the intervention period. Overall, 42% of all device interactions were coded as relevant in phase 1 (ie, related to diet, physical activity, or well-being). Reminders or prompts involved the family setting their own reminders. Examples of *Skills (diet)* and *Skills (physical activity)* used by families across the intervention period included fitness skills, recipe skills, and active game skills. During phase 1, the prompts or reminders provided by the research team aligned with the topics and tasks the families were covering in the SWEET project.

In phase 2, families did not attend the SWEET project, but the intervention content largely reflected the prompts or reminders provided to families in phase 1. Device interactions across phase 2 of the intervention were much higher, with families interacting with the device 312 times across the intervention period (equivalent to 31.3 interactions per week). Only 11% of interactions were coded as relevant (related to diet, physical activity, or well-being). Of the interactions that were coded as relevant, the most frequent interactions were when families asked questions about nutrition (healthy eating) or used *Skills* related to healthy eating, for example, recipes or healthy eating tips.

Intervention Acceptability

In total, 7 parents took part in focus groups and semistructured interviews to discuss their experiences of the IPAP project. At

the offset of these discussions, parents acknowledged the prominent role of technology in their family's everyday lives and the need to use it in a positive way:

> *Technology is there, and it can be used for good and evil. And it's not going to go away. The way they are growing up, they can't avoid it really so might as well try and use it for good.* [Family 4, male]
>
> *...they are probably more motivated by it [technology], so it probably is the future for the younger generation...* [Family 6, female]

Parents commented that the intelligent personal system motivated the child to engage with the intervention:

> *It actually motivated her quite a bit, because she was saying "mummy, we need to go for a family walk now...or I need to eat my fruit or...*" [Family 6, female]

Families found the intervention content acceptable and discussed how the prompts or reminders encouraged them to change their behaviors in a fun way (Table 5). Families also highlighted how they used other features, such as the skills for recipes or home workouts (Table 5).

Families highlighted several ways to increase engagement with the intervention, including further suggestions on how to use the device within the home and more personalization of the prompts or reminders. The timing of prompts or reminders was a key component of the intervention delivery, and families noted practical issues with this, in addition to the importance of ensuring that families were at home when the device was interacting with them (Table 5). Parents suggested incorporating other technologies alongside the intelligent personal system to facilitate this:

> *If it was connected to your phone, like a phone reminder as well, because Alexa's in the house.* [Parent 2, female]

In addition, families felt that the device needed to be linked to some type of feedback to increase accountability and provide families with opportunities to log their healthy eating or physical activity (Table 5).

Parents felt that the intelligent personal assistants played an additive role in encouraging children to be healthier and could work alongside other types of intervention:

> *I still think you need the traditional ways of activity rather than reliance on a device.* [Family 6, female]
>
> *...if there was an intervention or like, if there were a, a class or some sort of, erm, programme that was with, sent home with families and Alexa reminded you to do it...* [Family 6, female]

In terms of concerns about having a smart speaker within the home, most parents commented that they were cautious of both increasing engagement with technology and the potential issues with social media and young people (Table 5). These concerns regarding internet access or social media were more prominent from parents than issues specific to the intelligent personal assistants themselves:

> *...he's downloading games and I don't know what they are—I would be quite worried; not so much that it's listening, I wouldn't worry about Alexa listening, it's not gonna hear anything in my house.* [Parent 7, female]

Table 5. Supporting quotes from family focus groups and semistructured interviews.

Subheading	Supporting quotes
Findings related to intervention delivery	
Device setup	"It was easy to set up and easy to use. Quite interesting but, and the prompts were very good." [Family 1, female]
Prompts or reminders from the research team	• "We got a prompt, quick do 10 sit ups, and I'm like come on children, everyone on the floor, let's do it! It was some craic [fun] like, and everybody just downed the phones and going to do that challenge. They loved it." [Family 3, female] • "The whole jist of it was brilliant, like the wee prompts it tells you...try this or try that, you know it's just planting that wee seed in your head and when that wee seed's planted, obviously you are gonna try aren't you, so I think it is a great thing." [Family 2, female]
Using other device features	• "...the easy access to the workouts so that you could just do it at a time whenever it suited you." [Family 6, female] • "There were a couple of occasions where we asked Alexa for a healthy recipe to make something so we made a chilli one day and we asked Alexa for a recipe 'cos we were prompted by the device about, you know, healthy, eating healthily and stuff..." [Family 5, male] • "[Child name] was new-fangled with it, she was more into the music in it, bopping about but it got her active too, she was asking me how to do this, and will you do this 'Flossing'...it was good from that point of view you know." [Family 2, female] • "Even her homeworks, she was going out and asking, she was asking me how to spell this, I said ask Alexa, just to get her doing things for herself." [Family 3, male]
Overall device usage	• "We probably could have utilised it much more but it's just the pure fact if we had more time. Erm, and the fact that we were away from it all day long and then we came in, in the evening, it's usually kind of a race, get the dinner made and..." [Family 6, female] • "...but after, like, a week or so they kind of almost forgot it was there and maybe that was our fault, we didn't encourage them to use it as much, erm, but the prompts I think were a good idea." [Family 5, male]
Findings related to intervention optimization	
Timing of prompts or reminders	• "I think there was a couple of technical glitches where the timing wasn't right because we didn't seem to get the prompts and we used Alexa a lot like, we do ask a lot of questions and stuff but, erm, it didn't seem to prompt us; maybe we were out at the time." [Family 4, male] • "You know, if we weren't at home..., I don't know how many prompts there were." [Family 7, female]
Lack of feedback provided	• "...but what it would say to me, 'Have you had your five a day?' Do I shout back, 'Yes,' or, 'Alexa, yeah,' I don't know what way to answer..." [Family 7, female] • "If you had to log what you did, you know, because it's fair enough, erm, you could say, 'Right, go for a family walk,' but you know, then they come back and say, 'Well how many kilometres did you do?' or whatever...to close the loop." [Family 6, female]
Concerns	• "I just worry about that whole side of technology, erm, never mind Alexa but all social media, erm, in terms of how, how that can be utilised against them and I suppose that's a worrying thing for me as a parent..." [Family 6, female] • "I think if you find the right balance where, you know, I don't like the idea of my kids being constantly engaged to technology but I can see the benefit of, of that via a prompt or something like that but, you know, I wouldn't want them to be constantly going to Alexa..." [Family 3, male]
Increasing device usage	• "You know, I think they would maybe be set challenges to do because I think if they've, just can get an app and do so much, I'm not sure that they'll benefit from it." [Family 6, female] • "I think if it was maybe a wee bit more personalised...I don't actually know what I was supposed to be doing with Alexa, you know...and maybe it was in the documentation somewhere, maybe there was a letter written somewhere that I didn't see, that I didn't read." [Family 7, female]

Practical Considerations

Most families were able to set up their user accounts and link these to the smart speaker device. Overall, 2 families did not set up their devices in phase 1 of the study. Of these families, one parent noted that they could not set up the device because they shared the house with another family who did not want the device used, and the other family failed to respond to follow-up instructions from the research team, meaning they did not receive the intervention content. All families in phase 2 successfully set up and used the device.

The smart speakers had to be *online* to allow the research team to set up reminders or prompts and refresh information on the family's interactions with the device. The 2 families in phase 2 had their devices set to *Offline* for extended periods, limiting

the volume of interaction managed by the research team. A further family in phase 2 registered the device with their own personal Amazon account for 2 weeks during the intervention period; therefore, the research team was unable to set prompts or reminders or access information on the family's interactions with the device over this period. A protocol was also put in place to cover the potential issue of disclosure of information and unintended collection of data; however, no scenarios arose within this study.

Discussion

Principal Findings

To our knowledge, this is the first study to outline the development and usage of intelligent personal assistants to promote positive health-related behaviors within the home setting. Given the constraints that exist within current family-based interventions, including time and travel restraints [10], moving toward novel interventions that incorporate web-based learning may help improve engagement and attrition [10]. Within this pilot feasibility study, we assessed the acceptability and feasibility of using intelligent personal assistants alongside more traditional intervention approaches or as a stand-alone intervention tool. This feasibility study demonstrated that using intelligent personal assistants to deliver health-related messages and information within the home was feasible, with high levels of engagement from participating families. This work also highlighted methodological considerations and opportunities for intervention improvement moving forward.

To date, there is a paucity of research on both the development of interventions using this technology and the potential effectiveness of such interventions. An ongoing study is examining the role of a voice coach intervention (Amazon Alexa/Echo) on increasing levels of physical activity among overweight and obese cancer survivors [24]. In addition, Public Health England has used intelligent personal assistants (Amazon Echo) to encourage parents to adopt healthy behaviors around breastfeeding [34] by providing parents with general information and tailored advice based on the age of their child. This study highlighted for the first time that families found this type of intervention approach acceptable and feasible within the home setting. Most families assigned to the intervention were able to set up and initialize their devices and engage with the intelligent personal assistant across the intervention period. Focus groups and interviews with parents highlighted that the prompts or reminders were particularly useful and commented that the intervention encouraged the family to be healthier in a fun way.

Recent research has highlighted the plethora of *Health and Fitness*–related apps available for smart speakers [25], with health education, fitness and training, and nutrition the most frequently occurring of these apps. For the purposes of the IPAP intervention, prompts or reminders provided by the research team were based on the devices' existing functionality, and families were instructed to use the features already developed for these smart speakers. High levels of interaction were observed across the intervention period, with a higher volume of interactions in phase 2. Setting reminders or prompts, asking questions about nutrition, and using physical activity and nutrition apps (Skills) were the most common relevant interactions across the intervention period.

The mean frequency of device interactions across phase 2 was much greater (312 vs 65), but a higher proportion of interactions were coded as *relevant* in phase 1 (42% vs 11%). This provides important insights into how families used the devices and suggests that linking the devices to an ongoing intervention, as with phase 1, may be more directive in terms of prompting families to use the device for health-related interactions. The issue of families not adequately implementing intervention components has been highlighted in similar feasibility work evaluating the use of a web-based intervention to encourage families to increase their physical activity [23]. Within this study, families were provided with written instructions and reminders on how to interact and engage with the intelligent personal assistants. Parents highlighted several ways to improve engagement with the intervention, including incorporating challenges, providing feedback, and clearer guidance from the intervention facilitators on how to use the device within the home. Within this feasibility study, the intervention facilitators were members of the research team. Given the important role of facilitators in terms of intervention outcomes [35], providing families with more guidance and training before the intervention, and ongoing support during the intervention, may improve the family's utilization of the device [23].

Given the small sample size in this study, it was not possible to statistically compare the effectiveness of these 2 intervention approaches. As the families in phase 1 were already attending the SWEET project, the results from phase 1 and phase 2 could not be combined. A recent systematic review highlighted that most family-based eHealth interventions combined technology with other types of delivery, for example, face-to-face counseling, nutrition lessons, and so on, and from this literature, it is difficult to ascertain the exact effect of the eHealth component versus other approaches [16].

The development and feasibility testing of the intervention identified several important methodological considerations. First, the research team was not able to control the content, or indeed validity, of the responses families received when they asked for information on healthy eating or physical activity. At present, there is limited insight on whether these apps are developed based on evidence-based guidelines or available materials [25]; therefore, assessing the accuracy of educational information provided by these devices would be an important methodological consideration moving forward. Indeed, a previous study examining the provision of medical advice from these devices highlights the importance of cautioning users not to use such technologies in place of medical advice without consulting with their health care provider first [36]. Second, families noted that the intervention in its current format did not provide any opportunities for feedback or accountability with limited options for families to log their healthy eating or physical activity. Moving forward, studies should explore the potential of linking these intelligent personal assistants with other technologies to monitor behaviors, set goals, and provide feedback [37,38], which may help improve the effectiveness of technology-based interventions [39].

The implementation of the intervention was dependent on a few factors. An important practical consideration was the capacity of the research team to access the family's device remotely. If the device was switched off or the family had Wi-Fi connection issues, the delivery of the intervention was affected, as the research team was unable to set new reminders and prompts during these periods. During the focus group and interview discussions, parents highlighted how the timing of the prompts or reminders may have affected their adherence to the intervention. Although attempts were made to tailor the intervention to suit the schedules of individual families, future studies using similar intervention components should seek to provide families with further guidance and ownership in relation to managing the devices themselves.

Strengths and Limitations

The IPAP study adopted a cross-sectoral, interdisciplinary approach to explore the role of intelligent personal assistants within the home environment to promote and maintain physical activity and other health-related behaviors in families. The intervention development and evaluation used novel methods to capture intervention engagement, addressing key recommendations for research in this field to adopt appropriate methodologies that enable interventions to be effectively evaluated [17]. This study developed the intervention content and tested its feasibility in line with the best practice for intervention development [40]. Owing to the small sample size, no statistical analysis was undertaken at this stage to evaluate the effectiveness of the intervention. Accelerometer compliance was low during phase 1 of the study, despite the use of incentives to encourage adherence. In addition, device usage was much lower across phase 1. Given that these families were already taking part in the SWEET project at the time, they may have felt overburdened with data collection.

Conclusions

This study demonstrates the feasibility and acceptability of a family-based intervention using intelligent personal assistants. This novel intervention has highlighted important methodological considerations and provides important suggestions to further optimize the potential of intelligent personal assistants to promote positive health-related behaviors in the home setting. This work will inform future pilot and fully powered studies to build upon this feasibility work and test whether such interventions are effective at changing health-related behaviors, including physical activity and healthy eating.

Acknowledgments

The authors thank the families that participated in the study. The authors are grateful to the SWEET project team (George McGowen, Julie White, and Aoibhin Kelly) for their involvement in the study and for their assistance with the project management. The authors thank Jordan Fleming, Naomi Bell, and Sofia Marini for their assistance with data entry. This project was funded by the GetAMoveOnNetwork+ (Engineering and Physical Sciences Research Council grant EP/NO27299/1). The funder had no role in the study design, data collection and analysis, decision to publish, or preparation of the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT EHEALTH checklist V1.6.1.
[PDF File (Adobe PDF File), 1732 KB - formative_v5i1e17501_app1.pdf]

References

1. NatCen Social Research and UCL. Health Survey for England 2016. Government Statistical Service. Surrey: NHS Digital; 2017 Dec 13. URL: https://files.digital.nhs.uk/pdf/s/q/hse2016-summary.pdf [accessed 2020-12-21]
2. Health Survey Northern Ireland: First Results 2016/17. Health Survey Northern Ireland. Belfast: Department of Health; 2017 Oct 24. URL: https://www.health-ni.gov.uk/publications/health-survey-northern-ireland-first-results-201617 [accessed 2020-12-21]
3. Wilkie H, Standage M, Sherar L, Cumming S, Parnell C, Davis A, et al. Results from England's 2016 Report Card on physical activity for children and youth. J Phys Act Health 2016 Nov;13(Suppl 2):S143-S149. [doi: 10.1123/jpah.2016-0298] [Medline: 27848746]
4. Woods C, Powell C, Saunders JA, O'Brien W, Murphy MH, Duff C, et al. The Children's Sport Participation and Physical Activity Study 2018 (CSPPA 2018). Research Gate 2019 Sep [FREE Full text] [doi: 10.13140/RG.2.2.22091.95521]
5. Childhood Obesity: A Plan for Action. Department of Health and Social Care. Downing Street: Government Digital Service; 2017 Jan 20. URL: https://www.gov.uk/government/publications/childhood-obesity-a-plan-for-action/childhood-obesity-a-plan-for-action [accessed 2020-12-21]
6. Brown T, Moore TH, Hooper L, Gao Y, Zayegh A, Ijaz S, et al. Interventions for preventing obesity in children. Cochrane Database Syst Rev 2019 Jul 23;7(7):CD001871 [FREE Full text] [doi: 10.1002/14651858.CD001871.pub4] [Medline: 31332776]

7. Jones M, Defever E, Letsinger A, Steele J, Mackintosh KA. A mixed studies systematic review and meta-analysis of school-based interventions to promote physical activity and/or reduce sedentary time in children. J of Sport and Health Sci. Amsterdam: ScienceDirect; 2020 Jan. URL: https://doi.org/10.1016/j.jshs.2019.06.009 [accessed 2020-12-21]
8. Moore LL, Lombardi DA, White MJ, Campbell JL, Oliveria SA, Ellison RC. Influence of parents' physical activity levels on activity levels of young children. The J of Pedia 1991 Feb;118(2):215-219. [doi: 10.1016/s0022-3476(05)80485-8]
9. Scaglioni S, Salvioni M, Galimberti C. Influence of parental attitudes in the development of children eating behaviour. Br J Nutr 2008 Feb 01;99(S1):S22-S25. [doi: 10.1017/s0007114508892471]
10. Knowlden AP, Sharma M. Systematic review of family and home-based interventions targeting paediatric overweight and obesity. Obes Rev 2012 Jun;13(6):499-508. [doi: 10.1111/j.1467-789X.2011.00976.x] [Medline: 22221298]
11. Berge JM, Everts JC. Family-based interventions targeting childhood obesity: a meta-analysis. Child Obes 2011 Apr;7(2):110-121 [FREE Full text] [doi: 10.1089/chi.2011.07.02.1004.berge] [Medline: 26182126]
12. Brown HE, Atkin AJ, Panter J, Wong G, Chinapaw MJM, van Sluijs EMF. Family-based interventions to increase physical activity in children: a systematic review, meta-analysis and realist synthesis. Obes Rev 2016 Apr;17(4):345-360 [FREE Full text] [doi: 10.1111/obr.12362] [Medline: 26756281]
13. Kriemler S, Meyer U, Martin E, van Sluijs EM, Andersen LB, Martin BW. Effect of school-based interventions on physical activity and fitness in children and adolescents: a review of reviews and systematic update. Br J Sports Med 2011 Sep 11;45(11):923-930 [FREE Full text] [doi: 10.1136/bjsports-2011-090186] [Medline: 21836176]
14. Love R, Adams J, van Sluijs E. Are school-based physical activity interventions effective and equitable? A meta-analysis of cluster randomized controlled trials with accelerometer-assessed activity. Obes Rev 2019 Jun 09;20(6):859-870 [FREE Full text] [doi: 10.1111/obr.12823] [Medline: 30628172]
15. Niemeier BS, Hektner JM, Enger KB. Parent participation in weight-related health interventions for children and adolescents: a systematic review and meta-analysis. Prev Med 2012 Jul;55(1):3-13. [doi: 10.1016/j.ypmed.2012.04.021] [Medline: 22575353]
16. Hammersley ML, Jones RA, Okely AD. Parent-focused childhood and adolescent overweight and obesity ehealth interventions: a systematic review and meta-analysis. J Med Internet Res 2016 Jul 21;18(7):e203 [FREE Full text] [doi: 10.2196/jmir.5893] [Medline: 27443862]
17. Tate DF. Application of innovative technologies in the prevention treatment of overweight in children adolescents. In: Handbook of Childhood and Adolescent Obesity. Boston, MA: Springer; 2008:404.
18. Nguyen B, Kornman KP, Baur LA. A review of electronic interventions for prevention and treatment of overweight and obesity in young people. Obes Rev 2011 May;12(5):298-314. [doi: 10.1111/j.1467-789X.2010.00830.x] [Medline: 21348921]
19. Quelly SB, Norris AE, DiPietro JL. Impact of mobile apps to combat obesity in children and adolescents: a systematic literature review. J Spec Pediatr Nurs 2016 Jan;21(1):5-17. [doi: 10.1111/jspn.12134] [Medline: 26494019]
20. Communications Market Report. UK Communications Market Report 2018. 2018 Aug 2. URL: https://www.ofcom.org.uk/__data/assets/pdf_file/0022/117256/CMR-2018-narrative-report.pdf [accessed 2020-12-21]
21. An J, Hayman LL, Park Y, Dusaj TK, Ayres CG. Web-based weight management programs for children and adolescents. Advances in Nursing Science 2009;32(3):222-240. [doi: 10.1097/ans.0b013e3181b0d6ef]
22. National Institute for Health and Care Excellence. Obesity in Children and Young People: Prevention and Lifestyle Weight Management Programmes. Quality standard [QS94]. 2015 Jul 23. URL: https://www.nice.org.uk/guidance/qs94 [accessed 2020-12-22]
23. Guagliano JM, Brown HE, Coombes E, Hughes C, Jones AP, Morton KL, et al. The development and feasibility of a randomised family-based physical activity promotion intervention: the Families Reporting Every Step to Health (FRESH) study. Pilot Feasibility Stud 2019 Feb 9;5(21):21 [FREE Full text] [doi: 10.1186/s40814-019-0408-7] [Medline: 30788135]
24. Hassoon A, Schrack J, Naiman D, Lansey D, Baig Y, Stearns V, et al. Increasing physical activity amongst overweight and obese cancer survivors using an Alexa-based intelligent agent for patient coaching: protocol for the Physical Activity by Technology Help (PATH) trial. JMIR Res Protoc 2018 Feb 12;7(2):e27 [FREE Full text] [doi: 10.2196/resprot.9096] [Medline: 29434016]
25. Chung AE, Griffin AC, Selezneva D, Gotz D. Health and fitness apps for hands-free voice-activated assistants: content analysis. JMIR Mhealth Uhealth 2018 Sep 24;6(9):e174 [FREE Full text] [doi: 10.2196/mhealth.9705] [Medline: 30249581]
26. Smart Speakers: The Next Step for Modern Survey Design? Lightspeed Research. 2018 Apr 5. URL: https://blog.lightspeedresearch.com/smart-speakers-the-next-step-for-modern-survey-design [accessed 2020-12-21]
27. Rao G. Childhood obesity: highlights of AMA Expert Committee recommendations. Am Fam Physician 2008 Jul 01;78(1):56-63 [FREE Full text] [Medline: 18649611]
28. Department of Health and Social Care. UK Chief Medical Officers' Physical Activity Guidelines. Department of Health. 2019 Sep 7. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/832868/uk-chief-medical-officers-physical-activity-guidelines.pdf [accessed 2020-12-21]
29. Nutrition Science Team, Public Health England. Government recommendations for energy and nutrients for males and females aged 1-18 years and 19+ years. Public Health England (2016) Government Dietary Recommendations. London: Public Health England; 2016 Aug. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/618167/government_dietary_recommendations.pdf [accessed 2020-12-21]

30. 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]
31. Evenson KR, Catellier DJ, Gill K, Ondrak KS, McMurray RG. Calibration of two objective measures of physical activity for children. Journal of Sports Sciences 2008 Dec 15;26(14):1557-1565. [doi: 10.1080/02640410802334196]
32. Golan M, Weizman A. Reliability and validity of the family eating and activity habits questionnaire. Eur J Clin Nutr 1998 Oct 30;52(10):771-777 [FREE Full text] [doi: 10.1038/sj.ejcn.1600647] [Medline: 9805227]
33. Braun V, Clarke V. Using thematic analysis in psychology. Quali Res in Psychol 2006 Jan 21;3(2):77-101 [FREE Full text] [doi: 10.1191/1478088706qp063oa]
34. Public Health England. Latest technology supports new mums to breastfeed. Start4Life. London: Government Digital Service, UK; 2018 Mar 1. URL: https://www.gov.uk/government/news/new-technology-supports-new-mums-to-breastfeed [accessed 2020-12-21]
35. Morgan PJ, Young MD, Smith JJ, Lubans DR. Targeted health behavior interventions promoting physical activity. Exercise and Sport Sciences Reviews 2016 Apr;44(2):71-80 [FREE Full text] [doi: 10.1249/jes.0000000000000075]
36. Bickmore TW, Trinh H, Olafsson S, O'Leary TK, Asadi R, Rickles NM, et al. Patient and consumer safety risks when using conversational assistants for medical information: an observational study of Siri, Alexa, and Google Assistant. J Med Internet Res 2018 Sep 04;20(9):e11510 [FREE Full text] [doi: 10.2196/11510] [Medline: 30181110]
37. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. Health Psychol 2009 Nov;28(6):690-701 [FREE Full text] [doi: 10.1037/a0016136] [Medline: 19916637]
38. Pearson ES. Goal setting as a health behavior change strategy in overweight and obese adults: a systematic literature review examining intervention components. Patient Educ Couns 2012 Apr;87(1):32-42 [FREE Full text] [doi: 10.1016/j.pec.2011.07.018] [Medline: 21852063]
39. Bondaronek P, Alkhaldi G, Slee A, Hamilton FL, Murray E. Quality of publicly available physical activity apps: review and content analysis. JMIR Mhealth Uhealth 2018 Mar 21;6(3):e53 [FREE Full text] [doi: 10.2196/mhealth.9069] [Medline: 29563080]
40. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, Medical Research Council Guidance. Developing and evaluating complex interventions: the new medical research council guidance. BMJ 2008 Sep 29;337(sep29 1):a1655-a1655 [FREE Full text] [doi: 10.1136/bmj.a1655] [Medline: 18824488]

Abbreviations

FEAHQ-R: Revised Family Eating and Activity Habits Questionnaire
IPAP: Intelligent Personal Assistant Project
RCT: randomized controlled trial
SWEET: Safe Wellbeing Eating & Exercise Together

Edited by G Eysenbach; submitted 09.01.20; peer-reviewed by A Griffin, D Thompson; comments to author 29.06.20; revised version received 18.08.20; accepted 31.08.20; published 28.01.21.

Please cite as:
Carlin A, Logue C, Flynn J, Murphy MH, Gallagher AM
Development and Feasibility of a Family-Based Health Behavior Intervention Using Intelligent Personal Assistants: Randomized Controlled Trial
JMIR Form Res 2021;5(1):e17501
URL: http://formative.jmir.org/2021/1/e17501/
doi:10.2196/17501
PMID:33507155

Original Paper

Smartphone-Detected Ambient Speech and Self-Reported Measures of Anxiety and Depression: Exploratory Observational Study

Daniel Di Matteo[1], MASc; Wendy Wang[1], BASc; Kathryn Fotinos[2], BSc (Hons); Sachinthya Lokuge[2], BSc (Hons); Julia Yu[2], MSc; Tia Sternat[2,3], MPsy; Martin A Katzman[2,3,4,5], MD; Jonathan Rose[1], PhD

[1]The Centre for Automation of Medicine, The Edward S Rogers Sr Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada

[2]START Clinic for Mood and Anxiety Disorders, Toronto, ON, Canada

[3]Department of Psychology, Adler Graduate Professional School, Toronto, ON, Canada

[4]Department of Psychology, Lakehead University, Thunder Bay, ON, Canada

[5]The Northern Ontario School of Medicine, Thunder Bay, ON, Canada

Corresponding Author:
Daniel Di Matteo, MASc
The Centre for Automation of Medicine
The Edward S Rogers Sr Department of Electrical and Computer Engineering
University of Toronto
DL Pratt Building
6 King's College Road
Toronto, ON, M5S 3H5
Canada
Phone: 1 416 978 6992
Fax: 1 416 946 8734
Email: dandm@ece.utoronto.ca

Abstract

Background: The ability to objectively measure the severity of depression and anxiety disorders in a passive manner could have a profound impact on the way in which these disorders are diagnosed, assessed, and treated. Existing studies have demonstrated links between both depression and anxiety and the linguistic properties of words that people use to communicate. Smartphones offer the ability to passively and continuously detect spoken words to monitor and analyze the linguistic properties of speech produced by the speaker and other sources of ambient speech in their environment. The linguistic properties of automatically detected and recognized speech may be used to build objective severity measures of depression and anxiety.

Objective: The aim of this study was to determine if the linguistic properties of words passively detected from environmental audio recorded using a participant's smartphone can be used to find correlates of symptom severity of social anxiety disorder, generalized anxiety disorder, depression, and general impairment.

Methods: An Android app was designed to collect periodic audiorecordings of participants' environments and to detect English words using automatic speech recognition. Participants were recruited into a 2-week observational study. The app was installed on the participants' personal smartphones to record and analyze audio. The participants also completed self-report severity measures of social anxiety disorder, generalized anxiety disorder, depression, and functional impairment. Words detected from audiorecordings were categorized, and correlations were measured between words counts in each category and the 4 self-report measures to determine if any categories could serve as correlates of social anxiety disorder, generalized anxiety disorder, depression, or general impairment.

Results: The participants were 112 adults who resided in Canada from a nonclinical population; 86 participants yielded sufficient data for analysis. Correlations between word counts in 67 word categories and each of the 4 self-report measures revealed a strong relationship between the usage rates of death-related words and depressive symptoms (r=0.41, P<.001). There were also interesting correlations between rates of word usage in the categories of reward-related words with depression (r=–0.22, P=.04) and generalized anxiety (r=–0.29, P=.007), and vision-related words with social anxiety (r=0.31, P=.003).

Conclusions: In this study, words automatically recognized from environmental audio were shown to contain a number of potential associations with severity of depression and anxiety. This work suggests that sparsely sampled audio could provide relevant insight into individuals' mental health.

(JMIR Form Res 2021;5(1):e22723) doi:10.2196/22723

KEYWORDS

mobile sensing; passive sensing; psychiatric assessment; mood and anxiety disorders; mobile apps; linguistics; speech recognition; speech content; lexical choice

Introduction

Background

Depression and anxiety disorders are mental health conditions that can, and do, impact people from all geographic and socioeconomic areas of life. Those who suffer from these disorders experience a lower quality of life [1], and many people unknowingly suffer from these disorders due to lack of sufficient access to mental health care or misdiagnoses [2]. The challenge presented by these disorders requires efforts in many areas, including improvements to policy, funding, outreach, treatment, and pharmacotherapy, among others. The diagnosis and assessment of depression and anxiety disorders is also an area where improvements may reduce suffering and improve quality of life for those living with the disorders. In this paper, we explore how fine-grained technology-enhanced observation of patients might give insights into their mental health state.

Modern smartphones are ubiquitous devices that are equipped with a number of sensors that can sense physical activity, geolocation, communication patterns, and the speech of their owners as they go about their day-to-day lives. This sensing capability offers a potential new paradigm for diagnosis and assessment, where instead of asking patients to report their feelings and behaviors relevant to their mental health, it might be possible to infer this information passively and objectively from smartphone-collected data [3]. Given enough data over time, these inferences may prove sufficient to act as a novel severity measure for depression and anxiety disorders. A key advantage of this approach would be that these severity measures would not require expensive, unavailable, or otherwise inaccessible mental health professionals. This study focused specifically on how the linguistic content of speech, recognized from ambient audio recorded by participants' smartphones, may be used as correlates of severity of depression, anxiety, and impairment due to poor mental health.

Prior Work

Our prior efforts explored audio (nonlinguistic) features and correlates with mental health scales [4].

The link between the words spoken by an individual and anxiety or depression has been investigated in 2 major subdomains. The first is the acoustic features of words, that is, the qualities and characteristics of the sounds produced independent of the meaning of the words spoken. While not the focus of this work, prior work has demonstrated numerous quantifiable differences in the acoustic properties of speech in depressed individuals [5]. The literature also shows links between voice acoustics and anxiety [6,7].

The second subdomain upon which this work focused, linguistic analysis, encompasses how an individual's choice of words may relate to symptoms of depression and anxiety. Given this focus, the analysis of the written word and its relationship to anxiety and depression is just as relevant as the spoken word, as the methods employed in this study ignore the additional acoustic information present in the spoken word.

The analysis of speech content and word selection, sometimes referred to as content analysis in the literature, has been studied extensively in psychotherapy contexts [8]. Oxman et al [9] demonstrated that the analysis of speech transcripts of free-form speech could be used to classify psychiatric patients into their respective diagnostic groups with accuracy on par with psychiatric raters. Similar analysis of linguistic style has also been shown to discern between psychiatric inpatients and healthy controls—psychiatric patients used fewer words pertaining to optimism compared to controls (among other differences) [10].

In the linguistic analysis of depression, it has been widely reported that first-person singular pronoun use is correlated with depression severity. A meta-analysis of 21 studies of these correlations confirmed this relationship, where the studies performed analyses of multiple media, including writing, speech, and Facebook status updates [11]. It is believed that this relationship is as a result of the link between depression and self-focused attention [12]. A link between first-person singular pronoun use and social anxiety disorder was also demonstrated [13]. Another linguistic analysis of social anxiety disorder showed that individuals with social anxiety disorder used more positive emotion words than individuals in the control group [14]; the authors hypothesized that such behavior may be a result of the desire to appease others in the effort avoid scrutiny, which is a key fear of social anxious individuals. A number of studies [15] have mined data from social media networks (eg, Twitter) to extract linguistic features which have then been showed to capable to distinguish individuals with mental disorder (eg, depression) from neurotypical controls.

Goal of This Study

While studies [11-15] have demonstrated links between the choice of participants' words and mental health state, the linguistic content of their entire audio environment may shed even more light into mental states, since the environment also contains words spoken by others, such as members of conversations or speakers in news or entertainment media present in the auditory environment. The goal of this exploratory study was to determine if spoken words in recordings of participants' environments may be used to find correlates of

depression, social anxiety disorder, generalized anxiety disorder, and general psychiatric impairment.

Methods

Overview

This study used data collected in a previous study [4]. Participants were recruited from a web-based recruitment platform (Prolific [16]). Participants were not screened for the presence of any psychiatric diagnoses. The study inclusion criteria were the following: participants must (1) reside in Canada, (2) be fluent in English, (3) own an Android phone, (4) have completed at least 95% of their previous Prolific studies successfully, and (5) have previously participated in at least 20 Prolific studies. The final criterion was used to ensure that participants were proficient in using the Prolific system and were generally technology-literate. There were no exclusion criteria for the study. Participants were paid £11 (approximately US $13.37) for participating in the study.

Participants entered a 2-week observational study in which a custom app was installed onto their personal Android phone. Self-report measures of anxiety, depression, and general quality of life were collected at the beginning and end of the study. Throughout the duration of the study, the smartphone app passively collected audiorecordings of the environment (15-second recordings approximately every 5 minutes). The study was approved by the University of Toronto Health Sciences Research Ethics Board (protocol 36687).

Materials and Data

Participants completed 4 self-report measures, in digital form within the study app, at the beginning and end of the 14-day study. A review [17] found that self-administered survey scores do not differ when deployed by app versus other delivery modes. These surveys were completed by participants on their own, with no supervision by clinicians. Participants completed the following 4 self-report measures of mental health: the Liebowitz Social Anxiety Scale (LSAS), which is a 24-item self-report scale used in the assessment of social anxiety disorder [18]; the Generalized Anxiety Disorder 7-item scale (GAD-7), which is an assessment tool for generalized anxiety disorder [19]; the Patient Health Questionnaire 8-item scale (PHQ-8), which is an assessment tool for depression [20]; and the Sheehan Disability Scale, which is a 3-item scale that assesses general impairment due to mental health [21].

The self-report scores collected at the end of the study were used for analysis because the self-report measures ask respondents to evaluate symptoms over the past 2 weeks; therefore, the window of symptom assessment would coincide with the window of electronic data collection.

To assess the severity of the exit scores, we also used the LSAS, GAD-7, and PHQ-8 scores to screen participants for social anxiety, generalized anxiety, and depression, respectively, using diagnostic thresholds found in the literature. A cutpoint of 60 [22] was used with the LSAS scores to screen for social anxiety disorder (generalized subtype). A cutpoint of 10 [19] was used with the GAD-7 scores to screen for generalized anxiety disorder. A cutpoint of 10 [20] was used with the PHQ-8 scores to screen for depression.

Spoken words detected in the participants' environments were collected by the smartphone app. To do so, audiorecordings were collected every 5 minutes for a duration of 15 seconds by the app. These audiorecordings were captured consistently throughout the study at all hours of the day. Transcripts of the audiorecordings were generated using automatic speech recognition software (Google Speech-to-Text [23]). Transcripts of recordings were not checked for correctness by human auditors to preserve participant privacy. Words from each participants' transcripts were stored in randomized order, without any timestamps, to prevent reconstruction of their transcripts, and the audiorecordings were destroyed after transcripts were generated to maintain privacy.

Analysis

A software tool, Linguistic Inquiry and Word Count (LIWC; version 2015; Text Analysis Portal for Research, University of Alberta) was used to analyze participants' words along a number of linguistic and psychological dimensions [24]. LIWC is a tool which was developed to categorize words according to both their linguistic function (ie, what part of speech a word is functioning as a noun, adverb, etc) and according to the words' meanings with respect to psychologically-relevant concepts such as emotions, social concerns, and other constructs. Some of these categories are organized hierarchically, for example, the *affect* category contains the subcategories of positive and negative emotion, and the *negative emotion* category is further broken down into anxiety, sadness, and anger. Examples of these psychological categories, and some of the words within, are given in Table 1.

Participants' environmental words were analyzed using all possible LIWC categories except summary dimensions, punctuation marks, and informal language. This resulted in 67 total categories that were tested, including the top-level categories of function words (ie, parts of speech), other grammar (ie, more parts of speech), affect, social, cognitive processes, perceptual processes, biological processes, drives, time orientation, relativity, and personal concerns.

Participants who completed all study tasks were included in the analysis if the total number of words detected in their ambient audiorecordings was greater than a minimum of 769 words. This minimum threshold was determined by noting that LIWC was built from a corpus of words, and the least frequently observed word category in the corpus (the sexual words category) had a mean frequency of 0.13% [25]. This implies that, on average, 1 in 769 words in the corpus fell within this category. Assuming that the word data collected from participants are similarly distributed, we would require an expected value of 769 words to detect any words in this category; hence, 769 was the minimum threshold.

Table 1. Sample of Linguistic Inquiry and Word Count word categories.

Category	Example words
Personal pronouns	I, them, her
Common verbs	eat, come, carry
Positive emotion	love, nice, sweet
Social processes	mate, talk, they
Death	bury, coffin, kill

The resulting 67 category counts (expressed as the percentage of total words counted which fell within that category) were then tested as correlates of the 4 self-report measures by computing the Pearson correlation coefficient between each category and each measure. Significance of the correlations were tested by computing 2-sided *P* values using the exact distribution of *r*. Due to the exploratory nature of this study, we wished to concisely highlight potentially interesting associations from the large number of correlations measured; therefore, only correlations with an associated *P* value less than .05 are presented. However, due to the large number of comparisons being performed (4 scales × 67 word categories = 268 comparisons), we considered a result statistically significant at a Bonferroni-corrected significance level of α=.0002.

Results

Participant Demographics

Of the 112 participants who completed the study, 86 participants yielded sufficient data for analysis. The study sample consisted of 43% females (37/86) and 57% males (49/86), and the average participant age was 30.1 years (SD 8.5). Participant employment status was as follows: 63% (54/86) were employed in full-time work, 16% (14/86) were employed part-time, 12% (10/86) were unemployed and job seeking, 3% (3/86) were not engaged in paying work (eg, retired or homemaker), and 6% (5/86) reported some other employment status. The 86 participants included in analysis and 26 participants excluded from analysis did not differ in mean age, gender distribution, or mean score of any of the 4 self-report measures.

Self-Report Measures

Table 2 summarizes the self-report measures of the study sample collected at study exit. Intake and exit scores on the LSAS, GAD-7, PHQ-8, and SDS were significantly correlated with r=0.90 (P<.001), r=0.81 (P<.001), r=0.86 (P<.001), and r=0.78 (P<.001), respectively. We interpreted these strong correlations as indicating the reliability of these measures.

Table 2. Results of screening the study sample for depression and anxiety disorders.

Measure	Score, mean (SD)	Diagnostic threshold	Participants over diagnostic threshold (n=86), n (%)
Liebowitz Social Anxiety Scale	53.5 (25.3)	60	32 (37)
Generalized Anxiety Disorder–7	6.5 (4.6)	10	21 (24)
Patient Health Questionnaire–8	8.5 (5.5)	10	30 (35)
Sheehan Disability Scale	10.9 (7.8)	N/A[a]	N/A

[a]N/A: not applicable.

Environmental Audiorecordings

Within the 86-participant sample, the mean number of audiorecordings captured was 3647 (SD 802), and the mean number of recordings that contained speech was 579 (SD 257). On average, 16% of recorded ambient audio contained intelligible speech. This low percentage is reasonable given that recordings were performed throughout all hours of the day. The average number of detected environmental words per participant was 4379 (SD 2625). While the original transcripts were destroyed after generation, the total number of recordings that contained detected speech was recorded for each participant. The mean number of words was 7.4, which seems reasonable given that the audiorecordings were 15 seconds long. All summary statistics for the total number of recordings captured, number of recordings found to contain speech, total detected words, and average word length of the transcripts are presented in Table 3.

Table 3. Summary statistics for word counts of the transcripts of environmental audiorecordings (n=86).

Statistic	Mean (SD)	Minimum	First quartile	Second quartile	Third quartile	Maximum
Total recordings captured	3646 (802)	330	3764	3908	4001	4271
Recordings containing speech	579 (257)	91	390	574	725	1288
Total detected words	4379 (2625)	841	2470	3842	5720	14882
Average number of words in recordings with speech detected	7.4 (2.0)	3.7	6.2	6.8	8.0	15.5

Correlation Analysis

Table 4 presents the correlations between word counts of the LIWC word categories with each of the 4 self-report measures (LSAS, GAD-7, PHQ-8, and SDS) whose *P* values were less than .05. All 67 categories are presented in Multimedia Appendix 1.

Of the correlations presented in Table 4, only the correlation between the *death* category and PHQ-8 scores was statistically significant ($P<.001$) at a Bonferroni-corrected significance level of $\alpha=.0002$. This positive correlation shows that higher rates of death-related words detected in the environment are associated with stronger self-reported symptoms of depression.

Interestingly, the rates of words detected in the *positive emotion* and *negative emotion* categories were both measured as having very low associations with all self-report measures, with the absolute value of the Pearson *r* measured under 0.2 in all cases. The rates of words detected in the *negative emotion* category were most strongly correlated with the PHQ-8 ($r=0.15$, $P=.17$). The rates of words detected in the *positive emotion* category were also most strongly correlated with the PHQ-8 ($r=–0.18$, $P=.09$). Correlations and *P* values for all associations, including word rates in the *positive emotion* and *negative emotion* categories, are presented in Multimedia Appendix 2.

Table 4. Top correlations between Linguistic Inquiry and Word Count categories and Liebowitz Social Anxiety Scale, Generalized Anxiety Disorder–7, Patient Health Questionnaire–8, and Sheehan Disability Scale scores.

Word category	Percentage of total words, mean (SD)	Correlation, *r*	*P* value
Liebowitz Social Anxiety Scale			
death	0.16 (0.10)	0.32	.002
home	0.45 (0.14)	–0.31	.003
see	1.26 (0.28)	0.31	.003
sexual	0.22 (0.29)	–0.24	.02
Generalized Anxiety Disorder–7			
reward	1.61 (0.30)	–0.29	.007
death	0.16 (0.10)	0.27	.01
friend	0.35 (0.15)	0.26	.02
prep	11.75 (1.10)	0.24	.03
bio	2.07 (0.59)	–0.23	.04
relativ	13.57 (1.10)	–0.22	.04
Patient Health Questionnaire–8			
death	0.16 (0.10)	0.41	<.001
function	55.31 (3.13)	0.24	.02
home	0.45 (0.14)	–0.24	.03
reward	1.61 (0.30)	–0.22	.04
Sheehan Disability Scale			
death	0.16 (0.10)	0.28	.009
friend	0.35 (0.15)	0.24	.03
negate	2.29 (0.52)	0.23	.03

Discussion

Key Findings

A key finding is the correlation between the proportion of detected words within the concept of death and all self-reported measures. This correlation was positive in all cases, meaning individuals who had more death-related words detected in their ambient audio displayed worse self-reported symptoms of social anxiety, generalized anxiety, depression, and mental health-related functional impairment. The association between the use of death-related words and depression is in line with previous studies [26,27] showing that depressed individuals tend to use more death-related words. It is important to note that these prior studies [26,27] analyzed only words that were spoken or written by participants, whereas we included all the words detected in the participants' environments.

Other Interesting Findings

In light of the fact that only the correlation between rates of death-related words and the PHQ-8 was statistically significant, it is important to note that the Bonferroni correction is known to be conservative and can cause important relationships to be deemed nonsignificant [28]. That being said, this work has also revealed other interesting potential relationships between different environmental words and mental health.

The first was the positive correlation between vison-related words (the *see* category, including words such as "view," "saw," and "seen") and self-reported symptoms of social anxiety (r=0.31, P=.003). Higher rates of these words being associated with worse symptoms of social anxiety may be related to a known feature of the disorder. Specifically, individuals with social anxiety disorder fear the scrutiny of others, and socially anxious individuals will attempt to detect this scrutiny by visually attending to the others, especially the faces of others [29]. It may be that individuals verbalize this concern about observing this scrutiny throughout their days.

Another interesting relationship was the negative correlation between the rates of the reward-related words in the environment and self-reported symptoms of generalized anxiety (r=–0.29, P=.007) and depression (r=–0.22, P=.04). Lower rates of words in this category, such as "take," "prize," and "benefit" were associated with stronger symptoms of generalized anxiety and depression. In the case of depression, this observed association may be linked to the known deficit in reward processing, and therefore, low hedonic tone noted in depressed individuals [30,31]. If the rates of reward-related words can be used as a proxy for reward-seeking, then lower usage rates of reward-related works might be a result of this diminished capacity to focus or search out and respond to rewards. The link between reward and anxiety is less well-understood, but Gray and McNaughton [32] posited that a key feature of anxiety is related to failure or loss of reward. In this sense, anxious individuals may avoid reward-seeking to avoid triggering anxiety related to potential loss of reward. Again, if rates of reward-related works can be used as a proxy for reward seeking, this may shed some light on the observed relationship between reward-related words and symptoms of generalized anxiety.

Ambient Versus Participant-Only Content Analysis

A key feature of the methodology employed in our study is that the environmental audio recorded for each participant contained speech from any speaker in the environment—the participants themselves but also other humans and recordings (eg, television, radio, music, etc). To the best of our knowledge, no other studies have performed linguistic analysis of audio transcripts containing speech from all ambient sources. This is important to keep in mind when we discuss previous studies that focus only upon speech or writing produced by the participant.

To provide some insight into the impact of other voices in the ambient audio and this study, it is useful to first have an estimate of how much ambient speech is typically produced by the participant and how much comes from other sources. One study [33], which employed a similar audiorecording technology (with wrist-worn smart watches), determined that, of the detected speech in the environment, roughly 18% was produced by the participant, another 18% came from other present people, and 54% from TV and radio. While the presence of other sources of speech in the audio, and therefore in the transcripts, is a confounding factor, it may also contain relevant information. While other individuals will be thought of as polluting the data, the individuals with whom one chooses to associate with may influence one's own state of mind and mental health, especially with regard to depression [34]. Similarly, the presence of words produced by TV or other media in the environmental audio could be a confound but may also contain useful information. As with the company they keep, participants' choices of media may be reflective of their state of mind and mental health. For instance, one study [35] of film preference and mental health showed an association between preference for film noire movies and depression.

Comparisons With Other Studies

The most reported association between participant-only word categories and mental health in the literature is the association between the use of first-person personal pronouns and depression. A meta-analysis [11] estimated the correlation to be small (r=0.13, 95% CI 0.10-0.16). This correlation was also measured to be quite weak in our study of ambient speech (r=0.11, P=.30) but with weaker confidence due to a much smaller sample size.

Several studies [36,37] have investigated associations between participant-only linguistic content in social media posts and self-reported measures of anxiety and depression; these same studies have also used LIWC in their analyses and so can be compared with our work. The comparison has the caveat that our work explored speech from other parties in addition to the participant. A linguistic analysis of Facebook posts revealed positive correlations between the *sadness* self-speech word category and self-reported anxiety (r=0.34, P<.01) [36], whereas our study measured the ambient speech correlation to be much weaker (r=0.07, P=.51). They also measured the correlation between the *sadness* word category and self-reported symptoms of depression (r=0.22, P<.01) [36], which corresponds more closely to our results (r=0.17, P=.13). Another linguistic analysis of Facebook data also found the *sadness* LIWC word category to be a significant predictor of depression diagnosis (standardized regression coefficient β=0.17, P<.001) [37].

Limitations

One technical limitation of this study was the sampling technique used to capture ambient audio. Ambient audiorecordings were produced quite frequently, once every 5 minutes, but for a short duration (only 15 seconds). The short duration of recording helps to preserve smartphone battery life, but it is likely that some conversations or utterances were not captured in full. A more sophisticated sampling technique would record for a variable duration, extending the recording window until silence was detected, so that complete conversations or utterances were captured.

A fundamental limitation is due to the manner in which the environmental audio is used to generate transcripts. Automatic speech recognition software does not perform as well as human transcribers for audio recorded in noisy environments or for audio containing multiple speakers who may be interrupting one another. Furthermore, this software is often being updated and improved; therefore, reproducibility and the ability to do direct comparisons is a key concern for future studies. While this limitation is significant, it is important to also note that the accuracy of Google's Speech-to-Text API (which was used in this study) has been evaluated in clinical talk-therapy settings and demonstrating 83% sensitivity and 83% positive predictive

value in detecting death-related words [38], which implies acceptable validity for the use of this type of data in our analyses.

A final limitation is related to the use of LIWC to perform the linguistic analysis of the transcripts of environmental audio. LIWC is a dictionary-based tool, and as such, categorizes words without looking at contextual information that is key to human language, ignoring sarcasm, metaphor, and analogy.

Conclusion

This study has explored how the proportions of detected words in ambient speech audio across different grammatical and psychological categories may be associated with self-reported symptoms of social anxiety, generalized anxiety, depression, and general psychiatric impairment. We have highlighted several potential relationships, including associations between death-related words, reward-related word, and words related to vision being potentially associated with self-reported measures of social anxiety, generalized anxiety, depression, and general psychiatric impairment.

Acknowledgments

This research was funded in part by the Natural Sciences and Engineering Research Council of Canada Discovery Grants program (grant number RGPIN-2019-04395).

Conflicts of Interest

MK has been a consultant or advisory board member for GlaxoSmithKline, Lundbeck, Eli Lilly, Boehringer Ingelheim, Organon, AstraZeneca, Janssen, Janssen-Ortho, Solvay, Bristol-Myers Squibb, Shire, Sunovion, Pfizer, Purdue, Merck, Astellas, Tilray, Bedrocan, Takeda, Eisai, and Otsuka. MK has undertaken research for GlaxoSmithKline, Lundbeck, Eli Lilly, Organon, AstraZeneca, Jannsen-Ortho, Solvay, Genuine Health, Shire, Bristol-Myers Squibb, Takeda, Pfizer, Hoffman La Rosche, Biotics, Purdue, Astellas, Forest, and Lundbeck. MK has received honoraria from GlaxoSmithKline, Lundbeck, Eli Lilly, Boehringer Ingelheim, Organon, AstraZeneca, Janssen, Janssen-Ortho, Solvay, Bristol-Myers Squibb, Shire, Sunovion, Pfizer, Purdue, Merck, Astellas, Bedrocan, Tilray, Allergan, and Otsuka. MK has received research grants from the Canadian Institutes of Health Research, Sick Kids Foundation, Centre for Addiction and Mental Health Foundation, Canadian Psychiatric Research Foundation, Canadian Foundation for Innovation, and the Lotte and John Hecht Memorial Foundation.

Multimedia Appendix 1
Anonymized study data set including scale scores, audiorecording metadata, and LIWC word category percentages for study participants.
[XLSX File (Microsoft Excel File), 35 KB - formative_v5i1e22723_app1.xlsx]

Multimedia Appendix 2
All tested correlations (Pearson *r* and *P* values) of LIWC word category usage rates and self-report measures.
[XLSX File (Microsoft Excel File), 18 KB - formative_v5i1e22723_app2.xlsx]

References

1. Williams SZ, Chung GS, Muennig PA. Undiagnosed depression: a community diagnosis. SSM Popul Health 2017 Dec;3:633-638 [FREE Full text] [doi: 10.1016/j.ssmph.2017.07.012] [Medline: 29349251]
2. Kasper S. Anxiety disorders: under-diagnosed and insufficiently treated. Int J Psychiatry Clin Pract 2006;10 Suppl 1:3-9. [doi: 10.1080/13651500600552297] [Medline: 24931537]
3. Vaid S, Harari G. Smartphones in Personal Informatics: A Framework for Self-Tracking Research with Mobile Sensing. Cham: Springer International Publishing; 2019:65-92.
4. Di Matteo D, Fotinos K, Lokuge S, Yu J, Sternat T, Katzman MA, et al. The relationship between smartphone-recorded environmental audio and symptomatology of anxiety and depression: exploratory study. JMIR Form Res 2020 Aug 13;4(8):e18751 [FREE Full text] [doi: 10.2196/18751] [Medline: 32788153]
5. Cummins N, Scherer S, Krajewski J, Schnieder S, Epps J, Quatieri TF. A review of depression and suicide risk assessment using speech analysis. Speech Communication 2015 Jul;71:10-49. [doi: 10.1016/j.specom.2015.03.004]
6. Pope B, Blass T, Siegman AW, Raher J. Anxiety and depression in speech. J Consult Clin Psychol 1970;35(1, Pt.1):128-133. [doi: 10.1037/h0029659]
7. Laukka P, Linnman C, Åhs F, Pissiota A, Frans Ö, Faria V, et al. In a nervous voice: acoustic analysis and perception of anxiety in social phobics' Speech. J Nonverbal Behav 2008 Jul 18;32(4):195-214. [doi: 10.1007/s10919-008-0055-9]
8. Gottschalk LA. The application of computerized content analysis of natural language in psychotherapy research now and in the future. Am J Psychother 2000;54(3):305-311. [doi: 10.1176/appi.psychotherapy.2000.54.3.305] [Medline: 11008629]
9. Oxman TE, Rosenberg SD, Schnurr PP, Tucker GJ. Diagnostic classification through content analysis of patients' speech. Am J Psychiatry 1988 Apr;145(4):464-468. [doi: 10.1176/ajp.145.4.464] [Medline: 3348449]

10. Junghaenel DU, Smyth JM, Santner L. Linguistic dimensions of psychopathology: a quantitative analysis. J Soc Clin Psychol 2008 Jan;27(1):36-55. [doi: 10.1521/jscp.2008.27.1.36]
11. Edwards T, Holtzman NS. A meta-analysis of correlations between depression and first person singular pronoun use. J Res Personal 2017 Jun;68:63-68 [FREE Full text] [doi: 10.1016/j.jrp.2017.02.005]
12. Smith TW, Greenberg J. Depression and self-focused attention. Motiv Emot 1981 Dec;5(4):323-331. [doi: 10.1007/bf00992551]
13. Anderson B, Goldin PR, Kurita K, Gross JJ. Self-representation in social anxiety disorder: linguistic analysis of autobiographical narratives. Behav Res Ther 2008 Oct;46(10):1119-1125 [FREE Full text] [doi: 10.1016/j.brat.2008.07.001] [Medline: 18722589]
14. Hofmann SG, Moore PM, Gutner C, Weeks JW. Linguistic correlates of social anxiety disorder. Cogn Emot 2012;26(4):720-726 [FREE Full text] [doi: 10.1080/02699931.2011.602048] [Medline: 21851248]
15. Guntuku SC, Yaden DB, Kern ML, Ungar LH, Eichstaedt JC. Detecting depression and mental illness on social media: an integrative review. Curr Opin Behav Sci 2017 Dec;18:43-49. [doi: 10.1016/j.cobeha.2017.07.005]
16. Online participant recruitment for surveys and market research. Prolific. URL: https://www.prolific.co/ [accessed 2020-01-21]
17. Marcano Belisario JS, Jamsek J, Huckvale K, O'Donoghue J, Morrison CP, Car J. Comparison of self-administered survey questionnaire responses collected using mobile apps versus other methods. Cochrane Database Syst Rev 2015 Jul 27(7):MR000042. [doi: 10.1002/14651858.MR000042.pub2] [Medline: 26212714]
18. Liebowitz MR. Social phobia. Mod Probl Pharmacopsychiatry 1987;22:141-173. [doi: 10.1159/000414022] [Medline: 2885745]
19. 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]
20. Kroenke K, Strine TW, Spitzer RL, Williams JBW, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. J Affect Disord 2009 Apr;114(1-3):163-173. [doi: 10.1016/j.jad.2008.06.026] [Medline: 18752852]
21. Leon AC, Olfson M, Portera L, Farber L, Sheehan DV. Assessing psychiatric impairment in primary care with the Sheehan Disability Scale. Int J Psychiatry Med 1997;27(2):93-105. [Medline: 9565717]
22. Mennin DS, Fresco DM, Heimberg RG, Schneier FR, Davies SO, Liebowitz MR. Screening for social anxiety disorder in the clinical setting: using the Liebowitz Social Anxiety Scale. J Anxiety Disord 2002 Jan;16(6):661-673. [doi: 10.1016/S0887-6185(02)00134-2]
23. Speech-to-text: automatic speech recognition. Google Cloud. URL: https://cloud.google.com/speech-to-text/ [accessed 2020-01-12]
24. Tausczik YR, Pennebaker JW. The psychological meaning of words: LIWC and computerized text analysis methods. J Lang Soc Psychol 2009 Dec 08;29(1):24-54. [doi: 10.1177/0261927X09351676]
25. Pennebaker JW, Boyd RL, Jordan K, Blackburn K. The development and psychometric properties of LIWC2015. University of Texas at Austin. 2015. URL: https://repositories.lib.utexas.edu/handle/2152/31333 [accessed 2020-01-23]
26. Veltman BR. Linguistic analysis of the semantic content of the rorschach inkblot test. PhD thesis. Fuller Theological Seminary. 2006. URL: https://search.proquest.com/docview/304953299 [accessed 2021-01-15]
27. Stirman SW, Pennebaker JW. Word use in the poetry of suicidal and nonsuicidal poets. Psychosom Med 2001;63(4):517-522. [doi: 10.1097/00006842-200107000-00001] [Medline: 11485104]
28. Perneger TV. What's wrong with Bonferroni adjustments. BMJ 1998 Apr 18;316(7139):1236-1238 [FREE Full text] [doi: 10.1136/bmj.316.7139.1236] [Medline: 9553006]
29. McTeague LM, Shumen JR, Wieser MJ, Lang PJ, Keil A. Social vision: sustained perceptual enhancement of affective facial cues in social anxiety. Neuroimage 2011 Jan 15;54(2):1615-1624 [FREE Full text] [doi: 10.1016/j.neuroimage.2010.08.080] [Medline: 20832490]
30. Cooper JA, Arulpragasam AR, Treadway MT. Anhedonia in depression: biological mechanisms and computational models. Curr Opin Behav Sci 2018 Aug;22:128-135 [FREE Full text] [doi: 10.1016/j.cobeha.2018.01.024] [Medline: 29503842]
31. Sternat T, Katzman MA. Neurobiology of hedonic tone: the relationship between treatment-resistant depression, attention-deficit hyperactivity disorder, and substance abuse. Neuropsychiatr Dis Treat 2016;12:2149-2164 [FREE Full text] [doi: 10.2147/NDT.S111818] [Medline: 27601909]
32. Gray JA, McNaughton N. The Neuropsychology of Anxiety: An Enquiry into the Function of the Septo-Hippocampal System. Oxford, UK: Oxford University Press; 2003.
33. Liaqat D, Wu R, Gershon A, Alshaer H, Rudzicz F, de Lara E. Challenges with real-world smartwatch based audio monitoring. In: WearSys '18: Proceedings of the 4th ACM Workshop on Wearable Systems and Applications.: Association for Computing Machinery; 2018 Presented at: 4th ACM Workshop on Wearable Systems and Applications; June 10; Munich, Germany p. 54-59. [doi: 10.1145/3211960.3211977]
34. Joiner TE, Katz J. Contagion of depressive symptoms and mood: meta‐analytic review and explanations from cognitive, behavioral, and interpersonal viewpoints. Clinical Psychology: Science and Practice 1999;6(2):149-164. [doi: 10.1093/clipsy.6.2.149]

35. Till B, Tran US, Voracek M, Sonneck G, Niederkrotenthaler T. Associations between film preferences and risk factors for suicide: an online survey. PLoS One 2014;9(7):e102293 [FREE Full text] [doi: 10.1371/journal.pone.0102293] [Medline: 25028966]
36. Settanni M, Marengo D. Sharing feelings online: studying emotional well-being via automated text analysis of Facebook posts. Front Psychol 2015;6:1045 [FREE Full text] [doi: 10.3389/fpsyg.2015.01045] [Medline: 26257692]
37. Eichstaedt JC, Smith RJ, Merchant RM, Ungar LH, Crutchley P, Preoţiuc-Pietro D, et al. Facebook language predicts depression in medical records. Proc Natl Acad Sci U S A 2018 Dec 30;115(44):11203-11208 [FREE Full text] [doi: 10.1073/pnas.1802331115] [Medline: 30322910]
38. Miner AS, Haque A, Fries JA, Fleming SL, Wilfley DE, Terence Wilson G, et al. Assessing the accuracy of automatic speech recognition for psychotherapy. NPJ Digit Med 2020;3:82 [FREE Full text] [doi: 10.1038/s41746-020-0285-8] [Medline: 32550644]

Abbreviations

GAD: Generalized Anxiety Disorder
LIWC: Linguistic Inquiry and Word Count
LSAS: Liebowitz Social Anxiety Scale
PHQ: Patient Health Questionnaire
SDS: Sheehan Disability Scale

Edited by G Eysenbach; submitted 21.07.20; peer-reviewed by L Ungar, L Castro; comments to author 17.09.20; revised version received 13.10.20; accepted 24.12.20; published 29.01.21.

Please cite as:
Di Matteo D, Wang W, Fotinos K, Lokuge S, Yu J, Sternat T, Katzman MA, Rose J
Smartphone-Detected Ambient Speech and Self-Reported Measures of Anxiety and Depression: Exploratory Observational Study
JMIR Form Res 2021;5(1):e22723
URL: http://formative.jmir.org/2021/1/e22723/
doi:10.2196/22723
PMID:33512325

Original Paper

The Effect of a Name-Based Mask Rationing Plan in Taiwan on Public Anxiety Regarding a Mask Shortage During the COVID-19 Pandemic: Observational Study

Yu-Lin Tai[1], MD; Hsin Chi[2,3], MD; Nan-Chang Chiu[2,3], MD; Cheng-Yin Tseng[1,4], MD; Ya-Ning Huang[2], MD; Chien-Yu Lin[1,3], MD

[1]Hsinchu MacKay Memorial Hospital, Department of Pediatrics and Infectious Disease, Hsinchu, Taiwan

[2]MacKay Children's Hospital, Department of Pediatrics and Infectious Disease, Taipei, Taiwan

[3]MacKay Medical College, Department of Medicine, New Taipei, Taiwan

[4]China Medical University, Graduate Institute of Chinese Medicine, School of Chinese Medicine, Taichung, Taiwan

Corresponding Author:
Chien-Yu Lin, MD
Hsinchu MacKay Memorial Hospital
Department of Pediatrics and Infectious Disease
690 Guanfu Rd, East District
Hsinchu, 300
Taiwan
Phone: 886 6119595 ext 2582
Email: mmhped.lin@gmail.com

Abstract

Background: The COVID-19 pandemic is a severe global health crisis. Wearing a mask is a straightforward action that can be taken, but shortage of stock and equity of allocation were important issues in Taiwan. Furthermore, increased anxiety leading to the stockpiling of masks has been common during the pandemic.

Objective: We aim to summarize the name-based mask rationing plan implemented in Taiwan and explore the public's perceived anxiety about mask shortages.

Methods: The government of Taiwan took action to control the supply and allocation of face masks. We summarize the timeline and important components of the mask rationing plan. A survey that aimed to investigate the overall response to the mask rationing plan was answered by 44 participants.

Results: The mask rationing plan was implemented in late January 2020. Daily production capacity was increased from 2 million masks to 16 million masks in April 2020. People could buy 9 masks in 14 days by verification via their National Health Insurance card. Digital face mask availability maps were created. Moreover, the mask plan safeguarded the purchase of masks and resulted in decreased anxiety about a mask shortage (4.05 [SD 1.15] points; 72.7% [n=32] of participants answered "agree" or "strongly agree"). The majority of people felt that the mask plan was satisfactory (4.2 [SD 0.92] points; 79.5% [n=35] of participants answered "agree" or "strongly agree").

Conclusions: We found that the unique name-based mask rationing plan allowed for control of the production and supply of masks, and contributed to the appropriate allocation of masks. The mask rationing plan not only provided the public with physical protection, but also resulted in reduced anxiety about mask shortages during the pandemic.

(JMIR Form Res 2021;5(1):e21409) doi:10.2196/21409

KEYWORDS

coronavirus; COVID-19; novel coronavirus; SARS-CoV-2; mask; rationing; Taiwan; anxiety; mental health; observational; crisis; plan

Introduction

The COVID-19 pandemic has become a severe global crisis and there were more than 20 million cases as of late August 2020 [1]. This disease is highly contagious with protean clinical manifestations, making it difficult to prevent disease spread [2-4]. The Taiwanese government implemented several strategies early on and Taiwan had a relatively controllable situation [1,5]. As of late August, Taiwan had a total of 496 cases (approximately 20 cases per million residents), with 7 mortalities [6]. Taiwan's success in combating COVID-19 captured our attention and mask use in the public was believed to play a crucial role in the battle against COVID-19 [7-9].

Wearing masks to protect against viral transmission is straightforward but attitudes toward mask use varied across countries [10]. The recommendations regarding mask wearing varied across time and as the severity of the pandemic changed [6,10]. Some people felt discomfort due to having something covering their faces and were afraid of asphyxia; as a result, some people were not willing to wear a mask. Furthermore, the protective effects of wearing masks were doubted in some areas. The protective effectiveness of wearing a mask during a mass gathering was investigated and a relative risk of 0.89 was found [11]. Wearing a mask was found to result in a large reduction of infection risk (adjusted odds ratio 0.15) in a recent systematic review [7]. Although controversies related to wearing masks existed, wearing a mask was believed to be protective during the pandemic [6].

Wearing a mask is common in Asian countries; prior to the COVID-19 pandemic, people wore masks in public places to prevent infection. Due to the sudden, unexpected, and overwhelming pandemic, people panic-bought masks in Taiwan. Increased anxiety about mask shortages led to people stockpiling masks. Thus, the shortage of mask storage, soaring prices, and equity of allocation were important issues in the early phase of the COVID-19 pandemic. The government of Taiwan took action to control the supply and allocation of face masks; this unique mask plan was believed to have greatly contributed to the success of the battle against COVID-19 in Taiwan [5,9,12]. Furthermore, mask-buying surged since the pandemic and anxiety about inadequate mask supply was also noted. In addition to medical illness, mental health issues might arise in response to the COVID-19 pandemic [13-16]. Reallocation and a guarantee of available masks may reduce panic buying during the pandemic. We aim to summarize the unique name-based mask rationing plan in Taiwan and study whether the mask rationing system might have contributed to a reduction in public anxiety about mask shortages during the COVID-19 pandemic.

Methods

Our study was approved by the ethical committee of MacKay Memorial Hospital, Taipei, Taiwan (registration number 20MMHIS140e). We summarize here the strategies undertaken by the government of Taiwan, as well as details of the name-based mask rationing plan, taken from the website of the Centers for Disease Control, Taiwan (CDC) [5]. As the face mask production rate increased, the rationing plan evolved over time.

To investigate the potentially psychogenic impacts of the mask plan, we also conducted a simple survey with a 5-point scale questionnaire that aimed to investigate the overall response to the mask rationing plan. Taiwanese residents aged >18 years were freely recruited at the entrance of our hospital. They were able to read and write Mandarin. The questionnaire was anonymous and not related to medical services. There were 10 simple questions using plain language and it took approximately 1-2 minutes to finish the questionnaire (Table 1). Question 1 explored the respondent's attitude toward mask use. Questions 3, 4, and 8 investigated the number of masks required by the respondent. Questions 5 and 9 were regarding the prices of masks. Questions 2 and 6 surveyed the respondent's perceived anxiety about a potential mask shortage. Finally, questions 7 and 10 investigated the waiting time required to buy masks and the respondent's satisfaction with the mask plan. Participants completed the questionnaire between April 24 and April 30, 2020.

Results

Figure 1 shows the timeline of Taiwan's confirmed cases and the evolution of the name-based mask rationing plan. Figures 2-4 show the different versions of the mask rationing plan. The first case of COVID-19 in Taiwan was diagnosed on January 21, 2020, and a "National Mask Team" was formed in late January [5]. All mask factories were recruited and mask machines were provided by the government to ensure Taiwan was able to produce masks quickly. All masks were allocated by the government and people could buy masks at local pharmacies using a unique name-based mask rationing plan. Verification during purchase was required to ensure every resident could buy the masks they needed and to reduce mask stockpiling. Purchases were verified using a national health insurance card and everyone was allowed to buy 2 masks in a 7-day span in early February 2020 (Figure 2). Initially, production capacity was 2 million masks per day in late January 2020. Production capacity increased to 10 million masks per day in late February 2020. The mask plan evolved to 2.0 and real-time mask maps were established to show the availability of masks. Figure 3 shows the user interface of one software application [17] for face mask availability, deployed on websites, social networking sites, and mobile apps in Taiwan [5,18]. As of April 2020, Taiwan had an adequate mask supply and could donate masks to help other countries. The daily production capacity was approximately 16 million masks and people could buy 9 masks during a 14-day period. The mask plan evolved to 3.0 and residents of Taiwan could make online reservations and payments and buy masks at convenience stores (Figure 4) [19].

Figure 1. Timeline of confirmed cases in Taiwan and the mask rationing plan. (A) January 21: The first confirmed case in Taiwan. (B) January 24: The first day of the export ban on masks. (C) January 28: The first local case of COVID-19. (D) January 31: Mask factories are identified and provided with mask-making machinery. (E) February 6: Implementation of the mask rationing plan begins. (F) Late February: Mask production was increased from 1 million per day in early February to 10 million per day. (G) March 12: Version 2.0 of the mask rationing plan begins. (H) April 8: Masks are donated to other countries. (I) April 9: Version 3.0 of the mask rationing plan begins.

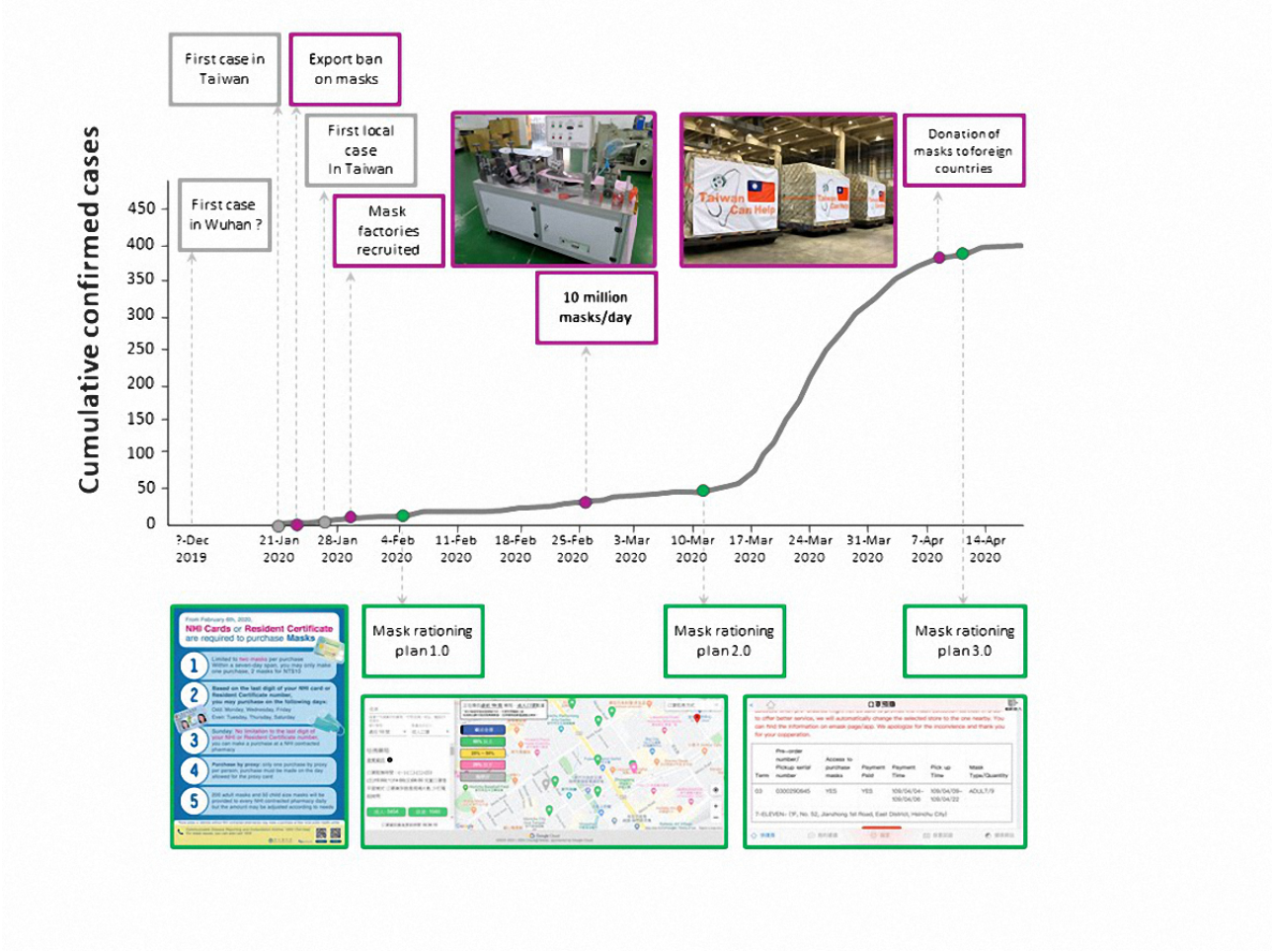
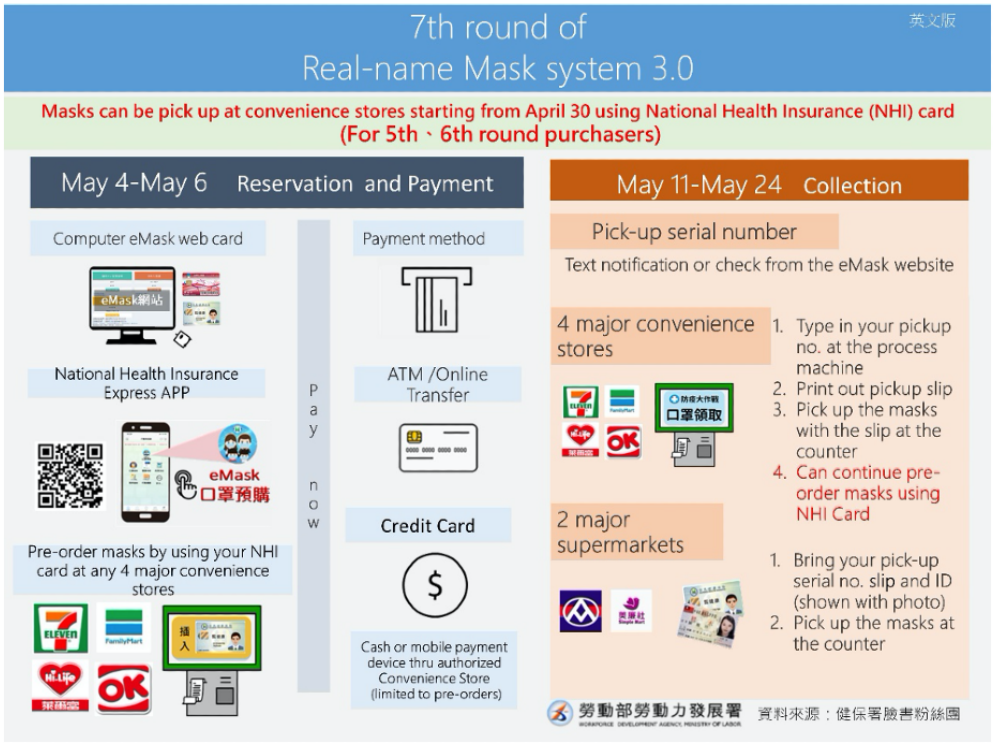


Figure 2. Mask rationing plan 1.0, published on the Ministry of Health and Welfare's Facebook page. Illustration of a new procedure for purchasing medical face masks, which was announced by the Ministry of Health and Welfare of Taiwan on February 6, 2020. Surgical masks can be purchased at local pharmacies upon presentation of a National Health Insurance card. Medical staff are permitted 1-2 masks per day and others can purchase 3 masks per week. A real-time mask map website provided information on mask availability.

Figure 3. Mask rationing plan 2.0. Screenshot of a map-view software application for face mask availability. Mask purchases can be made by reservation using a mobile app and are available for purchase at convenience stores. Digital mapping software displays information including the pharmacy name, available quantity of adult-sized masks, available quantity of child-sized masks, opening hours, pharmacy phone number, pharmacy address, update time, and opening hours of nearby pharmacies. The triangle colors correspond to the availability of different types of masks or data about masks.

Figure 4. Mask rationing plan 3.0. Adults may purchase 9 masks in a 2-week period, while 10 masks can be purchased for children in a 2-week period. Masks can be purchased via mobile phones or from machines located in convenience stores (the typical transaction time is 1 minute). Masks can be sent abroad.



The questionnaire was administered to 44 adults residing in Taiwan. Table 1 shows the responses of the 44 participants to questionnaires investigating their need for masks and their satisfaction with the mask plan. Most participants agreed that mask wearing is protective (average score of 4.8 [SD 0.47], with 42 participants [95.5%] answering “agree” or “strongly agree”). On average, people needed 1 mask per day (question 8) and the amount allocated per person might be inadequate (average score of 2.07 for question 3 and 2.98 for question 4). Some participants felt anxious if they could not buy an adequate number of masks (score 3.7 [SD 1.22], with 26 participants [59.1%] answering “agree” or “strongly agree”) and the mask plan may have contributed to decreasing this anxiety (score 4.05 [SD 1.15], with 32 participants [72.7%] answering “agree” or “strongly agree”). The average acceptable price was 5.23 NTD (US $0.19) per mask and the price of each mask distributed

through the mask plan was 6 NTD. The average amount of time participants were willing to stand in a queue to buy masks was 20.1 (SD 15.2) minutes. In addition, the majority of respondents felt that the mask rationing plan was satisfactory (average score of 4.2 [SD 0.92] points, with 35 participants [79.5%] answering "agree" or "strongly agree").

Table 1. Participant responses to questionnaires investigating the need for masks and their satisfaction with the mask plan (N=44).

Questions	Average response[a]	Participants answering disagree or strongly disagree, n (%)	Participants answering agree or strongly agree, n (%)
1. I believe that wearing a mask may decrease the risk of infection.	4.8 (0.47)	0 (0)	42 (95.5)
2. I feel anxiety if I am unable to buy masks.	3.7 (1.22)	7 (15.9)	26 (59.1)
3. I need only 3 masks for 1 week.	2.07 (1.18)	30 (68.2)	6 (13.6)
4. I need only 9 masks for 2 weeks.	2.98 (1.47)	19 (43.2)	19 (43.2)
5. I feel the price is too high.	2.75 (1.43)	19 (43.2)	11 (25)
6. A mask rationing plan will decrease my anxiety.	4.05 (1.15)	5 (11.4)	32 (72.7)
7. In general, I feel that the mask rationing plan in Taiwan is satisfactory.	4.2 (0.92)	3 (6.8)	35 (79.5)
8. On average, how many masks do you use per day?	1.03 (0.37)	N/A[b]	N/A
9. I think a reasonable price per mask is…	5.23 (2.71) NTD (US $0.19 [$0.1])	N/A	N/A
10. How many minutes would you stand in a queue to buy masks?	20.1 (15.2) minutes	N/A	N/A

[a]Responses to questions 1-7 are presented as mean (SD), where 1=strongly disagree and 5=strongly agree; questions 8-10 were open-ended questions.

[b]N/A: not applicable.

Discussion

In Taiwan, several strategies were implemented to reduce the spread of COVID-19 and the unique name-based mask rationing plan was believed to play a crucial role in Taiwan's success in the battle against the virus [5]. This mask plan was executed in January 2020. In this study, we summarized the timeline of the mask plan and introduced some details about clinical practice. Furthermore, we conducted a survey to investigate public responses to the mask plan and whether the mask plan may contribute to decreased anxiety about mask availability.

There have been controversies about the protective effectiveness of masks. In addition, attitudes toward wearing masks have varied across countries. Wearing a face mask is a straightforward and simple measure that the general public can use to prevent the spread of contaminated droplets and limit virus transmission. Specialized medical masks are a critical component of personal protective equipment in clinical settings. However, the effectiveness of mask use among those in the general community remains controversial [7,10,20]. Wearing masks is a common practice in Asian countries, although this is not the case among all cultures [10,21]. Taiwan and some other Asian countries mandated mask wearing in public places, while wearing masks was not mandatory in other countries, such as the United States, Canada, and some European countries. As time went by, mounting evidence demonstrated the benefit of mask wearing for reducing infection spread and it has since become a common recommendation [6]. Additionally, rather than simply wearing a mask due to its potential protective effectiveness, wearing a mask may also be a symbol of safety in some countries [21]. During the COVID-19 pandemic, despite soaring prices, supplies of face masks were rapidly depleted and they quickly became unavailable at stores in the community. Mask stockpiling was not beneficial for infection control and reducing disease spread. In addition, the shortage of personal protective equipment, including masks, among frontline health care personnel also increased the risk of nosocomial infection. To improve the situation, the government of Taiwan implemented a mask rationing plan. The recruitment of all mask factories and an increase in production capacity ensured adequate mask production. The name-based system ensured equitable allocation and people felt less anxious because they could purchase masks. The COVID-19 pandemic is ongoing and this nationwide name-based mask plan may serve as a reference for policy makers worldwide.

The pandemic can cause stress and approximately 1 in 5 COVID-19 survivors experienced mental health problems within 90 days of their COVID-19 diagnosis [22]. The use of masks at the community level may be associated with better mental health [16]. Our questionnaire asked respondents about anxiety related to the face mask shortage at the beginning of the COVID-19 pandemic. People agreed that masks were likely to provide them with effective protection (95.5% [n=42] of participants answered "agree" or "strongly agree") and more than half of the participants (59.1%, n=26) felt anxious if they were unable to obtain them. However, the required quantities

of masks differed among respondents and not everyone was satisfied with the number of masks allocated under the plan (the percentage of participants saying they "disagree" or "strongly disagree" that a given number of masks was adequate was 68.2% [n=30] for question 3 [3 masks in 1 week] and 43.2% [n=19] for question 4 [9 masks in 2 weeks]). The mask rationing plan safeguarded mask purchasing and reduced the public's perceived anxiety about mask shortages (among 72.7% [n=32] of participants). In a study performed in Poland, Maciaszek et al [23] also indicated an overall decrease in psychopathological symptoms after wearing face coverings in public spaces became obligatory. The mask plan ensured everyone had some masks to wear and this may have had psychological benefits. The implementation of this name-based mask plan and the high uptake of mask wearing may not only have prevented virus transmission, but also decreased anxiety about mask shortages. Although the specific price individuals were willing to pay per mask and acceptable waiting time in queues to purchase masks differed, 4 out of 5 people felt that the mask rationing plan was satisfactory (79.5% [n=35] answered "agree" or "strongly agree"). Further studies are warranted to elucidate the full impact of the mask plan.

While controversy regarding the effectiveness of mask use remains and our study was limited to a small group of responders, the results from this pilot study suggest that the nationwide strategy of mask rationing contributed to the appropriate allocation of masks and a reduction in anxiety about mask shortages. The diagnosis of psychiatric conditions is rigorous and based on the Diagnostic and Statistical Manual of Mental Disorders. We did not aim to confirm a causal relationship between the mask shortage and mental illness; rather, we conducted this preliminary study to indicate whether there is a potential relationship between the mask plan and the public's perceived anxiety during the pandemic. Further large-scale population-based studies are required to draw stronger conclusions.

In conclusion, the disease burden of the COVID-19 pandemic was still increasing when the name-based mask plan was implemented and the mask plan contributed to the success of Taiwan's battle against COVID-19. In this study, we highlighted the important timeline dates and components of this mask plan; this may serve as a reference for policy makers. The mask plan safeguarded mask allocation and may also have decreased perceived anxiety during the pandemic. Further studies are required.

Acknowledgments

We received no funding for this study. We thank everyone for their efforts to combat COVID-19.

Authors' Contributions

YLT and CYL were responsible for the conception of the study. HC, NCC, CYT, YNH, and YLT were involved in data collection. NCC and CYT designed the questionnaire. YLT created the figures and wrote the first draft. All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. World Health Organization. Coronavirus disease (COVID-19) outbreak situation. 2020. URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019 [accessed 2020-09-12]
2. Chi H, Chiu N, Tai Y, Chang H, Lin C, Sung Y, et al. Clinical features of neonates born to mothers with coronavirus disease-2019: A systematic review of 105 neonates. J Microbiol Immunol Infect 2020 Aug 14:1 [FREE Full text] [doi: 10.1016/j.jmii.2020.07.024] [Medline: 32847748]
3. Chi H, Chiu N, Peng C, Lin C, Tai Y, Lee M, et al. One-Seventh of Patients with COVID-19 Had Olfactory and Gustatory Abnormalities as Their Initial Symptoms: A Systematic Review and Meta-Analysis. Life (Basel) 2020 Aug 22;10(9):1 [FREE Full text] [doi: 10.3390/life10090158] [Medline: 32842563]
4. Wiersinga WJ, Rhodes A, Cheng AC, Peacock SJ, Prescott HC. Pathophysiology, Transmission, Diagnosis, and Treatment of Coronavirus Disease 2019 (COVID-19): A Review. JAMA 2020 Aug 25;324(8):782-793. [doi: 10.1001/jama.2020.12839] [Medline: 32648899]
5. Centers For Disease Control and Prevention. Coronavirus disease 2019 (COVID-19). 2020. URL: https://www.cdc.gov.tw/En [accessed 2020-09-12]
6. COVID-19 Coronavirus Pandemic. Worldometer. URL: https://www.worldometers.info/coronavirus/ [accessed 2020-09-12]
7. Chu DK, Akl EA, Duda S, Solo K, Yaacoub S, Schünemann HJ, COVID-19 Systematic Urgent Review Group Effort (SURGE) study authors. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. Lancet 2020 Jun 27;395(10242):1973-1987 [FREE Full text] [doi: 10.1016/S0140-6736(20)31142-9] [Medline: 32497510]
8. Chang C, Tan T, Ho T, Chen C, Su T, Lin C. COVID-19: Taiwan's epidemiological characteristics and public and hospital responses. PeerJ 2020;8:e9360 [FREE Full text] [doi: 10.7717/peerj.9360] [Medline: 32551205]

9. Chiu N, Chi H, Tai Y, Peng C, Tseng C, Chen C, et al. Impact of Wearing Masks, Hand Hygiene, and Social Distancing on Influenza, Enterovirus, and All-Cause Pneumonia During the Coronavirus Pandemic: Retrospective National Epidemiological Surveillance Study. J Med Internet Res 2020 Aug 20;22(8):e21257 [FREE Full text] [doi: 10.2196/21257] [Medline: 32750008]
10. Feng S, Shen C, Xia N, Song W, Fan M, Cowling BJ. Rational use of face masks in the COVID-19 pandemic. Lancet Respir Med 2020 May;8(5):434-436 [FREE Full text] [doi: 10.1016/S2213-2600(20)30134-X] [Medline: 32203710]
11. Barasheed O, Alfelali M, Mushta S, Bokhary H, Alshehri J, Attar AA, et al. Uptake and effectiveness of facemask against respiratory infections at mass gatherings: a systematic review. Int J Infect Dis 2016 Jun;47:105-111 [FREE Full text] [doi: 10.1016/j.ijid.2016.03.023] [Medline: 27044522]
12. Chen C, Tseng C, Choi W, Lee Y, Su T, Hsieh C, et al. Taiwan Government-Guided Strategies Contributed to Combating and Controlling COVID-19 Pandemic. Front Public Health 2020;8:547423 [FREE Full text] [doi: 10.3389/fpubh.2020.547423] [Medline: 33194949]
13. Huang H, Zhao W, Li G. Knowledge and Psychological Stress Related to COVID-19 Among Nursing Staff in a Hospital in China: Cross-Sectional Survey Study. JMIR Form Res 2020 Sep 18;4(9):e20606 [FREE Full text] [doi: 10.2196/20606] [Medline: 32640419]
14. Jacobson NC, Lekkas D, Price G, Heinz MV, Song M, O'Malley AJ, et al. Flattening the Mental Health Curve: COVID-19 Stay-at-Home Orders Are Associated With Alterations in Mental Health Search Behavior in the United States. JMIR Ment Health 2020 Jun 01;7(6):e19347 [FREE Full text] [doi: 10.2196/19347] [Medline: 32459186]
15. Ni MY, Yang L, Leung CMC, Li N, Yao XI, Wang Y, et al. Mental Health, Risk Factors, and Social Media Use During the COVID-19 Epidemic and Cordon Sanitaire Among the Community and Health Professionals in Wuhan, China: Cross-Sectional Survey. JMIR Ment Health 2020 May 12;7(5):e19009 [FREE Full text] [doi: 10.2196/19009] [Medline: 32365044]
16. Wang C, Chudzicka-Czupała A, Grabowski D, Pan R, Adamus K, Wan X, et al. The Association Between Physical and Mental Health and Face Mask Use During the COVID-19 Pandemic: A Comparison of Two Countries With Different Views and Practices. Front Psychiatry 2020;11:569981 [FREE Full text] [doi: 10.3389/fpsyt.2020.569981] [Medline: 33033485]
17. Kiang GitHub. URL: https://kiang.github.io/pharmacies/ [accessed 2021-01-10]
18. Yuan EJ, Hsu C, Lee W, Chen T, Chou L, Hwang S. Where to buy face masks? Survey of applications using Taiwan's open data in the time of coronavirus disease 2019. J Chin Med Assoc 2020 Jun;83(6):557-560 [FREE Full text] [doi: 10.1097/JCMA.0000000000000325] [Medline: 32304508]
19. Taiwan government website. URL: https://fw.wda.gov.tw/wda-employer/home/file/2c95efb371bafaef0171ca58b08b42f0 [accessed 2021-01-10]
20. Liang M, Gao L, Cheng C, Zhou Q, Uy JP, Heiner K, et al. Efficacy of face mask in preventing respiratory virus transmission: A systematic review and meta-analysis. Travel Med Infect Dis 2020;36:101751 [FREE Full text] [doi: 10.1016/j.tmaid.2020.101751] [Medline: 32473312]
21. Nakayachi K, Ozaki T, Shibata Y, Yokoi R. Why Do Japanese People Use Masks Against COVID-19, Even Though Masks Are Unlikely to Offer Protection From Infection? Front Psychol 2020;11:1918 [FREE Full text] [doi: 10.3389/fpsyg.2020.01918] [Medline: 32849127]
22. Taquet M, Luciano S, Geddes JR, Harrison PJ. Bidirectional associations between COVID-19 and psychiatric disorder: retrospective cohort studies of 62 354 COVID-19 cases in the USA. Lancet Psychiatry 2020 Nov 09 [FREE Full text] [doi: 10.1016/S2215-0366(20)30462-4] [Medline: 33181098]
23. Szczesniak D, Ciulkowicz M, Maciaszek J, Misiak B, Luc D, Wieczorek T, et al. Psychopathological responses and face mask restrictions during the COVID-19 outbreak: Results from a nationwide survey. Brain Behav Immun 2020 Jul;87:161-162 [FREE Full text] [doi: 10.1016/j.bbi.2020.05.027] [Medline: 32389696]

Abbreviations

CDC: Centers for Disease Control, Taiwan

Edited by G Eysenbach; submitted 14.06.20; peer-reviewed by C Leung, CF Yen; comments to author 28.10.20; revised version received 17.11.20; accepted 24.12.20; published 22.01.21.

Please cite as:
Tai YL, Chi H, Chiu NC, Tseng CY, Huang YN, Lin CY
The Effect of a Name-Based Mask Rationing Plan in Taiwan on Public Anxiety Regarding a Mask Shortage During the COVID-19 Pandemic: Observational Study
JMIR Form Res 2021;5(1):e21409
URL: http://formative.jmir.org/2021/1/e21409/
doi:10.2196/21409
PMID:33400678

Original Paper

COVID-19–Related Knowledge and Practices Among Health Care Workers in Saudi Arabia: Cross-sectional Questionnaire Study

Talal Ahmad Shaikhain[1,2], MBBS; Faisal Abdulmohsen Al-Husayni[1,2], MBBS; Essam Awadh Alhejaili[1], MBBS; Maha Nawaf Al-Harbi[3], BSN; Anas Abdullah Bogari[1], MBBS; Bayan Abdulaziz Baghlaf[1], MBBS; Mohammed Saeed Alzahrani[2,4,5], MD

[1]Department of Internal Medicine, National Guard Hospital, Jeddah, Saudi Arabia

[2]King Abdullah International Medical Research Center, King Abdulaziz Medical City, Jeddah, Saudi Arabia

[3]Department of Nursing, National Guard Hospital, Jeddah, Saudi Arabia

[4]Department of Infectious Disease, National Guard Hospital, Jeddah, Saudi Arabia

[5]King Saud bin Abdulaziz University for Health Sciences, Jeddah, Saudi Arabia

Corresponding Author:
Mohammed Saeed Alzahrani, MD
Department of Infectious Disease
National Guard Hospital
Old Makkah Road
Jeddah, 22384
Saudi Arabia
Phone: 966 556655712
Email: mzahrani@gmail.com

Abstract

Background: Health care workers are at the front line against COVID-19. The risk of transmission decreases with adequate knowledge of infection prevention methods. However, health care workers reportedly lack a proper attitude and knowledge of different viral outbreaks.

Objective: This study aimed to assess the knowledge and attitude of health care workers in Saudi Arabia toward COVID-19. Assessment of these parameters may help researchers focus on areas that require improvement.

Methods: A cross-sectional questionnaire study was conducted among 563 participants recruited from multiple cities in Saudi Arabia. An online questionnaire was shared via social media applications, which contained questions to health care workers about general information regarding COVID-19 and standard practices.

Results: The mean age of the study population was 30.7 (SD 8) years. Approximately 8.3% (47/563) of the health care workers were isolated as suspected cases of COVID-19, and 0.9% (n=5) were found positive. The majority agreed that social distancing, face masks, and hand washing are effective methods for preventing disease transmission. However, only 63.7% (n=359) knew the correct duration of hand washing. Almost 70% (n=394) strictly adhered to hand hygiene practices, but less than half complied with the practice of wearing a face mask. Significant differences in health care workers' attitudes were observed on the basis of their city of residence, their adherence to COVID-19 practices, and their compliance with the use of a face mask. Among the health care workers, 27.2% (n=153) declared that they will isolate themselves at home and take influenza medication if they experience COVID-19 symptoms.

Conclusions: The majority of health care workers in Saudi Arabia presented acceptable levels of general knowledge on COVID-19, but they lack awareness in some crucial details that may prevent disease spread. Intense courses and competency assessments are highly recommended. Prevention of disease progression is the only option for the time being.

(JMIR Form Res 2021;5(1):e21220) doi:10.2196/21220

KEYWORDS

COVID-19; health care workers; infection control; attitude; knowledge; Saudi Arabia

Introduction

SARS-CoV-2 is a novel virus of the large group of coronaviruses circulating in the environment and is thought to originate from bats [1]. Previous outbreaks such as severe acute respiratory syndrome (SARS) in 2003 and Middle East respiratory syndrome (MERS) in 2015 share similarities with COVID-19 [2]. This novel viral outbreak was epidemiologically linked to the Hua Nan seafood and wet animal wholesale market [3]. Moreover, SARS-CoV-2 was first discovered in Wuhan City, Hubei Province, China, by Chinese authorities. It was initially reported to manifest as pneumonia cases of unknown etiology on December 31, 2019 [4]. Later on, China officially announced the identification of a novel virus, which caused the pneumonia. Shortly after, the World Health Organization (WHO) had declared the outbreak of a novel coronavirus [5]. In February 2020, the disease was named COVID-19 [6].

People infected with COVID-19 may experience a wide range of symptoms, from mild to severe illness. These symptoms include cough, shortness of breath, fever, muscle pain, chills, sore throat, and loss of the sense of taste or smell [7]. However, these symptoms are not universal, as other studies have reported patients with gastrointestinal symptoms such as nausea, vomiting, or diarrhea [7].

According to the WHO, approximately 80% of COVID-19 patients in China experienced mild symptoms and recovered without any medical intervention [8], while 14% of them had experienced severe illness, and 5% were critically ill. However, the risk of having severe illness is higher in the elderly and individuals with underlying chronic diseases such as cancer, diabetes, and lung diseases [8].

Regarding the current state of COVID-19 in Saudi Arabia, the government imposed a curfew from March 23 to June 20, 2020. Mosques, schools, and businesses were closed during that period, and travel was restricted. At the time of writing, Saudi Arabia has reported approximately 49,176 COVID-19 cases, which is lesser than those reported in western countries [9]. In health care settings, all COVID-19 patients were initially hospitalized regardless of disease severity and treated free of charge, including visa violators [10]. Similar to the rest of the world, Saudi Arabia had experienced a shortage of personal protective equipment (PPE), prompting recommendations from the Saudi Center for Disease Prevention and Control on the use and reuse of available PPE [11]. Furthermore, outpatient clinics started seeing most patients virtually, and nonurgent consultations were rescheduled.

According to the Saudi Ministry of Health (SMOH)'s statistical yearbook of 2018, the health care workforce includes 36,717 physicians, 83,616 nurses, 3277 pharmacists, and over 50,000 allied health personnel [12]. Furthermore, health care workers are at the front line and directly come in contact with COVID-19 patients. Consequently, they are always at high risk of infection. The transmission of any disease among health care workers is mainly associated with overcrowding, the absence of isolation facilities, and environmental contamination [13]. However, the transmission risk might also be related to inadequate knowledge of methods for infection prevention [14]. Consequently, health care workers need to have adequate awareness of proper infection prevention practices. In a study conducted at District 2 Hospital, Ho Chi Minh City, Vietnam, the majority (88.4%) of health care workers had adequate knowledge of COVID-19, and 90% of participants have a positive attitude toward COVID-19 [2].

It is essential to have infection control guidelines with the best available evidence to deal with COVID-19 in every health care setting and maximally avoid exposure to the virus. Emphasis should be placed on hand hygiene, which is known to be the best way to prevent the spread of microorganisms and microbial infections in health care facilities [15]. Education on proper PPE, patient screening, and mask use should be provided in accordance with the guidelines of the WHO and the Centers for Disease Control and Prevention (CDC) [16-18]. Previous studies have reported that health care workers might lack a proper attitude and knowledge toward SARS and MERS [19-21]. Therefore, this study aimed to assess the knowledge and attitude toward COVID-19 among health care workers in Saudi Arabia. This assessment may help prevent disease transmission by identifying areas requiring intervention.

Methods

Study Design

A cross-sectional questionnaire-based study was performed with health care workers in Saudi Arabia to assess their level of awareness, knowledge, and perception of COVID-19, their level of adherence to the applied curfew, and their understanding of methods for infection prevention. Convenience sampling was carried out by sending the questionnaire through social media platforms (Twitter and WhatsApp), as face-to-face interviews were unavailable owing to curfew regulations. Considering this data collection method, the number of health care workers who received the questionnaire could not be identified because they were encouraged to share the questionnaire within their social circle of health care workers; however, the initial number of health care workers among whom the questionnaire was shared was 1068. The study included health care workers within Saudi Arabia, while those who did not complete the questionnaire or those who worked abroad were excluded. A self-administered questionnaire was developed and distributed from April 30 to May 14, 2020. The questionnaire covered the following items: sociodemographic data such as age, nationality, city of residence, and employment status during the curfew.

Cities were divided as large (population >300,000), medium (population ranging 100,000-300,000), and small (population <100,000) cities. The categorization of cities sizes was based on the measures of the Saudi General Authority for Statistics [22]. The questionnaire also assessed the level of knowledge using "agree," "neutral," and "disagree" statements, which also included questions about the duration of hand washing, COVID-19 symptoms, and the timing for COVID-19 testing. Regarding symptoms, the respondents were provided with a list of established COVID-19 symptoms and asked to choose items related to the disease. The Saudi guidelines recommend COVID-19 testing when individuals experience severe respiratory symptoms or flu-like symptoms, or if they come in

contact with positive individuals or those with flu-like symptoms. These options were provided to the participants in addition to "any time." The complete questionnaire is available as Multimedia Appendix 1. After explaining the study objectives to the participants and assuring their confidentiality, the participants were asked to complete the questionnaire. At the end of the questionnaire survey, an email regarding any inquiries was sent to the participants. Informed consent was obtained before data collection, and no identifiers were requested. None of the responders was compensated, and the data were only accessible to the authors to assure confidentiality. The study received ethical approval from the King Abdullah International Medical Research Center (RJ20/079/J).

Statistical Analysis

Data were entered and analyzed using SPSS (version 25, IBM Corp). Data are presented as ranges, means, SD, medians, and IQR for quantitative variables and frequencies and percentages for qualitative variables. Between-group comparisons were performed using χ^2 or Fisher exact tests. Results are also expressed as odds ratio (OR) and 95% CI values. *P* values less than .05 were considered statistically significant.

Results

A total of 563 health care workers completed the questionnaire survey. As indicated in Table 1, the participants' ages ranged from 21 to 69 years. The majority of participants (n=537, 95.4%) were Saudi nationals. Furthermore, 47 (8.3%) health care workers were isolated as suspected COVID-19 cases, and 5 (0.9%) of them tested positive.

Table 2 summarizes the levels of knowledge among the participants, indicated through "agree," "neutral," and "disagree" questions. Most of the cohort (n=542, 96.3%) agreed that COVID-19 is a pandemic, while 71.2% (n=401) thought it is more dangerous than seasonal influenza. The highest percentage of agreement (n=547, 97.2%) was obtained for social distancing being an effective method to prevent COVID-19 transmission, followed by hand washing (n=544, 96.6%) and impending curfew (n=542, 96.3%). Furthermore, 33.6% (n=189) of health care workers agreed that COVID-19 transmission could be prevented by wearing gloves.

Table 1. Sociodemographic characteristics of health care workers in Saudi Arabia (N=563).

Criterion	Value
Age (years)	
Range	21-69
Mean (SD)	30.7 (8)
Median (IQR)	28 (25-33)
Sex, n (%)	
Male	322 (57.2)
Female	241 (42.8)
Nationality, n (%)	
Saudi	537 (95.4)
Non-Saudi	26 (4.6)
City of residence, n (%)	
Large	459 (81.5)
Medium	78 (13.9)
Small	26 (4.6)
Participants with chronic diseases, n (%)	80 (14.2)
Living with people older than 65 years, n (%)	179 (31.8)
Diagnosed with COVID-19, n (%)	5 (0.9)
Isolated as a suspected case of COVID-19, n (%)	47 (8.3)
Working status during curfew, n (%)	
Yes, I go to work daily	329 (58.4)
Yes, I work online	142 (25.2)
No	92 (16.3)

Table 2. Levels of knowledge of COVID-19 among health care workers in Saudi Arabia (N=563).

Item	Agree, n (%)	Neutral, n (%)	Disagree, n (%)
COVID-19 is a pandemic	542 (96.3)	14 (2.5)	7 (1.2)
COVID-19 is more dangerous than seasonal influenza	401 (71.2)	105 (18.7)	57 (10.1)
COVID-19 is only dangerous among the elderly and patients with chronic diseases	142 (25.2)	102 (18.1)	319 (56.7)
Hand washing is effective to prevent transmission of COVID-19	544 (96.6)	17 (3)	2 (0.4)
Social distancing is effective to prevent transmission of COVID-19	547 (97.2)	13 (2.3)	3 (0.5)
Wearing face masks is effective to prevent transmission of COVID-19	438 (77.8)	100 (17.8)	25 (4.4)
Wearing hand gloves is effective to prevent transmission of COVID-19	189 (33.6)	172 (30.6)	202 (35.9)
Impending curfew is effective to prevent transmission of COVID-19	520 (92.4)	33 (5.9)	10 (1.8)

When asked about the recommended duration of hand washing to prevent COVID-19 transmission, only 359 (63.8%) of the health care workers selected 40-60 s, while 180 (31.9%) selected 20-30 s, and 24 (4.3%) selected 10-15 s.

Health care workers were provided a list of symptoms and asked to select those related to COVID-19. As shown in Figure 1, the top selected symptoms were cough or shortness of breath (552/563, 98.1%) and fever (n=533, 94.7%). The lowest percentage (n=199, 35.4%) was for a runny nose.

Figure 2 shows the responses to the question "when should a person seek testing for COVID-19?" The most frequent response (509/563, 90.4%) was when contacting someone positive for COVID-19, followed by when experiencing severe respiratory symptoms (n=455, 80.8%). Few (n=62, 11%) health care workers chose to test for COVID-19 at any time, even if asymptomatic. Furthermore, 561 (99.6%) health care workers answered "Yes" when asked about the probability of COVID-19 patients being asymptomatic. Moreover, 532 (94.5%) health care workers were aware of the absence of an established therapy for COVID-19.

Figure 1. Knowledge of COVID-19 symptoms among health care workers in Saudi Arabia (N=563).

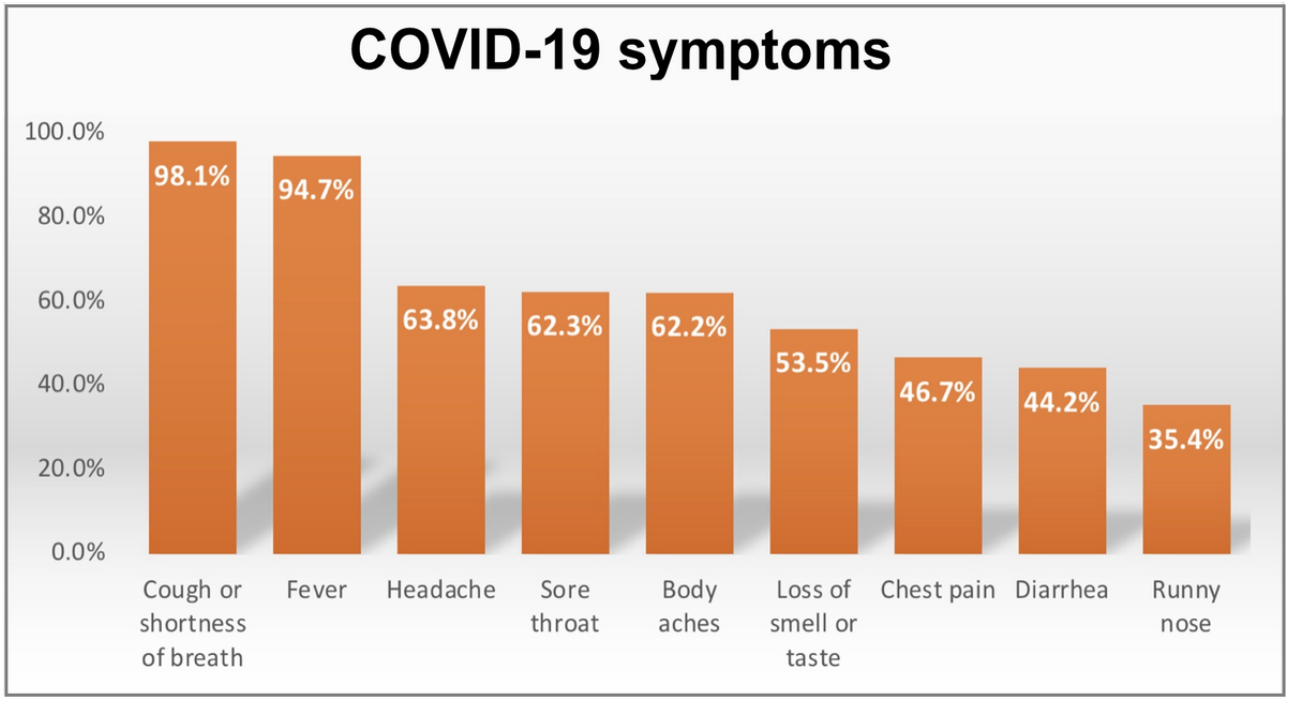


Figure 2. Responses of health care workers in Saudi Arabia regarding the timing for seeking a COVID-19 test (N=563).

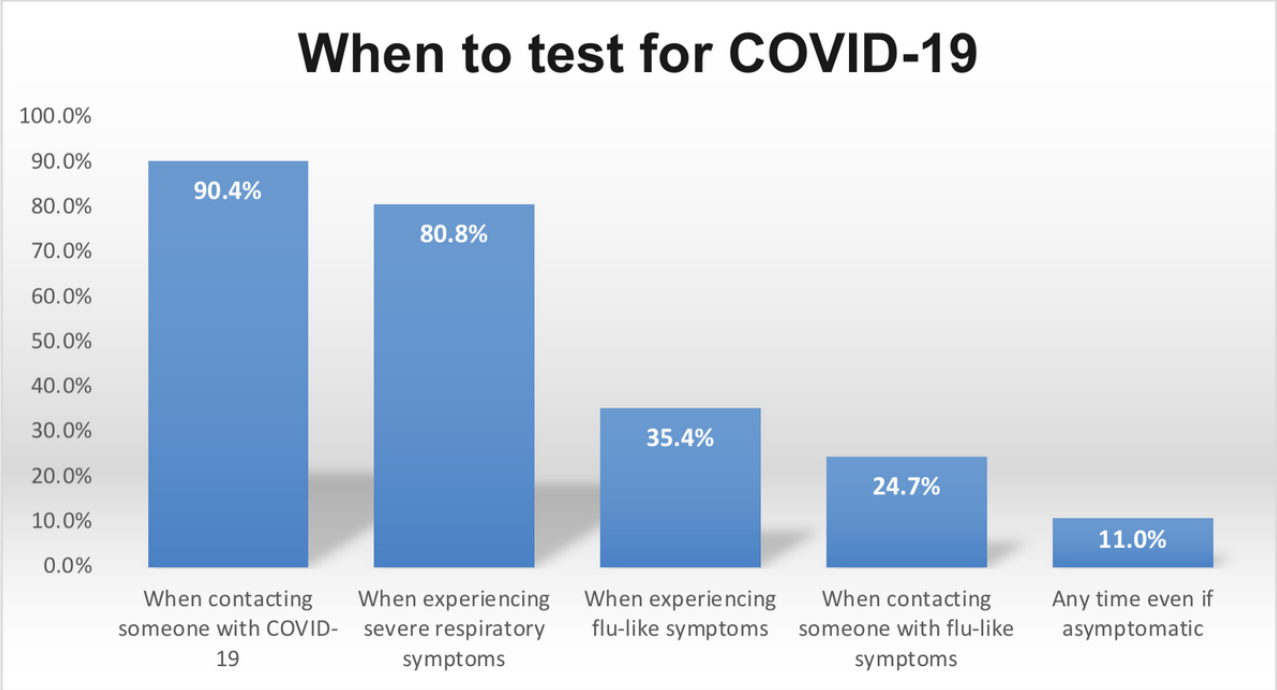


Table 3 summarizes the attitude of health care workers toward COVID-19. Among them, 461 (81.9%) were always compliant with curfew regulations. Most (n=409, 72.6%) health care workers were always compliant with hand washing. With regards to wearing face masks in public places, 264 (46.9%) health care workers were always compliant, and 116 (20.6%) were not compliant. Furthermore, 290 (51.5%) health care workers frequently followed up with COVID-19–related news.

Table 3. Attitudes of health care workers in Saudi Arabia toward COVID-19 (N=563).

Questions	Participants, n (%)
Are you compliant with curfew regulations?	
Always	461 (81.9)
Most of the time	94 (16.7)
Sometimes	5 (0.9)
No	3 (0.5)
Are you compliant with hand washing?	
Always	409 (72.6)
Most of the time	138 (24.5)
Sometimes	12 (2.1)
No	4 (0.7)
Are you compliant with wearing face masks in public places?	
Always	264 (46.9)
Most of the time	114 (20.2)
Sometimes	69 (12.3)
No	116 (20.6)
Do you follow COVID-19 news?	
Always	290 (51.5)
Most of the time	156 (27.7)
Sometimes	93 (16.5)
No	24 (4.3)

When asking the participants what they would do if they experience flu-like symptoms, 350 (62.2%) responded that they would call the SMOH hotline for advice. In comparison, 153 (27.2%) health care workers responded that they would stay at home and take flu medication. Fifty (8.9%) participants responded that they would go to the hospital to test for COVID-19, and 10 (1.8%) would not take any action.

When comparing health care workers living in large, medium, or small cities (Table 4), a significant difference was observed in their compliance with wearing face masks in public places (*P*=.04). The larger the city, the more compliant the participant. Furthermore, health care workers in medium and small cities followed COVID-19 news more than their peers in large cities (*P*=.02).

Finally, 479 (85.1%) health care workers followed COVID-19 news from official authorities including the SMOH and the WHO, while 61 (10.8%) followed news from social media platforms, 19 (3.4%) followed daily news, and only 4 (0.7%) obtained COVID-19 news from their friends.

Table 4. Comparison of the attitude toward COVID-19 among health care workers in large, medium, and small cities in Saudi Arabia (N=563).

Questions	City		*P* value	OR[a] (95% CI)
	Large (N=459), n (%)	Medium/small (N=104), n (%)		
Are you compliant to curfew regulations?			.66	0.627 (0.076-5.151)
Always/most of the time	452 (98.5)	103 (99)		
Sometimes/no	7 (1.5)	1 (1)		
Are you compliant to hand washing?			.49	1.49 (0.471-4.716)
Always/most of the time	447 (97.4)	100 (96.2)		
Sometimes/no	12 (2.6)	4 (3.8)		
Are you compliant to wearing face masks in public places?			.04[b]	1.574 (1.016-2.438)
Always/most of the time	317 (69.1)	61 (58.7)		
Sometimes/no	142 (30.9)	43 (41.3)		
Do you follow COVID-19 news?			.02[b]	0.488 (0.262-0.907)
Always/most of the time	355 (77.3)	91 (87.5)		
Sometimes/no	104 (22.7)	13 (12.5)		

[a]OR: odds ratio.

[b]Statistically significant (*P*<.05).

Discussion

Principal Findings

This study illustrates the knowledge and practices of health care workers in Saudi Arabia at the early stages of the pandemic during a period of significant uncertainty and rapidly changing policies and practices. Among our study participants, marked consensus was observed in their responses to hand hygiene, social distancing, and curfew regulations for effectively preventing disease transmission. Responses to questions on masks and gloves were widely distributed, probably owing to unclear information during the early stages of the pandemic from both the literature and local policies. Moreover, when asked about the timing for COVID-19 testing, most responded with “on experiencing severe symptoms” or “on coming in contact with positive cases,” reflecting the local messaging at that time. Furthermore, their compliance with general hand hygiene and universal masking was concerning and represents an area of improvement.

When faced with a novel viral pandemic, particularly one with no vaccine or effective treatment at the time of writing, other aspects of disease control become increasingly important. The SMOH implemented daily televised briefings with relevant statistics and discussions regarding the best practices for the current time, and any inquiries usually made by the press were addressed. Practices including hand hygiene and social distancing had the most robust emphasis, while messages regarding the worldwide use of masks were inconsistent owing to their shortage in hospitals and the need to reserve them for frontline health care workers. While the public should be preferentially informed of the best available practices to reduce disease transmission, a higher emphasis should be placed on health care workers, since they constitute a high-risk group for contracting COVID-19, and by the nature of their occupation, they have direct contact with an especially vulnerable part of our community. Hence, it is essential to assess their knowledge and practice and compare them to those of their peers elsewhere. This study also provides an insight into the early stages of the knowledge, attitudes, and practices for disease management among the health care workers, which are expected to change as the pandemic evolves or when more information becomes available.

Multiple outbreaks were reported in health care settings, emphasizing the need for infection control and prevention [23,24]. Risk perception reportedly enhances compliance with protective measures [25]. Approximately 71% of individuals believed that COVID-19 is more dangerous than seasonal

influenza, and slightly more than half were aware that COVID-19 could be hazardous to individuals other than the elderly, indicating an area of improvement. Moreover, a study on Egyptian health care workers reported that almost 90% of them believed that the virus is more dangerous in the elderly [26]. Furthermore, Bhagavathula et al [27] reported that only 11.4% of health care workers agreed that COVID-19 is a fatal disease.

In this study, approximately 8.3% (47/563) of health care workers were isolated as suspected cases of COVID-19; fortunately, only 0.9% (n=5) tested positive, and this number is likely to increase as the spread of the pandemic progresses.

Most of our study participants believe in adopting nationwide protective measures, including social distancing, maintenance of regular hand hygiene, and universal use of face masks during public activities. If these beliefs translate into practice, it could help decrease transmission by decreasing the reproductive number or "flattening the curve," allowing for better utilization of health care facilities or buying time until vaccine or treatment availability [28,29]. Interestingly, a study form Uganda [30] reported that 55% of health care workers do not believe that face masks may help prevent disease transmission, while almost all of them agreed that avoiding crowded places decreases the risk of acquiring COVID-19. Social distancing proved to be one of the most effective methods of preventing disease transmission during the initial COVID-19 outbreak in Wuhan [31].

Regarding hand hygiene, almost all our study participants agreed on the importance of hand washing, which is higher than reported in other studies [27,32], but only 63.8% (n=359) were aware of the correct duration of washing, which is at least 40 s [33].

The primary source of the participants' knowledge was the SMOH daily press briefings and its updates about COVID-19, which contained evidence-based information when available in different areas, including the best infection control practices, policies, and regulations to be implemented and various misconceptions and misinformation about COVID-19. This reflects a drastic improvement in the spread of information alongside practical knowledge through a simple, widely accessible tool such as the television, as opposed to that reported by Khan et al [34] during the MERS outbreak. In their study, the participants faced difficulty following news updates about the disease on the internet from the SMOH website and in looking for new emerging studies. In another study by Albarrak et al [35], the sources of information for the study participants during the MERS outbreak were almost equally distributed among seminars, pamphlets, articles, radio, and television.

We believe that the SMOH performed an admirable job in handling the pandemic and provided transparency and continuous information regarding changes to policies as new data emerged or as the pandemic evolved. Of particular note is the high uniformity in the responses to the messaging, and areas of uncertainty included low levels of knowledge and practices in our study population. We believe that complete transparency and clear messaging are needed for maximum benefits during such events. This study provides a cross-sectional insight into a relatively early stage of the pandemic, and comparisons can potentially be made with the emergence of more data from other countries.

Limitations

Our study did not define the specialty of the health care workers (eg, nurse, physician, or pharmacist). We also believe that the categorization of health care settings by type (eg, outpatient department, rural hospital, or polyclinic) would have provided more context to the participants' responses.

Furthermore, our study is limited by its convenience sampling method, which might have introduced a potential selection bias. Furthermore, the self-reporting nature of the study questionnaire might have introduced its own set of biases, such as social desirability.

Conclusion

In conclusion, the majority of our questionnaire respondents had acceptable general knowledge of COVID-19, based on their responses to our questions. Knowledge of decreased disease transmission with the use of face masks was not as uniform as we expected, perhaps reflecting the unclear messaging at that time. Furthermore, approximately half of the study participants disagreed with the statement that COVID-19 is only dangerous in the elderly. Other areas of improvement include the knowledge of the recommended duration of hand washing. Compliance with precautions for infection prevention still need to be emphasized; this can be achieved through intense educational programs and competency assessments to promote positive preventive practices. This study provides a cross-sectional insight into the relatively early stages of the COVID-19 pandemic in Saudi Arabia, and if additional similar studies from other countries become available, comparisons can be made between different populations.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The questionnaire used in the study.
[DOCX File , 17 KB - formative_v5i1e21220_app1.docx]

References

1. Tang X, Wu C, Li X, Song Y, Yao X, Wu X, et al. On the origin and continuing evolution of SARS-CoV-2. Natl Sci Rev 2020;7(6):1012-1023. [doi: 10.1093/nsr/nwaa036]
2. Huynh G, Nguyen TH, Tran V, Vo K, Vo V, Pham L. Knowledge and attitude toward COVID-19 among healthcare workers at District 2 Hospital, Ho Chi Minh City. Asian Pac J Trop Med 2020;13(6):260-265. [doi: 10.4103/1995-7645.280396]
3. Mackenzie JS, Smith DW. COVID-19: a novel zoonotic disease caused by a coronavirus from China: what we know and what we don't. Microbiol Aust 2020 Mar 17:MA20013 [FREE Full text] [doi: 10.1071/MA20013] [Medline: 32226946]
4. Pneumonia of unknown cause – China. World Health Organization. 2020 Jan 5. URL: https://www.who.int/csr/don/05-january-2020-pneumonia-of-unkown-cause-china/en/ [accessed 2020-05-14]
5. Novel Coronavirus (2019-nCoV): Situation Report 11. World Health Organization. 2020 Jan 31. URL: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200131-sitrep-11-ncov.pdf?sfvrsn=de7c0f7_4 [accessed 2020-05-10]
6. Rolling Updates on Coronavirus Disease (COVID-19). World Health Organization. 2020 Jul 31. URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen [accessed 2020-05-10]
7. Symptoms of Coronavirus. Centers for Disease Control and Prevention. 2020 Dec 22. URL: https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html [accessed 2020-05-11]
8. Coronavirus disease 2019 (COVID-19): Situation Report 41. World Health Organization. 2020 Mar 01. URL: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200301-sitrep-41-covid-19.pdf?sfvrsn=6768306d_2 [accessed 2020-05-10]
9. (COVID-19) Disease Interactive Dashboard. Saudi Center for Disease Prevention and Control. URL: https://covid19.cdc.gov.sa/daily-updates/ [accessed 2020-05-15]
10. Coronavirus: Saudi's King Salman orders treatment for all, including visa violators. Al Arabiya News. URL: https://english.alarabiya.net/en/News/gulf/2020/03/ [accessed 2020-09-26]
11. Guidance for Extended Use and Limited Reuse of N95 Respirators. Saudi Center for Disease Prevention and Control. URL: https://covid19.cdc.gov.sa/professionals-health-workers/guidance-for-extended-use-and-limited-reuse-of-n95-respirators/ [accessed 2020-09-26]
12. Statistical Yearbook 2018. Ministry of Health. URL: https://www.moh.gov.sa/en/Ministry/Statistics/book/Pages/default.aspx [accessed 2020-09-12]
13. Zhang M, Zhou M, Tang F, Wang Y, Nie H, Zhang L, et al. Knowledge, attitude, and practice regarding COVID-19 among healthcare workers in Henan, China. J Hosp Infect 2020 Jun;105(2):183-187 [FREE Full text] [doi: 10.1016/j.jhin.2020.04.012] [Medline: 32278701]
14. McEachan R, Taylor N, Harrison R, Lawton R, Gardner P, Conner M. Meta-Analysis of the Reasoned Action Approach (RAA) to Understanding Health Behaviors. Ann Behav Med 2016 Aug 11;50(4):592-612 [FREE Full text] [doi: 10.1007/s12160-016-9798-4] [Medline: 27169555]
15. Hand Hygiene Recommendations. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/coronavirus/2019-ncov/hcp/hand-hygiene.html [accessed 2020-05-14]
16. Using Personal Protective Equipment (PPE). Centers for Disease Control and Prevention. URL: https://www.cdc.gov/coronavirus/2019-ncov/hcp/using-ppe.html [accessed 2020-05-17]
17. Rational use of personal protective equipment for coronavirus disease (COVID-19) and considerations during severe shortages: interim guidance, 6 April 2020. World Health Organization. URL: https://apps.who.int/iris/handle/10665/331695 [accessed 2020-05-17]
18. Infection prevention and control during health care when novel coronavirus (nCoV) infection is suspected. World Health Organization. URL: https://www.who.int/publications-detail/infection-prevention-and-control-during-health-care-when-novel-coronavirus-(ncov)-infection-is-suspected-20200125 [accessed 2020-05-17]
19. Deng J, Olowokure B, Kaydos-Daniels S, Chang H, Barwick R, Lee M, SARS International Field Team. Severe acute respiratory syndrome (SARS): knowledge, attitudes, practices and sources of information among physicians answering a SARS fever hotline service. Public Health 2006 Jan;120(1):15-19 [FREE Full text] [doi: 10.1016/j.puhe.2005.10.001] [Medline: 16298404]
20. Althomairy S, Baseer M, Assery M, Alsaffan A. Knowledge and Attitude of Dental Health Professionals about Middle East Respiratory Syndrome in Saudi Arabia. J Int Soc Prev Community Dent 2018;8(2):137-144 [FREE Full text] [doi: 10.4103/jispcd.JISPCD_9_18] [Medline: 29780739]
21. Alsahafi A, Cheng A. Knowledge, Attitudes and Behaviours of Healthcare Workers in the Kingdom of Saudi Arabia to MERS Coronavirus and Other Emerging Infectious Diseases. Int J Environ Res Public Health 2016 Dec 06;13(12):1214 [FREE Full text] [doi: 10.3390/ijerph13121214] [Medline: 27929452]
22. Population Characteristics surveys 2017. Saudi General Authority for Statistics. URL: https://www.stats.gov.sa/en/854-0 [accessed 2020-05-17]
23. McMichael TM, Currie DW, Clark S, Pogosjans S, Kay M, Schwartz NG, et al. Epidemiology of Covid-19 in a Long-Term Care Facility in King County, Washington. N Engl J Med 2020 May 21;382(21):2005-2011 [FREE Full text] [doi: 10.1056/NEJMoa2005412] [Medline: 32220208]

24. Pierce M. A short preliminary report on nursing homes and Covid-19: Measures introduced in Ireland. LTC Responses to COVID-19. 2020 Apr 03. URL: https://ltccovid.org/2020/04/03/a-short-preliminary-report-on-nursing-homes-and-covid-19-measures-introduced-in-ireland/ [accessed 2020-05-19]
25. Prati G, Pietrantoni L, Zani B. Compliance with recommendations for pandemic influenza H1N1 2009: the role of trust and personal beliefs. Health Educ Res 2011 Oct 25;26(5):761-769. [doi: 10.1093/her/cyr035] [Medline: 21613380]
26. Abdel Wahed WY, Hefzy EM, Ahmed MI, Hamed NS. Assessment of Knowledge, Attitudes, and Perception of Health Care Workers Regarding COVID-19, A Cross-Sectional Study from Egypt. J Community Health 2020 Dec 07;45(6):1242-1251 [FREE Full text] [doi: 10.1007/s10900-020-00882-0] [Medline: 32638199]
27. Bhagavathula AS, Aldhaleei WA, Rahmani J, Mahabadi MA, Bandari DK. Knowledge and Perceptions of COVID-19 Among Health Care Workers: Cross-Sectional Study. JMIR Public Health Surveill 2020 Apr 30;6(2):e19160 [FREE Full text] [doi: 10.2196/19160] [Medline: 32320381]
28. Kissler S, Tedijanto C, Lipsitch M, Grad YH. Social distancing strategies for curbing the COVID-19 epidemic. medRxiv. Preprint posted online March 24, 2020. [doi: 10.1101/2020.03.22.20041079]
29. Dalton CB, Corbett SJ, Katelaris AL. COVID-19: implementing sustainable low cost physical distancing and enhanced hygiene. Med J Aust 2020 Jun;212(10):443-446.e1 [FREE Full text] [doi: 10.5694/mja2.50602] [Medline: 32356573]
30. Olum R, Chekwech G, Wekha G, Nassozi DR, Bongomin F. Coronavirus Disease-2019: Knowledge, Attitude, and Practices of Health Care Workers at Makerere University Teaching Hospitals, Uganda. Front Public Health 2020 Apr 30;8:181 [FREE Full text] [doi: 10.3389/fpubh.2020.00181] [Medline: 32426320]
31. Wilder-Smith A, Freedman D. Isolation, quarantine, social distancing and community containment: pivotal role for old-style public health measures in the novel coronavirus (2019-nCoV) outbreak. J Travel Med 2020 Mar 13;27(2) [FREE Full text] [doi: 10.1093/jtm/taaa020] [Medline: 32052841]
32. Saqlain M, Munir M, Rehman S, Gulzar A, Naz S, Ahmed Z, et al. Knowledge, attitude, practice and perceived barriers among healthcare workers regarding COVID-19: a cross-sectional survey from Pakistan. J Hosp Infect 2020 Jul;105(3):419-423 [FREE Full text] [doi: 10.1016/j.jhin.2020.05.007] [Medline: 32437822]
33. World Health Organization. 2009 Aug. URL: https://www.who.int/gpsc/5may/Hand_Hygiene_Why_How_and_When_Brochure.pdf [accessed 2020-06-01]
34. Khan MU, Shah S, Ahmad A, Fatokun O. Knowledge and attitude of healthcare workers about Middle East Respiratory Syndrome in multispecialty hospitals of Qassim, Saudi Arabia. BMC Public Health 2014 Dec 16;14(1):1281 [FREE Full text] [doi: 10.1186/1471-2458-14-1281] [Medline: 25510239]
35. Albarrak AI, Mohammed R, Al Elayan A, Al Fawaz F, Al Masry M, Al Shammari M, et al. Middle East Respiratory Syndrome (MERS): Comparing the knowledge, attitude and practices of different health care workers. J Infect Public Health 2019 Aug 17;14(1):89-96 [FREE Full text] [doi: 10.1016/j.jiph.2019.06.029] [Medline: 31431424]

Abbreviations

CDC: Centers for Disease Control and Prevention
MERS: Middle East respiratory syndrome
OR: odds ratio
SARS: severe acute respiratory syndrome
SMOH: Saudi Ministry of Health
WHO: World Health Organization

Edited by G Eysenbach; submitted 08.06.20; peer-reviewed by W Al-Chetachi, C Jacob; comments to author 06.08.20; revised version received 08.10.20; accepted 17.01.21; published 25.01.21.

Please cite as:
Shaikhain TA, Al-Husayni FA, Alhejaili EA, Al-Harbi MN, Bogari AA, Baghlaf BA, Alzahrani MS
COVID-19–Related Knowledge and Practices Among Health Care Workers in Saudi Arabia: Cross-sectional Questionnaire Study
JMIR Form Res 2021;5(1):e21220
URL: http://formative.jmir.org/2021/1/e21220/
doi:10.2196/21220
PMID:33460390

Original Paper

A Novel Artificial Intelligence-Powered Emotional Intelligence and Mindfulness App (Ajivar) for the College Student Population During the COVID-19 Pandemic: Quantitative Questionnaire Study

Ronda Sturgill[1*], PhD; Mary Martinasek[1*], PhD; Trine Schmidt[2*], MA; Raj Goyal[2*], MD

[1]Department of Health Sciences and Human Performance, The University of Tampa, Tampa, FL, United States

[2]Ajivar, Tarpon Springs, FL, United States

[*]all authors contributed equally

Corresponding Author:
Ronda Sturgill, PhD
Department of Health Sciences and Human Performance
The University of Tampa
401 West Kennedy Blvd
Box 30F
Tampa, FL, 33606
United States
Phone: 1 8132573445
Email: rsturgill@ut.edu

Abstract

Background: Emotional intelligence (EI) and mindfulness can impact the level of anxiety and depression that an individual experiences. These symptoms have been exacerbated among college students during the COVID-19 pandemic. Ajivar is an app that utilizes artificial intelligence (AI) and machine learning to deliver personalized mindfulness and EI training.

Objective: The main objective of this research study was to determine the effectiveness of delivering an EI curriculum and mindfulness techniques using an AI conversation platform, Ajivar, to improve symptoms of anxiety and depression during this pandemic.

Methods: A total of 99 subjects, aged 18 to 29 years, were recruited from a second-semester group of freshmen students. All participants completed the online TestWell Wellness Inventory at the start and end of the 14-week semester. The comparison group members (49/99, 49%) were given routine mental wellness instruction. The intervention group members (50/99, 51%) were required to complete Ajivar activities in addition to routine mental wellness instruction during the semester, which coincided with the onset of the COVID-19 pandemic. This group also completed assessments to evaluate for anxiety, using the 7-item Generalized Anxiety Disorder (GAD-7) scale, and depression, using the 9-item Patient Health Questionnaire (PHQ-9).

Results: Study participants reported a mean age of 19.9 (SD 1.94) years; 27% (27/99) of the group were male and 60% (59/99) identified as Caucasian. No significant demographic differences existed between the comparison and intervention groups. Subjects in the intervention group interacted with Ajivar for a mean time of 1424 (SD 1168) minutes. There was a significant decrease in anxiety, as measured by the GAD-7: the mean score was 11.47 (SD 1.85) at the start of the study compared to 6.27 (SD 1.44) at the end (P<.001). There was a significant reduction in the symptoms of depression measured by the PHQ-9: the mean score was 10.69 (SD 2.04) at the start of the study compared to 6.69 (SD 2.41) at the end (P=.001). Both the intervention and comparison groups independently had significant improvements in the TestWell Wellness Inventory from pretest to posttest. The subgroups in the social awareness and spirituality inventories showed significant improvement in the intervention group. In a subgroup of participants (11/49, 22%) where the GAD-7 was available during the onset of the COVID-19 pandemic, there was an increase in anxiety from the start of the study (mean score 11.63, SD 2.16) to mid-March (ie, onset of the pandemic) (mean score 13.03, SD 1.48; P=.23), followed by a significant decrease at the end of the study period (mean score 5.9, SD 1.44; P=.001).

Conclusions: It is possible to deliver EI and mindfulness training in a scalable way using the Ajivar app during the COVID-19 pandemic, resulting in improvements in anxiety, depression, and EI in the college student population.

(JMIR Form Res 2021;5(1):e25372) doi:10.2196/25372

KEYWORDS

mindfulness; COVID-19; college students; emotional intelligence

Introduction

Background

Students who do not perform well at college often have low emotional intelligence (EI), even students with a high IQ. EI involves a set of learned noncognitive skills, which helps individuals foster better relationships, improve time and stress management, maintain better impulse control, and improve problem-solving abilities [1-3]. Mindfulness has been defined as the awareness that arises by "paying attention to the present moment on purpose and non-judgmentally" [4]. EI and mindfulness can impact the level of anxiety and depression an individual experiences. This is particularly important during stressful times, such as during the COVID-19 pandemic. Ajivar is an app that utilizes artificial intelligence (AI) and machine learning (ML) to deliver personalized mindfulness and EI training to college students. Through ongoing engagement of the user, the app can decrease stress and anxiety while improving performance and emotional well-being.

Anxiety and depression are ever prominent among college students. More than 50% of the 20 million college students in the United States reported symptoms of anxiety and depression that prevented them from performing well at college [5]. The symptoms associated with anxiety and depression often persist over time. More than 75% to 85% of college students have reported feeling stressed and overwhelmed [6-8]. There are several reasons college students experience disproportionately higher levels of stress than the general population, including first-time independence leading to having to manage their academic responsibilities, financial responsibility, and planning for their future [9,10]. Despite the prevalence and persistence of anxiety and depression, up to 75% of college students do not access the needed resources and needed assistance, primarily because of the underlying stigma associated with a diagnosis and mental health services [11]. Additionally, colleges and universities have limited resources to aid the large number of students who suffer from mental health conditions. These conditions are further exacerbated by extenuating circumstances, such as the COVID-19 pandemic.

COVID-19

On March 11, 2020, the World Health Organization declared COVID-19, an acute respiratory syndrome caused by SARS-CoV-2, a global pandemic [12]. Not only has this novel pandemic resulted in severe health consequences, but it has also impacted the mental well-being of college students who have struggled with immediate displacement from college, learning online models of course delivery, and having to socially distance from their friends and social support. In a self-reported survey by the Healthy Minds Network during the COVID-19 crisis, college students reported increased mental health concerns (ie, anxiety, depression, suicidal ideation, etc), which resulted in poor academic performance [13]. Increased frequency of anxiety and depression have been reported among undergraduate students during the pandemic [14,15].

Initial findings from the pandemic suggest that the mental health consequences will be far reaching. Time will tell the true impact of the social and behavioral changes experienced by college students. Providing novel, scalable, and disruptive tools to help this vulnerable section of the population is going to be important if we are to prevent long-term mental health sequelae from the pandemic.

Emotional Intelligence and Mindfulness

EI is "a set of skills hypothesized to contribute to the accurate appraisal and expression of emotion in oneself and in others, the effective regulation of emotion in self and others, and the use of feelings to motivate, plan, and achieve in one's life" [16]. The ability to accurately identify, manage, and express emotions is the foundation of good self-esteem, healthy relationships, and good work and academic performance [17]. EI is not commonly taught in colleges and universities, which can inhibit students' full potential of learning due to lack of coping skills in managing stress, anxiety, and interpersonal relationships. The inability to emotionally self-regulate often leads to psychopathology due to the avoidance of, or preoccupation with, negative emotion [18] and, as a result, may have negative consequences on the individual's health, relationships, and work and school performance [19]. EI acts as a protective factor for mental health and is positively associated with adaptive coping styles, peer relationships, and socioemotional competence [20]. EI also mediates the relationship between mindfulness and depression and anxiety in adolescents [21].

Mindfulness is a state of mind, a moment-to-moment awareness of one's experience without judgment [22]. Being fully focused on the present moment allows for the absence of rumination and negative thoughts that lead to stress and anxiety. Mindfulness positively correlates with the capacity to be more emotionally aware, including the ability to identify and change emotional states [23]. Several mindfulness techniques, such as meditation, have proven efficacy and have shown a reduction in depression, anxiety, and rumination [24]. Mindfulness has also been linked to emotional stability, such as calmness, clarity, and concentration, as well as to emotion regulation [25]. Bridging mindfulness and EI is positive psychology, the science of happiness and flourishing [26]. It is a strength-based therapeutic approach to mental health that highlights well-being, resiliency, and compassion, which are also key factors in Ajivar.

Previous studies utilizing different mobile apps to deliver mindfulness and EI have proven effective in a myriad of populations. Improvements in psychosocial outcomes, an increase in the effectiveness of mindfulness treatment, and an increase in positive psychological interventions were all found in adult populations following the use of mindfulness apps [27-29].

Study Objectives

The main objective of this research study was to determine the effectiveness of delivering an EI curriculum and mindfulness techniques using an AI conversation platform, Ajivar. Several

validated and self-reported measurements of mental wellness and EI were tracked, including the TestWell Wellness Inventory, the 7-item Generalized Anxiety Disorder (GAD-7) scale, the 9-item Patient Health Questionnaire (PHQ-9), and the emotional quotient (EQ).

The study was launched at the end of January 2020; then in March, the COVID-19 pandemic began to impact students, staff, and faculty. Another objective that presented itself because of the pandemic was to evaluate how EI and mindfulness techniques taught during Ajivar interactions impacted student mental wellness and EI. The timing of the research study offered us an opportunity to evaluate the app during a pandemic in this population.

Methods

Study Design

Subjects included students attending a midsized liberal arts institution in Florida, United States. The students in the sample were attending the university as second-semester freshmen. This study consisted of comparison (49/99, 49%) and intervention (50/99, 51%) groups. All participants were college-age freshmen, who were 18 to 29 years of age, and were given routine mental wellness instructions as part of the classes.

Students in the intervention group were required to complete Ajivar activities as part of the course requirement. The requirement included a minimum of two Ajivar interactions per week. The comparison group members did not utilize Ajivar during the semester. All subjects completed the TestWell Wellness Inventory at the beginning and end of the 14-week semester. Completion of the inventory was a course requirement in the intervention and comparison groups. During interactions with Ajivar, the students completed assessment questionnaires as outlined below.

Assessment Questionnaires

The TestWell Wellness Inventory

The TestWell Wellness Inventory [30,31] for college students consists of 100 questions on a 5-point scale. The questions included assessment of physical fitness, self-care and safety, social awareness, emotional management, occupational wellness, nutrition, environmental wellness, emotional wellness, emotional awareness, intellectual wellness, spirituality, and values.

The 9-Item Patient Health Questionnaire

The PHQ-9 is a valid, reliable, 9-item self-administered questionnaire that evaluates the past 2 weeks' depressive symptoms using a scale from 0 to 3 per item [32]. The total score is divided into four outcome ranges of depression severity: 0-4 (no symptoms), 5-9 (mild depression), 10-14 (moderate depression), 15-20 (moderate-severe depression), and 21-27 (severe depression) [32].

The 7-Item Generalized Anxiety Disorder Scale

The GAD-7 scale is a valid, reliable, 7-item self-administered questionnaire that is used as a screening tool and severity measure for generalized anxiety disorder [33]. The scores range from 0 to 21 and are categorized according to symptom severity: 0-4 (minimal anxiety), 5-9 (mild anxiety), 10-14 (moderate anxiety), and 15-21 (severe anxiety) [33].

Emotional Quotient

There are several measures of EI [17]. As part of the Ajivar EI curriculum that is delivered via the app, there is a proprietary self-reported EQ assessment. This EQ assessment evaluates the user across several categories, which include self-awareness, self-regulation, self-efficacy, empathy, and social skills. The score is normalized to 100, and scores are characterized as follows: ≤39 (low EI), 40-49 (low-average EI), 50-69 (average EI), and ≥70 (high EI).

Ajivar: An EI- and AI-Powered Life Coach

Using AI and ML, Ajivar delivers personalized EI training and mindfulness techniques through brief conversations, videos, and activities founded in self-help practices from positive psychology and mindfulness. The Ajivar platform interacts with the user in a text-based conversational format, similar to interactions users can have with Alexa, Google Assistant, Cortana, etc. During these interactions, the information gathered by Ajivar helps it learn about the individual and the community the individual belongs to. The Ajivar app provides fun and uplifting ways to engage the user, such as positive affirmations (ie, "Posimations"), journaling, and out-of-zone (ie, "Ooz") challenges that help people get outside of their comfort zone to increase self-esteem and acceptance.

Ajivar delivers a personalized EI curriculum and mindfulness techniques to students that are approved by the Ajivar clinical consultation team. The app acts as a life coach that helps the user improve self-esteem and social and emotional awareness. Ajivar responds in a text-based conversation to emotions and underlying beliefs that the user expresses during the interactions with the app. Using emotionally intelligent conversations, the app supports and guides the user with empathetic responses, individualized content, and feedback and encourages the individual to apply their knowledge outside of the app. Apart from the psychoeducational content of positive psychology and mindfulness, the app also utilizes the following components to improve EI and resiliency:

1. Personalization (ie, automated tailoring). Ajivar utilizes ML that is based on all users' interactions, in addition to the specific user's interactions, to deliver personally relevant information and techniques based on what would benefit the user in the specific moment.
2. Resiliency and EI training. The app evaluates the user's current level of EI (ie, EQ score) based on built-in assessment tools and the user's responses during interactions. Ajivar utilizes an EI curriculum developed by clinicians to build coping skills and resiliency that are used to form healthy habits in the real world.
3. Reflection. The journaling component of the app is used for gratitude practice and reflecting on insights and feelings. The sentiment analysis function provides the user with feedback on their past and current emotional state. This insight is imperative to reinforce learning and growth.

4. Positive affirmations. These are personalized through the app interactions and are stored in a list for the user to review on a regular basis to combat negative self-talk.
5. Engagement. Ajivar uses real-time engagement, such as notifications, challenges, app mentors (ie, avatars that are collected), access to videos, and other gamification tools to keep the user engaged. The AI also gives feedback on mood and emotions, which provides the users with real-time feedback on the dashboard.

The app can be downloaded from the Apple App Store and Google Play. All interaction data are encrypted and stored on a separate server from any identifiable personal information. Only anonymized data were used for research analysis. Ajivar is not a medical device and is designed to be a support tool used for mental well-being. Anonymized data regarding the user interactions were provided to the researchers. This included the duration of user interactions with the different components of the Ajivar app and the number of interactions that took place during the study time interval.

Quantitative Measurements

The data were analyzed using the statistical software SPSS, version 26 (IBM Corp) [34].

User Engagement and Attrition

There were a total of 93 students who completed the TestWell pretest and posttest: 49 participants in the comparison group (53%) and 44 participants in the intervention group (47%).

In the intervention group, 49 out of 50 participants (98%) downloaded and interacted with the Ajivar app. If a participant carried out a greater level of activity than the course requirements, they were categorized as *high engagers* (30/49, 61%). Participants that carried out the required level of activity or less than what was required by the course were identified as *low engagers* (19/49, 39%). The mean total time on the app for the high engagers was 2069 (SD 1076) minutes versus 480 (SD 321) minutes among the low engagers. Mean total engagement for the group was 1424 (SD 1168) minutes.

Analysis of Assessment Questionnaires

An analysis of covariance (ANCOVA) was used to evaluate the TestWell pre- and posttest scores for the intervention group, the comparison group, and the individual subgroups. Pre- and posttest scores for the GAD-7, the PHQ-9, and the EQ were evaluated for differences over time for the high engagers. A GAD-7 time series of pretest, midpoint (ie, March 2020), and posttest scores was also evaluated.

Informed Consent and Institutional Review Board Approval

The study was reviewed and approved by the Institutional Review Board at the University of Tampa. Participation was voluntary, and the participants indicated their consent to the study protocol via the informed consent form. All study data were collected by the academic institution with the exception of the Ajivar app usage data. All usage data were anonymized before being given to the research team. Usage data were not linked to specific research participants.

Results

Demographic Data

This study consisted of comparison (49/99, 49%) and intervention (50/99, 51%) groups. The mean age of the study participants who reported was 19.9 (SD 1.94) years; 27% (27/99) were male and 69% (68/99) were female. Over half the students reported their ethnicity as Caucasian (59/99, 60%) and the rest reported their ethnicity as Hispanic (13/99, 13%), African American (7/99, 7%), Asian Pacific Islander (7/99, 7%), and Other (13/99, 13%). No significant demographic differences existed between the comparison and intervention groups.

Quantitative Analysis

Analysis of Assessment Questionnaires

To answer the first research question regarding the effectiveness of the TestWell Wellness Inventory, ANCOVA was performed to assess whether statistically significant differences occurred between the intervention and comparison groups based on their pre- and posttest scores on the Wellness Inventory. The results were not statistically significant (P=.41). However, both the intervention group (P=.04) and the comparison group (P=.002) independently showed statistically significant improvements from pretest to posttest on the Wellness Inventory.

The Wellness Inventory was assessed further using ANCOVA for statistically significant differences between the intervention group and the comparison group pretest and posttest on subgroups. In the intervention group, the inventory values for the social awareness subgroup significantly improved from pretest to posttest, as did the values for the spirituality subgroup. The results are shown in Table 1.

Table 1. Analysis of covariance of TestWell Wellness Inventory scores for each subgroup.

Subgroup	TestWell Wellness Inventory scores, mean (SD)		*P* value
	Pretest	Posttest	
Social awareness	82.32 (12.56)	88.91 (10.89)	.04
Spirituality	83.32 (13.86)	88.23 (12.74)	.04

Mental Wellness Indicators

In the intervention group's high-engagers subset (30/49, 61%), we assessed mean differences between the pre- and posttest scores for GAD-7, PHQ-9, and EQ. The results were found to be statistically significant as noted in Table 2. In particular, mean GAD-7 scores decreased from 11.47 (SD 1.85) to 6.27 (SD 1.44) (P<.001), indicating a statistically significant decline in anxiety from the beginning to the end of the study. Mean PHQ-9 scores decreased significantly from 10.69 (SD 2.04) to

6.69 (SD 2.41) (P<.001), indicating improved mental health in this aggregate data. EQ scores increased from pre- to poststudy, with initial mean aggregate scores of 62.87 (SD 10.12) that increased to 71.17 (SD 8.46) (P<.001).

Table 2. Mean differences between the 7-item Generalized Anxiety Disorder (GAD-7) scale, the 9-item Patient Health Questionnaire (PHQ-9), and the emotional quotient (EQ).

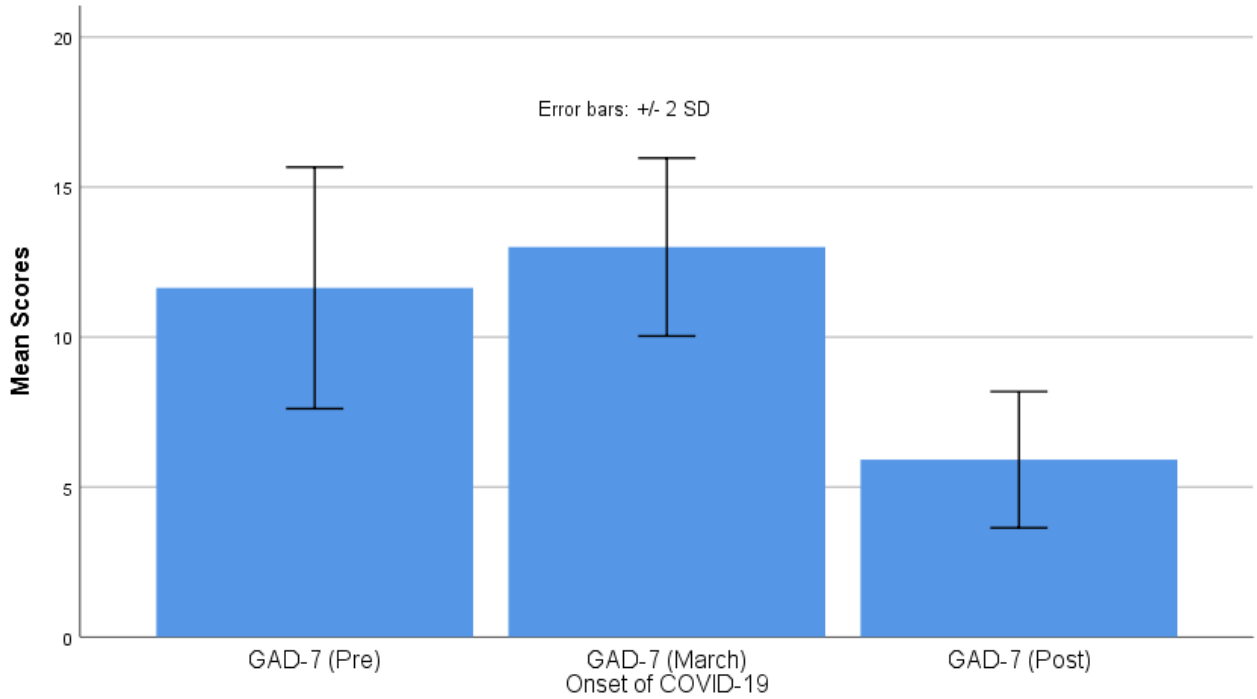
Measure	Mean (SD)	Mean difference	*P* value
GAD-7 pretest	11.47 (1.85)	5.2	<.001
GAD-7 posttest	6.27 (1.44)	N/A[a]	N/A
PHQ-9 pretest	10.69 (2.04)	4.0	<.001
PHQ-9 posttest	6.69 (2.41)	N/A	N/A
EQ pretest	62.87 (10.12)	–8.3	<.001
EQ posttest	71.17 (8.46)	N/A	N/A

[a]N/A: not applicable.

COVID-19 Impact

We collected midpoint data in the spring semester on GAD-7 scores. We conducted repeated-measures analyses of variance to assess within-subjects differences in this intervention subgroup (11/49, 22%). The GAD-7 results indicated a mean pretest (time 1) score of 11.64 (SD 2.16), a mid-March (time 2) mean score of 13.03 (SD 1.48), and an end-of-spring (time 3) mean score of 5.9 (SD 1.44) (see Figure 1). The Greenhouse-Geisser correction indicated statistically significant (P=.001) within-subjects differences on the continuous variable. Bonferroni pairwise comparison corrects for multiple comparisons; the results indicated statistical significance between times 1 and 3 and times 2 and 3, but not between times 1 and 2.

Figure 1. Repeated-measures analysis of variance for within-subjects differences in 7-item Generalized Anxiety Disorder (GAD-7) scale scores.



Discussion

Principal Findings

The findings from our study suggest that the use of Ajivar had many positive benefits for the college student participants. These results were especially impactful given that the students experienced the pandemic during the course of the study. We found that students in the intervention group had less depression over time as measured by the PHQ-9, less anxiety as measured by the GAD-7, and improvement in EI as measured by the EQ. Student anxiety increased around mid-March and then decreased significantly by the end of the semester, despite the upheaval caused by the pandemic. Mid-March was the time during the semester when learning was shifted for all students at the university to a remote format. This transition occurred over the course of a few days. Further testing is suggested as the pandemic continues. The subgroups of social awareness and

spirituality were significantly improved with the use of the app. There were also areas of no significance, including the pretest and posttest scores of the TestWell Wellness Inventory. We also noted a high engagement rate, with over 60% of users interacting with the Ajivar app for a greater amount of time than was required.

Comparison With Prior Work

Previous studies utilizing different mobile apps to deliver mindfulness and EI techniques have proven effective in a myriad of populations. Improvements in psychosocial outcomes were noted with mindfulness training and positive psychological interventions in adult populations following the use of mindfulness apps [27-29]. Champion et al showed a significant improvement of self-reported stress, resilience, and satisfaction with life among adults after 10 days of using a mindfulness-based smartphone app [27]. The magnitude of benefits was found to increase further following 30 days of use, with the rate of benefit greatest between baseline and day 10. The study did include a small sample size, an inactive control group, and inclusion of self-selected participants. A study by Ly et al showed that behavioral activation and mindfulness treatment helped adult participants who suffered from major depression [28]. Behavioral activation was found to be more effective for participants with a higher severity of depression, and treatment with mindfulness was more effective for patients who initially had a lower severity of depression. The study concluded that effective treatment of mild to moderate major depression through the use of smartphone app was feasible.

A study of college students with elevated stress levels used a mobile app daily; the majority of the intervention subjects indicated that the app helped with their stress and that they would use the app in the future [35]. The study concluded that delivering mindfulness meditation via an app is a more practical approach to reduce stress. This approach requires fewer resources, involves fewer time constraints, and allows students to participate remotely. A study by Bostock et al on healthy adult employees found long-term improvement in well-being, distress, job strain, and perceptions of workplace social support using mindfulness meditation delivered by a phone app [36].

Studies have shown the impact of EI and emotional regulation on individuals' experiences of life and how changing the way a situation is perceived decreases its emotional impact [18]. The decreases in emotional experience and behavioral expression, which have no impact on memory, lower stress and improve relationships.

This study confirms the findings of previous studies that showed that mindfulness and EI training can be delivered via an app rather than requiring in-person instruction. To our knowledge, this is the first study to demonstrate that a text-based conversational app, Ajivar, can deliver personalized mindfulness and EI training with resulting improvements in anxiety, depression, and EI. In contrast to the previous work, EI and mindfulness training was delivered using AI and ML algorithms to personalize the user journey and experience. This study also demonstrated high user engagement with an app when learning techniques for mental wellness.

COVID-19 Effect

Previous studies showed that college students experienced an uptick of anxiety and depression related to COVID-19, which were greater than the levels experienced previously during college [15,37,38]. In the Huckins [15] study, participants had almost a 50% increase in anxiety, as measured by the 2-item Generalized Anxiety Disorder (GAD-2) scale. The GAD-2 is a brief version of the GAD-7 questionnaire used in this study. When comparing GAD-7 scores in March with those at the start of the study, we showed an 11% increase in anxiety levels, though they did not reach statistical significance. With the use of Ajivar, these levels decreased to 55% (P=.001) below the GAD-7 scores at the start of the study among these individuals.

Strengths and Limitations

Strengths of this study were the ability to use validated instruments to measure anxiety, depression, and EI in this sample. The study found positive results in a very limited amount of time. Limitations are evident in the study, including a small sample size. Some measurements were not available, as they were based on the usage of the app. For instance, students had to reach a certain level of usage to advance to the next area that may have provided additional data. There were no data available to control for influential variables external to Ajivar use. Lastly, the data are not generalizable.

Future Studies

This was a small initial study, and increased participant numbers for future studies are anticipated. Future plans include a year-long study during the next academic year. The target population will be broader in scope and include students across the university, not restricted to specific classes.

Conclusions

EI has proven key to long-term success. There is unlimited potential for using apps for mental health, including for issues related to the current pandemic. Also, similar studies should be replicated and completed to determine additional results of using mindfulness technology for mental health in a variety of populations, including college-age students. Technology apps can be used for mental health conditions along with other behavioral issues. Bakker et al found specific recommendations for mental health apps [39], including the aim to prevent emotional mental health problems. Also, to maximize engagement, it is recommended to use gamification, as was carried out in Ajivar, so that habit formation of self-care practices can take place [39]. Mental health apps have proven to provide positive results and success in a variety of populations. Additional testing is recommended as the pandemic continues. However, the results of this study demonstrate positive outcomes for mental and emotional health, along with a decrease in anxiety and depression levels. We are going to need scalable resources like Ajivar if we are going to meet the demands that we are likely to experience in the near future.

Acknowledgments

The authors would like to thank Ajivar for providing access to the anonymized data from their mobile app during the study period for research purposes. Ajivar partially funded the publication fees for the paper.

Conflicts of Interest

TS and RG are cofounders of Ajivar and, therefore, have a financial interest in that company.

References

1. Romanelli F, Cain J, Smith K. Emotional intelligence as a predictor of academic and/or professional success. Am J Pharm Educ 2006 Jun 15;70(3):69 [FREE Full text] [doi: 10.5688/aj700369] [Medline: 17136189]
2. Kanoy K. Emotional Intelligence Learning: What the Research Teaches Us About its Importance to Students. San Francisco, CA: Jossey-Bass URL: https://bit.ly/2KRAwGr [accessed 2020-12-17]
3. Claros E, Sharma M. The relationship between emotional intelligence and abuse of alcohol, marijuana, and tobacco among college students. J Alcohol Drug Educ 2012 Jun;56(1):8-37 [FREE Full text]
4. Williams JMG, Kabat-Zinn J, editors. Mindfulness: Diverse Perspectives on its Meaning, Origins and Applications. Abingdon, UK: Routledge; 2013.
5. Zivin K, Eisenberg D, Gollust SE, Golberstein E. Persistence of mental health problems and needs in a college student population. J Affect Disord 2009 Oct;117(3):180-185. [doi: 10.1016/j.jad.2009.01.001] [Medline: 19178949]
6. American College Health Association. American College Health Association-National College Health Assessment II: Reference Group Executive Summary Spring 2018. Silver Spring, MD: American College Health Association; 2018. URL: https://www.acha.org/documents/ncha/NCHA-II_Spring_2018_Reference_Group_Executive_Summary.pdf [accessed 2020-12-18]
7. Pierceall EA, Keim MC. Stress and coping strategies among community college students. Community Coll J Res Pract 2007 Oct 08;31(9):703-712. [doi: 10.1080/10668920600866579]
8. Almojali AI, Almalki SA, Alothman AS, Masuadi EM, Alaqeel MK. The prevalence and association of stress with sleep quality among medical students. J Epidemiol Glob Health 2017 Sep;7(3):169-174 [FREE Full text] [doi: 10.1016/j.jegh.2017.04.005] [Medline: 28756825]
9. Neely ME, Schallert DL, Mohammed SS, Roberts RM, Chen Y. Self-kindness when facing stress: The role of self-compassion, goal regulation, and support in college students' well-being. Motiv Emot 2009 Jan 6;33(1):88-97. [doi: 10.1007/s11031-008-9119-8]
10. Anastasiades MH, Kapoor S, Wootten J, Lamis DA. Perceived stress, depressive symptoms, and suicidal ideation in undergraduate women with varying levels of mindfulness. Arch Womens Ment Health 2017 Feb;20(1):129-138. [doi: 10.1007/s00737-016-0686-5] [Medline: 27778150]
11. Hunt J, Eisenberg D. Mental health problems and help-seeking behavior among college students. J Adolesc Health 2010 Jan;46(1):3-10. [doi: 10.1016/j.jadohealth.2009.08.008] [Medline: 20123251]
12. 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]
13. The Healthy Minds Network, American College Health Association. The Impact of COVID-19 on College Student Well-Being. Ann Arbor, MI: The Healthy Minds Network; 2020. URL: https://healthymindsnetwork.org/wp-content/uploads/2020/07/Healthy_Minds_NCHA_COVID_Survey_Report_FINAL.pdf [accessed 2020-12-18]
14. Fried EI, Papanikolaou F, Epskamp S. Mental health and social contact during the COVID-19 pandemic: An ecological momentary assessment study. PsyArXiv. Preprint posted online on April 24, 2020. [doi: 10.31234/osf.io/36xkp]
15. Huckins JF, daSilva AW, Wang W, Hedlund E, Rogers C, Nepal SK, et al. Mental health and behavior of college students during the early phases of the COVID-19 pandemic: Longitudinal smartphone and ecological momentary assessment study. J Med Internet Res 2020 Jun 17;22(6):e20185 [FREE Full text] [doi: 10.2196/20185] [Medline: 32519963]
16. Salovey P, Mayer JD. Emotional intelligence. Imagin Cogn Pers 2016 Nov 28;9(3):185-211. [doi: 10.2190/DUGG-P24E-52WK-6CDG]
17. Goleman D, Boyatzis RE. Emotional intelligence has 12 elements. Which do you need to work on? Harvard Business Review. 2017 Feb 06. URL: https://hbr.org/2017/02/emotional-intelligence-has-12-elements-which-do-you-need-to-work-on [accessed 2020-12-18]
18. Gross JJ. Emotion regulation: Affective, cognitive, and social consequences. Psychophysiology 2002 May;39(3):281-291. [doi: 10.1017/s0048577201393198] [Medline: 12212647]
19. John JD, Salovey P, Caruso DR. Emotional intelligence: Theory, findings, and implications. Psychol Inq 2004;15(3):197-215 [FREE Full text] [doi: 10.1207/s15327965pli1503_02]
20. Frederickson N, Petrides K, Simmonds E. Trait emotional intelligence as a predictor of socioemotional outcomes in early adolescence. Pers Individ Dif 2012 Feb;52(3):323-328. [doi: 10.1016/j.paid.2011.10.034]

21. Foster B, Lomas J, Downey L, Stough C. Does emotional intelligence mediate the relation between mindfulness and anxiety and depression in adolescents? Front Psychol 2018;9:2463 [FREE Full text] [doi: 10.3389/fpsyg.2018.02463] [Medline: 30618918]
22. Davis DM, Hayes JA. What are the benefits of mindfulness? Monitor on Psychology. Washington, DC: American Psychological Association; 2012. URL: https://www.apa.org/monitor/2012/07-08/ce-corner [accessed 2020-12-18]
23. Brown KW, Ryan RM. The benefits of being present: Mindfulness and its role in psychological well-being. J Pers Soc Psychol 2003 Apr;84(4):822-848. [doi: 10.1037/0022-3514.84.4.822] [Medline: 12703651]
24. Kabat‐Zinn J. Mindfulness‐based interventions in context: Past, present, and future. Clin Psychol Sci Pract 2003 Jun;10(2):144-156 [FREE Full text] [doi: 10.1093/clipsy.bpg016]
25. Walsh R, Shapiro SL. The meeting of meditative disciplines and Western psychology: A mutually enriching dialogue. Am Psychol 2006 Apr;61(3):227-239. [doi: 10.1037/0003-066X.61.3.227] [Medline: 16594839]
26. Compton WC, Hoffman E. Positive Psychology: The Science of Happiness and Flourishing. 2nd edition. Belmont, CA: Wadsworth Publishing; 2012.
27. Champion L, Economides M, Chandler C. The efficacy of a brief app-based mindfulness intervention on psychosocial outcomes in healthy adults: A pilot randomised controlled trial. PLoS One 2018;13(12):e0209482 [FREE Full text] [doi: 10.1371/journal.pone.0209482] [Medline: 30596696]
28. Ly KH, Trüschel A, Jarl L, Magnusson S, Windahl T, Johansson R, et al. Behavioural activation versus mindfulness-based guided self-help treatment administered through a smartphone application: A randomised controlled trial. BMJ Open 2014 Jan 09;4(1):e003440 [FREE Full text] [doi: 10.1136/bmjopen-2013-003440] [Medline: 24413342]
29. Howells A, Ivtzan I, Eiroa-Orosa FJ. Putting the 'app' in happiness: A randomised controlled trial of a smartphone-based mindfulness intervention to enhance well-being. J Happiness Stud 2014 Oct 29;17(1):163-185. [doi: 10.1007/s10902-014-9589-1]
30. Using TestWell: Overview. TestWell. URL: https://www.testwell.org/index.php?id=1717&id_tier=3530 [accessed 2020-12-18]
31. Owen TR. The reliability and validity of a wellness inventory. Am J Health Promot 1999;13(3):180-182. [doi: 10.4278/0890-1171-13.3.180] [Medline: 10351545]
32. 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]
33. Löwe B, Decker O, Müller S, Brähler E, Schellberg D, Herzog W, et al. Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. Med Care 2008 Mar;46(3):266-274. [doi: 10.1097/MLR.0b013e318160d093] [Medline: 18388841]
34. IBM SPSS Statistics 26 documentation. IBM Knowledge Center. URL: https://www.ibm.com/support/knowledgecenter/SSLVMB_26.0.0/statistics_kc_ddita/spss/product_landing.html [accessed 2020-12-18]
35. Huberty J, Green J, Glissmann C, Larkey L, Puzia M, Lee C. Efficacy of the mindfulness meditation mobile app "Calm" to reduce stress among college students: Randomized controlled trial. JMIR Mhealth Uhealth 2019 Jun 25;7(6):e14273 [FREE Full text] [doi: 10.2196/14273] [Medline: 31237569]
36. Bostock S, Crosswell AD, Prather AA, Steptoe A. Mindfulness on-the-go: Effects of a mindfulness meditation app on work stress and well-being. J Occup Health Psychol 2019 Feb;24(1):127-138 [FREE Full text] [doi: 10.1037/ocp0000118] [Medline: 29723001]
37. Liu X, Liu J, Zhong X. Psychological state of college students during COVID-19 epidemic. Preprints with The Lancet. Preprint posted online on March 10, 2020. [doi: 10.2139/ssrn.3552814]
38. Cao W, Fang Z, Hou G, Han M, Xu X, Dong J, et al. The psychological impact of the COVID-19 epidemic on college students in China. Psychiatry Res 2020 May;287:112934 [FREE Full text] [doi: 10.1016/j.psychres.2020.112934] [Medline: 32229390]
39. 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]

Abbreviations

AI: artificial intelligence
ANCOVA: analysis of covariance
EI: emotional intelligence
EQ: emotional quotient
GAD-2: 2-item Generalized Anxiety Disorder
GAD-7: 7-item Generalized Anxiety Disorder
ML: machine learning
PHQ-9: 9-item Patient Health Questionnaire

Edited by G Eysenbach; submitted 30.10.20; peer-reviewed by C Parsons, M Majurul Ahsan; comments to author 09.11.20; revised version received 13.11.20; accepted 08.12.20; published 05.01.21.

Please cite as:
Sturgill R, Martinasek M, Schmidt T, Goyal R
A Novel Artificial Intelligence-Powered Emotional Intelligence and Mindfulness App (Ajivar) for the College Student Population During the COVID-19 Pandemic: Quantitative Questionnaire Study
JMIR Form Res 2021;5(1):e25372
URL: http://formative.jmir.org/2021/1/e25372/
doi:10.2196/25372
PMID:33320822

Original Paper

A Relaxation App (HeartBot) for Stress and Emotional Well-Being Over a 21-Day Challenge: Randomized Survey Study

Laya Iyer[1*]; Ranjani B Iyer[2*], PhD; Vetriliaa Kumar[3*]

[1]Code for Nonprofit Inc, Novi, MI, United States

[2]Heartfulness Program for Schools, Heartfulness Institute, Novi, MI, United States

[3]Code for Nonprofit Inc, Austin, TX, United States

[*]all authors contributed equally

Corresponding Author:
Ranjani B Iyer, PhD
Heartfulness Program for Schools
Heartfulness Institute
24977 Glenda St
Novi, MI, 48375
United States
Phone: 1 2482559635
Email: ranjani.heartfulness@gmail.com

Abstract

Background: HeartBot is an app designed to enable people 14 years and older to use relaxation tools offered by Heartfulness Institute to deal with daily stress and anxiety in a healthy, productive manner. These tools have proven effective in stress management and mental wellness when administered in a controlled environment by a certified proctor.

Objective: This study aimed to explore the app's effectiveness and evaluate the implementation of the tools.

Methods: In this study, 88 participants were recruited and randomly sorted into 2 groups, the HeartBot intervention group (n=46) and the waitlist control group (n=42). Pre- and postsurveys measured participants' stress levels using the Perceived Stress Scale (PSS) and their social-emotional well-being using the EPOCH (Engagement, Perseverance, Optimism, Connectedness, and Happiness) Measure of Adolescent Well-Being before and after they used the app for 21 days for 30 minutes every day.

Results: The study received institutional review board approval on August 18, 2019. Participant recruitment lasted from the approval date until September 30, 2019. The 21-day challenge started on October 1, 2019. Of the 135 people who signed up, 88 completed the study. There was a statistically significant difference in the mean PSS scores before and after the intervention (from 18.3 to 7.89; $P<.001$). The paired Wilcoxon rank sum test on the EPOCH scores indicated a significant difference in the medians of the total scores ($W=411.5$, $P<.001$).

Conclusions: Evidence from this study shows that HeartBot is an effective app that can be used to manage stress and improve positive characteristics of emotional wellness. Future research and widespread usage of the app under this study are encouraged based on this preliminary evidence of its effectiveness.

Trial Registration: ClinicalTrials.gov NCT04589520; https://clinicaltrials.gov/ct2/show/NCT04589520

(JMIR Form Res 2021;5(1):e22041) doi:10.2196/22041

KEYWORDS

Heartfulness; stress management; iOS app; social-emotional; mental health

Introduction

Stress is a major health problem around the world and is one of the main causes of early death and disease in the United States [1]. According to statistics from the American Institute of Stress [2], adults experience stress that is mainly caused by job pressure, money, health, relationships, poor nutrition, media overload, and sleep deprivation. About 77% of adults regularly experience physical symptoms caused by stress, 73% regularly experience psychological symptoms caused by stress, and 33% report living with extreme stress. Research shows that nearly half (49%) of all students report feeling a great deal of stress on a daily basis, and 31% report feeling somewhat stressed [3]. Childhood and adolescence are crucial formative developmental

stages that lay the groundwork for an individual's capacity to maintain their emotional well-being and mental health in adulthood [4,5]. These statistics show the urgent need for accessible stress management techniques that are effective for individuals of all ages.

Technology, especially in a portable form such as an app, is unique in introducing practical, accessible mental wellness tools to members of the general population. A previous study examining the effectiveness of mindfulness apps in improving users' well-being found that mindfulness-based positive intervention can be delivered via a smartphone app successfully [6]. Additionally, a study examining the effect of mindfulness apps versus traditional intervention stress techniques found that participants that used the app showed marginally more compassion satisfaction and marginally less burnout [7]. Relative to in-person interventions, digital technologies can reach a broader audience in less time, are cost-effective, and are more personalized to the individual [8].

There is a wide variety of apps that provide users with the tools to manage their stress, including apps such as Headspace and Calm. Such apps represent a convenient and cost-effective technology that can easily be scaled up to address barriers to implementing traditional mindfulness-based stress reduction programs but may require supplemental support to promote their use [9].

Digital mediums, therefore, have incredible potential for improving public health [10]. In addition, those using the app Calm for 8 weeks reported significant differences in outcomes of stress, mindfulness, and self-compassion [11]. This study aims to explore a stress management app to mitigate stress and promote overall well-being.

Heartfulness is a heart-based practice of meditation that focuses on the relationship between the heart and mind. It promotes further discovery in the science of yoga, as it relates to the body-mind complex and plays a vital role in expanding upon mindfulness practice. It is a journey to the center of the heart, a place of inner silence. Tuning into the heart develops calmness from within and uncovers every individual's brilliant self.

HeartBot, an iOS app, aims to provide its users with a convenient means of managing stress by providing personalized audios and guided experience of Heartfulness relaxation tools. Figure 1 presents the user interface of HeartBot, and the app features 6 guided tools. Previous studies concerning Heartfulness programs have established Heartfulness tools' effectiveness in a structured environment [12]. This study expands upon previous research and investigates whether the HeartBot app is associated with a decrease in stress levels and an increase in emotional well-being in a 21-day challenge.

Figure 1. HeartBot app.

Methods

Study Design and Setting

This study used a pre- and postsurvey completed online. The intervention was completed remotely in a virtual setting.

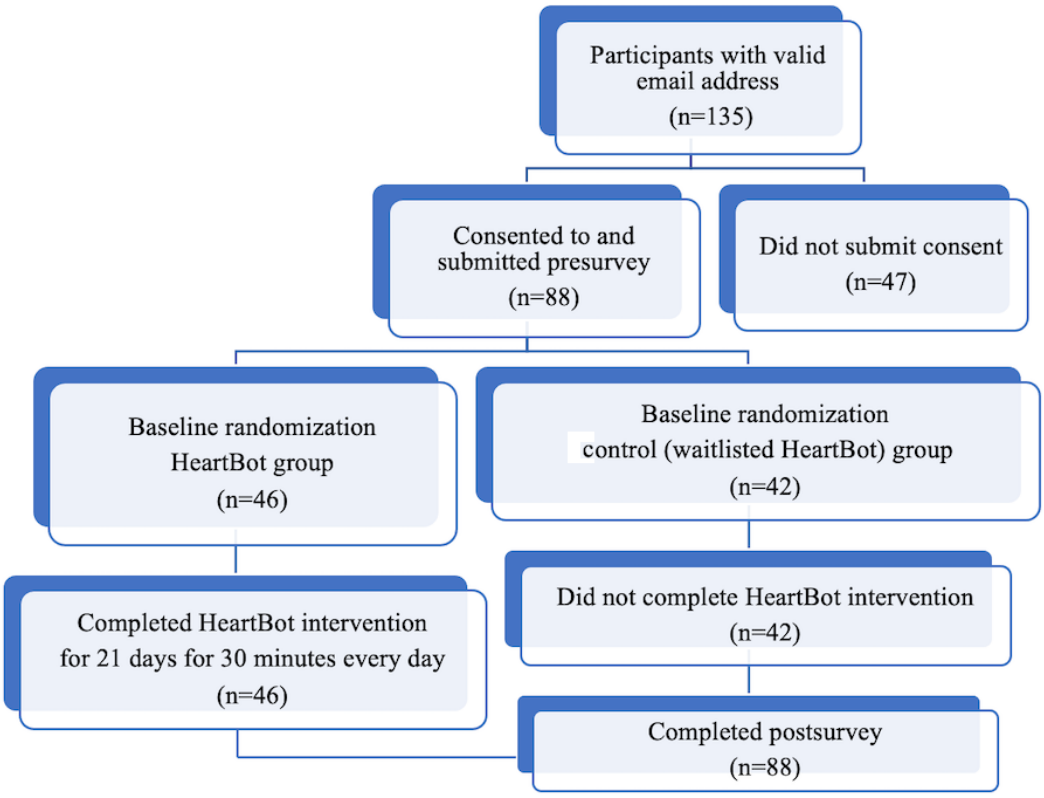
Ethics Approval

The study was approved by Solutions IRB (registration No. IORG0007116) in August 2019. All participants provided electronically signed consent and assent prior to participation in the study.

Recruitment

The participants' recruitment was through convenience and snowball sampling because of their convenient accessibility and proximity to the researchers [13]. Posters were placed in schools and venues such as libraries to raise awareness about the study. Inclusion criteria to participate in the study were (1) being 14 years or older and (2) having an iOS device that allowed the user to download and use the app. Recruitment occurred from August 2019 through September 2019. We anticipated about 100 participants and expected a dropout rate of about 50%. We got responses from 135 interested participants, 88 of whom continued with the study and 47 of whom did not submit their consent. Figure 2 shows the study flowchart.

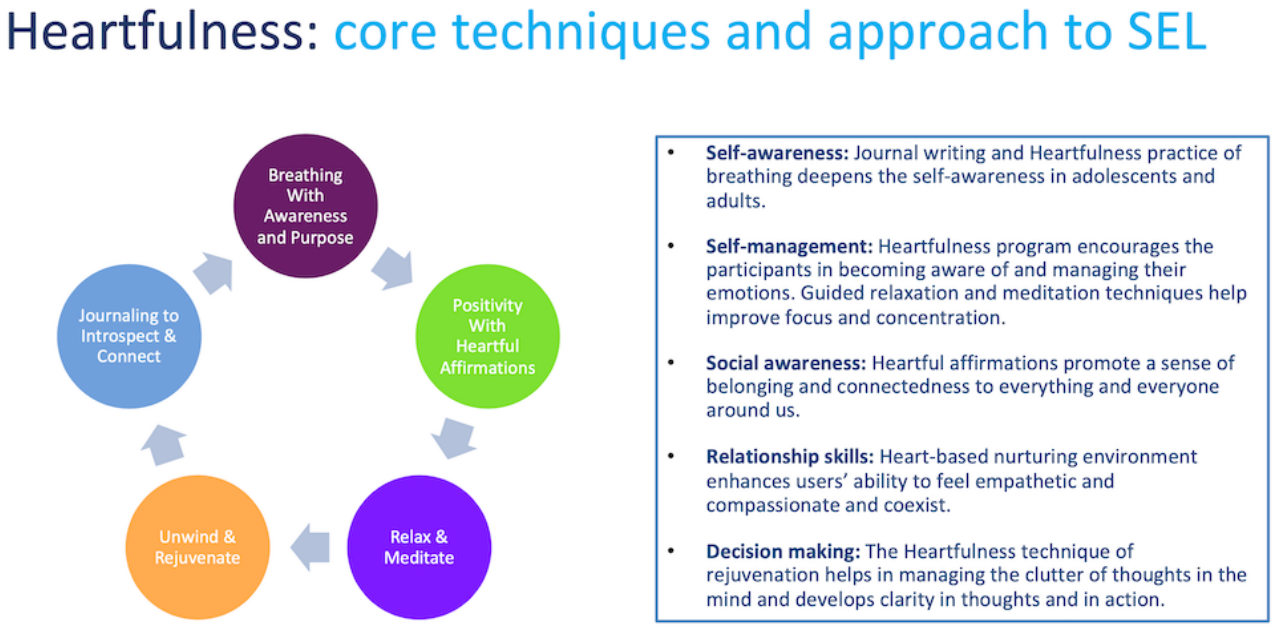
Figure 2. Study flowchart.



HeartBot App and Intervention

HeartBot is an iOS app developed by high school students for users 14 years and older to learn and practice the 5 practical Heartfulness tools: relaxation, meditation, rejuvenation, affirmations, and breathing [14]. This app expands upon previous social-emotional wellness apps by implementing these tools and journal writing to promote the development of the 5 core competencies of social and emotional learning, as charted by the Collaborative for Academic, Social, and Emotional Learning [15] (Figure 3).

Figure 3. Heartfulness approach to social-emotional learning [15]. SEL: social-emotional learning.



Procedures

A total of 88 participants of mixed ages who were 14 years and older took part in the study. Participants, who were screened to fit the inclusion criteria, were asked to provide a written consent (18 years or older) or assent form (younger than 18 years). They were randomly divided into the HeartBot and the waitlist control groups using an online random name generator to balance the two groups by age. There was no significant difference in the number of people in the two groups. Participants randomized to the waitlist group received the HeartBot intervention after the waitlist period.

The HeartBot group downloaded the app and used it daily for 21 days based on a calendar that provided them with step-by-step outlines of the tools for each day (Multimedia Appendix 1). The control group did not download the app and experienced no change to their daily routines. Presurveys were sent to all participants via email before and after their 21-day challenge. Data collected gave a baseline and final quantitative measure of participants' stress and emotional wellness. Participants were not individually monitored or targeted during their participation. Those who completed the challenge were entered in a raffle and had a chance to win one of 10 US $50 gift cards.

Sample Size

A sample size larger than 30 and less than 500 is appropriate for most research. With a 95% confidence level and a 9.5% margin of error, the sample size came to 100. We also noted that estimating a margin of error for sample sizes ranging from 10 to 10,000 reduces to 10% [16].

Measures

Participant stress levels were measured using the Perceived Stress Scale (PSS). This 10-item scale measures the degree to which a situation in someone's life is considered stressful [17]. The survey consists of a 5-point scale for each question, where 0=never and 4=very often, to indicate how often respondents felt a certain way about a certain stimulus or event. The minimum score is 0 and the maximum score is 40, with higher scores suggesting higher stress levels. This scale has been proven valid through its use in other studies, which found that higher PSS scores are associated with a greater vulnerability to stressful life event–elicited depressive symptoms [17]. Similarly, reliability between the collected PSS scores and PSS data norms was about 0.78, indicating a very reliable measure.

The emotional wellness of the participants was measured using the EPOCH (Engagement, Perseverance, Optimism, Connectedness, and Happiness) Measure of Adolescent Well-Being. This is a 20-item scale used to measure 5 different positive characteristics: engagement, perseverance, optimism, connectedness, and happiness [18]. Each question on this survey consists of a 5-point scale, where "almost never" or "not at all like me" is a 1 and "almost always" or "very much like me" is a 5. On this scale, the minimum score is 20 and the maximum score is 100, and higher scores correlate with a higher level of emotional wellness. The test has demonstrated very high validity and reliability, with a Cronbach α of .90, Guttman $\lambda 6$ of 0.91, minimum split-half reliability of 0.75, and maximum split-half reliability of 0.93 [18].

Data Analysis

The data collected online from the pre- and postsurveys were cleaned and adjusted to eliminate data errors or corruption. Then, the dependent variables were examined for normality and missing values. In examining the relationship between variables, researchers can use a *t* test or analysis of variance (ANOVA) to compare the means of two groups on the dependent variable [19]. Baseline differences from the presurvey scores between the HeartBot and the control groups on the PSS and EPOCH were measured. Descriptive statistics and the difference between the pre- and postsurvey scores from each of the outcome measures (PSS, EPOCH) were analyzed. Statistical significance was determined by $P<.05$.

Results

Overview

A total of 135 participants were recruited for the study. Of these, 88 participants (65.2%) completed the full 21-day challenge, and the data were analyzed. A total of 46 participants in the HeartBot group and 42 participants in the control waitlisted group completed the pre- and postsurveys. Out of the 46 participants in the intervention group, 20 (43%) were aged between 14 and 17 years and 26 (57%) were older than 18 years.

Baseline Equivalences

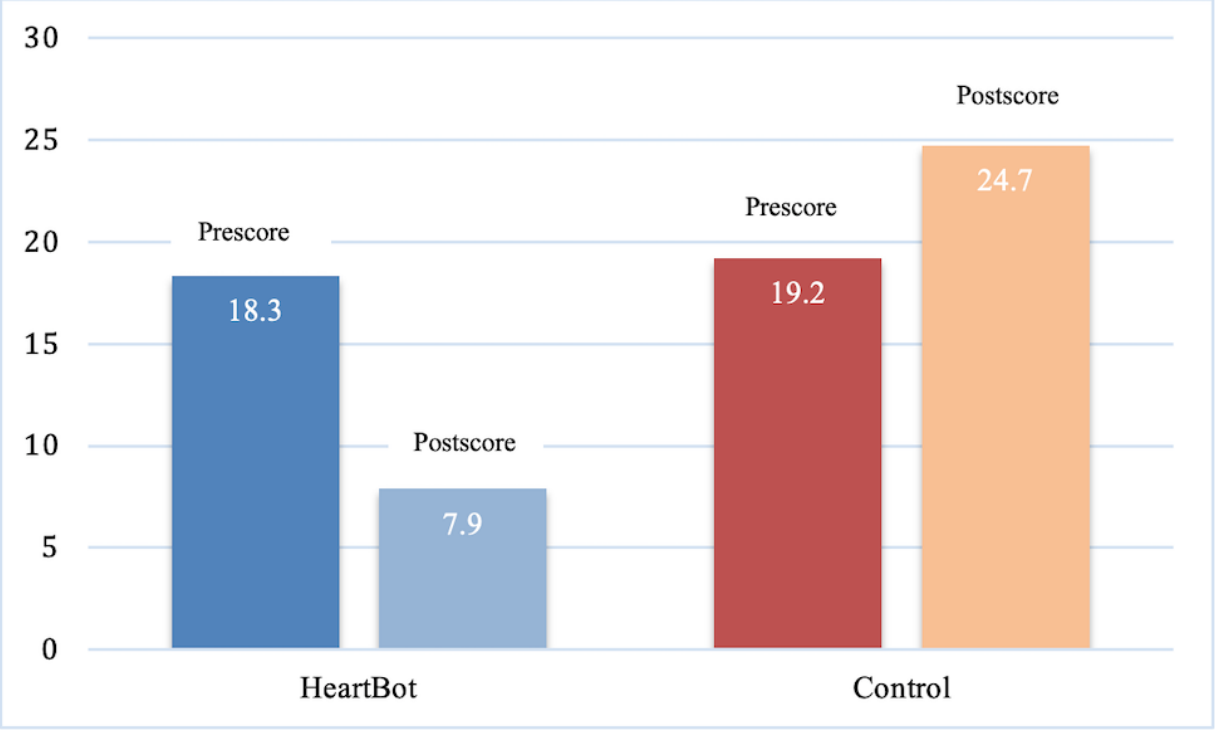
Baseline level as prescores and postscores after the 21-day challenge reflected the outcome measures for changes in the mean scores in the perceived stress levels and the 5 characteristics for well-being in the EPOCH scores. The initial analysis examined the differences in the PSS and EPOCH baseline scores in the HeartBot and control group. The baseline equivalence compared the average baseline characteristics for the HeartBot and control groups.

There was no significant difference between the baseline mean PSS scores in the HeartBot group and the control group at the beginning of the study, as shown by 2-tailed paired *t* tests. Although the participants' selection in the two groups was randomized, there were significant differences at baseline in the EPOCH scores between the HeartBot and the control groups (Multimedia Appendix 2).

Descriptive Analysis

The mean PSS pre- and postscores in the HeartBot group showed a significant decrease from 18.3 to 7.89 (P<.001). In contrast, the mean scores increased from 19.2 to 24.7 (P<.001) in the control group (Figure 4). This finding suggests a significant decrease in the perceived stress in the HeartBot group compared with the participants in the control group (Multimedia Appendix 3).

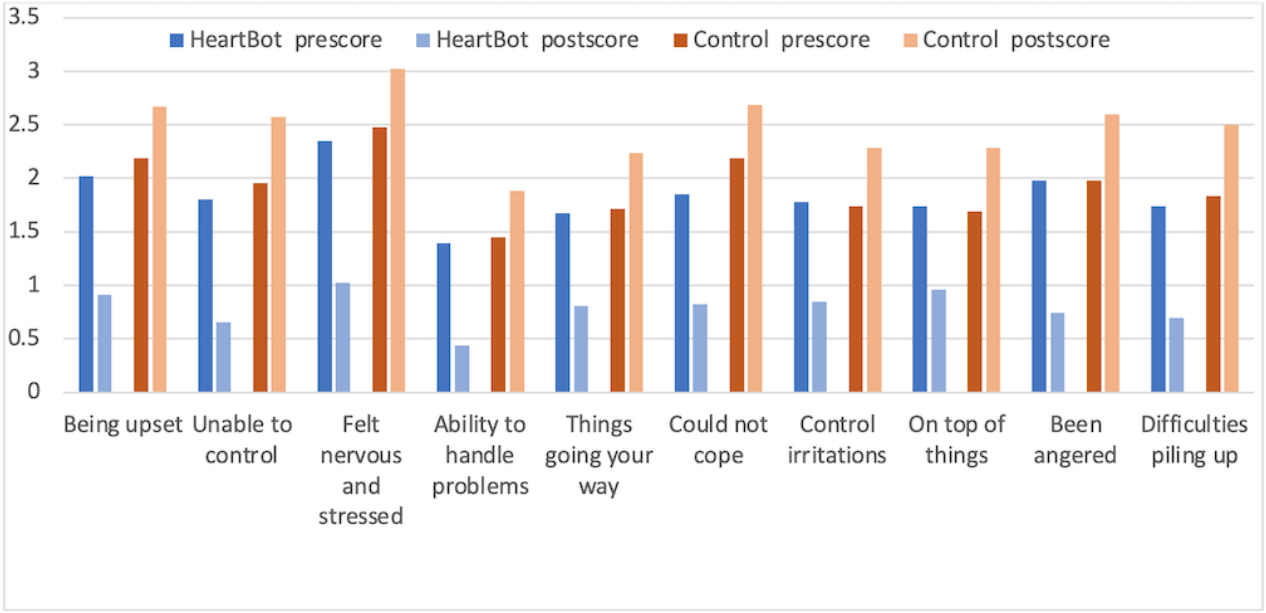
Figure 4. Change in Perceived Stress Scale mean scores on pre- and postsurveys for the HeartBot and control groups.



One-way ANOVA examined the difference between the mean PSS score in the HeartBot and control groups. The ANOVA was significant ($F_{1,82}$=125.76, P<.001). This result supports the conclusion that there is a statistically significant and robust relationship between the two groups and the perceived stress scales (Multimedia Appendix 3).

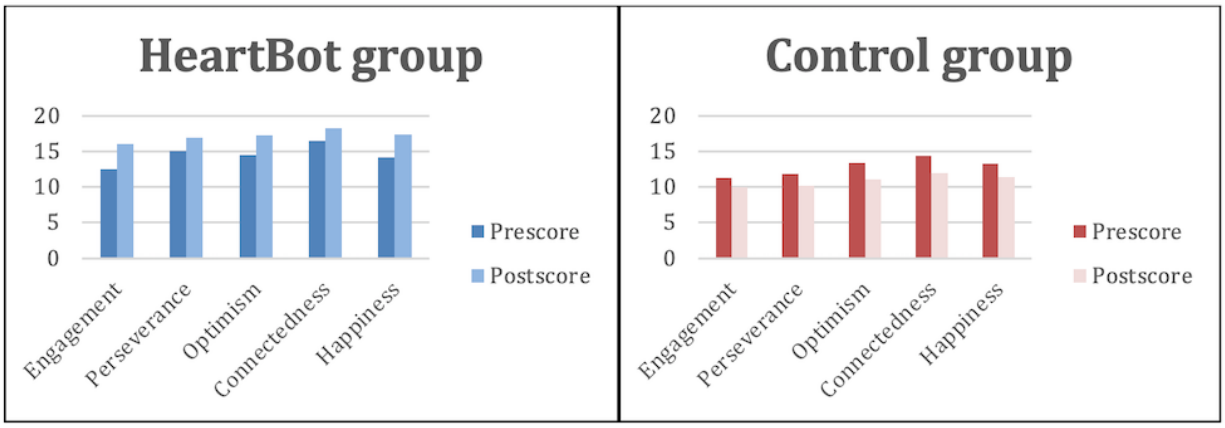
The questions (items 1-10) on the PSS, which ask participants if they were feeling nervous and stressed (question 3), could not cope (question 6), and have been angered (question 9), showed a decreased score in HeartBot participants, indicating that the app helped in reducing stress. Figure 5 shows the responses to the questions on participants' ability to handle problems (question 4) and things going their way (question 5), indicating an improvement in coping skills in the HeartBot group. As noted, for the HeartBot group, PSS scores increased for items that showed stress management and coping, showing an improvement. This finding indicates that the app worked on the participants in that group and that HeartBot helped the participants reduce perceived stress.

Figure 5. Summary of statistics on items on the Perceived Stress Scale survey.



The descriptive data and comparison of the means in the HeartBot group for the 5 characteristics of EPOCH showed an increase in all 5 positive characteristics (Figure 6). The findings indicate that the positive characteristics of perseverance, optimism, and connectedness increased significantly in the HeartBot group. There was also a decrease noted in all 5 positive characteristics in the control group.

Figure 6. Change in EPOCH total mean scores on the pre- and postsurveys for the two groups. EPOCH: Engagement, Perseverance, Optimism, Connectedness, and Happiness.



For the postintervention total EPOCH scores, the median total EPOCH score was 89.0 for the HeartBot group and 60.5 for the control group. There was a significant difference in the medians of the postintervention total EPOCH scores for the HeartBot and control groups, with Kruskal-Wallis χ^2_1=51.0 (*P*<.001) and Bonferroni correction 9.5×10^{-13}.

In comparing the baseline and postintervention total EPOCH scores for the HeartBot group, there was a significant difference in the medians of the total scores (*W*=411.5, *P*<.001). In comparing the baseline and postintervention total EPOCH scores for the control group, there was a significant difference in the medians of the total scores (*W*=1145, *P*=.01). Paired 2-tailed *t* tests on baseline pre-EPOCH and post-EPOCH scores between the HeartBot and control groups showed a significant difference (Multimedia Appendix 4).

Age Categories

An adult is any participant aged 18 years or older, and a minor is any participant younger than 18 years (14-17 years for this study).

For adults, a statistically significant difference was found on the PSS at baseline in contrast to the postdata (HeartBot group: t_{26}=7.14, *P*<.001; control group: t_{27}=–4.07, *P*<.001). When comparing the data for minors, there was a statistically significant difference found on the PSS at baseline in contrast to the postdata (HeartBot group: t_{20}=8.61, *P*<.001; control group: t_{15}=–2.8, *P*<.001). This finding suggests that there was a

significant decrease in perceived stress in the HeartBot group in comparison with the control group.

For adults, comparing the baseline and postintervention total EPOCH scores for the HeartBot group showed that there was a significant difference in the medians of the total scores (W=130.5, P<.001). When comparing the baseline and postintervention total EPOCH scores for adults in the control group, there was no significant difference in the medians of the total scores (W=434.5, P=.23). For minors, comparing the baseline and postintervention total EPOCH scores revealed a significant difference in the medians of the total scores for the HeartBot group (W=75, P<.001) and for the control group (W=186, P=.002).

Discussion

Principal Findings

This study's objectives were to assess changes in scores measuring stress levels (using PSS) and emotional wellness (using EPOCH). This study showed that there was a significant decrease in the perceived stress levels and an increase in the emotional well-being of participants who used the app for 21 days for 30 minutes every day based on a 21-day calendar. Apps have been shown to provide more efficient delivery of health care and increase treatment effectiveness [20]. This app provides users with a convenient way to practice something that would be harder for them to do otherwise. This study demonstrated that the HeartBot app reduces stress levels and improves emotional well-being for its users.

A previous study on Heartfulness in schools showed a significant decrease in stress levels and a significant increase in participants' overall well-being [12]. Another study showed that practicing Heartfulness techniques for 12 weeks demonstrated improved wellness and amelioration of burnout [21]. The results of this study are consistent with these studies.

The findings from this study contribute to the growing movement of stress management apps and their effectiveness. Small to moderate effects on global well-being and positive affect were seen over 1 working day of using a mindfulness app [22]. Combining technology and mindfulness techniques has been shown to elicit meaningful benefits by increasing the accessibility and efficacy of mindfulness training [23]. Teenagers prefer using a digital medium for help rather than a face-to-face approach [24]. Studies have shown that computerized platforms such as apps are a comparable and valid means of delivering mindfulness training compared with face-to-face interactions, showing that apps can be just as effective a method of practicing mindfulness [25]. This study adds to the evidence that the HeartBot app can help significantly decrease stress and improve emotional wellness.

This study addresses how much time a user needs to spend on the app to reap its benefits. All participants in the HeartBot group spent 30 minutes a day for 21 days and saw a significant difference in their stress levels. EPOCH showed a significant increase in the HeartBot group for perseverance (P=.009), optimism (P=.005), and connectedness (P<.001). Future research could further explore the social benefits of Heartfulness and the long-term effects of using the HeartBot app. Notably, studies in the future could explore the connection between loneliness and the practice of Heartfulness tools.

Going into this study, we expected to see a significant result for the group of participants aged 14 to 18, as teenagers are more likely to prefer using technological mediums than are older participants [24]. The app was designed specifically with this age cohort in mind. To conclude, the results suggest a significant decrease in the perceived stress levels and a significant increase in social-emotional well-being for all users aged 14 years or older. Future studies can be designed to investigate the effect of HeartBot after 3 months and then after 6 months to provide further evidence.

Limitations

The current study is the first study to explore the effectiveness of the HeartBot app. Some limitations of this study were identified. First, participants used the app for 21 days based on a calendar, but the specific data reporting the exact length of time that each participant spent on the app were not measured. Second, this study had restrictions on time and resources. Future studies could use a randomized controlled design with a larger sample and longer duration to establish the app's effectiveness. Third, although groups were randomized and no self-selection was involved, there were some baseline differences in the EPOCH scores between the two groups. Fourth, 27% (24/88) of the sample participants were men and 73% (64/88) were women, limiting generalizability to men. Fifth, as the study included minors and adults, further studies could focus on one target group for more substantial data. Sixth, EPOCH has demonstrated good reliability, validity, and sensitivity to change in adolescents but not in adults. Future work should extend the findings of this feasibility study and consider these limitations to lead to broader generalizability.

Implications for Mental Health

The authors examined the HeartBot app's effectiveness in reducing perceived stress and improving overall well-being in participants 14 years and older. These findings provide evidence that HeartBot enables users who use the app for 21 days for 30 minutes every day to manage stress effectively by providing personalized guided audios and Heartfulness tools. Results from the quantitative analyses provide further evidence supporting the use of this app for providing a convenient way for adolescents and adults to learn and practice Heartfulness tools for stress management and social-emotional well-being. More widespread usage of the app under this study could be encouraged based on this preliminary evidence of its effectiveness.

Acknowledgments

The authors are grateful for all the participants in the study. They also thank Kamlesh D Patel, Dr Jayaram Thimmapuram, and Victor Kannan of Heartfulness Institute for their immense help and guidance with the Heartfulness relaxation techniques. Finally, the authors thank Dr Sandra Batista for her support in data analyses for this study.

This study received a grant for funding support for the institutional review board fee from the Heartfulness Institute in the United States.

Conflicts of Interest

None declared. The HeartBot app is available free for its users, and there is no financial benefit to the authors from this study.

Multimedia Appendix 1
Calendar for HeartBot study.
[PDF File (Adobe PDF File), 47 KB - formative_v5i1e22041_app1.pdf]

Multimedia Appendix 2
Statistical data for baseline equivalence.
[DOCX File , 29 KB - formative_v5i1e22041_app2.docx]

Multimedia Appendix 3
Statistical data for ANOVA- Single Factor- HeartBot and Control- Post-PSS scores.
[DOCX File , 30 KB - formative_v5i1e22041_app3.docx]

Multimedia Appendix 4
Statistical data for HeartBot and Control- Post-EPOCH scores.
[DOCX File , 28 KB - formative_v5i1e22041_app4.docx]

References

1. Dillon A, Kelly M, Robertson I, Robertson DA. Smartphone Applications Utilizing Biofeedback Can Aid Stress Reduction. Front Psychol 2016;7:832 [FREE Full text] [doi: 10.3389/fpsyg.2016.00832] [Medline: 27378963]
2. Stress research. American Institute of Stress. URL: https://www.stress.org/stress-research [accessed 2019-11-13]
3. Leonard N, Gwadz M, Ritchie A, Linick JL, Cleland CM, Elliott L, et al. A multi-method exploratory study of stress, coping, and substance use among high school youth in private schools. Front Psychol 2015;6:1028 [FREE Full text] [doi: 10.3389/fpsyg.2015.01028] [Medline: 26257685]
4. Luthar SS. Vulnerability and Resilience: A Study of High-Risk Adolescents. Child Dev 1991 Jun;62(3):600. [doi: 10.2307/1131134]
5. Ford T, Goodman R, Meltzer H. The British Child and Adolescent Mental Health Survey 1999: the prevalence of DSM-IV disorders. J Am Acad Child Adolesc Psychiatry 2003 Oct;42(10):1203-1211. [doi: 10.1097/00004583-200310000-00011] [Medline: 14560170]
6. Howells A, Ivtzan I, Eiroa-Orosa FJ. Putting the 'app' in Happiness: A Randomised Controlled Trial of a Smartphone-Based Mindfulness Intervention to Enhance Wellbeing. J Happiness Stud 2014 Oct 29;17(1):163-185. [doi: 10.1007/s10902-014-9589-1]
7. Morrison Wylde C, Mahrer NE, Meyer RM, Gold JI. Mindfulness for Novice Pediatric Nurses: Smartphone Application Versus Traditional Intervention. J Pediatr Nurs 2017;36:205-212. [doi: 10.1016/j.pedn.2017.06.008] [Medline: 28888505]
8. Wahbeh H, Svalina MN, Oken BS. Group, One-on-One, or Internet? Preferences for Mindfulness Meditation Delivery Format and their Predictors. Open Med J 2014;1:66-74 [FREE Full text] [doi: 10.2174/1874220301401010066] [Medline: 27057260]
9. Rung AL, Oral E, Berghammer L, Peters ES. Feasibility and Acceptability of a Mobile Mindfulness Meditation Intervention Among Women: Intervention Study. JMIR Mhealth Uhealth 2020 Jun 02;8(6):e15943 [FREE Full text] [doi: 10.2196/15943] [Medline: 32442147]
10. Economides M, Martman J, Bell MJ, Sanderson B. Improvements in Stress, Affect, and Irritability Following Brief Use of a Mindfulness-based Smartphone App: A Randomized Controlled Trial. Mindfulness (N Y) 2018 Mar;9(5):1584-1593 [FREE Full text] [doi: 10.1007/s12671-018-0905-4] [Medline: 30294390]
11. Huberty J, Green J, Glissmann C, Larkey L, Puzia M, Lee C. Efficacy of the Mindfulness Meditation Mobile App "Calm" to Reduce Stress Among College Students: Randomized Controlled Trial. JMIR Mhealth Uhealth 2019 Jun 25;7(6):e14273 [FREE Full text] [doi: 10.2196/14273] [Medline: 31237569]

12. Iyer RB, Iyer BN. The Impact of Heartfulness-based Elective on Middle School Students. Am J Health Behav 2019 Jul 01;43(4):812-823. [doi: 10.5993/ajhb.43.4.14]
13. Dörnyei Z. Research methods in applied linguistics: quantitative, qualitative, and mixed methodologies. Oxford, United Kingdom: Oxford University Press; 2007.
14. Heartful schools. Heartfulness. URL: https://www.heartfulnessinstitute.org/education [accessed 2019-12-03]
15. SEL: What Are the Core Competence Areas and Where are they Promoted? Collaborative for Academic, Social, and Emotional Learning (CASEL). URL: https://casel.org/sel-framework/ [accessed 2019-12-03]
16. Sample Size: How Many Survey Participants Do I Need? Science Buddies. URL: https://www.sciencebuddies.org/science-fair-projects/references/sample-size-surveys [accessed 2019-12-03]
17. Cohen S. Contrasting the Hassles Scale and the Perceived Stress Scale: Who's really measuring appraised stress? Am Psychol 1986;41(6):716-718. [doi: 10.1037/0003-066x.41.6.716]
18. Kern ML, Benson L, Steinberg EA. The EPOCH Measure of Adolescent Well-being. Peggy Kern website. 2014. URL: http://www.peggykern.org/uploads/5/6/6/7/56678211/epoch_measure_of_adolescent_well-being_102014.pdf [accessed 2019-12-03]
19. Green SB, Salkind NJ. Using SPSS for Windows and Macintosh: Analyzing and Understanding Data. 5th ed. Upper Saddle River, NJ: Pearson Education Inc; 2010.
20. Vital Wave Consulting. mHealth for Development: The Opportunity of Mobile Technology for Healthcare in the Developing World. United Nations Foundation and Vodafone Foundation Technology Partnership. 2009. URL: http://www.globalproblems-globalsolutions-files.org/unf_website/assets/publications/technology/mhealth/mHealth_for_Development_full.pdf [accessed 2021-01-21]
21. Thimmapuram J, Pargament R, Sibliss K, Grim R, Risques R, Toorens E. Effect of heartfulness meditation on burnout, emotional wellness, and telomere length in health care professionals. J Community Hosp Intern Med Perspect 2017 Jan;7(1):21-27 [FREE Full text] [doi: 10.1080/20009666.2016.1270806] [Medline: 28634520]
22. Bostock S, Crosswell AD, Prather AA, Steptoe A. Mindfulness on-the-go: Effects of a mindfulness meditation app on work stress and well-being. J Occup Health Psychol 2019 Feb;24(1):127-138 [FREE Full text] [doi: 10.1037/ocp0000118] [Medline: 29723001]
23. Mrazek AJ, Mrazek MD, Cherolini CM, Cloughesy JN, Cynman DJ, Gougis LJ, et al. The future of mindfulness training is digital, and the future is now. Curr Opin Psychol 2019 Aug;28:81-86. [doi: 10.1016/j.copsyc.2018.11.012] [Medline: 30529975]
24. Lungu A, Sun M. Time for a Change: College Students' Preference for Technology-Mediated Versus Face-to-Face Help for Emotional Distress. Telemed J E Health 2016 Dec;22(12):991-1000. [doi: 10.1089/tmj.2015.0214] [Medline: 27223266]
25. Tunney C, Cooney P, Coyle D, O'Reilly G. Comparing young people's experience of technology-delivered v. face-to-face mindfulness and relaxation: two-armed qualitative focus group study. Br J Psychiatry 2017 Apr;210(4):284-289. [doi: 10.1192/bjp.bp.115.172783] [Medline: 27765773]

Abbreviations

ANOVA: analysis of variance
EPOCH: Engagement, Perseverance, Optimism, Connectedness, and Happiness
PSS: Perceived Stress Scale

Edited by G Eysenbach; submitted 02.07.20; peer-reviewed by D Saleh, J Green, M Mullarrkey; comments to author 23.07.20; revised version received 03.08.20; accepted 17.01.21; published 29.01.21.

Please cite as:
Iyer L, Iyer RB, Kumar V
A Relaxation App (HeartBot) for Stress and Emotional Well-Being Over a 21-Day Challenge: Randomized Survey Study
JMIR Form Res 2021;5(1):e22041
URL: http://formative.jmir.org/2021/1/e22041/
doi:10.2196/22041
PMID:33455902

Review

Identifying Health Economic Considerations to Include in the Research Protocol of a Randomized Controlled Trial (the REDUCE-RISK Trial): Systematic Literature Review and Assessment

Mattias Neyt[1], MSc, PhD; Annick Christiaens[2], BNurs; Marina Aloi[3], MD, PhD; Lissy de Ridder[4], MD, PhD; Nicholas M Croft[5], MBBS, PhD; Sibylle Koletzko[6,7], MD; Arie Levine[8], MD; Dan Turner[9], MD, PhD; Richard K Russell[10], MD, PhD; Frank M Ruemmele[11,12], MD, PhD; Gigi Veereman[13], MD, PhD

[1]Medical Evaluation and Technology Assessment (ME-TA), Merendree, Belgium

[2]Pediatric Gastroenterology and Nutrition, University Hospital Brussels, Brussels, Belgium

[3]Department of Maternal and Child Health, Gastroenterology Unit, Sapienza University of Rome, Rome, Italy

[4]Erasmus MC Sophia Childrens' Hospital Rotterdam, Rotterdam, Netherlands

[5]Blizard Institute, Barts and the London School of Medicine, Queen Mary University of London, London, United Kingdom

[6]Department of Pediatrics, Dr von Hauner Children's Hospital LMU Klinikum, University of Munich, Munich, Germany

[7]Department of Pediatrics, Gastroenterology and Nutrition, School of Medicine Collegium Medicum University of Warmia and Mazury, Olsztyn, Poland

[8]Pediatric Gastroenterology and Nutrition Unit, Wolfson Medical Center, Holon, Israel

[9]Institute of Paediatric Gastroenterology, Shaare Zedek Medical Center, The Hebrew University of Jerusalem, Jerusalem, Israel

[10]Department of Paediatric Gastroenterology, The Royal Hospital for Sick Children, Edinburgh, United Kingdom

[11]Hôpital Necker Enfants Malades, Assistance Publique - Hôpitaux de Paris, Service de Gastroentérologie Pédiatrique, Paris, France

[12]Faculté de Médecine, Université de Paris, Paris, France

[13]Pediatric Gastroenterology and Nutrition, Vrije Universiteit Brussel, University Hospital Brussels, Brussels, Belgium

Corresponding Author:
Mattias Neyt, MSc, PhD
Medical Evaluation and Technology Assessment (ME-TA)
Jan Burssensstraat 23
Merendree, 9850
Belgium
Phone: 32 477901317
Email: mattias.neyt@me-ta.eu

Abstract

Background: The REDUCE-RISK trial was set up to compare the effectiveness of weekly subcutaneously administered methotrexate with daily oral azathioprine or 6-mercaptopurine in low-risk Crohn disease (CD) or subcutaneously administered adalimumab (ADA) in high-risk CD in a pediatric population (age 6-17 years).

Objective: The aim of this study is to perform a systematic review to provide input into the research protocol to gather the necessary information to improve the performance of an evidence-based economic evaluation when the trial is finished.

Methods: The Centre for Reviews and Dissemination (CRD) Health Technology Assessment (HTA) database, websites of HTA institutes, CRD's National Health Service Economic Evaluation Database, MEDLINE (OVID), and Embase databases were consulted to retrieve (reviews of) relevant economic evaluations. Studies were eligible if they included a pediatric or adult population with inflammatory bowel diseases (CD and ulcerative colitis [UC]) treated with ADA (Humira). There were no restrictions on the comparator. Only economic evaluations expressing outcomes in life years gained or quality-adjusted life years gained were selected.

Results: A total of 12 primary studies were identified. None of these studies included a pediatric population because of a lack of supporting trials. The economic evaluations identified in our systematic review indicate that ADA is an appropriate intervention for inclusion in such a trial. From a health economic point of view, it is important to make an incremental analysis comparing such an intervention with standard care and not immediately versus another (expensive) biological treatment. Information on the

impact of children's school attendance and parents' productivity is currently lacking in economic evaluations, and none of the underlying trials measured quality of life (QoL) using a generic utility instrument.

Conclusions: The review of the economic literature on ADA for the treatment of patients with CD supports the performance of a trial with biologicals in pediatric patients, including making a distinction according to disease severity. Conducting an economic literature review enabled us to decide which variables should be added to the research protocol from an economic point of view. Measurements for children's and parents' QoL (EuroQol 5-Dimension questionnaires), children's school attendance, and parents' productivity (WPAI-CD-CG questionnaire) were added to the research protocol. This will provide support for the calculation of the cost-effectiveness of the interventions evaluated in the REDUCE-RISK trial.

Trial Registration: ClinicalTrials.gov NCT02852694; https://clinicaltrials.gov/ct2/show/NCT02852694

(JMIR Form Res 2021;5(1):e13888) doi:10.2196/13888

KEYWORDS

Crohn disease; cost-benefit analysis; adalimumab; clinical trial; protocol; technology assessment, biomedical

Introduction

The REDUCE-RISK Trial

Immunomodulators such as thiopurines (azathioprine [AZA] or 6-mercaptopurine [6-MP]), methotrexate (MTX), and biologicals such as adalimumab (ADA) are well established for the maintenance of remission in pediatric Crohn disease (CD). However, it remains unclear which maintenance medication should be used first line in specific patient groups. The REDUCE-RISK trial (*Risk-stratified randomized controlled trial in paediatric Crohn's Disease: Methotrexate versus azathioprine or adalimumab for maintaining remission in patients at low or at high risk for aggressive disease course, respectively – a treatment strategy*) aims to compare the efficacy of maintenance therapies in newly diagnosed CD based on stratification into high- and low-risk groups for severe CD evolution: (1) MTX versus AZA/6-MP in low-risk patients and (2) MTX versus ADA in high-risk patients. Patients are eligible if aged 6 to 17 years with new-onset (<6 months) treatment-naïve active and/or perianal fistulizing CD and receiving steroids or exclusive enteral nutrition (EEN) for induction of remission. They are stratified into low- and high-risk groups based on phenotype and disease response to induction therapy. Individual informed consent is obtained before participation in the study. Patients are followed up for 12 months post randomization at prespecified intervals. The primary endpoint is sustained steroid or EEN-free remission at 12 months [1].

The REDUCE-RISK trial, an international multicenter open-label prospective randomized controlled trial (RCT), has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement number 668023. This trial has been reviewed and approved by the National Ethics Services of participating countries and is prospectively registered (ClinicalTrials.gov Identifier: NCT02852694; date of registration: June 9, 2016; EudraCT Number: 2016-000522-18).

Background on Health Technology Assessment and Economic Evaluations

When setting up the protocol for REDUCE-RISK, the research team prepared to allow the performance of a full health technology assessment (HTA). The European Network for HTA (EUnetHTA) describes HTA as "a multidisciplinary process that summarizes information about the medical, social, economic and ethical issues related to the use of a health technology in a systematic, transparent, unbiased, robust manner. Its aim is to inform the formulation of safe, effective, health policies that are patient focused and seek to achieve best value. Despite its policy goals, HTA must always be firmly rooted in research and the scientific method." In HTA, an economic evaluation is performed to determine whether an intervention offers value for money in comparison with other alternatives. This economic consideration might provide support to policy makers when making decisions while trying to make efficient use of limited resources.

An economic evaluation is "the comparative analysis of alternative courses of action in terms of both their costs and consequences" [2]. In an economic evaluation, the incremental cost-effectiveness ratio (ICER) is calculated using the following general formula:

$$ICER=IC/IE=(C_{Int}-C_{Comp})/(E_{Int}-E_{Comp})$$

with C being the costs, Comp the comparator, E the effects, IC the incremental cost, IE the incremental effect, and Int the intervention.

This formula shows that the focus should be on the incremental elements, that is, those that differ between the compared alternatives. Economic evaluations can be performed from different perspectives. As mentioned in the EUnetHTA guideline on methods for health economic evaluations [3], "economic evaluations should at minimum be conducted from a health care perspective. However, several countries require a societal perspective." To make the results of the international REDUCE-RISK trial useful for researchers and policy makers in different countries, incremental elements for both the health care payer and societal perspective will be taken into account.

A Review of the Literature to Provide Input to the Research Protocol

In preparation of a future economic evaluation, we determine the most important incremental elements. The International Society for Pharmacoeconomics and Outcomes Research (ISPOR) guidelines state that "assessing relative value is rarely the primary purpose of an experimental study. Nevertheless, when the decision is made to conduct an economic evaluation

alongside a clinical trial, it is important that the economic investigator contributes to the design of the study to ensure that the trial will provide the data necessary for a high-quality economic evaluation" [4]. Our research question is *which additional elements should we include in the research protocol of the REDUCE-RISK trial to provide support to a high-quality economic evaluation?* Therefore, a systematic search for economic literature about the cost-effectiveness of ADA (Humira) for the treatment of inflammatory bowel disease (IBD) was performed.

The purpose of this systematic review is to obtain more useful insights and knowledge from previous economic studies [5]. These previous economic evaluations guide us in finding the key variables that enable us to provide well-directed input for the research protocol. In this paper, the review of the economic literature is transparently presented. No official review protocol was established for the systematic review. The findings are used to provide input for the research protocol from a health economic point of view (eg, to decide which questionnaires should be added to the research protocol). On the other hand, we also want to avoid overloading the research protocol and only focus on the incremental elements that influence an intervention's cost-effectiveness. The results of this systematic review help us to focus on gathering the right information in the REDUCE-RISK trial, which will support researchers at the end of the trial to make a high-quality economic evaluation.

Methods

A systematic review of the literature was conducted using predefined selection criteria that included considerations of population, intervention, comparator, and design. As the goal of this study is to provide input for the research protocol, the applied selection criteria were not too restrictive. Studies were included if (1) the population included children or adults with IBD (CD and ulcerative colitis [UC]); (2) ADA was one of the included interventions; and (3) the design reflected a full economic evaluation, that is, studies comparing at least two alternative treatments in terms of costs and outcomes and expressing outcomes in life years gained or quality-adjusted life years (QALYs) gained. No restrictions were applied to the comparator. Studies were excluded if they only considered other treatments than ADA at the moment of randomization. Studies that only included switching to ADA in case of no response to the interventions under consideration (ie, not including ADA at the moment of randomization) were not selected. Cost analysis or cost-of-illness studies did not fulfill the aforementioned definition of an economic evaluation and were excluded. As summarized in an EUnetHTA guideline providing an overview of national guidelines for 25 countries, "all countries except four specify that the preferred outcome measure is QALYs or both QALYs and life years. Of the four countries with guidelines that do not announce QALYs as a preferred method, at least three accept QALYs in special circumstances or in complementary analyses" [3]. In this study, we focus on these preferred outcomes. Studies expressing results in disease-specific outcomes (eg, cost per remission [6,7], cost per responder [8], or cost per mucosal healing [9]) are thus excluded. *Before-after* analyses [10] comparing the costs before and after the start of treatment with ADA were also excluded, as they also do not fulfill the definition of an economic evaluation (ie, lack of a comparative intervention). Abstracts were excluded because of a lack of sufficient details to allow for a proper evaluation. No time restriction was imposed. English, French, German, and Dutch papers were eligible.

Various databases were consulted. In February 2016, before the final protocol was set up, reviews on this topic were searched by consulting the Centre for Reviews and Dissemination (CRD) HTA database and HTA institute websites listed on the International Network of Agencies for Health Technology Assessment website. The websites of ex- or nonmember HTA institutes such as the National Institute for Health and Care Excellence (NICE) were also consulted. In September 2016, CRD's National Health Service Economic Evaluation Database, MEDLINE (OVID), and Embase databases were searched to retrieve both full economic evaluations and reviews of full economic evaluations of ADA for IBD treatment. To ensure reproducibility, further details of the search strategy are provided in Multimedia Appendix 1.

The selection of relevant papers was performed in a 2-step procedure: an initial assessment of the title, abstract, and keywords, followed by a full-text assessment of the selected references. When no abstract was available and the citation was unclear or ambiguous, consideration of the citation was made directly on the basis of a full-text assessment. Reference lists of the selected studies were checked for additional relevant citations. The procedure was performed by a health economist (MN), and in case of doubt for medical reasons, a medical specialist (GV) provided support.

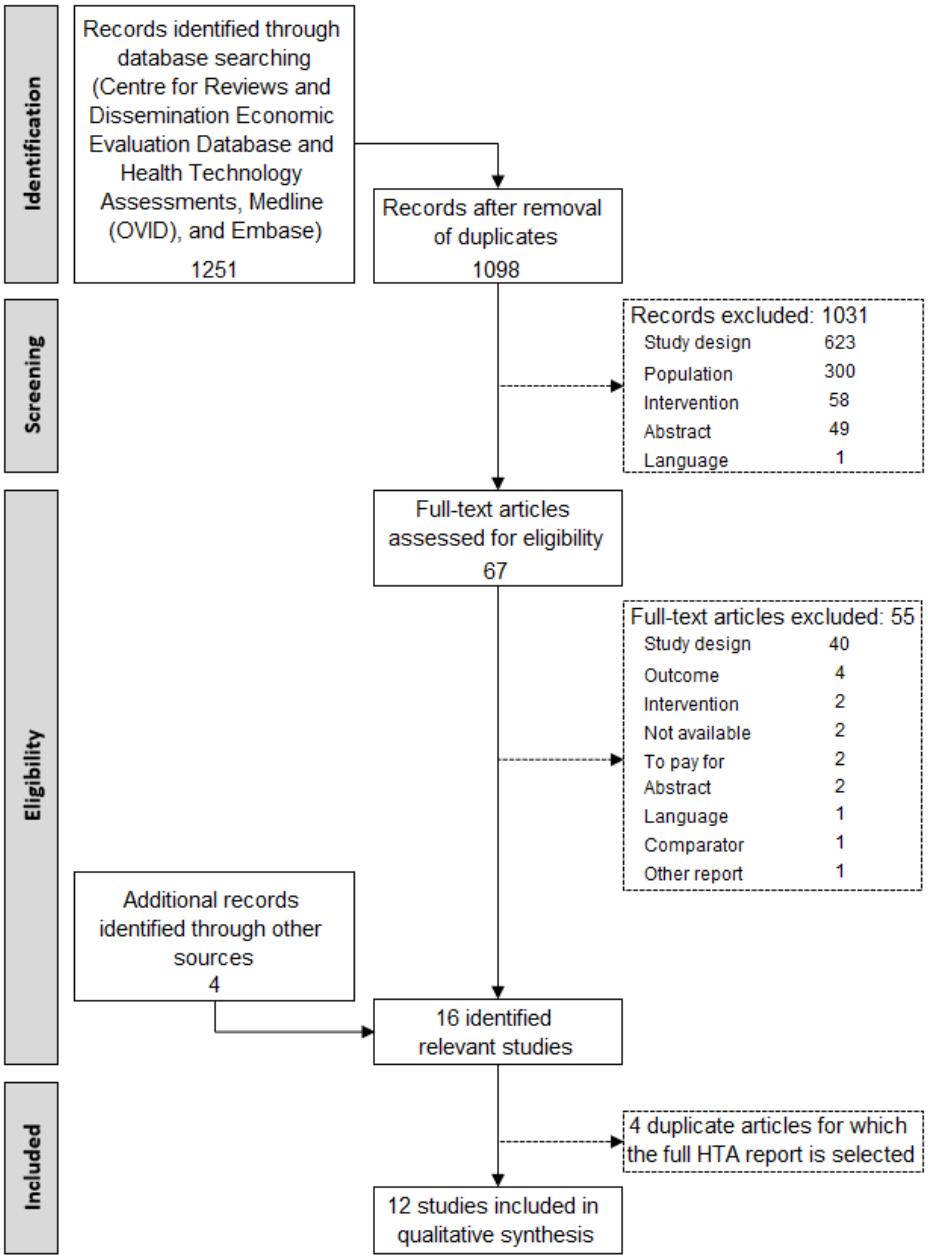
The primary full economic evaluations were summarized in an in-house data extraction sheet listing all variables (eg, population, intervention, comparators, and quality of life [QoL] input) and summary measures (eg, ICERs and results of sensitivity analyses) for which data were sought (Multimedia Appendix 1). The information gathered in these sheets reflects the elements that are usually reported in an economic evaluation (eg, according to the Consolidated Health Economic Evaluation Reporting Standards [CHEERS] guidelines [11,12]). This information was used to set up summary tables that form the basis for a further critical assessment. On the basis of the results of this assessment, we judged whether from an HTA and economic perspective, elements in the research protocol of the REDUCE-RISK trial could be added.

Results

Article Selection

Figure 1 presents the flowchart of the selection process. A total of 12 papers were identified in the electronic databases. Four additional references were identified by searching websites of HTA institutes. Information from 3 journal papers [13-15] and 1 report [16] were already published in an HTA report. To avoid overlap, only the 12 primary studies [17-28] will be further discussed. The list of the 16 identified economic evaluations and information on duplicates is provided in Multimedia Appendix 2.

Figure 1. Selection of relevant articles.



General Information

Half of the studies were performed for the United Kingdom (n=6; Table 1). Two studies conducted an analysis for Canada, another 3 for the United States, and 1 for Poland. All but one of the studies explicitly declared the presence or absence of conflicts of interest. All studies were cost-utility analyses. Most short-term models (1 year) applied a decision tree, whereas long-term analyses (5, 10, or 30 years or lifetime) are Markov models or a combination of an initial decision tree and a Markov component. For the long-term models, applied discount rates are in agreement with national recommendations in all but one of the analyses. In this study, the manufacturer assumed an annual discount rate of 3% for both health and cost outcomes, although the Canadian Agency for Drugs and Technologies in Health (CADTH) guidelines recommend a 5% discount rate [19]. However, a 5% discount was applied in the sensitivity analysis. For further details, we refer to section 2.3.1 of the Multimedia Appendix 1.

Table 1. General information on the identified economic evaluations.

Study	Country	CoI[a]	Analytic technique	Design	Time horizon	Discount rate[b] (%)
Archer et al [17]	United Kingdom	No	CUA[c]	Markov	Lifetime	3.5
AbbVie submission[d]	N/A[e]	Yes	CUA	Markov	10 years	3.5
MSD[f] submission[d]	N/A	Yes	CUA	Decision tree+Markov	10 years	3.5
Assasi et al [18]	Canada	No	CUA	Markov	5 years	5
Bodger et al [23]	United Kingdom	No	CUA	Markov	Lifetime (60 years)	3.5
CADTH [19]	Canada	Yes or no[g]	CUA	Markov	10 years	3
Dretzke et al [20]	United Kingdom	No	CUA	Markov	1 year	—[h]
Abbott submission[i]	N/A	Yes	CUA	Markov	1 year (56 weeks)	—
Essat et al [21]	United Kingdom	Yes or no[j]	CUA	Decision tree+Markov	10 years	3.5
Kaplan et al [24]	United States	Yes	CUA	Decision tree	1 year	—
Loftus et al [25]	United Kingdom	Yes	CUA	Regression model	1 year	3.5
Rafia et al [22]	United Kingdom	Yes or no[j]	CUA	Decision tree+Markov	10 years	3.5
Stawowczyk et al [26]	Poland	No	CUA	Markov	30 years	Costs: 5; Effects: 3.5
Tang et al [27]	United States	Not declared	CUA	Decision tree	1 year (54 weeks)	—
Yu et al [28]	United States	Yes	CUA	Decision tree	1 year (56 weeks)	—

[a]CoI: conflict of interest.

[b]Discount rate for both costs and effects, unless otherwise mentioned.

[c]CUA: cost-utility analysis.

[d]The AbbVie and MSD submissions are part of the report published by Archer et al [17].

[e]N/A: not applicable. This submission is part of the UK report in which it was published.

[f]MSD: Merck Sharp & Dohme.

[g]Submission by the manufacturer reviewed by the CADTH team (Common Drug Review Analyses).

[h]No discount rate is applied because of the short time horizon of 1 year.

[i]Abbott submission is part of the report published by Dretzke et al [20].

[j]The manufacturer submitted a model-based health economic analysis as part of their submission, which was then evaluated by a team of researchers from the School of Health and Related Research.

Population and Compared Interventions

None of the studies included a pediatric population. The primary economic evaluations investigated treatment strategies for adult patients (average age 35-40 years and average weight 69-77 kg) with moderate-to-severe UC or CD. In two studies, a secondary analysis is considered for the pediatric population [17,20]. The authors consider this as an exploratory analysis, as the efficacy data are drawn from trials undertaken within an adult UC population [17].

Most studies explicitly mention that patients failed (intolerance, inadequate response, or loss of response) to respond to standard therapy before ADA was considered. In all but 3 studies [24,27,28] and the MSD submission [17], conventional nonbiological therapy is considered as a comparator. This usually exists as a mix of 5-aminosalicylates (5-ASAs), corticosteroids, and immunosuppressants. In 2 studies, only biologicals are included [24,28]. The study by Kaplan et al [24] considered whether dose escalation of infliximab (IFX; to 10 mg/kg every 8 weeks) is a cost-effective strategy compared with ADA initiation after loss of response to 5 mg/kg of IFX. In addition, Yu et al [28] compared IFX and ADA. This study was also part of the Abbott submission, which contained 2 models: one comparing the cost-effectiveness of ADA as a maintenance therapy against standard care (SC) and the other comparing the cost-effectiveness of ADA and IFX as maintenance therapies [20]. The report of Dretzke et al [20], which made a critical assessment of Abbott's submission, concentrates on the model including SC as a treatment option (refer to the *Discussion* section).

In 2 studies [17,21] and the MSD submission [17], surgery (colectomy) is considered as an initial treatment option. In these studies, surgery is included as an alternative treatment strategy and a downstream component of the pathway for patients in the other treatment strategies. In other models, such as the models

discussed in the CADTH report [19] and from the AbbVie submission [17], surgery is not considered a direct comparator but only included as a treatment for patients who failed both biological and nonbiological drug treatments.

Next to ADA, the most frequently included biological treatments are IFX, golimumab, and vedolizumab. Certolizumab pegol and natalizumab are also included in individual studies. In most studies, ADA was administered as follows: induction—160 mg (week 0), 80 mg (week 2); maintenance—40 mg every other week (starting from week 4) [17-19,21,24,26-28]. In other studies, the following treatment schedule was applied: induction—80 mg (week 0), 40 mg (week 2); maintenance—40 mg every other week [20,22,23,25]. The duration of treatment might also be different but is not always clearly stated. Bodger et al [23] included 1 or 2 years of treatment with ADA or IFX, after which patients return to SC. For further information on the treatment schedule of the other interventions and dose escalation, we refer to section 2.3.2 of Multimedia Appendix 1.

Costs

Most economic evaluations are performed from the perspective of the health care payer. Tang et al [27] took the perspective of a managed care organization in the United States and excluded patient co-payments. Only 2 studies applied a broader societal perspective, including costs related to lost productivity. Loftus et al [25] assumed that each CD-related hospitalization corresponds to a missed interval of work equal to the average duration of serious adverse events (AEs) leading to hospitalization (on average 16.55 days based on the Crohn's Trial of the Fully Human Antibody Adalimumab for Remission Maintenance (CHARM) trial [29]). This was then multiplied by an 8-hour workday and an average hourly wage in the United Kingdom of £13.00 (US $17.6) [25]. Stawowczyk et al [26] included indirect costs based on an unpublished study carried out in Poland on 202 patients with UC. Indirect costs included absenteeism, presenteeism, and leaving the labor market earlier. Yearly indirect costs for remitted patients counted as PLN 6524 (US $1767) [26]. For patients with active disease, this was PLN 22,935 (US $6211). An overview of the perspective, currency, and year of costing is provided in Multimedia Appendix 1.

Costs of Biological Treatments

An overview of the unit costs of biological treatments is provided in Multimedia Appendix 1. It is remarkable that although the unit cost for 40 mg of ADA is lower than 100 mg of IFX in the studies for the United Kingdom [17,21,22] and Canada [18,19], this is the opposite in all US studies [24,27,28]. However, IFX is assumed to be administered in a day-case setting, whereas ADA can be self-administered subcutaneously. As a result, the total treatment costs with IFX are not always lower if administration costs are also taken into account. For example, in the study of Yu et al [28], the total therapy cost for ADA equals the drug costs of US $17,176. For IFX, the total therapy cost of US $18,214 consists of the drug costs (US $14,663), the drug administration costs (US $1605), and excess uninfused drug costs (US $1946). Differences in start-up dose or dose escalation probabilities further influence total treatment costs (Multimedia Appendix 1).

Costs of Standard Care

The costs of SC are substantially lower than those of biological treatment costs. In the study by Archer et al [17], SC in the induction phase of 8 weeks, consisting of 5-ASAs, AZA, 6-MP, and prednisolone, costs £167.6 (US $227.4). In the maintenance phase of 26 weeks, this is £343.8 (US $466.4). The same use of background therapies is assumed for all biological treatment arms. Standard care cost differences were small between the different treatment strategies. For example, in their model comparing golimumab with ADA, IFX, or SC, the background therapy costs were £251.43 (US $341.1) per cycle for the standard nonbiological treatment group versus £200.03 (US $271.4) per cycle for the biological treatments during the induction treatment (cycle=8 weeks). During maintenance treatment (cycle=2 months), this was £121.15 (US $164.4) versus £120.98 (US $164.1), respectively [17]. Assasi et al [18] include a total non-anti-TNF outpatient drug costs per cycle (8 weeks) of CAD 116.30 (US $91.6) for drug responsive patients and CAD 85.95 (US $67.7) for drug refractory patients. In the report of Essat et al [21], conventional treatment (balsalazide, mesalazine, olsalazine, sulfasalazine, budesonide, prednisolone, AZA, 6-MP, and MTX) costs £153.6 (US $208.3) per induction cycle (6 weeks) and £204.8 per maintenance cycle (8 weeks). The authors assume that while patients are receiving biological therapy, the costs of conventional therapies are halved (£102.4 [US $138.9]). The same logic is applied in the study of Rafia et al [22], with a cost of £52.62 (US $71.4) per induction cycle and £70.16 (US $95.2) per maintenance cycle, which is halved (£35.08 [US $47.6]) for patients while receiving biological treatment. In the Polish study, standard treatment costs per cycle (8 weeks) are PLN 204 (US $55.3) [26]. Thus, the incremental impact of these costs is minimal in comparison with the biological treatment costs.

Costs of Colectomy or Surgery

From the 3 studies including colectomy as an alternative treatment strategy, Archer and Essat refer to information from the study of Buchanan et al [30] to include a cost of £13,452 (US $18,242) [17] and £13,577 (US $18,412) [21] for surgery. In the MSD submission [17], the cost for colectomy is £8968 (US $12,160). Other studies include surgery as an event in their model, without providing further details on the type of hospitalization. A wide range of costs is mentioned: the surgery cost is PLN 12,480 (US $3380) in the Polish study of Stawowczyk et al [26] up to $31,923 for a hospital unit cost in the US study of Yu et al [28]. The cost of surgery was $11,341 in the US study of Kaplan et al [24], £10,581 (US $14,351) in the UK study of Rafia et al [22], and CAD 19,269 (US $15,171) in the CADTH study [19].

Incremental Costs Related to AEs

Archer et al [17] report that serious and severe AEs were not considered in the AbbVie model. The manufacturer notes that most AEs experienced by patients were nonserious and considered to be unrelated to the study drugs (based on results from the ULTRA2 trial [31]) [17]. In addition, the manufacturer highlights that the exclusion of these events represents a conservative assumption as "the ULTRA2 trial reported slightly higher incidences of serious and severe AEs in the placebo arm

than in the adalimumab arm of the trial; therefore, considering serious and severe AEs in the model would have increased medical costs and reduced health gains within the conventional management group" [17]. In addition, the Polish study refers to the ULTRA2 trial [31] to justify that certain AEs were not included in the model because ADA treatment was generally well tolerated and the overall safety profile of ADA was comparable with that of placebo [26]. Rafia et al [22] also indicate that the impact of AEs on the ICER is minimal. Finally, Tang et al [27] mention that they are not aware of evidence that demonstrates large differences in the proportion of adverse drug reactions across the 4 biological treatments (ADA, IFX, certolizumab pegol, and natalizumab), and the frequency of these complications is low. On the basis of their clinical judgment, they conclude that adverse drug reactions should not be included in the structure of the model. Progressive multifocal leukoencephalopathy occurrence with natalizumab was considered a rare but significant AE with a treatment cost between $14,544 (lower limit) and $22,725 (upper limit) [27].

QoL

Archer et al [17] performed a systematic literature search for utility values. A total of 10 studies reported EuroQol 5-Dimension (EQ-5D) estimates for one or more health states relevant to their model (Multimedia Appendix 1). The authors considered the values reported by Woehl et al [32] and Swinburn et al [33] to be the most useful as "they are UK based, included a fairly large number of patients (n=180 and n=230, respectively) and have the greatest coverage of the health states in the model" [17]. Unfortunately, both studies are only published as an abstract. Swinburn et al [33] included 230 UC patients (postsurgery [n=30], remission [n=78], mild disease [n=47], moderate disease [n=31], and severe disease [n=44]). The EQ-5D utility scores were collected via an online survey. The results are presented in a figure in the abstract, without mentioning the exact utility values. Archer et al [17] extracted utility values from this graph: for patients who had not undergone surgery, the utility scores for each disease severity were as follows: remission 0.91 (95% CI 0.87-0.95), mild disease 0.80 (95% CI 0.70-0.85), moderate disease 0.68 (95% CI 0.58-0.78), and severe disease 0.45 (95% CI 0.35-0.55). Archer et al [17] extracted the following mean EQ-5D scores from the study of Woehl et al [32]: remission 0.87 (SD 0.15), mild disease 0.76 (SD 0.18), and moderate-to-severe disease 0.41 (SD 0.34).

None of the identified economic evaluations is based on an RCT, including a head-to-head comparison of the relevant intervention and comparator, in which utilities are measured with a generic utility instrument. Next to the previously mentioned studies of Woehl et al [32] and Swinburn et al [33], a variety of other sources and assumptions were used in the other identified economic evaluations. We refer to section 2.3.5 of Multimedia Appendix 1 for further details. Most economic evaluations refer to the study of Gregor et al [34] to retrieve relevant utility values. They used the time trade-off (TTO), Standard Gamble (SG), and Visual Analog Scale (VAS) methods in 180 consecutive patients with CD to obtain utilities. All but one of the studies referring to Gregor et al [34] mentioned the use of values from the SG approach [18,24,25,27,28]. Only Dretzke et al [20] used the TTO values. The results for the SG technique were as follows: mild disease, 0.82; moderate disease, 0.73; and severe disease, 0.54 [34]. For the TTO technique, these were 0.96, 0.88, and 0.71, respectively. A second table presented the following mean utility scores for the SG technique at the initial visits: chronically active therapy resistant, 0.74; chronically active therapy responsive, 0.86; acute disease exacerbation, 0.77; remission, 0.88; and overall, 0.81. With the TTO approach, these were 0.88, 0.98, 0.89, 0.96, and 0.92, respectively [34].

Bodger et al [23] mapped the midpoint CD Activity Index (CDAI) scores to EQ-5D utility scores. An algorithm developed by Buxton et al [35] was used (EQ-5D=0.9168−0.0012×CDAI). This algorithm was based on multiple observations from 905 patients with moderate-to-severe CD who participated in the Efficacy of Natalizumab as Active Crohn's Therapy (ENACT-1) and Evaluation of Natalizumab as Continuous Therapy (ENACT-2) clinical trials [36]. We refer to our discussion on QoL for some critical remarks from the authors who developed this algorithm.

Finally, large differences are observed in the postsurgery remission utility values. Several studies assign a value equal [18] or similar [24,27] to the utility for (medical) remission. In contrast, several other studies assign a much lower value for postsurgery remission. For example, in the MSD submission, the utility value for postcolectomy remission (0.60) was assumed to be equal to the utility value for late complications (postcolectomy). Similar values for postsurgery remission were assumed in the study discussed in the CADTH report (0.67) [19] and in the studies by Essat et al (0.60) [21], Rafia et al (0.57) [22], and Stawowczyk et al (0.61) [26], whereas remission utility values were much higher—0.82, 0.86, 0.82, and 0.88, respectively. Multimedia Appendix 1 provides further details on the utility values assigned to (post-)surgery health states. We come back to the observed differences in the *Discussion* section.

Treatment Effect

The treatment effect of the included studies was based on a wide range of sources. An overview of the trials is provided in Multimedia Appendix 1. Several authors conducted a network meta-analysis for both the induction and maintenance phases [17,21,22]. Other studies performed an indirect comparison. For example, in the study of Assasi et al [18], the initial remission and response rates for IFX were derived from the 12-week results of the 5 mg/kg arm that was reported by Targan et al [37]. For ADA, the 4-week results of the 160 mg and 80 mg arm of the CLASSIC 1 study were used [38]. For the usual care strategy, pooled rates from the placebo arms of these 2 trials were used to estimate remission and response rates. Bodger et al [23] and Dretzke et al [20] selected the ACCENT I trial [39] to model the IFX arm and the CHARM trial [29] for ADA. This was also the case in the study by Kaplan et al [24]. However, in the latter study, the initial response rate to ADA was retrieved from the GAIN study [40] that evaluated ADA induction following IFX failure. In addition, the CADTH report [19] referred to an indirect treatment comparison conducted by the manufacturer to estimate the efficacy of treatments for

inducing response or remission. Although most studies refer to the CHARM trial [29] for information on ADA, Stawowczyk et al [26] referred to the ULTRA 2 study [31] for estimates of response and remission with ADA or SC.

Yu et al [28] relied on data from the CHARM [41] and ACCENT I [42] trials to model results for the ADA and IFX treatment arms. The authors remark that patient samples were not equivalent at baseline. The CHARM trial included patients with a maximum baseline CDAI score of 450 versus 400 for ACCENT I [28]. Therefore, the sample of 234 ADA-treated patients was weighted to have the same baseline median and the same 25th and 75th percentile CDAI values, sex distribution, and median age as those in the IFX arm of ACCENT I [28]. No such adjustments were made in the other indirect comparisons. Furthermore, strong assumptions are made. For example, Archer et al [17] point out that in the MSD submissions, "patients who have previously achieved a response can either maintain or lose that response, but they cannot improve (i.e. they cannot subsequently transit to the remission state). ...no additional patients can achieve remission after induction and no patients with remission can completely lose response during any given model cycle."

Incremental Cost-Effectiveness Ratios

In this overview, we focus on the results of treatment with ADA and other biologicals (IFX, golimumab, and vedolizumab). For a more detailed overview, we refer to sections 2.3.8 and 2.3.9 of Multimedia Appendix 1. Section 2.3.8 of Multimedia Appendix 1 also includes overview tables presenting the results of the identified economic evaluations.

In the study by Archer et al [17], when colectomy is an alternative, colectomy is expected to dominate IFX, ADA, golimumab, and conventional nonbiological treatments. When elective colectomy is not an acceptable or preferred option, IFX and golimumab are expected to be ruled out because of dominance. The ICER of ADA versus conventional nonbiological treatment is expected to be approximately £50,300 (approximately US $68,200) per QALY gained [17]. In the AbbVie submission (marketing ADA–Humira), ADA had an ICER of £34,590 (approximately US $46,900) per QALY gained. In contrast, in the MSD submission (marketing IFX or Remicade and golimumab or Simponi), ADA is expected to be dominated by golimumab [17].

Assasi et al [18] calculated that usual care created the lowest expected QALYs. However, the costs associated with ADA and IFX could be perceived as high with ICERs of about CAD 193,000 (approximately US $152,000) per QALY and CAD 451,000 (approximately US $355,000) per QALY, respectively.

According to Bodger et al [23], IFX was always more expensive and less effective than ADA, for which their model suggests acceptable ICERs of less than £14,000 (approximately US $19,000) per QALY [23].

CADTH evaluated the manufacturer's pharmacoeconomic evaluation. According to the manufacturer's calculations, golimumab has an ICER of about CAD 42,000 (approximately US $33,000) per QALY and other biosimilars are (extendedly) dominated. However, the Common Drug Review identified several issues in the indirect treatment comparison and assumptions that might bias the results in the favor of golimumab [19].

Dretzke et al [20] did not mutually compare ADA and IFX. For induction therapy, both biosimilars dominated SC in the management of severe CD. For moderate CD, only ADA was dominant relative to SC. Neither drug was considered cost-effective as maintenance therapy with ICERs of about £5 million (approximately US $6.8 million) per QALY and £14 million (approximately US $19 million) per QALY for ADA and IFX, respectively [20].

According to the manufacturer's analysis reviewed by Essat et al [21], vedolizumab dominates surgery, IFX, and golimumab. Compared with ADA, the ICER of vedolizumab is estimated at £6634 (US $8990) per QALY. In contrast, according to the Evidence Review Group (ERG), surgery is likely to dominate all medical treatments. If surgery is not an option, the review group indicates that vedolizumab is expected to be dominated by ADA [21].

Kaplan et al [24] estimated that IFX dose escalation yielded 0.03 extra QALYs compared with the ADA strategy. However, in combination with an extra cost of more than $10,000, this results in an ICER of more than $330,000 per QALY [24].

In comparison with nonbiological pharmacotherapy, Loftus et al [25] calculated ADA has an ICER of about £16,000 (approximately US $21,700) per QALY and £34,000 (approximately US $46,000) per QALY in the treatment of severe or moderate-to-severe CD, respectively [25].

In the study by Rafia et al [22], ADA provided 0.21 additional QALYs in comparison with conventional nonbiological therapy for an additional cost of £4000 (approximately US $5400), resulting in an ICER of about £19,000 (approximately US $25,700) per QALY. Vedolizumab was extendedly dominated. IFX provides 0.0383 additional QALYs in comparison with ADA for an additional cost of approximately £4400 (approximately US $6000), leading to an ICER of almost £116,000 (approximately US $157,000) per QALY [22].

In addition, the study of Stawowczyk et al [26] indicated that ADA is more effective and more costly than SC. One-year ADA treatment results in an ICER of €71,000 (approximately US $86,700) or €76,000 (approximately US $92,800) per QALY, depending on the perspective [26].

Tang et al [27] compared several treatments after SC failed. No significant differences in efficacy were calculated between the 4 biological treatments (IFX, ADA, certolizumab pegol, and natalizumab). They produce similar QALYs with overlapping 95% CIs. On the basis of Monte Carlo simulations, IFX had the highest probability of being the most cost-effective therapy compared with the other biological treatment options [27].

In contrast with the previous study, Yu et al [28] calculated that ADA delivers more QALYs and saves approximately $4900 in comparison with IFX. On the basis of the probabilistic analysis, ADA dominates IFX in approximately 94% of the simulations and is the preferred biological treatment option [28].

Uncertainty

Almost all studies performed both probabilistic sensitivity analysis and scenario analyses or one-way sensitivity analyses to estimate the uncertainty surrounding estimates of incremental costs, incremental effects, and ICERs. Only Kaplan et al [24] did not perform a probabilistic sensitivity analysis.

Multimedia Appendix 1 gives an overview of the most determining variables, as indicated by the authors of the original economic evaluations. The most often mentioned variables are the treatment effect [18,20,21,24,27], utilities [17,21,22,27], and time horizon [18-23,25]. Some authors also highlight the importance of the treatment duration [21,23,26], drug treatment costs [24,27], health state costs [17,21], and patient weight [18].

Discussion

An overview of the economic literature allows us to identify important issues related to (the calculation of) the cost-effectiveness of ADA. A major strength is that this exercise was performed before the trial was started. This way, we avoid that important information to allow the performance of a high-quality economic evaluation was not measured in the trial. In the first part of this discussion, issues identified from the review of the economic literature relevant to the REDUCE-RISK trial are discussed. On the basis of these issues and expert opinion, input is provided to ameliorate the protocol of the REDUCE-RISK trial from a health economic point of view. The added questionnaires are discussed in the second part of this discussion.

Issues Identified in the Economic Literature Review Related to the REDUCE-RISK Trial

Pediatric Population

All the identified economic evaluations performed an analysis for an adult population. Two studies also included a secondary analysis for the pediatric population [17,20]; however, efficacy data still relied on trials that included only an adult population. The analysis also did not include youngest children. Archer et al [17] reported that patients' starting age in their pediatric population was 15 years. The lack of information related to the treatment effect of biologicals in pediatric patients means that the results of such secondary analyses should be interpreted with caution. Archer et al [17] suggested RCTs assessing the clinical effectiveness of biologicals in pediatric patients as a research priority. This makes the conduct of the REDUCE-RISK trial, which includes patients aged 6 to 17 years, very worthwhile.

Severity of Disease

Almost all studies explicitly include a population with moderate-to-severe CD or UC disease. Only two reports differentiate results according to disease severity. Dretzke et al [20], inclusive of the Abbott submission discussed in the same report, distinguish between severe and moderate disease. Loftus et al [25] performed calculations for severe CD and moderate-to-severe CD. They did not perform a separate analysis for the moderate CD patients. Such a distinction is important in economic evaluations as applying the same relative treatment effect to a higher baseline risk for a specific event results in a larger absolute treatment effect. The severity of disease might have a significant impact on an intervention's ICER. Making an explicit distinction in the REDUCE-RISK trial between patients at low or high risk for aggressive disease course is desirable from both clinical and economic points of view.

Adalimumab Versus Other Biological Treatment Options

In most of the identified economic evaluations, ADA has a better cost-effectiveness than the other biologicals. In the study by Archer et al [17], IFX and golimumab are expected to be ruled out in the economic analysis because of dominance (less effective and more expensive), whereas the ICER of ADA versus conventional nonbiological treatment is expected to be approximately £50,300 (approximately US $68,200) per QALY gained [17]. In the study by Assasi et al [18], the ICER of ADA versus usual care is relatively high (approximately CAD 193,000 [US $152,000] per QALY); however, this is even higher for IFX versus ADA (CAD 451,000 [approximately US $355,000] per QALY). In the study by Bodger et al [23], IFX is dominated by ADA.

In the study by Dretzke et al [20], IFX and ADA are not mutually compared. However, the findings of the economic model were in favor of ADA: for induction, both ADA and IFX were cost-effective (dominant relative to SC) in the management of severe CD, and ADA was cost-effective for moderate CD (dominant relative to SC) [20].

In the study by Essat et al [21], according to the ERG group, vedolizumab is expected to be dominated by ADA if surgery is not an option [21].

The results of the study by Rafia et al [22] indicate that, assuming a cost-effectiveness threshold of £30,000 (approximately US $40,000) per QALY, ADA has the highest probability of being the most cost-effective intervention (78%). Similarly, in the study by Yu et al [28] based on the probabilistic analysis, ADA dominates IFX in approximately 94% of the simulations.

In the US study of Tang et al [27], IFX and ADA are about equally effective and IFX is cheaper. Finally, Loftus et al [25] and Stawowczyk et al [26] only compared ADA with conventional nonbiological treatment.

On the basis of the aforementioned information, most of the economic studies were in favor of ADA in comparison with other biologicals. Only in the manufacturer's submissions, the conclusion is different. For example, in the MSD submission [marketing IFX or Remicade and golimumab or Simponi), ADA is expected to be dominated by golimumab [17]. However, the manufacturer's submission includes a discount for the drug. If this discount is not taken into account, golimumab is ruled out because of extended dominance. Similarly, in a Canadian study, according to the manufacturer's calculations, golimumab has an ICER of approximately CAD 42,000 (approximately US $33,000) per QALY and IFX and ADA are (extendedly) dominated [19].

Most of the identified studies indicate that ADA has a better cost-effectiveness than the other biologicals included in the

analyses, and thus, from a health economic point of view, it seems to be a justified biological intervention in future trials. Nevertheless, in economic evaluations, it is important to work on the efficiency frontier, that is, comparing treatments with the next best non-(extendedly) dominated intervention. From a health economic point of view, it is important to make an incremental analysis and include SC in the analysis and not immediately compare biologicals with each other. As mentioned by Dretzke et al [20], this would only be relevant "where both adalimumab and infliximab have been first justified as maintenance therapies versus standard care (SC). Where one or both maintenance therapies are not cost-effective versus SC, this comparison provides no information to decision-makers." Therefore, from a health economic point of view, we considered it justified to compare methotrexate with ADA in the high-risk patient group of the REDUCE-RISK trial.

Treatment Effect

The input for the conventional nonbiological, ADA, and IFX treatment is often based on the CLASSIC I [43], CHARM [29], and ACCENT I [39,42] trials, respectively. However, comparing outcomes from individual treatment arms of separate trials might bias the results, and the direction of this bias is unknown. Head-to-head RCTs are needed to allow an unbiased comparison of biological therapy with SC. It is also necessary to set up reliable health economic models to estimate the intervention's cost-effectiveness. The REDUCE-RISK trial is an example of such a head-to-head RCT.

QoL

The EUnetHTA guideline for methods for health economic evaluations recommends that results be presented in terms of both a cost-effectiveness analysis and a cost-utility analysis [44]. The primary outcome measures should, where appropriate, be presented as natural units (including life years) and as QALYs [44]. The health-related quality of life (HRQoL) aspects of the QALY were captured in a HRQoL weight. On the basis of the review of guidelines used by EUnetHTA partners, EQ-5D is the most commonly recommended instrument for deriving HRQoL weights, although other instruments are also mentioned (eg, Health Utility Index, Short-Form 6-Dimension, or 15-dimensional) [44].

A major limitation is that none of the underlying trials measured QoL with a generic utility instrument. As a result, the authors of the economic evaluations have to make many assumptions in their models. Previous reviewers also noticed strange assumptions regarding utility values that are linked to health states. For example, Essat et al [21] reported that the utility value in postsurgical remission was lower than for moderate or severe disease (0.60 vs 0.68), which appears to be inconsistent. Bodger et al [23] transformed CDAI scores to utilities, based on an algorithm developed by Buxton et al [35]. In this study, the correlation between CDAI and EQ-5D is −0.62, and 29% of the variability in EQ-5D scores is explained by CDAI [35]. However, Buxton et al [35] mention in their discussion that "based on the variance explained, the relationships between the CDAI and utilities in the simple models are weaker than those for the IBDQ [Inflammatory Bowel Disease Questionnaire] and suggest that the CDAI provides a poorer basis for estimating utilities. Again its relatively poor performance as a predictor of utility reflects its main role as clinical indicator of disease activity, rather than of HRQoL." In the absence of utility values for surgery, Dretzke et al [20] assumed that this health state is represented by the EQ-5D state 22222, which has a UK utility weight of 0.516. Such assumptions are arbitrary and not very reliable. As the model results are sensitive to such utility assumptions, better evidence-based input is desirable. In the REDUCE-RISK trial, this is taken into account by including the EQ-5D questionnaire in the research protocol (see the second part *EQ-5D* of this discussion).

Indirect Costs

Finally, the studies are performed from a health care payer's perspective, which excludes indirect nonhealth care related costs, such as costs related to lost productivity. In contrast, indirect costs would represent a substantial portion of the costs of CD. A US study indicates that this accounts for 28% of the total CD cost in the United States [45]. Only 2 studies [25,26] included a scenario with the inclusion of these costs. In the Polish study, based on an unpublished study, yearly indirect costs for remitted patients counted to PLN 6524 (US $1767). This was PLN 22,935 (US $6211) for patients with active disease. Loftus et al [25] indicate that "including indirect costs related to lost productivity due to hospitalization improved the cost-effectiveness of adalimumab therapy. However, the estimate of indirect costs was likely substantially underestimated because only work missed during hospitalization was included. Other indirect costs, such as decreased productivity at work and labor force nonparticipation, were not included." Assasi et al [18] also reported that if a societal perspective was taken and indirect costs were included in the model, the cost-effectiveness of anti-TNFs compared with that of usual care likely would have been lower. In addition, Yu et al [28] claim that reliable data sources to include the impact on indirect costs are lacking. Efforts should be taken to gather reliable information about the impact of different treatments on indirect costs. In the next part of this discussion, we discuss how measures for school attendance and parents' productivity are included in the research protocol of the REDUCE-RISK trial.

Added Elements in the REDUCE-RISK Trial Research Protocol

As recommended by the EUnetHTA guidelines on HRQoL [46], future studies should include a generic utility instrument in complement to disease-specific questionnaires to adequately capture the impact of a disease on daily life. Including a generic utility instrument in further research is also suggested as a research priority by the reviewers in the study by Archer et al [17] and Dretzke et al [20] and the underlying NICE report [16]. The disease-specific IMPACT-3 HRQoL measure was already included in the protocol. On the basis of the aforementioned data, a generic utility instrument (EQ-5D) is added to the research protocol. The ISPOR guidelines also recommend "prioritization of high-cost resources as well as those that are expected to differ between treatment arms, without distinction as to whether they are related to disease or intervention [47]. The scope of resources considered should include direct medical and nonmedical resources and indirect or productivity costs

across patients and caregivers." [4]. Similar to previous economic evaluations, differences in treatment costs and costs related to AEs will be taken into account. In addition, in the REDUCE-RISK trial, children's school attendance and parents' productivity will also be measured.

EQ-5D

The economic literature review identified the lack of QoL data that could be expressed as utilities and also indicated this as a research priority. Following the EUnetHTA guidelines on HRQoL [46], such information will be included in the REDUCE-RISK trial through the inclusion of the generic EQ-5D questionnaire.

In patients with CD and UC, a study by Stark et al [48] showed that both the EQ-VAS (EuroQol-visual analog scale) and EQ index scores correlate well with disease activity indices and differ significantly between active disease and remission groups. The authors concluded that the EQ-5D generates valid, reliable, and responsive preference-based evaluations of health in CD and UC. The EQ-VAS scores were more responsive than EQ-5D index scores, and thus, small health differences that are important from the patient's perspective may not be reflected in the EQ index [48]. This is in line with the results from a previous study from this research group that also concluded the EQ-5D to be "reasonably valid, reliable and responsive in patients with inflammatory bowel disease. It can be used to generate preference-based valuations of health-related quality of life in inflammatory bowel disease." [49].

From a practical point of view, the time for completion is less than 2 min for the EQ-5D [50]. From a financial point of view, the EQ-5D could be used free of charge for this study.

Three EQ-5D questionnaires are included: the EQ-5D-Y (youth version), EQ-5D-Y proxy1, and EQ-5D-5L.

1. The EQ-5D-Y was administered to measure the children's QoL. Following the user's guidelines, the youth version is used for all patients included in the REDUCE-RISK trial: "A study only with children up to 18 years, in this case EQ-5D-Y for older children would be recommended in order to have only one EQ-5D version in the study. The switch-over to the adult version could bring discontinuity as the adult and child versions are two different instruments." [51].
2. In the youngest children (<8 years), it is not possible to apply a self-completing questionnaire. We ask one of the parents to fill in the proxy version. The proxy rates how he or she rates the child's health. "The use of proxies, such as caregivers or family, should be avoided where possible. However, the use of proxies for the measurement of HRQoL is unavoidable in some cases, e.g. cognitively impaired patients, small children." [46]. By asking this for all patients, we will be able to evaluate the agreement between self- and proxy-reports of the EQ-5D-Y questionnaire.
3. The EQ-5D-5L was used to measure parents' QoL. It is important to find ways of incorporating relatives' costs and effects when these might be substantial and may influence the ICERs [52]. Parents' QoL of children with CD or UC might be such an example. This has not been included in any of the identified studies, and thus, the impact is unclear. Davidson et al [52] stated that the most relevant outcome measure to use for relatives' effects would be their affected utility. Therefore, we also included the measurement of parents' QoL through the use of the EQ-5D questionnaire. There is a choice between the 3L and 5L versions. The EQ-5D-5L version might be more sensitive to changes in health status in comparison with the 3L version [53,54]. Schwenkglenks et al [55] expect that the 5L version will gradually replace the 3L version, because of reduced ceiling effects and more appropriate responsiveness. Goldsmith et al [56] also referred to the increased ability to discriminate health states, which may improve the prediction of EQ-5D index values. Therefore, the EQ-5D-5L version was used.

The QoL measurements were made at baseline and all following planned study visits (months 2, 4, 6, 9, and 12). More information about the EQ-5D questionnaires, the included language version, the available value sets, and a sample version is available in section 3.1 of Multimedia Appendix 1.

School Attendance

A review of the economic literature indicates that indirect costs might represent a substantial portion of costs related to CD and UC but that the impact of different treatment options on such costs is lacking. As the patient population in the REDUCE-RISK trial is restricted to children and adolescents aged 6 to 17 years, indirect costs do not immediately relate to the patient's productivity. Instead, we try to measure the impact on a patient's school attendance.

Three studies were identified measuring the impact of CD or UC on school performance [57-59]. One study used a semistructured questionnaire for both children and parents and found that significant psychosocial and academic difficulties are faced by children with chronic diseases such as IBD [57]. Children with CD and UC missed significantly more school days than age-matched healthy controls [57]. Another study [58] created an online survey that included a Student Adaptation to College Questionnaire (SACQ). The results show that "disease activity in students with CD was associated strongly with their self-reported ability to keep up with academic work (P<.0089) and confidence in their ability to meet future academic challenges (P<.0015). Students with active IBD reported feeling as if they were not academically successful (P<.018), and students with ulcerative colitis reported irregular class attendance (P<.043)." [58]. The third study obtained report cards and school absence information from schools. Children with IBD had poorer school functioning and significantly more absences [59]. None of these studies used a structured questionnaire that was validated for use in children with CD or UC. The SACQ questionnaire is a 67-item, self-report questionnaire that is for college students and is mainly used at universities for routine freshman screening. This is considered inappropriate for our research.

An additional nonsystematic Google search was performed to identify other potentially relevant questionnaires. However, these questionnaires are very general. For example, the *School Attendance Questionnaire* mentions that these questionnaires are generally designed by school authorities to find out the

reasons for missing school. However, the questions posed clearly indicate that this questionnaire of school attendance is not well placed to apply in a population of sick children (eg, “Are your parents aware of this attendance percentage?” or “Are you aware that ...can lead to your suspension from school?”). Other researchers proposed a novel method for measuring class attendance by using location and Bluetooth data collected from smartphone sensors [60]. This is not applicable for the youngest children in our population because they usually do not have a smartphone.

No well-suited questionnaire was thus identified that can be used for the international REDUCE-RISK trial. Therefore, a de novo school attendance questionnaire was set up. Limitations of this questionnaire are that it is nonvalidated and that we cannot rely on the experience of other researchers with this questionnaire. Nevertheless, the choice was made to use this new instrument because we preferred to take the initiative to try to measure the impact with a nonvalidated instrument instead of not trying to measure this important aspect. The school attendance questionnaire consists of a version that is used at the first visit and a version to be used at the follow-up visits.

The parents filled in the questionnaire. First, we asked them to give a general picture of a typical school week to be able to have a view on the number of days the child goes to school in a typical week (exclusive home education) and the presence of home schooling (or home education). The aim of the questionnaire is to estimate the impact of IBD (CD and UC) and its treatment on school attendance and home education. The questions are related to the following: the presence and amount of home education; whether home education is because of IBD; the percentage of school days that children could not attend; in the case of home education, the percentage of home schooling days that children could not attend; and for both school days and home education, the part of absence that is because of IBD (in the opinion of the parents).

To assist participants with accurate recall, the ISPOR guidelines [4] recommend economic investigators to consider using memory aids such as diaries to record medical visits and events. Investigators should inform participants that they will be asked to report this information throughout the trial [61]. In line with this recommendation, the last page of the questionnaire, entitled *Information for parents to take home to help in collecting information for the next follow-up visit*, contains an overview of the questions.

Further details on the timing of the measurements and a sample version of these questionnaires are available in section 3.2.1 of Multimedia Appendix 1.

Parents' Productivity

To improve the quality and uniformity of data generated from trials, the ISPOR guidelines [4] recommend using validated instruments when incorporating productivity costs [62-64]. The Work Productivity and Activity Impairment (WPAI) questionnaire [65] is a self-administered questionnaire assessing the impact of a disease on a patient's ability to work and/or perform nonwork activities. A version exists specifically for CD (WPAI:CD), and a caregiver version (WPAI:CD-CG) exists in which the effect of a child's specific health problem on the parent's work productivity is measured. This WPAI:CD-CG questionnaire is included in the REDUCE-RISK trial. The included questions are related to the following: Q1: current employment, Q2: hours missed because of problems associated with the child’s CD, Q3: hours missed for other reasons, Q4: hours actually worked, Q5: degree child’s CD affected productivity while working, and Q6: degree child’s CD affected regular activities. The questionnaire is available at no charge in several languages. Details about the scores that will be calculated from these questions and the timing of measurements are available in section 3.2.2 of Multimedia Appendix 1.

Conclusions

This paper addresses an important and progressive issue: including health economic considerations in the design of clinical trials. At the end of the trial, when all information on the intervention’s efficacy and safety has been gathered, the important incremental variables will be combined in a trial-based economic evaluation calculating the intervention’s incremental costs, effects, and ICERs, both from a health care payer and societal perspective. Guidelines for performing economic evaluations will be followed. For example, parameter uncertainty will be included by performing a probabilistic sensitivity analysis. Following the ISPOR guidelines, reporting of the methods and results of the economic evaluation will be performed according to the CHEERS guidelines [11,12].

In conclusion, we are of the opinion that performing a systematic literature review supports researchers in setting up a research protocol. In our case, the results of the literature review helped us to identify important variables for which evidence should be gathered in the REDUCE-RISK trial to allow the performance of a high-quality economic evaluation.

Conflicts of Interest

LR received a consultancy fee or a grant from Pfizer, Abbvie, Celltrion, ZonMw in the previous 3 years. NC’s institution received speaker fees, advisory board fees, and research funding from Abbvie, Eli Lilly, Takeda, Shire, Pfizer, Jansenn, Allergan, 4D Pharma. SK reports grants from Mead Johnson, Nestec Nutrition, BioGaia, personal fees from Nestle, Danone, Biocodex, Shire, Abbvie, R-Biopharm, Vifor, Pharmacosmos, Celgene, ThermoFisher, Pfizer, Janssen, outside the submitted work. AL report to have IP or received travel grants, speakers’ honoraria, research grants from Nestle, Janssen, AbbVie, Megapharm and Takeda. DT received consultation fee, research grant, royalties, or honorarium from Janssen, Pfizer, Hospital for Sick Children, Ferring, Abbvie, Takeda, Atlantic Health, Shire, Celgene, Lilly, Roche, ThermoFisher, BMS during the last 3 years. FR received consultation fee, research grant, or honorarium from Janssen, Pfizer, Abbvie, Takeda, Celgene, Nestlé Health Science, Nestlé Nutrition Institute during the last 3 years.

Multimedia Appendix 1
Paediatric inflammatory bowel diseases network for safety, efficacy, treatment and quality improvement of care (PIBD-SETQUALITY): economic evaluation considerations.
[PDF File (Adobe PDF File), 1656 KB - formative_v5i1e13888_app1.pdf]

Multimedia Appendix 2
Included and excluded studies (and reasons for exclusion).
[DOCX File , 35 KB - formative_v5i1e13888_app2.docx]

Multimedia Appendix 3
PRISMA-2009-Checklist (literature review economic part REDUCE-RISK trial).
[DOC File , 194 KB - formative_v5i1e13888_app3.doc]

References

1. Harris RE, Aloi M, de Ridder L, Croft NM, Koletzko S, Levine A, PIBD SETQuality consortiumPIBDnet. Protocol for a multinational risk-stratified randomised controlled trial in paediatric Crohn's disease: methotrexate versus azathioprine or adalimumab for maintaining remission in patients at low or high risk for aggressive disease course. BMJ Open 2020 Jul 1;10(7):e034892 [FREE Full text] [doi: 10.1136/bmjopen-2019-034892] [Medline: 32611737]
2. Drummond M, O'Brien B, Stoddart G, Torrance G. Methods for the Economic Evaluation of Health Care Programmes. Oxford, UK: Oxford University Press; 1997.
3. EUnetHTA European Network for Health Technology Assessment. Methods for health economic evaluations - a guideline based on current practices in Europe. EUnetHTA 2015.
4. Ramsey SD, Willke RJ, Glick H, Reed SD, Augustovski F, Jonsson B, et al. Cost-effectiveness analysis alongside clinical trials II-An ISPOR Good Research Practices Task Force report. Value Health 2015 Mar;18(2):161-172 [FREE Full text] [doi: 10.1016/j.jval.2015.02.001] [Medline: 25773551]
5. Anderson R. Systematic reviews of economic evaluations: utility or futility? Health Econ 2010 Mar;19(3):350-364. [doi: 10.1002/hec.1486] [Medline: 19378354]
6. Lofland JH, Mallow P, Rizzo J. Cost-per-remission analysis of infliximab compared to adalimumab among adults with moderate-to-severe ulcerative colitis. J Med Econ 2013;16(4):461-467. [doi: 10.3111/13696998.2013.775134] [Medline: 23445401]
7. Toor K, Druyts E, Jansen JP, Thorlund K. Cost per remission and cost per response with infliximab, adalimumab, and golimumab for the treatment of moderately-to-severely active ulcerative colitis. J Med Econ 2015 Jun;18(6):437-446. [doi: 10.3111/13696998.2015.1012513] [Medline: 25629655]
8. Liu Y, Wu EQ, Bensimon AG, Fan CS, Bao Y, Ganguli A, et al. Cost per responder associated with biologic therapies for Crohn's disease, psoriasis, and rheumatoid arthritis. Adv Ther 2012 Jul;29(7):620-634. [doi: 10.1007/s12325-012-0035-7] [Medline: 22843208]
9. Yokomizo L, Limketkai B, Park KT. Cost-effectiveness of adalimumab, infliximab or vedolizumab as first-line biological therapy in moderate-to-severe ulcerative colitis. BMJ Open Gastroenterol 2016;3(1):e000093 [FREE Full text] [doi: 10.1136/bmjgast-2016-000093] [Medline: 27195130]
10. Saro C, Ceballos D, Muñoz F, De la Coba C, Aguilar MD, Lázaro P, et al. Resources utilization and costs the year before and after starting treatment with adalimumab in Crohn s disease patients. Inflammatory Bowel Diseases 2015;21(7):1631-1640. [doi: 10.1097/mib.0000000000000413]
11. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, CHEERS Task Force. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. BMC Med 2013 Mar 25;11:80 [FREE Full text] [doi: 10.1186/1741-7015-11-80] [Medline: 23531108]
12. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, ISPOR Health Economic Evaluation Publication Guidelines-CHEERS Good Reporting Practices Task Force. Consolidated Health Economic Evaluation Reporting Standards (CHEERS)--explanation and elaboration: a report of the ISPOR Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force. Value Health 2013;16(2):231-250 [FREE Full text] [doi: 10.1016/j.jval.2013.02.002] [Medline: 23538175]
13. Tappenden P, Ren S, Archer R, Harvey R, James MM, Basarir H, et al. A model-based economic evaluation of biologic and non-biologic options for the treatment of adults with moderately-to-severely active ulcerative colitis after the failure of conventional therapy. Pharmacoeconomics 2016 Oct;34(10):1023-1038. [doi: 10.1007/s40273-016-0409-9] [Medline: 27125898]
14. Blackhouse G, Assasi N, Xie F, Marshall J, Irvine EJ, Gaebel K, et al. Canadian cost-utility analysis of initiation and maintenance treatment with anti-TNF-α drugs for refractory Crohn's disease. J Crohns Colitis 2012 Feb;6(1):77-85. [doi: 10.1016/j.crohns.2011.07.007] [Medline: 22261531]

15. Essat M, Tappenden P, Ren S, Bessey A, Archer R, Wong R, et al. Vedolizumab for the treatment of adults with moderate-to-severe active ulcerative colitis: an evidence review group perspective of a nice single technology appraisal. Pharmacoeconomics 2016 Mar;34(3):245-257. [doi: 10.1007/s40273-015-0334-3] [Medline: 26477040]
16. National Institute for Health and Care Excellence (NICE). Infliximab and adalimumab for the treatment of Crohn's disease's (TA187). Technology appraisal guidance 2010. [doi: 10.1007/978-1-349-94186-5_836]
17. Archer R, Tappenden P, Ren S, Martyn-St James M, Harvey R, Basarir H, et al. Infliximab, adalimumab and golimumab for treating moderately to severely active ulcerative colitis after the failure of conventional therapy (including a review of TA140 and TA262): clinical effectiveness systematic review and economic model. Health Technol Assess 2016 May;20(39):1-326 [FREE Full text] [doi: 10.3310/hta20390] [Medline: 27220829]
18. Assasi N, Blackhouse G, Xie F, Gaebel K, Marshall J, Irvine E, et al. Anti-TNF-a drugs for refractory inflammatory bowel disease: clinical- and cost-effectiveness analyses. In: Technology report number 120. Ottawa: Canadian Agency for Drugs and Technologies in Health (CADTH); 2009:203.
19. Canadian Agency for Drugs and Technologies in Health (CADTH). golimumab (Simponi). Common Drug Review Pharmacoeconomic Review Report 2014:1-29.
20. Dretzke J, Edlin R, Round J, Connock M, Hulme C, Czeczot J, et al. A systematic review and economic evaluation of the use of tumour necrosis factor-alpha (TNF-α) inhibitors, adalimumab and infliximab, for Crohn's disease. Health Technol Assess 2011 Feb;15(6):1-244 [FREE Full text] [doi: 10.3310/hta15060] [Medline: 21291629]
21. Essat M, Tappenden P, Ren S, Bessey A, Archer R, Wong R, et al. Vedolizumab for the treatment of adults with moderately to severely active ulcerative colitis. In: A Single Technology Appraisal. Sheffield: School of Health and Related Research (ScHARR); 2014.
22. Rafia R, Scope A, Harnan S, Stevens J, Stevenson M, Sutton A, et al. Vedolizumab for the treatment of adults with moderately to severely active Crohn's disease. In: A Single Technology Appraisal. Sheffield: School of Health and Related Research (ScHARR); 2014.
23. Bodger K, Kikuchi T, Hughes D. Cost-effectiveness of biological therapy for Crohn's disease: Markov cohort analyses incorporating United Kingdom patient-level cost data. Aliment Pharmacol Ther 2009 Aug;30(3):265-274 [FREE Full text] [doi: 10.1111/j.1365-2036.2009.04033.x] [Medline: 19438428]
24. Kaplan G, Hur C, Korzenik J, Sands BE. Infliximab dose escalation vs. initiation of adalimumab for loss of response in Crohn's disease: a cost-effectiveness analysis. Aliment Pharmacol Ther 2007 Dec;26(11-12):1509-1520 [FREE Full text] [doi: 10.1111/j.1365-2036.2007.03548.x] [Medline: 17931345]
25. Loftus EV, Johnson SJ, Yu AP, Wu EQ, Chao J, Mulani PM. Cost-effectiveness of adalimumab for the maintenance of remission in patients with Crohn s disease. Eur J Gastroenterol Hepatology 2009;21(11):1302-1309. [doi: 10.1097/meg.0b013e32832a8d71]
26. Stawowczyk E, Kawalec P, Pilc A. Cost-utility analysis of 1-year treatment with adalimumab/standard care and standard care alone for ulcerative colitis in Poland. Eur J Clin Pharmacol 2016 Nov;72(11):1319-1325 [FREE Full text] [doi: 10.1007/s00228-016-2103-4] [Medline: 27497991]
27. Tang DH, Armstrong EP, Lee JK. Cost-utility analysis of biologic treatments for moderate-to-severe Crohn's disease. Pharmacotherapy 2012 Jun;32(6):515-526. [doi: 10.1002/j.1875-9114.2011.01053.x] [Medline: 22528603]
28. Yu AP, Johnson S, Wang S, Atanasov P, Tang J, Wu E, et al. Cost utility of adalimumab versus infliximab maintenance therapies in the United States for moderately to severely active Crohn's disease. Pharmacoeconomics 2009;27(7):609-621. [doi: 10.2165/11312710-000000000-00000] [Medline: 19663531]
29. Colombel J, Sandborn WJ, Rutgeerts P, Enns R, Hanauer SB, Panaccione R, et al. Adalimumab for maintenance of clinical response and remission in patients with Crohn's disease: the CHARM trial. Gastroenterology 2007 Jan;132(1):52-65. [doi: 10.1053/j.gastro.2006.11.041] [Medline: 17241859]
30. Buchanan J, Wordsworth S, Ahmad T, Perrin A, Vermeire S, Sans M, et al. Managing the long term care of inflammatory bowel disease patients: the cost to European health care providers. J Crohns Colitis 2011 Aug;5(4):301-316. [doi: 10.1016/j.crohns.2011.02.005] [Medline: 21683300]
31. Sandborn WJ, van Assche G, Reinisch W, Colombel J, D'Haens G, Wolf DC, et al. Adalimumab induces and maintains clinical remission in patients with moderate-to-severe ulcerative colitis. Gastroenterology 2012 Feb;142(2):257-65.e1. [doi: 10.1053/j.gastro.2011.10.032] [Medline: 22062358]
32. Woehl A, Hawthorne A, McEwan P. The relation between disease activity, quality of life and health utility in patients with ulcerative colitis. Gut 2008;57(Suppl. 1):A153.
33. Swinburn P, Elwick H, Bean K, Curry A, Patel S, Bodger K, et al. The Impact of Surgery on Health Related Quality of Life in Ulcerative Colitis. Digestive Disorders Federation Meeting. Liverpool: Gut Conference; 2012.
34. Gregor JC, McDonald JW, Klar N, Wall R, Atkinson K, Lamba B, et al. An evaluation of utility measurement in crohn s disease. Inflammatory Bowel Diseases 1997;3(4):265-276. [doi: 10.1097/00054725-199712000-00004]
35. Buxton MJ, Lacey LA, Feagan BG, Niecko T, Miller DW, Townsend RJ. Mapping from disease-specific measures to utility: an analysis of the relationships between the Inflammatory Bowel Disease Questionnaire and Crohn's Disease Activity Index in Crohn's disease and measures of utility. Value Health 2007;10(3):214-220 [FREE Full text] [doi: 10.1111/j.1524-4733.2007.00171.x] [Medline: 17532814]

36. Sandborn WJ, Colombel JF, Enns R, Feagan BG, Hanauer SB, Lawrance IC, et al. Natalizumab induction and maintenance therapy for crohn's disease. N Engl J Med 2005 Nov 3;353(18):1912-1925. [doi: 10.1056/nejmoa043335]
37. Targan SR, Hanauer SB, van Deventer SJ, Mayer L, Present DH, Braakman T, et al. A short-term study of chimeric monoclonal antibody ca2 to tumor necrosis factor α for Crohn's disease. N Engl J Med 1997 Oct 9;337(15):1029-1036. [doi: 10.1056/nejm199710093371502]
38. Sandborn W, Lukas M, Rutgeerts P, Pollack P, Kent J. Fully human anti-TNF adalimumab maintains remission for one year in patients with active Crohn's disease: a randomised, controlled cohort [abstract]. Gut 2006;55(Suppl 2):A27.
39. Hanauer SB, Feagan BG, Lichtenstein GR, Mayer LF, Schreiber S, Colombel JF, et al. Maintenance infliximab for Crohn's disease: the ACCENT I randomised trial. Lancet 2002 May;359(9317):1541-1549. [doi: 10.1016/s0140-6736(02)08512-4]
40. Sandborn WJ, Rutgeerts P, Enns RA, Hanauer SB, Colombel JF, Panaccione R, et al. Adalimumab rapidly induces clinical remission and response in patients with moderate to severe crohn s disease who had secondary failure to infliximab therapy. Am J Gastroenterol 2006;101:S448. [doi: 10.14309/00000434-200609001-01147]
41. Colombel J, Sandborn WJ, Rutgeerts P, Kamm MA, Yu AP, Wu EQ, et al. Comparison of two adalimumab treatment schedule strategies for moderate-to-severe Crohn's disease: results from the CHARM trial. Am J Gastroenterol 2009 May;104(5):1170-1179. [doi: 10.1038/ajg.2009.59] [Medline: 19352339]
42. Rutgeerts P, Feagan BG, Lichtenstein GR, Mayer LF, Schreiber S, Colombel JF, et al. Comparison of scheduled and episodic treatment strategies of infliximab in Crohn's disease. Gastroenterology 2004 Feb;126(2):402-413. [doi: 10.1053/j.gastro.2003.11.014] [Medline: 14762776]
43. Hanauer SB, Sandborn WJ, Rutgeerts P, Fedorak RN, Lukas M, MacIntosh D, et al. Human anti-tumor necrosis factor monoclonal antibody (adalimumab) in Crohn's disease: the CLASSIC-I trial. Gastroenterology 2006 Feb;130(2):323-33; quiz 591. [doi: 10.1053/j.gastro.2005.11.030] [Medline: 16472588]
44. EUnetHTA - European Network for Health Technology Assessment. Methods for health economic evaluations - A guideline based on current practices in Europe. EUnetHTA 2015.
45. Yu AP, Cabanilla LA, Wu EQ, Mulani PM, Chao J. The costs of Crohn's disease in the United States and other Western countries: a systematic review. Curr Med Res Opin 2008 Feb;24(2):319-328. [doi: 10.1185/030079908x260790] [Medline: 18067689]
46. EUnetHTA - European Network for Health Technology Assessment. Endpoints used for relative effectiveness assessment: health-related quality of life and utility measures. EUnetHTA 2015.
47. Glick H, Doshi J, Sonnad S, Polsky D. Economic Evaluation in Clinical Trials. New York, USA: Oxford University Press; 2015.
48. Stark RG, Reitmeir P, Leidl R, König HH. Validity, reliability, and responsiveness of the EQ-5D in inflammatory bowel disease in Germany. Inflamm Bowel Dis 2010 Jan;16(1):42-51. [doi: 10.1002/ibd.20989] [Medline: 19475674]
49. König HH, Ulshöfer A, Gregor M, von Tirpitz C, Reinshagen M, Adler G, et al. Validation of the EuroQol questionnaire in patients with inflammatory bowel disease. Eur J Gastroenterol Hepatol 2002 Nov;14(11):1205-1215. [doi: 10.1097/00042737-200211000-00008] [Medline: 12439115]
50. Shearer D, Morshed S. Common generic measures of health related quality of life in injured patients. Injury 2011 Mar;42(3):241-247. [doi: 10.1016/j.injury.2010.11.044] [Medline: 21163477]
51. van Reenen M, Janssen B, Oppe M, Kreimeier S, Greiner W. EQ-5D-Y User Guide: Basic information on how to use the EQ-5D-Y instrument (Version 1.0). EuroQoL 2014.
52. Davidson T, Levin L. Is the societal approach wide enough to include relatives? Appl Health Econ Health Policy 2012 Aug 6;8(1):25-35. [doi: 10.1007/bf03256163]
53. Janssen MF, Pickard AS, Golicki D, Gudex C, Niewada M, Scalone L, et al. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. Qual Life Res 2013 Sep;22(7):1717-1727 [FREE Full text] [doi: 10.1007/s11136-012-0322-4] [Medline: 23184421]
54. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011 Dec;20(10):1727-1736 [FREE Full text] [doi: 10.1007/s11136-011-9903-x] [Medline: 21479777]
55. Schwenkglenks M, Matter-Walstra K. Is the EQ-5D suitable for use in oncology? An overview of the literature and recent developments. Expert Rev Pharmacoecon Outcomes Res 2016;16(2):207-219. [doi: 10.1586/14737167.2016.1146594] [Medline: 26808097]
56. Goldsmith KA, Dyer MT, Buxton MJ, Sharples LD. Mapping of the EQ-5D index from clinical outcome measures and demographic variables in patients with coronary heart disease. Health Qual Life Outcomes 2010 Jun 4;8:54 [FREE Full text] [doi: 10.1186/1477-7525-8-54] [Medline: 20525323]
57. Assa A, Ish-Tov A, Rinawi F, Shamir R. School attendance in children with functional abdominal pain and inflammatory bowel diseases. J Pediatric Gastroenterol Nutr 2015;61(5):553-557. [doi: 10.1097/mpg.0000000000000850]
58. Almadani SB, Adler J, Browning J, Green EH, Helvie K, Rizk RS, et al. Effects of inflammatory bowel disease on students' adjustment to college. Clin Gastroenterol Hepatol 2014 Dec;12(12):2055-62.e1. [doi: 10.1016/j.cgh.2014.03.032] [Medline: 24726907]

59. Mackner LM, Bickmeier RM, Crandall WV. Academic achievement, attendance, and school-related quality of life in pediatric inflammatory bowel disease. J Develop Behav Pediatr 2012;33(2):1-111. [doi: 10.1097/dbp.0b013e318240cf68]
60. Kassarnig V, Bjerre-Nielsen A, Mones E, Lehmann S, Lassen DD. Class attendance, peer similarity, and academic performance in a large field study. PLoS One 2017;12(11):e0187078 [FREE Full text] [doi: 10.1371/journal.pone.0187078] [Medline: 29117190]
61. Goossens ME, Mölken MP, Vlaeyen JW, van der Linden SM. The cost diary. J Clin Epidemiol 2000 Jul;53(7):688-695. [doi: 10.1016/s0895-4356(99)00177-8]
62. Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. Pharmacoeconomics 1993 Nov;4(5):353-365. [doi: 10.2165/00019053-199304050-00006] [Medline: 10146874]
63. Lerner D, Amick 3rd B, Rogers W, Malspeis S, Bungay K, Cynn D. The work limitations questionnaire. Med Care 2001;39(1):72-85. [doi: 10.1037/t66378-000]
64. Prasad M, Wahlqvist P, Shikiar R, Shih YT. A review of self-report instruments measuring health-related work productivity: a patient-reported outcomes perspective. Pharmacoeconomics 2004;22(4):225-244. [doi: 10.2165/00019053-200422040-00002] [Medline: 14974873]
65. Work Productivity and Activity Impairment Questionnaire (WPAI). Reilly Associates. URL: http://www.reillyassociates.net/ [accessed 2020-12-28]

Abbreviations

6-MP: 6-mercaptopurine
ADA: adalimumab
AE: adverse event
AZA: azathioprine
CADTH: Canadian Agency for Drugs and Technologies in Health
CD: Crohn disease
CHEERS: Consolidated Health Economic Evaluation Reporting Standards
CRD: Centre for Reviews and Dissemination
EEN: exclusive enteral nutrition
EQ-VAS: EuroQol-visual analog scale
ERG: Evidence Review Group
HTA: health technology assessment
IBD: inflammatory bowel disease
ICER: incremental cost-effectiveness ratio
IFX: infliximab
ISPOR: International Society for Pharmacoeconomics and Outcomes Research
MTX: methotrexate
NICE: National Institute for Health and Care Excellence
QALY: quality-adjusted life years
QoL: quality of life
SC: standard care
UC: ulcerative colitis

Edited by G Eysenbach; submitted 30.06.19; peer-reviewed by J Banks, P Schnell-Inderst, Z Zrubka; comments to author 01.07.20; revised version received 25.08.20; accepted 11.11.20; published 25.01.21.

Please cite as:
Neyt M, Christiaens A, Aloi M, de Ridder L, Croft NM, Koletzko S, Levine A, Turner D, Russell RK, Ruemmele FM, Veereman G Identifying Health Economic Considerations to Include in the Research Protocol of a Randomized Controlled Trial (the REDUCE-RISK Trial): Systematic Literature Review and Assessment
JMIR Form Res 2021;5(1):e13888
URL: http://formative.jmir.org/2021/1/e13888/
doi:10.2196/13888
PMID:33492239

Corrigenda and Addenda

Correction: Factors Influencing Patients' Initial Decisions Regarding Telepsychiatry Participation During the COVID-19 Pandemic: Telephone-Based Survey

Jennifer Severe[1], MD; Ruiqi Tang[2], BA; Faith Horbatch[2], BSc; Regina Onishchenko[2], BSc; Vidisha Naini[2], BSc; Mary Carol Blazek[1], MD

[1]Department of Psychiatry, University of Michigan, Ann Arbor, MI, United States

[2]University of Michigan Medical School, Ann Arbor, MI, United States

Corresponding Author:
Jennifer Severe, MD
Department of Psychiatry
University of Michigan
4250 Plymouth Road
Rachel Upjohn Building, Floor 1, Door 1
Ann Arbor, MI, 48109
United States
Phone: 1 617 416 8275
Email: severej@med.umich.edu

Related Article:

Correction of: https://formative.jmir.org/2020/12/e25469/

(JMIR Form Res 2021;5(1):e27357) doi:10.2196/27357

In "Factors Influencing Patients' Initial Decisions Regarding Telepsychiatry Participation During the COVID-19 Pandemic: Telephone-Based Survey" (JMIR Formative Research 2020;4(12):e25469) the authors noted one error. In the originally published article, a value was incorrect in the Results section of the Abstract. The text read as follows:

> *Approximately half of the respondents (114/244, 46.7%) stated they were likely to continue with telepsychiatry even after in-person visits were made available.*

This text has been revised to:

> *Half of the respondents (132/244, 54.1%) stated they were likely to continue with telepsychiatry even after in-person visits were made available.*

The original Results section in its entierty is as follows:

> *A total of 244 patients whose original in-person appointments were scheduled within the first 3 weeks of the stay-at-home order in Michigan completed the telephone survey. The majority of the 244 respondents (n=202, 82.8%) initially chose to receive psychiatric care through video visits, while 13.5% (n=33) chose telephone visits and 1.2% (n=3) decided to postpone care until in-person visit availability. Patient age correlated with chosen visit type (P<.001; 95% CI 0.02-0.06). Patients aged ≥44 years were more likely than patients aged 0-44 years to opt for telephone visits (relative risk reduction [RRR] 1.2; 95% CI 1.06-1.35). Patient sex (P=.99), race (P=.06), type of insurance (P=.08), and number of previous visits to the clinic (P=.63) were not statistically relevant. Approximately half of the respondents (114/244, 46.7%) stated they were likely to continue with telepsychiatry even after in-person visits were made available. Telephone visit users were less likely than video visit users to anticipate future participation in telepsychiatry (RRR 1.08; 95% CI 0.97-1.2). Overall, virtual visits met or exceeded expectations for the majority of users.*

The revised Results section of the abstract in its entirety is as follows:

> *A total of 244 patients whose original in-person appointments were scheduled within the first 3 weeks of the stay-at-home order in Michigan completed the telephone survey. The majority of the 244 respondents (n=202, 82.8%) initially chose to receive psychiatric care through video visits, while 13.5% (n=33) chose telephone visits and 1.2% (n=3) decided to postpone care until in-person visit availability. Patient age correlated with chosen visit type (P<.001; 95% CI 0.02-0.06). Patients aged ≥44 years were more likely than patients aged 0-44 years to opt for telephone visits (relative risk reduction [RRR] 1.2; 95% CI 1.06-1.35). Patient sex (P=.99), race (P=.06), type*

of insurance (P=.08), and number of previous visits to the clinic (P=.63) were not statistically relevant. Half of the respondents (132/244, 54.1%) stated they were likely to continue with telepsychiatry even after in-person visits were made available. Telephone visit users were less likely than video visit users to anticipate future participation in telepsychiatry (RRR 1.08; 95% CI 0.97-1.2). Overall, virtual visits met or exceeded expectations for the majority of users.

The correction will appear in the online version of the paper on the JMIR Publications website on January 27, 2021, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories

Submitted 22.01.21; this is a non–peer-reviewed article;accepted 22.01.21; published 27.01.21.

Please cite as:
Severe J, Tang R, Horbatch F, Onishchenko R, Naini V, Blazek MC
Correction: Factors Influencing Patients' Initial Decisions Regarding Telepsychiatry Participation During the COVID-19 Pandemic: Telephone-Based Survey
JMIR Form Res 2021;5(1):e27357
URL: http://formative.jmir.org/2021/1/e27357/
doi:10.2196/27357
PMID:33502998

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org